# |CURIEUX|

---

# ACADEMIC REVIEW

## Economics & Business

2026 Issue Part 3

Editing Staff

**Chief-In-Editor**

George Hasnah

**Lead Editor**

Amanuel Geremew

Ashley Du

**Assisting Editors**

Jing Yan Huang

Nidhi Potineni

Sophia Kong

Tony Yuan

Ridit Aggarwal

Manu Sharma

Rishi Gadikota

Eric Wang

**Table Of Contents**

**Page 6: Technology Adoption by Small and Medium Businesses in Pakistan By Khazain Hassan Mansha**

**Page 23: Duty Without Liability: Bank Reporting Incentives and Elder Financial Exploitation By Nicholas Djoulia**

**Page 45: The Reliability Cost of Climate Volatility: A Comparison of EDES and IES Portfolios By David Park**

**Page 57: How Do Economic Systems Exploit Predictable Irrationality? A Game-Theoretic and Behavioral Analysis By Teddy Hsu**

**Page 70: Corporate Disclosure and Information Asymmetry in the Intangible Economy By Daniyal Kamruddin**

**Page 82: AI Exposure and Wage Inequality Amongst Skilled and Unskilled Workers: Evidence from Brazil By Henrique Gomes**

**Page 97: The Twenty-Five Dollar Rule: Emergence of Decision Heuristics in a Ten-Year Longitudinal Single-Subject Study of 644 Micro-Lending Choices (2016–2026) By Aaliyah Chaudhary**

**Page 110: Examining the Role of Socio-Economic Factors in Women Entrepreneurs' Access to Formal Business Credit By Anika Vivek**

**Page 123: Understanding Retirement Planning and Savings Behavior of Working People in Pakistan By Amaan Siddiqui**

**Page 132: A Systematic Review and Quantitative Synthesis of Trading Activity, Diversification, and Investment Horizon By Bhargava Janga**

**Page 144: The Effect of Bankruptcy-Law Reform on Entrepreneurship Varies According to National Culture By James Nichols**

**Page 156: Consumer Preferences Towards Emerging Digital Tools such as AR, VR & 3D in Real Estate Sales & Marketing By Manya Juvvadi [1], Dr. T P Singh[2]**

**Page 171: Knowledge Alone Is Not Enough: Behavioral and Motivational Predictors of Physical Activity Among Chinese Adolescents and Young Adults By Ziyue Ren**

**Page 182: Market opportunity analysis of creator-led sports broadcasting in Southeast Asia By Dia Dsilva**

**Page 195: To What Extent is It Feasible for Singapore to Adopt Green Policies used by Norway and China to Increase Consumption of EVs By Sin Eike**

**Page 213: What Drives Entrepreneurial Intention in Indonesia? The Role of Confidence Among Young Adults By Ahmad Nagara Latief**

**Page 226: Closing the Loop: The Role of Retailers in California's SB 554 By Siree Chawla**

**Page 238: Exploration of Central Bank Independence and its Ties to COVID-19 Related Fiscal Policy By Emma Goga**

**Page 246: The Impact of WTI Crude Oil Price Changes on Producer Prices Across U.S. Industries, 2006–2026 By Hridhaan Kataria**

**Page 267: Prior AI Access and Perceived Loss: An Exploratory Study of Loss Aversion and Psychological Ownership By Ahaana Bisarya**

**Page 279: Attack or Explain? Debate Rhetoric and Low-Dollar Giving in the 2020 Democratic Presidential Primary By Danial Yessentay**

**Page 289: Impact of Analytical AI on Equity Volatility By Aliyaan Noorani**

**Page 302: Does the Tax Code Push Workers Away from Independent Contracting? By Nihit Yeruva**

**Page 316: The Effects of AI in the Stock Market and How Machine Learning Models are Applied to Predict Future Prices By Daniel Hong**

**Page 328: Individual Differences and Financial Education: Examining the Role of Big Five Personality Traits and Financial Literacy among High School Students in Hong Kong By Anna Chandler**

**Page 341: Expanding Capabilities: The Role of AI in Global South Development By Kai Yui Lee**

**Page 356: Has the Bank of England Regained Credibility? By Vinoothna Bondalapati**

**Page 365: Examining Factors Affect Online Review Helpfulness By Elise Duka**

**Page 371: Why Small Food Businesses Remain: Community and Economic Endurance in the Mid Mon Valley By Owen Ryan**

**Page 391: Consequence Framing and Academic Dishonesty: A Behavioral Economics Approach By Evan Jin Rossmeisl**

**Page 407: The NBA's Anti-Tanking Reform: History, Current Problems, and the 3-2-1 Draft Lottery By MingCheng Yuan**

**Page 413: Canva's Critical Success Factors: The Study of Visual Design Tools and Platform Business Models By Weizi Liu**

**Page 428: Japanese Governmental Policy Efficacy and Intention in Addressing the Gender Wealth Gap By Ava Gilardi**

**Page 437: A Quality-Based Approach to Assessing Accessibility to Essential Destinations Across the San Francisco Metropolitan Area By Lucas Rhee**

**Page 454: AI-Powered Personal Finance in Kazakhstan and the United States: A Comparative Review of Digital Tools for Budgeting, Saving, and Financial Decision-Making By Damily Murat**

**Page 464: Trust in AI and Human Financial Advisors Richard Zhang**

**Page 474: The Role of Ephemeral Content in Snapchat's Growth and Success By Brandon Chen**

**Page 482: How Common Are Five Behavioural Biases in the Investment Decisions of Indian Respondents, and How Do They Vary by Age, Gender, and Self-Reported Confidence? By Karthik Paruchuri**

**Page 491: Economic Crisis, Political Response: Comparing the USSR's Decline and America's Great Recession Recovery By Fedor Klopotovskii**

**Page 496: Nike's Shift from Traditional Advertising to Digital Marketing: Impact on Consumer Engagement and Brand Loyalty By Vivaan Daruka (student)[1]***

**Page 510: Understanding Commonalities Across Six Financial Risk Episodes in the United States: A Behavioral Economics, Law and Economics, and Financial Media Approach By Katya Kong**

**Page 533: A Comparative Analysis of the Influence of Influencer Partnerships and Brand Authenticity on Purchase Intention Among Generation Z and Millennials By Zain Melwani**

**Page 545: Beyond Oil and Ore? Non-Resource Exports from Kazakhstan to China in the Trade Records of Both Countries, 2015-2025 By Inara Beibit**

**Page 556: Beyond the Fine: Measuring Corporate Risk Across Legal and Regulatory Events By Saksham Singh**

**Technology Adoption by Small and Medium Businesses in Pakistan**
**By Khazain Hassan Mansha**


## Abstract

The way payments, accounting, stock, marketing, and customer relations are managed in small and medium-sized enterprises (SME) is changing as a result of digitalization. However, the availability of digital technology does not necessarily translate into advanced use by SMEs. This study investigates the extent to which technology is used by small and medium-sized enterprises in Lahore, Pakistan, and its impact on sales, profits, costs, customer satisfaction, and daily operations. A cross-sectional survey was conducted with 54 owners and/or managers of businesses in Lahore, Pakistan. The respondents were asked questions regarding the availability and use of digital technology, including digital access, payment systems, administrative software, social media marketing, and online selling. Frequencies and percentages were calculated, and respondents ranked a list of items. Cross-tabulations were also conducted to investigate relationships between variables. The majority of respondents rated their digital knowledge as intermediate. Bank transfers were found to be the most widely accepted method of payment. The majority stated that digital payments had increased their sales. Second-generation or later businesses were found to be more active in adopting technology than first-generation businesses. Digitalization is essential for improving performance and competitiveness, yet the extent to which SMEs adopt it remains uneven and is affected by several cost, security, knowledge, and customer-related factors. Therefore, simple, affordable, and secure digital solutions, together with appropriate training, are required to encourage greater use of digital technology by SMEs.



## Keywords

Small and medium-sized enterprises; digitalization; digital payments; business performance; Pakistan


## Introduction

Across the globe, technology is rapidly changing the way that business and society operate. Businesses are increasingly relying on technology, including digital payments, POS systems, e-commerce platforms, social media marketing, and inventory management software. These tools allow businesses to improve customer experience, increase efficiency, and become more competitive and adaptable in rapidly changing markets (Soomro et al.)

Pakistan is becoming part of this digital transformation; this shift is important for retail businesses as customer expectations are changing swiftly. If businesses want to remain competitive, they have to adapt to customers' expectations. This includes customer preferences, which over the last few years have shifted from cash-only payments to online and card payments. Digital tools, such as e-commerce platforms and social media marketing, also enable businesses to reach customers more easily and offer more convenient shopping options (Soomro et al.) As a

result, businesses that adopt technology can manage sales more accurately, reach a wider audience, respond to customers faster, and improve their overall operations (Soomro et al.). Despite the benefits that technology can bring to businesses, many still rely on traditional business practices, such as cash transactions, manual inventory systems, and walk-in customers. This can make it harder for them to compete with businesses that use modern digital tools. In Pakistan, Information and Communications Technology (ICT) adoption among SMEs is strongly influenced by the owner's or manager's attitude, knowledge, and willingness to change (Nazir et al.) This means that some businesses may not fall behind because technology is unavailable, but because they continue to use older systems or are hesitant to adopt new digital practices.
While there is some research on technology use among SMEs in Pakistan, few ground-level studies focus specifically on small- and medium-sized retail businesses. Previous research has examined ICT adoption among Pakistani SMEs through interviews with business owners and managers, but it did not focus solely on small or medium retail businesses in Lahore (Nazir et al.). Retail businesses face their own challenges when adopting technology, including competition from online sellers, changing customer behavior, inventory management challenges, software costs, employee training, and trust in digital payment systems. Therefore, more focused research is important to understand how these businesses can overcome these issues.

The present study aims to understand the adoption of technology by small and medium retail businesses in Pakistan. It will explore whether digital tools actually help businesses grow, improve performance, and create more value, given that previous research has shown that digital technologies can enhance SME performance and value creation (Soomro et al.) However, it will also examine whether barriers such as cost, lack of knowledge, and resistance to change stop many retailers from benefiting from these tools. This is important because technology adoption is no longer just an optional extra for businesses but is becoming necessary for survival in a more digital economy.

## Methods

### Tools used

This study investigates the use of technology by small- to medium-sized retail shops in Lahore, Pakistan. For this purpose, a quantitative cross-sectional survey study was the most appropriate. A cross-sectional survey study is used to collect single-point-in-time data from the largest number of participants within the shortest period. In this study, a structured survey was used as the primary instrument to gather data from owners/managers of small to medium-sized retail shops.

The survey consists of several sections; section 1 contains information on the respondents' demographic backgrounds. This information includes the age, sex, preferred language, place of residence, and level of education of each respondent. This information enables analysis of the results for different demographic groups of businesses. Section 2 of the survey collected data on the businesses included in the study. In the first question, the respondents were

asked to select the type of retail business that they ran. Next, several questions were asked about the business the respondent ran, including how long it had been operating. Following on from this, several questions were asked about the generation of the business that the respondent was part of. In addition, questions were asked about the business's annual income and about its main customers.

For each technology listed, businesses were asked whether they currently used it and, if so, how. The next question asked businesses to list and rate the benefits they received from each technology they used. The final question asked the businesses to list and rate the barriers that prevented the use of technology or the use of additional technology.

**Sampling Method and Sample Characteristics**

The researcher has used convenience sampling and targeted sampling to collect data for the above study. This is because the views of owners or managers of small and medium-sized retail businesses were required for the above study. Thus, the above sampling is of the targeted and convenience types.

A total of 54 individuals from various businesses completed the survey. In terms of gender, 50 respondents (92.6%) were male, while 4 (7.4%) were female. The age of the respondents ranged from 25 to 34 years, with 14 respondents (25.9%) falling in this category. The largest number of respondents spoke Urdu as their first language, with 46 (85.2%). The largest number of respondents resided in urban areas, with 44 (81.5%) completing the survey. In terms of graduate studies, 30 respondents (55.6%) held a Master’s degree or higher. This was followed by 24 respondents (44.4%) who held a Bachelor’s degree. The largest number of organizations and industries that completed the survey were service-based businesses and retail stores, with 22 (40.7%) and 18 (33.3%), respectively. Regarding the number of years a business has been in operation, 35 businesses (64.8%) completed the survey and had been in operation for more than 10 years. In terms of ownership, 26 of 54 businesses (48.1%) were first-generation. This was followed by 12 or 22.2% second-generation businesses, 13 or 24.1% third-generation businesses, and 3 or 5.6% fourth-generation or later.

**Ethical Considerations**

All participants in the study have been informed of the nature of the study. In addition, participants have been made aware that participation is completely voluntary; therefore, no individual has been forced to take part. During the research process, the researcher was aware of the ethical implications. During the distribution of the survey to the owners and managers of small- and medium-sized retail businesses in Lahore, the researcher ensured it contained no personal information that could cause harm or discomfort to respondents. Information regarding the respondents' business income was collected in ranges to avoid obtaining sensitive information. The researcher also ensured that the data collected from the survey would be used only for this research and that the findings would be presented in aggregate form.

**Data Collection Procedure**

To gather primary data, a structured survey was prepared, with most questions being multiple-choice. The questions were distributed in two sections. Section 1 contained questions for gathering background information about the respondents. Section 2 consisted of questions for gathering information regarding the businesses (of the respondents). Questions related to technology adoption included, for example, digital payment systems, POS systems, online shop systems, social media marketing, and stock management software. The collected data were checked for completeness and relevance. In the end, 54 complete and relevant questionnaires were collected. Each completed questionnaire included a single business.

**Data Analysis Strategy**

For single variables (business type, ownership generation, customer groups, income groups, and usage of single technologies), the respective frequencies and corresponding percentages have been calculated and presented in tables, diagrams, and charts. Benefits and barriers to technology adoption for individual technologies were ranked by frequency. Where appropriate, respondents were asked to rate the importance of factors to them on a scale; the average score for each factor was calculated to compare its importance relative to the others. Cross-tabulations were also conducted between technology adoption and business characteristics of the participating businesses. This has enabled an in-depth investigation into technology adoption and barriers to further adoption among small and medium-sized retail businesses, as well as intergroup comparisons. However, given that the study used a non-probability sample of 54 small- and medium-sized retail businesses in Lahore, the results are limited to those businesses that participated in the research.

**Results**

This section presents the main findings collected from the 54 businesses that completed the survey. The results are organized into different areas to explain businesses' access to technology, use of digital tools and payments, customer activity, and the effects of digitalization on their operations.

***Digital access in businesses***

This section examines whether business owners have access to essential digital tools, including smartphones, WiFi, and mobile data. It also considers their level of digital knowledge and ability to use technology confidently for business purposes.

The survey results show that most business owners have strong access to digital tools. Of 54 respondents, all said they use a smartphone. Internet access was high: 46 respondents reported using both WiFi and mobile data, 5 used only WiFi, and 3 used only mobile data. While the responses regarding digital access were more straightforward, digital knowledge itself was not. Most people rated their digital knowledge as intermediate, with 34 choosing intermediate, 17 choosing advanced, and only 3 choosing beginner. This shows that while most businesses have

access to smartphones and the internet, not all business owners have the same level of confidence or skill when using digital tools for business purposes.

***Customers served daily***

This section uses the number of customers served each day to show the different sizes and activity levels of the businesses surveyed. It also examines recent business growth and the main economic and operational challenges respondents face.

In the survey, businesses vary in size. Of the 54 businesses, 24 serve more than 100 customers per day, while 16 serve fewer than 10 customers per day, indicating that the survey responses included both large and small businesses. When we look at growth, 42 respondents said their business had increased over the past three years, while 6 said it stayed the same and 6 said it decreased. This suggests that most businesses experienced growth despite the challenges they faced. The biggest challenge reported was high costs, which 25 businesses selected. Other major challenges included competition (selected by 11 respondents) and a skilled labor shortage (selected by 10 respondents). This shows that while technology is important, businesses are also under pressure from wider economic and operational problems.

***Digitalization of payments***

This section examines how businesses have adopted digital payment methods instead of relying entirely on cash. It considers the proportion and number of digital transactions, the platforms used, reasons for using them, barriers to adoption, and their effect on sales.

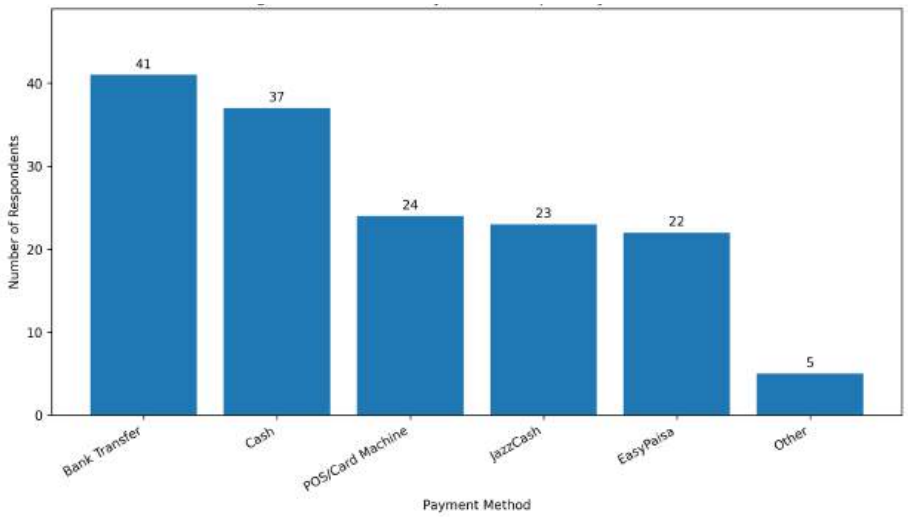



**Figure 1:** Different modes of payment used by businesses

We can see in Figure 1 that bank transfer was the most commonly accepted payment method among the businesses that took part in the survey, with 41 businesses accepting it. Cash was second, with 37 respondents accepting cash payments. This shows that even as digital payments become more common, traditional cash transactions remain important in Pakistani

businesses. About 24 respondents accepted POS/card machines; by comparison, 23 and 22 respondents accepted JazzCash and EasyPaisa, respectively, and 5 accepted other payment methods, indicating that businesses use a mix of payment methods.

For the percentage of transactions that are digital, the most common response was 41–60%, selected by 15 businesses, followed by 61–80% (13 businesses) and 21–40% (11 businesses). About 10 businesses said that 81–100% of their transactions were digital, while 5 businesses said that 0–20% were digital. This shows that most businesses have adopted digital payments to some extent, but the majority are not yet fully digital.

When looking at the number of digital transactions handled daily, there is a wide gap, with the largest group of 18 businesses reporting 0- 10 digital transactions. Another 14 respondents reported 11–20 digital transactions per day, while only 3 handled 21–50 and 2 handled 51–100. This suggests that digital payment usage is uneven across businesses.

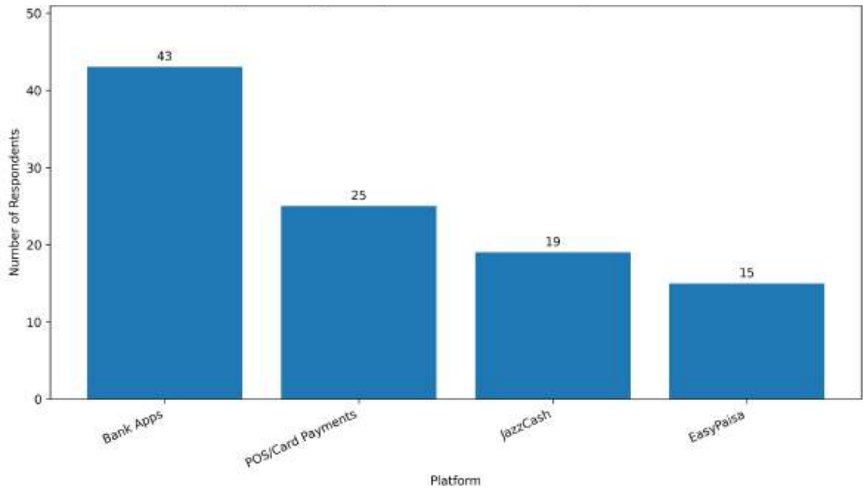



**Figure 2:** Digital platforms used by businesses

In Figure 2, we can see that bank apps were the most commonly used digital payment platform among the businesses surveyed, with 43 businesses using them, indicating that bank-based digital payments are more popular than other platforms. POS/card payments were the second-most-used option, with 25 businesses using them. JazzCash was used by 19 businesses, while EasyPaisa was used by 15 businesses. This suggests that although mobile wallet platforms are used by some businesses, bank apps and card-based payments are more common.

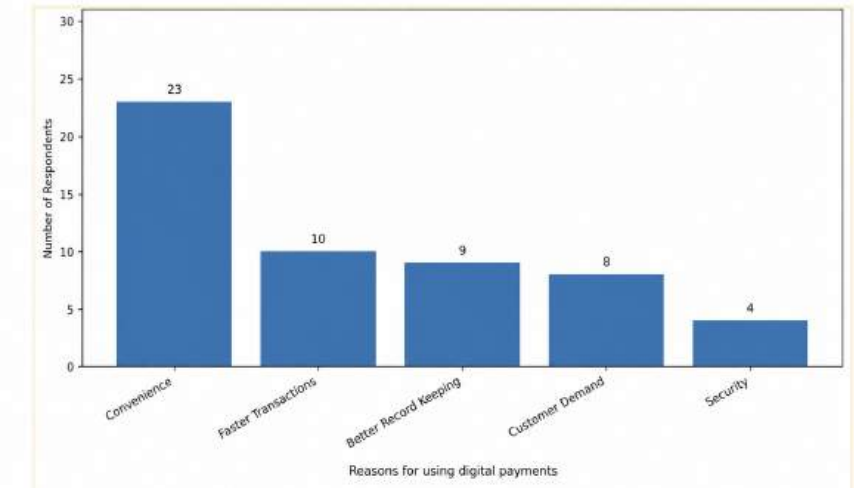



**Figure 3:** Reasons for using digital payments

In Figure 3, the most common reason businesses use digital payments is convenience, selected by 23 businesses. The second most common reason, selected by 10 respondents, was better record-keeping. Customer demand was also an important reason, with 8 respondents choosing it. Only 4 respondents selected security as their main reason for using digital payments. Overall, the use of technology is associated with businesses seeking more efficient and convenient methods of doing things.

The main factor limiting the use of digital payments was customers' preference for cash. This shows that even when businesses offer digital payment options, many customers still feel more comfortable using cash. Other important barriers included transaction fees, lack of internet access, security concerns, and lack of knowledge. This shows that even if businesses want to adopt technology, they are sometimes limited by their customers.

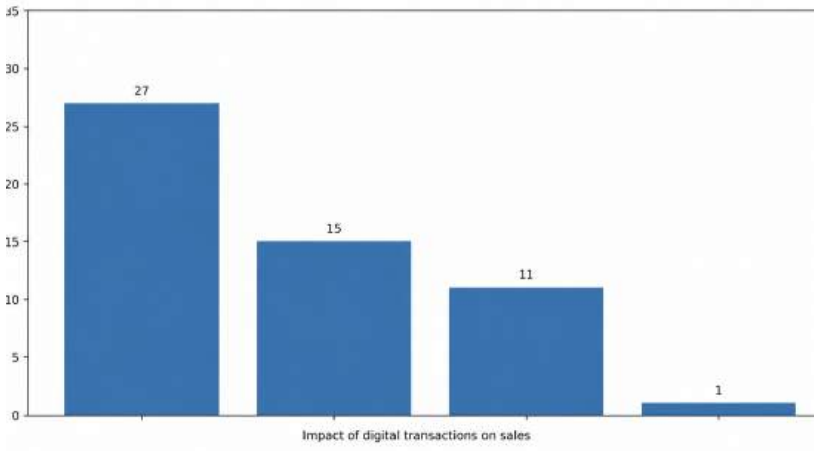



**Figure 4:** Impact of digital transactions on sales

Figure 4 shows that digital payments have had a mostly positive effect on business sales. Of the businesses, 27 said digital payments significantly increased their sales, while 15 said sales increased slightly. This means that most businesses experienced some improvement in sales after using digital payment methods. Only 11 respondents reported no change, and just 1 business reported that digital payments decreased sales.

***Digitalization of admin work***

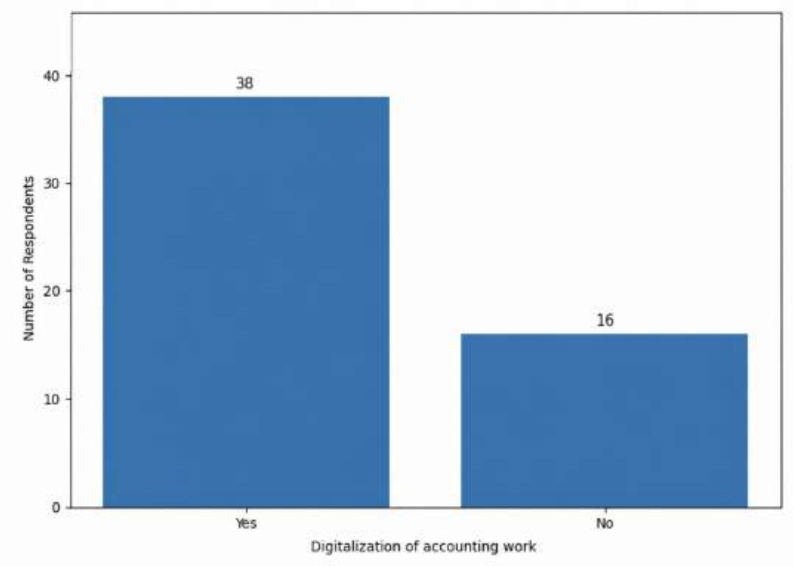



**Figure 5:** Digitalization of accounting work

In Figure 5, most surveyed businesses said they use accounting software. Out of the respondents, 38 said yes, while 16 said no. Many businesses have adopted digital tools to manage their financial records, but the 16 that responded "no" indicate some still rely on traditional or manual methods. So while software is common and continues to grow, the digitalization of admin work is still not complete across all businesses.

We also asked about the digital platforms they use to manage inventories. It was found that most businesses manage inventory using digital methods. A total of 30 respondents said they use software to manage inventory, while 20 respondents use Excel. Only 4 respondents said they manage inventory manually. This suggests that many businesses are moving away from fully manual inventory systems. However, the extensive use of Excel suggests that businesses are relying on basic digital tools rather than advanced inventory software.

Lastly, the amount of time spent on administrative work varied among the businesses when asked. The most common response was 1–2 hours daily, selected by 19 respondents. Another 16 respondents said they spend 3–4 hours on administrative work each day, while 12 respondents spend more than 4 hours. Only 7 respondents reported spending less than 1 hour daily on admin work, indicating that it takes up a significant amount of time.

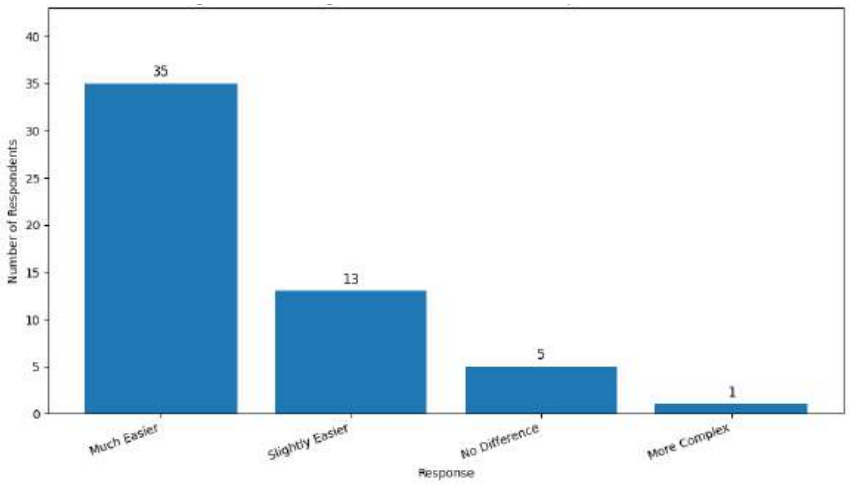



**Figure 6:** Impact of digitalization on business operations

Figure 6 shows that most respondents believe digitalization has made their business operations easier. Of the total businesses, 35 said digitalization made operations much easier, while 13 said it made operations slightly easier. Only 5 respondents said it made no difference, and just 1 respondent said it made operations more complex. This shows that digital tools are generally helping businesses.

***Technology adoption in businesses by different generations***

The next set of questions concerned the survey: 8 of 26 first-generation business owners said yes, while 8 said no when asked whether they had adopted technology in their business. In comparison, 26 out of 28 second-generation or later business owners said yes, while only 2 said no. This shows that second-generation and later business owners are more likely to use social media for marketing than first-generation business owners. This may be because younger businesses and business owners have more experience, resources, or a greater willingness to adopt digital tools, driven by the rapid global and Pakistani growth of technology.

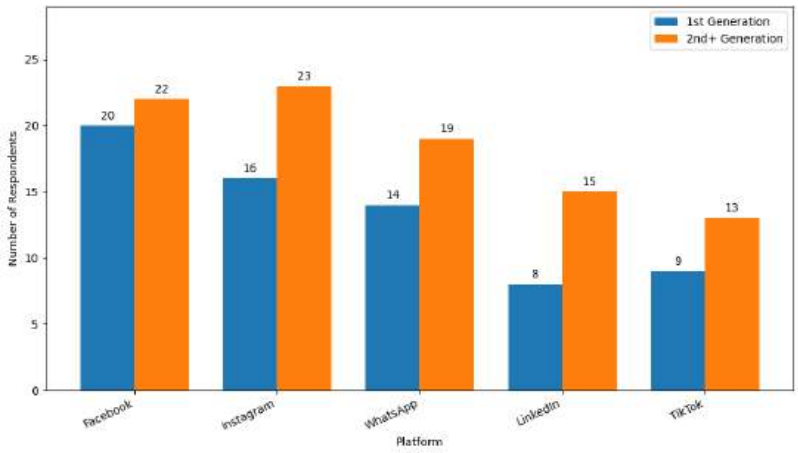



**Figure 7:** Social media platforms used by different generations

In Figure 7, we can see that Facebook was commonly used by both groups, with 20 first-generation and 22 second-generation-or-later business owners using it. However, second-generation or later businesses used most platforms more often than first-generation businesses. Instagram was used by 23 second-generation-or-later businesses, compared with 16 first-generation businesses. WhatsApp was used by 19 second-generation-or-later businesses, compared to 14 first-generation businesses. LinkedIn and TikTok also showed higher use among second-generation or later businesses, suggesting that technology adoption differed across generations, and that there was also a gap between generations on different platforms.

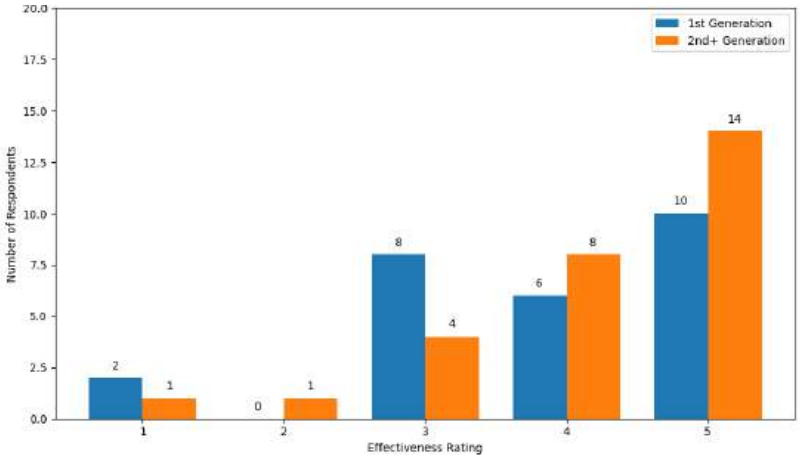



**Figure 8:** Perception of effectiveness of digital marketing by different generations

Figure 8 shows that both groups generally viewed digital marketing positively, but second-generation or later business owners gave higher ratings overall. Among first-generation business owners, 10 respondents rated digital marketing 5, 6 rated it 4, and 8 rated it 3. Among second-generation or later business owners, 14 respondents rated digital marketing as 5 and 8 rated it as 4. This shows that second-generation or later businesses may see digital marketing as more effective. Regarding whether businesses have ever run paid online advertisements, 14 of 26 first-generation business owners said yes, while 12 said no. In comparison, 22 out of 28 second-generation or later business owners said yes, while only 6 said no. This shows that second-generation or later businesses are more likely to invest money in paid online advertising.

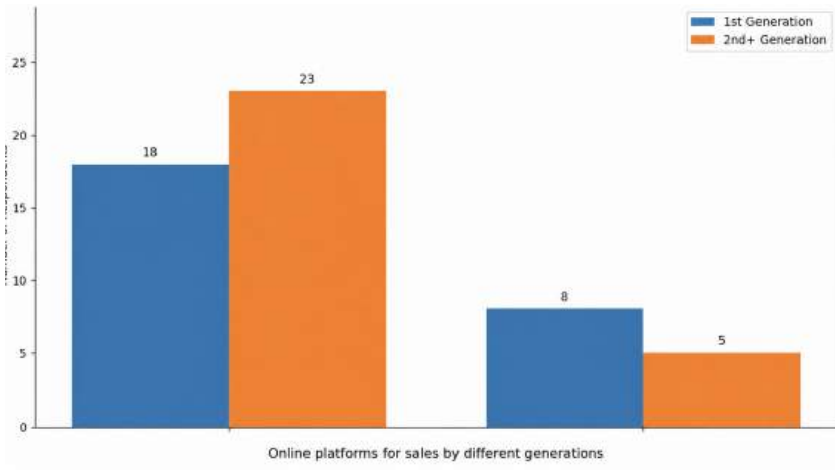



**Figure 9:** online platforms for sales by different generations

As shown in Figure 9, 18 first-generation business owners reported selling products or services online, while 8 reported not doing so. Among second-generation or later businesses, 23 said they sell online, while only 5 said they do not sell online. This shows that online selling is common in both groups. Still, it is more common among second-generation or later businesses, which aligns with the other responses regarding technology adoption in business.

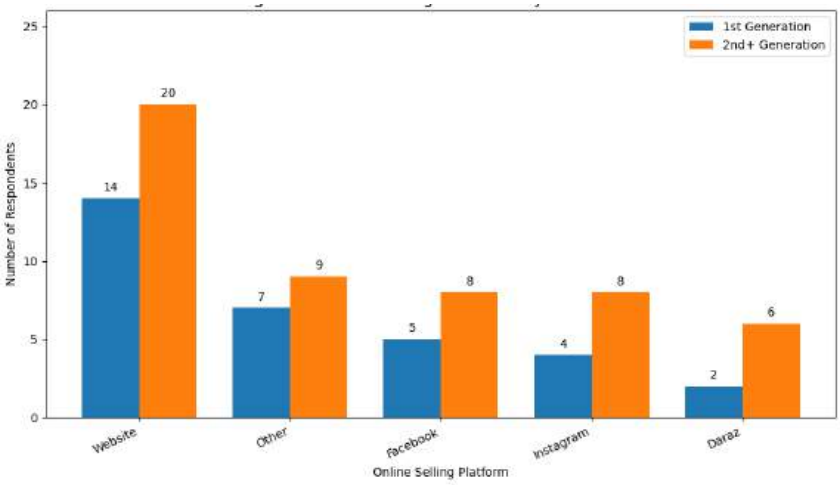



**Figure 10:** Online platforms used for sales

Figure 10 shows that websites were the most common online selling platform for both groups, with 14 first-generation and 20 second-generation-or-later businesses using them. Other platforms were used by 7 first-generation and 9 second-generation or later businesses. Facebook, Instagram, and Daraz were used more frequently by second-generation or later businesses, with Daraz showing a notable difference: 6 second-generation or later businesses used it, compared to only 2 first-generation businesses.

***Impact of digitalization on overall business***

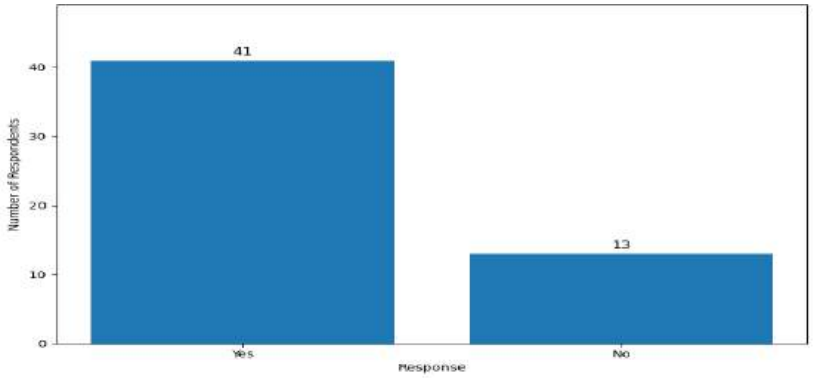



**Figure 11:** Impact of digitalization on overall business

Figure 11 shows that most businesses believe digitalization has improved their profits. Of the total respondents, 41 said yes and 13 said no. This shows that digital tools have had a positive financial impact on many businesses. Digitalization may improve profits by helping businesses reach more customers, reduce errors, manage records more effectively, and streamline transactions; however, the 13 respondents who said no show that digitalization does not increase profits for every business.

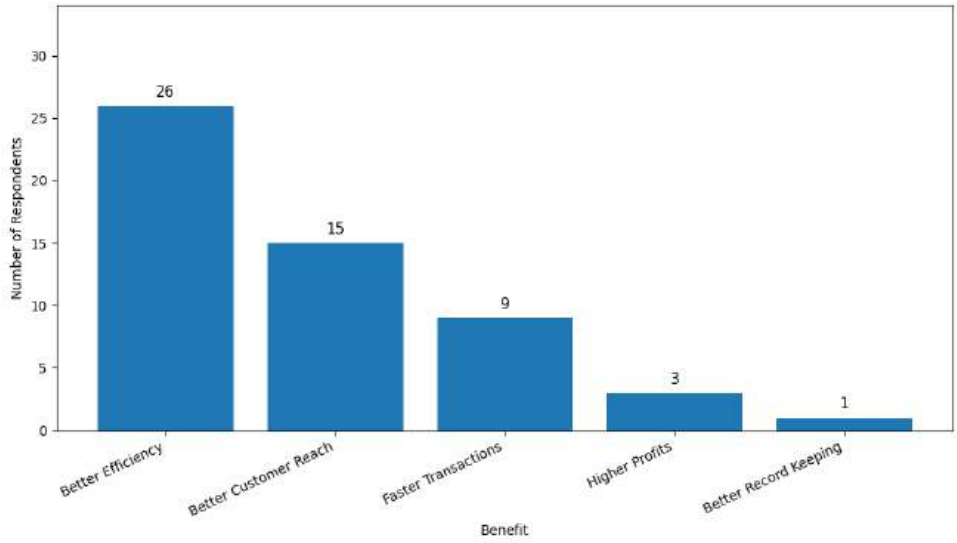



**Figure 12:** Benefits of digitalization on businesses

Figure 12 shows that the biggest benefit of digitalization was improved efficiency, chosen by 26 businesses, indicating that many businesses use digital tools to save time, improve organization, and streamline daily operations. Better customer reach was the second most common benefit, selected by 15 respondents, highlighting how digitalization helps businesses connect with more customers through online platforms and digital

marketing. 9 respondents noted faster transactions, while higher profits and better record-keeping were selected by 3 and 1 businesses, respectively.

Many businesses also shared that digitalization has reduced costs, but not for all. Of the total respondents, 41 said digitalization had reduced their costs. In contrast, 13 said it had not, depicting how digital tools can help businesses reduce expenses by saving time, reducing manual work, improving record-keeping, and streamlining operations. Moreover, a strong positive effect was found between digitalization and customer satisfaction. Of the respondents, 48 said digitalization improved customer satisfaction, while only 6 said it did not. This suggests that customers benefit from digital tools such as online ordering, faster communication, digital payments, and social media updates.

However, the biggest drawbacks of digitalization reported were security risks and cost. While businesses see the benefits of digital tools, they also worry about online fraud, data safety, software and equipment costs, and transaction fees. Other important drawbacks included technical difficulties, dependence on the internet, and the need for training. These responses suggest that digitalization can enhance business operations but also introduces new challenges that require careful management.

**Discussion**

In the present study, we found that access to technology among Lahore's small and medium retail businesses is high, but the depth of adoption remains uneven. All respondents had smartphones and internet access, most rated their digital knowledge as intermediate, and some businesses still relied on cash or basic tools such as Excel. This difference between access and actual use is supported by research involving 1,192 merchants in Punjab, which found that approximately 90% had smartphone access but cash remained their preferred payment method. This was due to fear of fraud, transaction costs, unreliable connectivity, and incorrect beliefs about customer preferences (Said and Perlik) This also helps explain why customer preference for cash was the largest payment barrier in the present study and why digital transactions varied greatly. Bank transfers and banking applications were more common because they are familiar and can be used without purchasing a separate POS system, and because of a somewhat established trust between banks and businesses, unlike with new technologies. (Matin et al.) The results of our survey also showed that digitalization often leads to stronger business performance. Businesses in our survey reported that technology adoption had increased profits and reduced costs; however, some reported improved customer satisfaction, and a few stated that business operations had become easier. These results are consistent with findings of (Soomro et al.), whose study of 305 Pakistani SME managers found that digital tools led to the creation of economic and social value, which was positively related to SME performance.

It was also found that digital payments increased sales by offering customers more choices, speeding up transactions, and preventing businesses from losing customers who do not carry cash. Accounting and inventory systems could reduce costs by limiting errors, improving stock control, and reducing repetitive and unnecessary work. Online selling and social media

also allow firms to reach customers beyond their physical stores at relatively low cost. (Qalati et al) similarly found a positive relationship between social media adoption and SME performance among Pakistani SME owners and managers.

Another finding was that second-generation or later businesses used social media, paid advertising, and online selling more than first-generation businesses. This shows greater access to technology, stronger resources, and a greater willingness to change established practices. (Nazir and Khan) found that the individual beliefs of SME owners and managers were the most important determinant of Information and Communications Technology (ICT) adoption in Pakistan. This also helps us explain why businesses with similar access to smartphones and the internet can still display different levels of digitalization.

***Implications and limitations***

The findings from this report could benefit several groups, such as retail owners, who can use them to prioritize practical technologies such as digital payments, accounting software, inventory systems, and social media marketing. Banks, fintech companies, and software providers can develop cheaper, simpler, and more secure products for smaller retailers and also help with the adoption of technology in many other areas. Moreover, understanding the experiences and views of business owners also provides a clear picture of the real challenges faced by retailers. Field-based research is useful because interviews with owners and managers can reveal how personal attitudes, business culture, and management decisions affect Information and Communications Technology (ICT) adoption (Nazir et al.). It can also help suggest practical ways for businesses, policymakers, and others to adopt technology in Pakistan. This connects to the concept of disruptive innovation, as digital tools such as social media, e-commerce, and digital payments can gradually reshape traditional retail practices and create new ways for businesses to compete (Qalati et al.).

While the paper's results are positive, they should be interpreted cautiously. The sample contained only 54 businesses from Lahore and may not represent retailers in rural areas, other cities, or every retail sector. Businesses were selected because they were relevant and accessible rather than through a clearly defined random process. On top of this, the data were self-reported, and this paper did not verify responses using financial records or control for business size, age, income, and sector. Despite the positive results, there are limitations. A total of 54 business owners and managers from Lahore were approached to complete the survey. Thus, the results of the study cannot be generalized to all retail businesses in Pakistan. However, all 54 respondents completed the survey, so the results were not affected by incomplete responses. The survey questions were designed to capture the perceptions of retail business owners and managers regarding various digital technologies and their impact on sales, profits, costs, and customer satisfaction. Thus, the results of the study are based on self-reported data, and there is a possibility of discrepancies between the actual changes in profits, costs, and customer satisfaction and the perceptions of the business owners and managers. A cross-sectional study was conducted to investigate the impact of digitalization on sales, profits, costs, and customer

satisfaction of retail businesses. Thus, the results of the study do not establish cause-and-effect relationships between various digital technologies and the sales, profits, costs, and customer satisfaction of retail businesses.

**Conclusion**

This study investigates the extent of adoption of digital technologies by small and medium-sized enterprises in the retail sector in Lahore. It examines the impact of this adoption on performance. The majority reported that customer preference for cash remained a barrier to adopting digital payments. A key finding of this research is the generational difference between first-generation businesses and those operated by second-generation or later owners. Second- and later-generation businesses were far more active in social media marketing, paid online advertising, and e-commerce than their first-generation counterparts. Thus, digital technologies offer a variety of benefits to retailers. However, the respondents in this study are also aware of several concerns, including security and costs. Thus, technology adoption creates opportunities as well as challenges that retailers must address.

There are many ways that future research can benefit from the current body of work on small and medium-sized enterprises (SMEs) in Pakistan. First, the sample could be drawn from across Pakistan, not just from cities. It would also be useful to draw a sample of retailers of different sizes and to compare them within particular sectors, such as food and non-food retail. A longitudinal study could also be used, in which the SMEs under investigation are examined before and after a period during which the business owner has introduced new digital tools. Such a study would derive greater benefit from the current body of work and provide evidence of the impact of digitalization on retail businesses' sales, profits, and costs. Qualitative research could also be used to gain greater insight into businesses' concerns about the security of online transactions, the costs of digital solutions, and the need for employee training, as well as to establish in greater detail why customers of particular businesses prefer to pay in cash.

**Works Cited**


Hassan, Masood Ul, et al. "Factors Affecting the Adoption of Internet Banking in Pakistan: An Integration of Technology Acceptance Model and Theory of Planned Behavior." International Journal of Business Information Systems, vol. 28, no. 3, 2018, pp. 342–370. https://doi.org/10.1504/IJBIS.2018.092530.

Matin, Raza, et al. Realizing the Promise of Responsible Digital Payments for Merchants in Pakistan. Better Than Cash Alliance, 2023, https://www.betterthancash.org/explore-resources/realizing-the-promise-of-responsible-digital-payments-for-merchants-in-pakistan.

Mubarak, M. F., et al. "The Impact of Digital Transformation on Business Performance: A Study of Pakistani SMEs." Engineering, Technology & Applied Science Research, vol. 9, no. 6, 2019, pp. 5056–5061. https://doi.org/10.48084/etasr.3201.

Mushtaq, Naila, et al. "Digital Transformation and Its Impact on Business Performance in SMEs of Pakistan: An Empirical Study." The Asian Bulletin of Big Data Management, vol. 3, no. 2, 2023, pp. 103–114. https://doi.org/10.62019/abbdm.v3i2.92.

Nazir, Muhammad Arsalan, and Mohsin Raza Khan. "Identification of Roles and Factors Influencing the Adoption of ICTs in the SMEs of Pakistan by Using an Extended Technology Acceptance Model (TAM)." Innovation and Development, vol. 14, no. 1, 2024, pp. 189–215. https://doi.org/10.1080/2157930X.2022.2116785.

Nazir, Muhammad Arsalan, et al. "Challenges of E-Commerce Adoption Experienced by Pakistani SMEs: A Qualitative Analysis." International Review of Entrepreneurship, vol. 20, no. 3, 2022, pp. 447–476. http://hdl.handle.net/2262/102680.

Nazir, Muhammad Arsalan, et al. "Identifying Prosperity Characteristics in Small and Medium-Sized Enterprises of Pakistan: Firm, Strategy and Characteristics of Entrepreneurs." Journal of Asia Business Studies, vol. 18, no. 1, 2024, pp. 21–43. https://doi.org/10.1108/JABS-09-2022-0309.

Said, Farah, and Lauren Perlik. "Why Pakistan's Shopkeepers and Customers Keep Choosing Cash over Digital Payments." Innovations for Poverty Action, 29 Apr. 2026, https://poverty-action.org/why-pakistans-shopkeepers-and-customers-keep-choosing-cash-over-digital-payments. Accessed 26 Aug. 2026.

Soomro, Raheem Bux, et al. "The Adoption of Digital Technologies by Small and Medium-Sized Enterprises for Sustainability and Value Creation in Pakistan: The Application of a Two-Staged Hybrid SEM-ANN Approach." Sustainability, vol. 16, no. 17, 2024, article 7351. https://doi.org/10.3390/su16177351.

Ullah, Ihsan, et al. "Factors Affecting Digital Marketing Adoption in Pakistani Small and Medium Enterprises." Logistics, vol. 7, no. 3, 2023, article 41. https://doi.org/10.3390/logistics7030041.

Qalati, Sikander Ali, et al. "A Mediated Model on the Adoption of Social Media and SMEs' Performance in Developing Countries." Technology in Society, vol. 64, 2021, article 101513. https://doi.org/10.1016/j.techsoc.2020.101513.

Qalati, Sikander Ali, et al. “Social Media and SMEs’ Performance in Developing Countries: Effects of Technological-Organizational-Environmental Factors on the Adoption of Social Media.” SAGE Open, vol. 12, no. 2, 2022, pp. 1–13. https://doi.org/10.1177/21582440221094594.

Yunis, Manal, et al. “The Role of ICT and Innovation in Enhancing Organizational Performance: The Catalysing Effect of Corporate Entrepreneurship.” Journal of Business Research, vol. 88, 2018, pp. 344–356. https://doi.org/10.1016/j.jbusres.2017.12.030.

**Duty Without Liability: Bank Reporting Incentives and Elder Financial Exploitation**
**By Nicholas Djoulia**


## Abstract

Elder financial exploitation (EFE) has grown from a containable, small-scale problem into a systemic economic and regulatory failure. Americans over 60 reported $2.4 billion in fraud losses to the Federal Trade Commission in 2024. Still, because most fraud goes unreported, the FTC estimates the true cost that year at between $10.1 billion and $81.5 billion (Federal Trade Commission). This paper argues that EFE persists not because seniors lack financial literacy, but because the institutions best positioned to detect exploitation, banks, face no legal duty to report it. Drawing on neuropsychological research, FinCEN's Bank Secrecy Act data, and New Jersey's statutory reporting framework, the paper shows that EFE is a market failure driven by three forces: wealth concentrated among older, less digitally literate Americans; AI-enabled scams that build convincing relationships with victims at scale; and a reporting regime built around the wrong kind of observer. Under New Jersey's permissive statute, confirmed by Lucca v. Wells Fargo Bank, N.A. (2015), reporting is strictly dominated: nondisclosure carries no penalty, so silence is costless while acting is not, and no level of confidence that a customer is being exploited makes reporting privately rational. That gap is not an oversight of an older statutory scheme: New Jersey imposed a mandatory exploitation-reporting duty on broker-dealers and investment advisers in 2019, four years after Lucca, while leaving depository institutions under the permissive standard. A formal model shows the result holds for arbitrarily small positive administrative costs, and that immunity cannot fix it, since immunity lowers the cost of acting but never below zero; only a mandatory duty closes the gap between social and private incentives. The paper concludes with illustrative, clearly hypothetical statutory language giving New Jersey depository institutions hold authority modeled on the 2019 SAFE Act, attaching a mandatory notification duty to it, and enforcing that duty through supervisory examination rather than a civil penalty, then explains why the design departs from the model's first-best.


## I. Introduction

Imagine receiving a call from your grandchild, panicked after a car accident and too afraid to ask their parents for help. The voice sounds exactly like them, kind and familiar, so you rush to your computer and wire the money. Days later, you learn the accident never happened: an AI-generated voice clone, built from a few seconds of audio lifted from a social media clip, had just taken your money. You have joined thousands of people your age who have been scammed the same way.

This method of stealing elders' wealth is real, increasingly common, and fairly new. For most of the digital age, EFE occurred on a small scale and was containable through law enforcement, with scammers sending bulk, poorly worded phishing emails to older, less tech-savvy targets. Around 2022, generative AI made fraud cheaper and easier to commit, producing growth that obsolete regulatory regimes have been unable to manage. The mismatch

between the scale of EFE and the policy response is not simply a lack of political will. It is a failure to ask why banks, which hold the most comprehensive view of a customer's financial activity, so rarely act on what they know, since EFE has been framed almost exclusively as a financial literacy problem.

This paper argues that EFE is a market failure with three causes: wealth concentrated among a generation less equipped for digital risk, AI-enabled scams that patiently cultivate relationships with victims at scale, and permissive reporting regimes that place no obligation on the party best able to intervene. The third cause is not theoretical: New Jersey's reporting statute, tested in Lucca v. Wells Fargo Bank, N.A., lets banks report suspected exploitation but does not require it. Faced with a statute that merely permits disclosure and attaches no consequence to silence, there is no level of confidence at which reporting is the best response; nondisclosure strictly dominates. That failure is built into the statute rather than a matter for normative debate, and diagnosing EFE as its logical consequence, rather than a failure of individuals or cognition, is what makes targeted reform possible.

Sections II through V discuss why EFE has been systemic rather than case-specific: the macroeconomic reality on the ground (II), the rising magnitude of AI-driven scams targeting elderly Americans (III), the neuropsychological basis for seniors' inherent vulnerability to those scams (IV), and the fragmented nature of the regulatory framework that has lagged behind both (V). Section VI discusses the particular legal mechanism causing that failure in New Jersey, and Section VII models it within a threshold framework calibrated to show that, under the current state reporting statute, reporting is strictly dominated whatever the bank's confidence that a customer is being exploited. Section VIII uses that result to propose a tailored statutory solution, Section IX evaluates that solution against the governance tradeoffs it generates, and Section X concludes.

## II. The Economics of an Aging America: Scale and Cost

This section and the three that follow establish the economic, technological, psychological, and regulatory context to which the reporting-duty model in Sections VI and VII responds.

### A. Wealth Concentration and the Great Wealth Transfer

To understand why EFE has become a systemic market issue, it is first crucial to understand where people's wealth is concentrated. Modigliani and Brumberg's life cycle hypothesis posits that people save money during their working years and access it to spend during retirement. This hypothesis suggests that wealth accumulation peaks for most just before retirement; as a result, most of the wealth in the US population is held by older adults.

This is a massive economic shift, not a small one. As Baby Boomers age and retire, the nation's wealth increasingly rests in the hands of older Americans, the Great Wealth Transfer. Kim et al. situate this within the broader relationship between population aging, financial development, and wealth inequality, finding that demographic aging shapes how the gains from

financial development are distributed: never before has so much wealth concentrated among generations now being forced to navigate a digital world they did not grow up with.

### B. What Fraud Actually Costs the Economy

There is also a major fiscal impact: as seniors' savings decrease due to fraud, applications for Medicaid, housing assistance, and SSI rise, converting private losses into public spending, a direct transfer to state and federal coffers that Gunther recognizes but does not quantify. Gunther, using a methodology developed with NORC at the University of Chicago, estimates annual EFE losses to Americans over 60 at $28.3 billion, of which roughly three-quarters, or $20.3 billion, is attributable to perpetrators known to the victim, and describes that figure as conservative once publicly funded programs for victims who have exhausted their own resources are counted. That source matters for Section V: most measurable losses stem from trusted-relationship transactions that SARs are least capable of identifying.

### C. The Compliance Cost Burden on Community Banks

EFE is causing negative institutional impacts in the banking sector, thereby increasing financial risk for seniors. Under the Bank Secrecy Act, FinCEN receives over 155,000 reports related to EFE annually, with approximately $27 billion in suspicious activity (FinCEN, "Elder Financial Exploitation"). The record $1.3 billion penalty against TD Bank for Anti-Money Laundering (AML) failures in 2024 prompted institutions to increase spending to combat this (FinCEN, "TD Bank Enforcement"). This financial burden falls primarily on small banks.

## III. How Fraud Techniques Outran the Framework

### A. From Phishing Emails to Synthetic Credibility

Technological advances have outpaced the infrastructure built to respond to them. For most of the digital age, elder fraud relied on mass, uncoordinated efforts: scammers blasted seniors with thousands of phishing emails knowing only a few would respond, and those letters were poorly worded and obvious to most recipients. No longer. Since late 2022, generative AI has been weaponized to script deepfake phone calls, personalized emails, and convincingly urgent technical-support chat interfaces, and the noisy fingerprints that once made scam attempts easy to spot have largely been scrubbed. Scammers can now establish deceitful-but-legitimate relationships with seniors at scale instead of one manpower-intensive victim (mark) at a time (Blackbird.AI). Voice cloning is one major reason: McAfee testing found cloned voices matched their source with about 85 percent accuracy using just three seconds of sampled audio (McAfee). That finding is vendor research, not peer-reviewed. Still, the underlying point, how little audio is now required, is not seriously contested. With audio routinely available from social media, deepfake impersonation has become one of the most lucrative scam techniques that existing law does little to combat.

### B. Pig-Butchering: Patience as a Fraud Strategy

The most significant and growing AI-powered scam technique in EFE beyond deepfakes is pig-butchering, also called a romance scam: named for fattening animals before slaughter, it has scammers spend days or weeks building relationships with seniors through social media before leading them into fraudulent, usually crypto, investments. Unlike traditional scams, it is slow-moving and emotionally driven, exploiting seniors' loneliness late in life (DeLiema).

The cost of this problem ranges from unbelievable to terrifying. Researchers at the University of Texas at Austin followed blockchain transfers from victims and discovered that pig-butchering groups funneled over $75 billion worldwide into scam crypto wallets from 2020 to 2024 alone (Griffin and Mei). Keep in mind that these statistics account for losses incurred by people of all ages; the researchers did not parse their findings by demographic, so consider the total more as a measure of how widespread the scheme has become than as an estimate of losses specific to seniors.

## IV. Why Senior Citizens Are Structurally Vulnerable

### A. What the Brain Science Shows

Seniors' vulnerability to financial fraud is not a matter of a lack of care or attention to detail. It can be rooted in biology. Older adults experience a steady cognitive decline, with thinning of brain structures that help regulate how the brain processes and responds to threats, which predisposes them to vulnerability in the digital age (Fenton et al.).

Timing matters for policy. Nicholas et al. looked at credit-report data for 81,364 Medicare beneficiaries. They found that patients later diagnosed with Alzheimer disease and related dementias missed bill payments up to six years before diagnosis and began accruing subprime credit scores about 2.5 years before diagnosis more often than non-diagnosed controls; no similar pattern preceded arthritis, glaucoma, or hip-fracture diagnoses, suggesting the effect is specific to cognitive decline. The analysis is observational, and the sample excludes many multi-person households. Still, the policy implication is clear: financial risk starts years before diagnosis, so protections triggered by a diagnosis, conservatorship, or age limit arrive after the period when they would matter most.

### B. Isolation, Trusted Perpetrators, and the Trust Trap

Social and demographic circumstances shape cognitive vulnerability in older adults, who commonly face increased isolation and loneliness that can bring on depression and mild cognitive impairment. DeLiema links this to EFE: scammers exploit that loneliness to draw seniors into online "relationships" with strangers who are, in fact, exploiting them, and shrinking social connections compound the risk.

Another piece of the puzzle is who the fraudster was. According to FinCEN's threat-pattern analysis, about 20 percent of elder-financial-exploitation BSA reports involved elder theft, defined as the misuse of an older adult's funds by someone the victim knows and

trusts. In nearly 40 percent of elder theft incidents, the adult child committed the crime (FinCEN, "Elder Financial Exploitation"). Betrayal trauma theory explains why this form of EFE is both unexpected and unusually damaging: exploitation by caregivers and relatives inflicts greater financial harm and worse psychological injury than the same loss at a stranger's hands, because the cognitive machinery people use to sustain close relationships resists recognizing abuse within them. The institutional deficiency is that standard regulatory action and fraud-awareness campaigns ignore this form of scam entirely, increasing risk for victims.

### C. Why Financial Education Is Not Enough

The instinct is to address EFE through financial literacy programs. Still, the evidence does not support this as an effective policy solution: financial knowledge and aptitude do not reduce the risk of online scams, while social advocacy and support for seniors is far more effective at preventing financial loss and protecting cognitive welfare (Gotelaere and Paoli; Button and Cross; Cross and Kelly; Prenzler).

Lachs and Han's concept of age-associated financial vulnerability supports this: they attribute financial vulnerability in older adults to four domains, cognitive/emotional, medical/functional, psychosocial, and environmental/societal, that financial literacy interventions can remedy only a small fraction of. Instruction in the recognizable markers of earlier-generation scams (implausible premises, poor spelling, unfamiliar senders) addresses a threat generative AI has already eliminated. The emotions today's scams exploit (urgency, family love, reciprocal trust) are not things money skills improve.

## V. The Regulatory Landscape and Where It Falls Short

### A. Too Many Agencies, No Clear Leader

Regulatory action against EFE falls short despite rising budgets, because no single agency is responsible for combating it: FinCEN handles Bank Secrecy Act reporting, the CFPB regulates financial products, the SEC covers financial advisers within BSA, the FTC handles fraud prosecutions, state agencies handle criminal enforcement, and Adult Protective Services handles the social side. In December 2024, seven federal and state regulators issued a joint statement setting out risk-management practices institutions "may consider." What is most telling about the statement is what it is not: a coordinated expression of supervisory expectations. It explicitly includes "no new mandates" and assigns "no primary agency responsibility" (Interagency Statement on Elder Financial Exploitation), underscoring how fragmented this landscape really is.

### B. The BSA's Blind Spot

The BSA's SAR system is our main systemic safeguard against EFE, and it was not designed for the problem we ask it to solve. The SAR detects suspicious account activity, such as large wire transfers to new businesses, which works well for outbound scamming and poorly for

exploitation inside trusted-relationship transactions. When the daughter of an older adult wires money out of her mother's account, the transaction passes every test on paper. That is the SAR's blind spot. FinCEN reports outbound scamming makes up 80% of EFE BSA filings while trusted-relationship fraud accounts for the remaining 20% (FinCEN, "Elder Financial Exploitation"). The gap is larger than that percentage indicates: elder theft accounts for only about a fifth of BSA reports but roughly three-quarters of total EFE losses, based on Gunther's back-of-the-envelope calculations. That is because only about 12.5 percent of elders scammed by someone they know report the crime, compared to 67 percent of otherwise similar victims tricked by a stranger.

In short, the category most likely to slip through the SAR pipeline is the category where most of the money is actually stolen, which is why it matters for the recommendation in Section VIII. A reporting duty acts only on suspicions that already exist; it cannot help a bank detect a suspicious transfer in the first place, but it does control behavior once exploitation is detected. For trusted-relational fraud specifically, reporting is the only way to reach someone who can intervene: if a bank witnesses this kind of fraud, the kind least likely to trigger a federal SAR, its only recourse is referral under state law. The mandate therefore hits the most expensive type of abuse the hardest.

### C. The Financial Exploitation Prevention Act's Permission Problem

The most notable federal legislative response to EFE is the Financial Exploitation Prevention Act (H.R. 2478 and S. 2840), which would let open-end investment companies and their transfer agents delay the redemption of securities when they reasonably believe a customer is being financially exploited. A prior version of the bill passed the House 419 to 0 in 2023 but expired without Senate action; the current H.R. 2478 passed the House 414 to 2 on June 25, 2026, and its Senate companion, S. 2840, remains pending in the Banking Committee. Even with that delay, the concept is promising: pausing a suspicious redemption buys time to identify exploitation before funds leave the account.

The bill's legality is still being negotiated, and for obvious reasons: it invades the rights of consenting adults to manage their own finances fully, and fails to define adequately what constitutes "suspicious activity." This same void, a reporting mechanism that is not designed to identify the kind of trusted-relationship abuse described above, exists at the state level in New Jersey's very own statutory language, which Section VI addresses head-on.

## VI. The Duty Gap: The New Jersey Case Study and the National Default

New Jersey is one of the handful of states that lacks an omnibus statute governing reports of suspected elder financial exploitation. Instead, we have four discrete reporting regimes, enacted at different times and addressing different actors. Together, however, they have the (perhaps unintended) effect of both particularizing and, we argue below, privatizing the problem: singling out the actor with the most frequent and comprehensive access to a customer's finances, the depository bank, as the only relevant observer without a duty to report what it witnesses.

Regime 1: N.J.S.A. 17:16T-1 to -4. New Jersey's original regime is this statute. It regulates banks directly. Passed in 1998, it allows but does not require "a financial institution [to] disclose the financial records about a customer's account to a law enforcement agency, a county adult protective services agency, or both" where the suspected victim is a bank's elderly customer or vulnerable adult customer. Immunity from suit attaches on disclosure or, in good faith, refusal to disclose. The statutory text does not itself specify which institutions the section covers, but the legislative history resolves that ambiguity: "financial institutions." This bill "would provide financial institutions the legal comfort necessary to make that decision."

Regime 2: The Adult Protective Services Act before its 2009 amendment, N.J.S.A. 52:27D-409(a). Passed in 1993 as P.L. 1993, c. 249, APSA preceded N.J.S.A. 17:16T-1 to -4 by five years, and its initial reporting provision used the same permissive construction later applied to the financial sector: every "person who has reason to believe that a vulnerable adult is an abuse victim," professional or otherwise, "may" report. Because APSA predates 17:16T-1 to -4, the borrowing runs the opposite direction from what a casual reader would assume: 17:16T quotes APSA's earlier permissive language.

Regime 3: APSA as amended in 2009, which introduced tiers: certain professionals ("health care professionals," "law enforcement officer," "firefighter," "paramedic," "emergency medical technician") got mandatory "shall report" language, while everyone else ("any other person") kept the original voluntary "may report." Banks were left out of the mandatory tier not because the Legislature overlooked them, but because the professions selected work closely with people at risk of immediate physical harm, not financial harm, and have less comprehensive customer information than a nurse or officer would.

Regime 4: The Safeguarding Against Financial Exploitation Act, N.J.S.A. 49:3-84 to -89. In 2019, the Legislature returned to financial exploitation a fourth time and, for the first time, imposed a mandatory duty on a financial-sector actor. Under the SAFE Act, P.L. 2019, c. 340 (effective April 2020), a qualified broker-dealer employee or agent who "reasonably believes that financial exploitation of an eligible adult... may have occurred, may have been attempted, or is being attempted shall promptly notify the bureau and the applicable county adult protective services provider" (N.J.S.A. 49:3-86(a)); "eligible adult" means a person 65 or older, or one covered by APSA. The notice requirement carries immunity for good-faith reports made with reasonable care and permission to delay a transaction or disbursement where exploitation is suspected (N.J.S.A. 49:3-88(a) to (c)).

The SAFE Act matters for three reasons. First, it shows the Legislature knows how to draft a mandatory exploitation-reporting obligation for a financial intermediary, and did so as recently as 2019. Second, it did so four years after Lucca publicly explained the permissive nature of N.J.S.A. 17:16T-1 to -4. The exclusion of depository institutions from the mandatory tier was therefore a choice the Legislature made and then kept, not an oversight. Third, and most useful here, the SAFE Act supplies a drafted, enacted New Jersey starting point for the amendment proposed in Section VIII: it need not be built from scratch, only adapted from Title 49 to fit Title 17.

Two caveats bear mentioning. The SAFE Act's duty runs to the Bureau of Securities and county APS, enforced through the Bureau's oversight powers rather than a private right of action, so its liability benchmark is regulatory rather than civil, a distinction Section VII addresses directly. And it applies to securities intermediaries, who see a narrower swath of a customer's finances than a depository institution does. Neither limitation undercuts the point that Section VIII's drafting problem has already been solved once in this jurisdiction.

Lucca v. Wells Fargo Bank, N.A., 441 N.J. Super. 301 (Law Div. 2015) litigated the real-world application of these two regimes to a depository institution. Wells Fargo processed twenty-seven wire transfers, totaling nearly $300,000 over six months, to overseas accounts at the instruction of an 82-year-old customer. A representative who suspected the transfers told the customer she was being scammed and logged an internal elder-abuse report after each transaction, but never notified police or APS. The customer sued for failure to report.

The trial court found no legal duty to report because N.J.S.A. 17:16T-1 to -4 "does not impose a duty on a financial institution to notify the authorities of a suspected scam"; the permissive language showed the Legislature intended this to be optional, and without a duty, there can be no breach. A bank that flagged the fraud internally and never referred to it externally was held not liable, since it owed no duty to release the funds. Lucca is not an outlier: the default rule of American commercial law is that a receiving bank that properly executes an authorized payment order owes no duty to any party beyond what statute or contract specifically imposes (U.C.C. section 4A-212). Courts nationwide have unanimously declined to impose a duty on depository institutions to stop authorized transfers even where fraud is known. See In re Cook, 2023 WL 3467209 (E.D. Va. 2023); Scott v. Branch Banking & Trust Co., 588 F. Supp. 2d 667 (W.D. Va. 2008). Lucca simply held that New Jersey's anti-exploitation statute did not implicitly repeal that default. That matters here because the model's independent variable is not one state's statutory choice, but the default legal treatment of authorized transfers across most states, which is exactly why Section VIII's proposed carve-out generalizes.

The court's reasoning is as significant as its holding. In the same passage, the court explains why the gap persists: the Legislature could have amended N.J.S.A. 17:16T-1 to -4 to require financial institutions to disclose suspected financial exploitation, but it did not. The Legislature opted instead to amend APSA to create a separate tier of professionals who are mandated to report abuse of vulnerable adults. "Other persons or entities who have reasonable cause to believe a vulnerable adult is subject to exploitation," the court tells us, "such as the bank here," "are not under that mandate. They may report the information, but are not required to do so."

That is New Jersey's duty gap, stated this time in clinical rather than statutory terms. FinCEN sees over 155,000 SAR filings related to EFE per year, totaling $27 billion in suspicious activity (FinCEN, "Elder Financial Exploitation"); no occupation newly designated as a mandatory reporter under APSA will encounter that many victims. Of the four mutually exclusive legal frameworks governing who must report suspected exploitation in New Jersey, the statute covering the entity best situated to detect it, the depository bank, imposes no duty at all.

Lucca's relationship to the model below should not be overstated: no single trial-level opinion could prove the model's prediction. Its role is narrower and antecedent: it establishes one of the model's parameters as a matter of law. Because a statute's grant of permission forecloses any cause of action for nondisclosure, the probability that silence creates liability is not merely low; it is zero, and everything that follows depends on that point of law together with the assumption that the bank minimizes expected cost.

**VII. Duty Without Liability: Modeling the Bank's Reporting Decision**

We can formalize this intuition about duties. Suppose a bank compliance officer receives an alert that triggers a suspicious activity report regarding the account of an elderly bank customer. She forms belief q in [0,1] that the transaction resulted from bona fide abuse. She may then either report that belief to a county adult protective services agency (or law enforcement) or not; acting here means notifying an outside agency. The model isolates notification because it is the least invasive intervention capable of reaching an actor empowered to investigate, which lets us show that not even the least-invasive response available to the bank is privately rational under a permissive regime. Section VIII bundles notification with intervention in the transaction itself; the model applies to the notification component of that design and need not price the intervention.

Acting is costly. Investigation will cost the bank some administrative expense $c_a$ per referral; each referral takes up employee time whether the suspicion is valid or otherwise. Further, the bank may face litigation. The bank is sued either because it placed a wrongful hold on an account not actually being exploited (probability $1 - q$) or because it improperly denied a customer access to her funds. Such a suit is actually filed with probability $s$, and in either case the bank must spend some amount $D$ defending against it. Inaction, by contrast, costs the bank nothing except if the customer was actually being exploited and the bank is sued for failure to report (probability $\pi$), in which case it pays a statutory penalty $H$. We can now write expected costs as:

$$E[\text{cost of acting}] = c_a + (1 - q) \times s \times D$$
$$E[\text{cost of silence}] = q \times \pi \times H$$

In general, a utility-maximizing bank chooses to report whenever the expected cost of acting is no greater than the expected cost of silence, and where $\pi \times H > 0$, that comparison pins down an interior private reporting threshold. The general expression is preserved below because it becomes meaningful once $\pi \times H > 0$, the range Section VIII makes the case we should be pushing the law towards operating in, and because it lets us later determine algebraically what magnitude of penalty is necessary to close the gap between private and social incentives:

$$q_{private} = (c_a + s \times D) / (\pi \times H + s \times D)$$

New Jersey's statute contains no enforcement provision whatsoever: $H = 0$ by the plain terms of N.J.S.A. 17:16T-1 to -4, so $\pi \times H = 0$ regardless of the value of $\pi$. Lucca provides the

other half of the equation, foreclosing any private right of action for failure to report to confirm that $\pi$ itself is zero on this channel. Grounding the result in statutory silence on penalties, rather than in Lucca by itself, also matters: Lucca was issued at the trial level, and the paper's main empirical claim should not hinge on one plausibly distinguishable decision.

Because $\pi \times H = 0$, E[cost of silence] = $q \times \pi \times H = 0$ for every value of q, while E[cost of acting] = $c_a + (1 - q) \times s \times D$, which is at least $c_a > 0$ for every value of q. Silence thus weakly dominates acting for every value of q in [0,1], not merely below some threshold, because the bank's belief about whether its customer is being cheated does not affect the comparison. The general threshold formula above offers no interior solution here. There is no $q_{private}$ for it to locate, since dominance holds identically at every point on [0,1] rather than pivoting on some cutoff belief.

This also deflates the response that comes to mind first. Grant banks absolute immunity and $s \times D$ falls to zero, collapsing the cost of acting from $c_a + (1 - q) \times s \times D$ to $c_a$. It cannot go lower: while the statute is permissive and the penalty zero, $c_a$ is the floor. Behavior therefore does not change, not through some perverse effect of immunity but because the friction immunity removes never mattered. Immunity lowers the cost of acting relative to inaction, and where no duty to act exists, inaction already costs nothing. Rules that impose an affirmative duty and attach a real sanction to its breach are therefore not one reform among many; they are the precondition for any of the others to have any effect.

One wrinkle remains, without which this conclusion may seem at odds with the aggregate reporting numbers in Section II. Banks are not silent: they file over 155,000 BSA reports relating to elder exploitation alone each year. The resolution is simple: that mandatory channel is not the reporting channel the model here describes. Banks must file SARs; they must report suspicious transactions that might evidence criminal conduct. Filing a SAR is a federal mandate backed by prudential supervision, and the penalties on that channel are neither nonexistent nor inconsequential, as the recent $1.3 billion settlement with TD Bank suggests. Except under unusually aggressive treatment of SAR reporting, the expected $\pi$ on that channel is very much greater than zero.

Here, the channel is different. Specifically, it is the state-mandated referral to law enforcement or a county APS agency under N.J.S.A. 17:16T-1 to -4. It is only on that channel where $\pi$, and by extension the cost of silence, equals zero according to Lucca. Lucca illustrates why this distinction matters. Wells Fargo detected the scam, notified the customer, and filed an internal elder-abuse report after each of the twenty-seven transfers, but never referred the matter to law enforcement or county APS, the only actors with authority to intervene. Against a backdrop where referral is mandatory on one channel and costless to omit on another, optimal strategy dictates silence: 155,000 referrals there, silence here.

If law enforcement already receives SAR data, what does state referral add? Three things: SAR filings arrive weeks after the transactions they describe, once funds have cleared; they go to federal law enforcement rather than the county APS agencies with statutory responsibility to evaluate capacity and existing relationships with vulnerable adults; and they involve no

interaction with the customer at all. Notification under the statute modeled here is the only currently available way to trigger intervention by an actor with both the authority and the relationship to investigate capacity while the transaction is still live.

To close: continuing to think about private action versus the social optimum, what would q be if a social planner decided? The social planner would act if the expected value of intervening exceeded its cost, acting whenever $q \times r \times L > c_a + (1 - q) \times c_f$. Notice the addition of r, the portion of L that would be saved by intervening early, and of $c_f$, the cost to the consumer of a false positive referral: the cost of losing privacy, of disrupting a transaction through an APS investigation predicated on nothing more than a suspicious transaction that was legitimate. Solving for q as above:

$$q_{social} = (c_a + c_f) / (r \times L + c_f)$$

Calibrating with L = $21,902, r = 0.72, $c_a$ = $150, and $c_f$ = $400 gives $q_{social}$ = 3.40%. Society wants the bank to act whenever it is at least 3.40% confident that exploitation is occurring. Recall that allowing the bank to choose lets it always prefer silence, regardless of its degree of confidence.

How much leeway is there in the 3.40% baseline number? Allowing L to range from roughly half to nearly seven times its baseline value, r from about one fifth to just above its baseline, and $c_a$ from about a quarter to three times as high produces social thresholds ranging from 0.51% to 14.92%. The private threshold cannot be breached, because given current law there is nothing to breach. Strictly positive $c_a$ guarantees silence dominates acting, and no values of L, r, s, or D can change that fact. The resulting gap between private and social thresholds is not sensitive to uncertainty about the parameters on which they both depend. It is an invariant of any legal regime where $\pi = 0$. Neither raising the cost to customers of a referral ($c_f$), nor lowering $c_a$ to a value still comfortably within the plausible range, nor both together, closes the gap.

The calibration below answers a different question than "what threshold does the bank use under current law," since the preceding paragraphs show there is none: dominance holds at every q. It instead asks what penalty would convert the bank's problem from a dominance game into a genuine interior-threshold game whose threshold matches $q_{social}$. That reframing matters beyond notation: it clarifies that a mandatory-duty-plus-penalty regime does not merely raise the cost of silence at the margin; it converts the decision's structure, from one in which the bank's belief about exploitation is irrelevant to one in which that belief becomes decision-relevant. The minimum sanction that accomplishes this conversion can be found by setting $q_{private} = q_{social}$ and solving for H:

$$H_{min} = [(c_a + s \times D) / q_{social} - s \times D] / \pi$$

Notice what happens when $\pi = 0$: the expression is undefined, because a penalty is divided by a probability of ever paying it that equals zero. No positive value of H seals the wedge, which is precisely why choosing a penalty structure cannot change a bank's incentives as long as New

Jersey stays in its current regime. Plug $\pi = 0.15$ into the expression and grant banks perfect impunity from wrongful-hold suits (set $s = 0$), and you get $H_{min} = \$29{,}399$. This is not a number I would suggest anyone put into a statute. It is meant only to convey scale.

That magnitude does not depend on the paper's least defensible parameter. When $s = 0$, the expression above reduces to $H_{min} = c_a / (q_{social} \times \pi)$, and since $q_{social}$ falls as r rises, a smaller recovery share raises $q_{social}$ and hence lowers the sanction needed to bring about that deterrence level. At $r = 0.15$, the lowest value examined in the paper, $q_{social} = 14.92\%$ and $H_{min}$ is approximately \$6,701. The paper concludes that the penalty needed to align private and social incentives perfectly is small and non-punitive, so it does not rely on the assumed value of r being accurate; it applies even more strongly if r is too high, as the Methods Appendix suggests.

As $\pi$ approaches zero, increasingly large penalties will be required to close the wedge, which is another way of stating a familiar result from the economics of law enforcement. Because violations that are unlikely to be detected can be expected to escape punishment more often than not, the optimal penalty for such violations must be very large to deter would-be offenders (Becker). The reverse is also true. If a jurisdiction wants the duty to remain effectively optional, it does not need to grant matching absolute immunity; it only needs to pair the duty with a sufficiently low and believable chance of enforcement.

## VIII. Toward a Statutory Fix

The Financial Exploitation Prevention Act (H.R. 2478/S. 2840), cited in Section V, allows open-end investment companies and transfer agents to delay redemptions when they reasonably believe a customer is being exploited. But it does not acknowledge or solve the reporting responsibility gap described in this paper. It provides extra authority to a party who is not lacking permission.

### A. What the model requires and what a workable design would supply

An implication of this model that bears on the enforcement design below is that a first-best would be a mandatory duty backed by an expected sanction large enough to exceed the cost of acting. Proposition 1: A mandate changes behavior if and only if $\pi \times H > c_a$. The design sketched below imposes no fixed civil penalty. The departure is not an oversight. Three facts about the world drive it.

First, the model leaves open the source or sources of $\pi$'s nonzero probability; it reflects only the probability that silence will be costly if attempted, not how many sources those costs stem from (see supra Section VII's discussion of the Bank Secrecy Act channel). Financial institutions already operate under regimes whose weight derives entirely from prudential supervision rather than a state-imposed penalty; an enforcement regime built the same way provides $\pi$ through routine examination. Notification records subject to examiner review and required to be retained have independent value to the institution, creating a reason to comply absent any posted penalty, since exam findings themselves cost money. Proposition 1 says

nothing about how high that cost must be beyond exceeding c_a, which is probably on the order of $150.

Second, depository institutions are partly regulated by state charter, splitting enforcement authority between state and federal regulators. A default civil penalty wielded by the Commissioner would apply directly only to state-chartered institutions; compliance by national banks would first require deference from their primary federal regulator, treating a non-uniform population of institutions as though it were uniform. Supervisory review avoids that problem by operating seamlessly across both charters, since each already has a designated primary examiner.

Third, any workable design must enlarge the action set available to institutions rather than reprice a fixed one. Notification alone does not stop funds from being paid out; intervention does. Pairing notification with hold authority gives institutions independent reason to comply with the notification duty, because the hold authority would be worth exercising on its own terms, a point outside this paper's static framework, which treats the action set as given.

**B. The Design and Its Statutory Language**

The design below picks up where the model leaves off. It has three elements, which the model shows must be added in the order listed.

Mandatory. Until the duty is mandatory, $\pi = 0$, and neither immunity nor enforcement consequences matter for the institution's decision-making.

Absolute. Until the enforcement consequence of notifying (or not) is stated in absolute terms, good faith remains in play. Leave the "s" term above in place, and you leave open the question of what happens to be litigated after the fact.

Some. The duty must carry some enforcement consequence, whether stated in terms of money (penalty) or job (supervisory jeopardy). Anything less is a mandate only in name: expected cost of silence remains zero, and we are right back to the baseline problem statement.

Examination-based enforcement can satisfy that third condition without posting a civil penalty. See discussion supra. But it must strictly exceed c_a, so Proposition 1 tells us the resulting expected sanction associated with failing to notify only needs to exceed $150 to change behavior. Below is calibrated but deliberately unstyled statutory language sketching what such a design would look like, cribbed loosely from the Legislature's own statutes imposing (but not calibrating) very similar duties on broker-dealers and investment advisers at N.J.S.A. 49:3-86 to -88.

a. Temporary hold. A financial institution may place a temporary hold on a transaction or disbursement involving a senior customer or vulnerable customer where a qualified employee reasonably believes the transaction may result in financial exploitation of the customer. A hold placed under this subsection need not be conditioned on the completion of an internal review process and may be placed immediately.

b. Notification to the Department and adult protective services. A financial institution that places a hold under subsection a. of this section shall, within [x] days of placing the hold, notify the Department of Banking and Insurance and the applicable county adult protective services

provider of the following information: (1) the fact that a hold has been placed; and (2) the reason suspected exploitation was believed to have occurred. Absent the placement of a hold pursuant to subsection a. of this section, any notification by a financial institution to adult protective services regarding suspected exploitation shall be permissive.

c. Notification to the customer. A financial institution that places a hold under subsection a. of this section shall notify the customer on whom the hold is placed of the following information within [x] days of the placement of the hold: (1) the fact that a hold has been placed on a transaction involving that customer; and (2) the reason the hold was placed. A financial institution need not provide notice under this subsection to any person that the financial institution reasonably believes is involved in the suspected exploitation.

d. Immunity. From and after the date of enactment of this act, a financial institution and any officer or employee thereof shall be immune from civil and administrative liability arising from the action of that institution or individual in placing a hold, making or declining to make a notification, or declining to place a hold under this section, provided that the individual completed training as prescribed by the Commissioner.

e. Condition of deposit. Without imposing a new condition of deposit, the Department of Banking and Insurance shall review compliance with this section in the ordinary course of examination.

f. No private right of action. Nothing in this section shall create any civil penalty or private right of action.

**C. What a design of this kind would not do**

Two caveats. First, this design preserves institutional discretion over the triggering act: nothing compels intervention or creates a duty to police exploitation, and notice remains optional wherever no hold is placed. On the facts of Lucca itself, an institution that detected exploitation, logged it internally, and placed no hold still would not have been required to refer, exactly as the model predicts; if the triggering act stays optional and unenforced, $\pi \times H$ on that channel remains zero, and Section VII's analysis carries through unchanged. What the design changes is the payoff to the optional act itself, by giving institutions an authority they lack today, one Wells Fargo had no lawful means to exercise in Lucca. How often that authority would actually be used is an empirical question this paper cannot answer.

Second, one response suggests itself since the Lucca defendant was a national bank: the National Bank Act preempts state laws that "impair" or "significantly interfere" with a national bank's powers, as construed by Barnett Bank of Marion County, N.A. v. Nelson, 517 U.S. 25 (1996), and narrowed by Dodd-Frank section 1044 to state laws whose discriminatory effect or interference is specific to that state law. A statute like the one sketched above is a poor preemption candidate: it reaches all financial institutions, not just national banks; grants authority rather than removing it; creates no private right of action; and yields to federal law and the primary federal regulator for federally chartered institutions, while preserving SAR confidentiality entirely. This jurisdictional asymmetry means that the Bureau of Securities can

supervise state intermediaries directly while depository supervision is split, which may also explain why the Legislature reached broker-dealers in 2019 but left depositories untouched.

## IX. Key Governance Tensions and the Future Trajectory

### A. Two Tensions That Must Be Resolved

Two core tensions underlie the future trajectory for EFE policy: how much to protect seniors versus preserve their autonomy, and how to regulate compliance costs without pushing seniors toward less-regulated, riskier alternatives.

Autonomy vs. Protection: The central question is whether to interfere with seniors' freedom of choice in order to protect them from fraud. Thaler and Sunstein propose a middle ground, using nudge theory to offer a protective service through government or healthcare providers that a senior may opt out of, with third-party monitoring of arrangements the government contracts. That preserves both protection for those who want it and freedom for those who find it intrusive. Disability rights advocates argue strenuously against any default protective system, contending that treating seniors as incompetent by default reinforces institutional ableism.

Compliance Costs vs. Financial Inclusion: Serving senior customers carries regulatory obligations that raise costs. If those costs mount, banks will curtail services to older people, leaving seniors with prepaid cards, wire services, and crypto, so regulation meant to protect them would leave them more exposed.

### B. How the Proposed Design Answers Them

Both tensions bear on the proposal in Section VIII. On autonomy, the proposal tilts further toward protection than the permissive status quo in two distinct ways: notice sends skepticism outside the institution for another actor to weigh, while intervention lets the institution act unilaterally on the customer's behalf, the more serious step. The design authorizes both, but bounds the second rather than leaving it open-ended: it terminates on a short, fixed schedule absent an agency request or court order, cannot be renewed indefinitely without judicial involvement, and is subject to prompt internal review.

Two features of the design respond directly to the autonomy objection: the institution must inform the customer of the hold and its reason within a short fixed period, so the person whose funds are restrained learns of the suspicion directly rather than through a failed transaction; and third-party notice is forbidden to anyone the institution reasonably believes is involved in the exploitation, addressing the trusted-relationship case Section IV identifies as the costliest and least visible form of abuse. Whether these safeguards are adequate in practice is a fair question; that the design confronts the objection rather than ignoring it is not.

On over-reporting, the model constrains the design only partially, and less precisely than a preset civil penalty would. Proposition 1 requires the expected cost of examination to exceed $c_a$ for any interior threshold to exist, but says nothing about where between $c_a$ and infinity it

falls. Reaching q_social would take an expected cost near $4,410 at $\pi = 0.15$, well above the roughly $150 the existence condition requires. Because an examination-based design sets $\pi$ through the intensity of supervisory inquiry rather than a figure legislators write down, over-reporting protection cannot be pinned down as tightly here. That limitation errs toward safety: across the range between $150 and $4,410, an incompletely specified examination regime is likely to generate too little expected cost than too much, so its dominant risk runs toward under-reporting.

Two further features bound whatever over-reporting risk remains. First, the social threshold already reflects the cost of a Type I error, since c_f enters q_social directly; even a regime that hit the social threshold precisely would not be "too protective" by this model's own accounting. Second, the design does not tie examination frequency to an institution's propensity to place holds, the mechanism by which over-reporting could compound within an institution over time, a point beyond this static framework. Supervisors implementing the design should nonetheless watch for it.

Regarding compliance cost, that concern is valid, and this paper is agnostic. Smaller firms will find the new requirement most burdensome, and, as Section II.C notes, may also be the firms whose customer relationships make detection most likely, so a clumsy mandate risks suppressing the very transactions it means to surface. Any enacted statute should pair the duty with a safe harbor or a tiered compliance schedule keyed to asset size, a design question this paper leaves open.

## X. Conclusion

EFE is not a failure by seniors to stay informed, nor simply a failure of financial guardians. It is a systemic failure with a known legal cause: banks have no duty to report elder financial abuse. New Jersey's statute, and the Lucca decision construing it, makes that concrete rather than hypothetical. The model shows this is not a near miss but a hole in the system itself, one no enforcement effort will plug on its own and that broader immunity will not fix. Any solution faces two tensions: protecting seniors against respecting their autonomy, and controlling compliance costs without pushing seniors toward unregulated predatory services. Closing the reporting-duty gap will not end EFE, but it does resolve the part lawmakers can solve outright.

Research into the biology of EFE explains individual harm; economic research shows broad economic losses; and regulatory analysis shows that existing policy and infrastructure require drastic reform. Taken together, these perspectives make clear that EFE is a systemic issue, not an individual one that financial literacy can fix.

## Methods Appendix

### Measuring Reporting Incentive Through a Calibrated Threshold Model

This paper models a bank's reporting decision as a function of its exposure to statutory liability that the reader can directly observe. I follow Becker in thinking of deterrence as a

function of the probability and magnitude of enforcement. Two threshold values were calculated for each of three counterfactual liability regimes, and are reported in Table A2: New Jersey's baseline permissive regime with good faith immunity and no penalty; New Jersey's baseline regime with absolute immunity replacing good faith immunity; and a mandatory-duty regime with absolute immunity and a statutory penalty. (These are variations on the parameters of the liability model, not the four statutory regimes described in Section VI. Those regimes are discussed there at length as a matter of law, rather than modeled separately.) It is the difference between the private and social thresholds, rather than either threshold value by itself, that is the subject of this paper.

**Parameter Selection**

Six inputs (L, r, c_a, c_f, s, D) and one legal fact (π) enter the model. Each is footnoted in the workbook's Sources tab with its citation, page number where available, and its status as sourced or assumed. The value of π is cited to the holding in Lucca. The other six were checked against primary sources (FinCEN reporting, FTC and AARP loss reporting, AML compliance-cost reporting) before being designated as assumptions.

**Sourcing Rule**

A parameter is marked sourced only if a primary source reports that exact number, not a loosely related or directionally consistent one; the rule was fixed before any parameter search began. Otherwise it is marked an assumption and bounded in the sensitivity analysis rather than left unlabeled. A directly reported figure still counts as an assumption when used as a proxy for something else. The recovery share r illustrates this: it is set to reported asset-recovery rates, the best available proxy for loss avoided through early intervention, but those rates condition on a recovery effort having started while r does not. That mismatch likely overstates r, so r is marked an assumption (proxy). A lower r would raise q_social and shrink the wedge, but leaves q_private and the central claim at $\pi = 0$ untouched, since r does not enter the private threshold.

**Sensitivity Rule**

L, r, and c_a were each swept independently over ranges fixed before the calibration was run: L from roughly half to nearly seven times its baseline value, r from about one-fifth to just over baseline, and c_a from about one-quarter to just over three times baseline. The wrongful-hold cost, c_f, was held at baseline rather than swept; because it appears in both the numerator and denominator of q_social, varying it shifts the social threshold without changing its sign, and it does not affect the paper's central finding at all, since that finding rests on the private threshold, whose ratio does not include c_f.

**General Limitations**

Four limitations bear on these results. First, litigation cost is assumed to fall privately on the bank rather than on society; if it carries a deadweight component, mandatory duty paired with

absolute immunity becomes more attractive still, so the assumption is conservative. Second, the compliance officer's belief q is taken as given: the paper addresses how a rational bank behaves once it forms that belief, not how well banks form it. Third, the model is static, so it omits reputational effects, repeat play with regulators, and supervisory pressure exercised outside this liability regime.

Fourth, and most consequential for the size of the wedge, the model assumes the bank captures none of the loss that early intervention would prevent. There is no $r \times L$ term in the private expected-cost expression, reflecting current law that a receiving bank owes no duty to absorb losses from a duly authorized payment order, even one procured by fraud. If banks instead internalized some fraction of $r \times L$, whether by statute or scheme rule, the private threshold would fall by that amount, closing part of the gap through a different mechanism than the reporting duty modeled here. That possibility is not modeled.

**Verification Limitations**

The parameters in this paper were compiled and coded by a single author, without independent cross-checking, which is a genuine methodological limitation relative to a multi-coder design. Readers validating a parameter should consult the original source in Table A1 directly rather than this paper's presentation. The author can provide the underlying workbook upon request.

**Falsification Condition**

The paper's central finding, that social and private reporting thresholds diverge under a permissive, no-liability regime, does not depend on the calibration: setting $\pi = 0$ in the expression for $q_{private}$ simplifies it to $1 + c_a / (s \times D)$, strictly greater than one for any strictly positive $c_a$ and $s \times D$. Because no plausible parameterization could close that gap while the statute remains permissive, the sensitivity analysis in Table A3 tests only the size of the social threshold, not the central result; s and D were not swept for the same reason.

Only two things could overturn this finding: a specification error in the derivation of $q_{private}$, or a demonstration that an omitted cost or benefit in the private expected-cost calculation is nonzero and large enough to change the reporting decision under a permissive regime. The likeliest candidate for the latter, discussed above, is the assumption that banks capture none of the loss L that early reporting would prevent.

**Appendix Tables**

Tables A1 through A3 below show, respectively, the model's parameter values and sources, the privately optimal reporting threshold for each of the three modeled liability schemes, and the sensitivity of the socially optimal threshold to its inputs. These are taken from the calibrated workbook provided in this appendix.

Table A1. Model parameters, values, and sources

| Parameter | Symbol | Value | Source | Status |
|---|---|---|---|---|
| Loss per case | L | $21,902 | FTC, Protecting Older Consumers 2024-2025 (Dec. 1, 2025) | SOURCED (derived) |
| Recovery share | r | 0.72 | FTC, Protecting Older Consumers 2024-2025 (Dec. 1, 2025), citing FBI IC3 Annual Report 2024 | ASSUMPTION (proxy) |
| Administrative cost | c_a | $150 | Bounded above by LexisNexis-adjacent secondary estimates | ASSUMPTION |
| Wrongful-hold cost | c_f | $400 | No primary source | ASSUMPTION |
| Probability of suit | s | 8.0% | No primary source | ASSUMPTION |
| Defense cost | D | $25,000 | Bounded below by secondary small-business litigation estimates | ASSUMPTION |
| Probability of liability | π | 0.0% | Lucca v. Wells Fargo Bank, N.A., 441 N.J. Super. 301 (Law Div. 2015) | SOURCED |
| Penalty if liable | H | $0 | N.J.S.A. 17:16T-1 to -4 (statutory text) | N/A (factual, not estimated) |

Note. Baseline values for the Section VII model; status follows the sourcing rule above. Loss per case is derived (aggregate reported losses divided by reports filed by persons 60+), likely biasing L and therefore q_social upward, which cuts against the paper's own argument. Recovery share is a proxy and likely biased upward as well, shrinking the wedge, though q_private is unaffected. Probability of liability is a legal certainty rather than an estimate: a permissive statute with no private right of action makes silence costless per Lucca, and H is zero for the same reason. The four remaining parameters could not be tied to a primary figure and are bounded in the sensitivity analysis.

Table A2. Private reporting threshold under three modeled liability configurations

| Regime | Duty to report | Immunity | Penalty | Private threshold | Bank reports? |
|---|---|---|---|---|---|
| New Jersey today | Permissive | Good-faith | None | No solution: silence dominates | Never |
| Absolute immunity only | Permissive | Absolute | None | Undefined | Never |
| Mandate + immunity + penalty | Mandatory | Absolute | $29,399 | 3.40% | Yes |

Note. Rows are counterfactual arrangements of the model's liability rules, not the four statutory regimes of Section VI; L, r, c_a, and c_f hold their Table A1 baselines throughout. New Jersey today sets $\pi = 0$ with good-faith immunity, so $\pi \times H = 0$ and the expected cost of silence is zero at every q while acting costs at least c_a, making acting strictly dominated rather than merely unlikely. Absolute immunity lowers the cost of acting to c_a but leaves that logic intact. The third row sets $\pi = 0.15$ and $s = 0$; H_min = $29,399 is the value at which q_private = q_social, offered to illustrate scale rather than as a recommendation. "Bank reports?" indicates whether the

private threshold is achievable on the unit interval, not observed behavior, and concerns only the state-law referral channel.

**Table A3. Sensitivity of the social threshold**

| Parameter varied | Low value | q* at low | High value | q* at high | Span |
|---|---|---|---|---|---|
| Loss per case (L) | $10,000 | 7.24% | $150,000 | 0.51% | 0.51% to 7.24% |
| Recovery share (r) | 0.15 | 14.92% | 0.85 | 2.89% | 2.89% to 14.92% |
| Administrative cost (c_a) | $40 | 2.72% | $500 | 5.57% | 2.72% to 5.57% |

Note. Each parameter was swept independently, all others held at their Table A1 baseline; thresholds follow q_social = (c_a + c_f) / (r × L + c_f). The private threshold is omitted because it depends only on c_a, ranging from 1.02 to 1.25 across the tested values. The wrongful-hold cost c_f was held at baseline, since it enters both the numerator and denominator of q_social and does not enter q_private at all.

**Works Cited**


Barnett Bank of Marion County, N.A. v. Nelson. 517 U.S. 25, Supreme Court of the United States, 1996.

Becker, Gary S. "Crime and Punishment: An Economic Approach." Journal of Political Economy, vol. 76, no. 2, 1968, pp. 169-217.

Blackbird.AI. "AI-Powered Obituary Scams: The New Frontier in Targeted Phishing." Blackbird.AI, 20 May 2025, blackbird.ai/blog/ai-powered-obituary-scams-targeted-phishing/.

Button, Mark, and Cassandra Cross. Cyber Frauds, Scams and Their Victims. Routledge, 2017.

Cross, Cassandra, and Michael Kelly. "The Problem of 'White Noise': Examining Current Prevention Approaches to Online Fraud." Journal of Financial Crime, vol. 23, no. 4, 2016, pp. 806-18.

DeLiema, Marguerite. "Elder Fraud and Financial Exploitation: Application of Routine Activity Theory." The Gerontologist, vol. 58, no. 4, 2018, pp. 706-18, doi.org/10.1093/geront/gnw258.

Federal Trade Commission. Protecting Older Consumers 2024-2025: A Report of the Federal Trade Commission. 1 Dec. 2025, www.ftc.gov/reports/protecting-older-consumers-2024-2025-report-federal-trade-commission.

Fenton, Laura, et al. "Lower Entorhinal Cortex Thickness Is Associated with Greater Financial Exploitation Vulnerability in Cognitively Unimpaired Older Adults." Cerebral Cortex, vol. 34, no. 9, 2024, article bhae360, doi.org/10.1093/cercor/bhae360.

Financial Crimes Enforcement Network (FinCEN). "Elder Financial Exploitation: Threat Pattern & Trend Information, June 2022 to June 2023." Financial Trend Analysis, U.S. Department of the Treasury, Apr. 2024.

Financial Crimes Enforcement Network (FinCEN). "TD Bank Enforcement: Assessment of Monetary Penalty Against TD Bank N.A. for Historic AML Failures." U.S. Department of the Treasury, 2024.

Gotelaere, Sien, and Letizia Paoli. "Prevention and Control of Financial Fraud: A Scoping Review." European Journal on Criminal Policy and Research, vol. 31, 2022, pp. 1-21, doi.org/10.1007/s10610-022-09532-8.

Griffin, John M., and Kevin Mei. "How Do Crypto Flows Finance Slavery? The Economics of Pig Butchering." Working paper, University of Texas at Austin, McCombs School of Business, 2024.

Gunther, Jilenne. The Scope of Elder Financial Exploitation: What It Costs Victims. AARP Public Policy Institute, 2023, doi.org/10.26419/ppi.00194.001.

Interagency Statement on Elder Financial Exploitation. Board of Governors of the Federal Reserve System, Consumer Financial Protection Bureau, Federal Deposit Insurance Corporation, Financial Crimes Enforcement Network, National Credit Union

Administration, Office of the Comptroller of the Currency, and state financial regulators, 4 Dec. 2024.

Kim, Dong-Hyeon, et al. "Population Aging and the Nexus between Financial Development and Wealth Inequality." International Review of Finance, vol. 25, no. 4, 2025, article e70048, doi.org/10.1111/irfi.70048.

Lachs, Mark S., and S. Duke Han. "Age-Associated Financial Vulnerability: An Emerging Public Health Issue." Annals of Internal Medicine, vol. 163, no. 11, 2015, pp. 877-78, doi.org/10.7326/M15-0882.

Lucca v. Wells Fargo Bank, N.A. 441 N.J. Super. 301, Law Div., 2015.

McAfee. Beware the Artificial Impostor: A McAfee Cybersecurity Artificial Intelligence Report. 1 May 2023, www.mcafee.com/content/dam/consumer/en-us/resources/cybersecurity/artificial-intelligence/rp-beware-the-artificial-impostor-report.pdf.

Modigliani, Franco, and Richard Brumberg. "Utility Analysis and the Consumption Function: An Interpretation of Cross-Section Data." Post-Keynesian Economics, 1954, pp. 388-436.

Nicholas, Lauren H., et al. "Financial Presentation of Alzheimer Disease and Related Dementias." JAMA Internal Medicine, vol. 181, no. 2, 2021, pp. 220-27, doi.org/10.1001/jamainternmed.2020.6432.

Prenzler, Tim. "What Works in Fraud Prevention: A Review of Real-World Intervention Projects." Journal of Criminological Research, Policy and Practice, vol. 6, no. 1, 2020, pp. 83-96.

Thaler, Richard H., and Cass R. Sunstein. Nudge: Improving Decisions about Health, Wealth, and Happiness. Yale University Press, 2008.

**The Reliability Cost of Climate Volatility: A Comparison of EDES and IES Portfolios**
**By David Park**


## Abstract

As climate change increases weather volatility, the reliability and cost-effectiveness of traditional Environment-Dependent Energy Systems (EDES), such as solar, wind, and hydro, are being called into question. This paper introduces the concept of Risk-Adjusted Delivered Cost (RADC) to evaluate power portfolios by incorporating not only the initial Levelized Cost of Energy (LCOE) but also the necessary costs of system integration (storage, firming) and the financial tail risk from extreme weather events. We hypothesize that while EDES possess a lower "sticker price" (LCOE), Environment-Independent Energy Systems (IES) (such as nuclear and advanced geothermal) become more cost-effective for achieving required grid reliability targets: 1. When backup/storage costs exceed twice the LCOE difference. 2. Extreme weather events, which cause massive economic losses (e.g., Winter Storm Uri's $80–$130B in Texas), occur more frequently than once per decade. By analyzing metrics like the Environmental Dependence Index (EDI) and Climate Volatility Exposure (CVE), and conducting case studies on events like the Texas Winter Storm Uri (2021), this research aims to provide an evidence-based direction for future renewable investment, arguing for increased deployment of IES and firmed EDES to ensure a stable, cost-efficient, and resilient net-zero power grid.


## Introduction

The effectiveness of power systems revolves around averages: average load, average seasonal effects, average costs, and average environmental conditions. However, due to the growing impacts of climate change and overall environmental volatility, planners of such power systems often find that significant deviations from the norm become increasingly common. Across the Electric Reliability Council of Texas (ERCOT), the technological systems with the lowest advertised cost of energy (LCOE) are also the very systems whose output swings most sharply in the wake of extreme weather events such as storms and droughts (Bernecker).
This tension between seemingly low initial sticker prices and unreliable output has become increasingly evident, as the implications of climate change continue to increase the frequency and severity of environmental disasters. In recent years, heavy storms and rainfall have increased notably in once-stable environments, while some warmer regions have seen virtually no rain. For example, the state of New York has experienced a 10% increase in overall precipitation from 1901 to 2022, with precipitation levels fluctuating to a much greater degree and estimates of up to a 17% increase by the end of the 21st century ("Precipitation"). Currently, with New York aiming to deploy 1,000,000 solar panels in the near future to reduce rising carbon dioxide emissions, questions will surely arise about their effectiveness given the state's current climate trends (Anke). Such illustrations highlight the sheer magnitude of climate change and the danger it poses if governments continue to pour massive amounts of money into inefficient energy systems. Already, a record $386 billion has been invested in renewable infrastructure in the first

half of 2025, with further investment expected (“Global Renewable Energy Investment”). Yet, there is an understated but increasingly critical problem that goes largely unnoticed: not all renewable energy systems are effective at sustainably providing ample power.
Take Lake Mead, for example, which powers the Hoover Dam: it once held a maximum capacity of 26 million acre-feet of water, making it an integral source of energy, whereas it now holds only 33% of that capacity (Edwards). As a result, the dam's energy output has been significantly reduced, forcing individuals in the surrounding region to pay more per megawatt-hour.

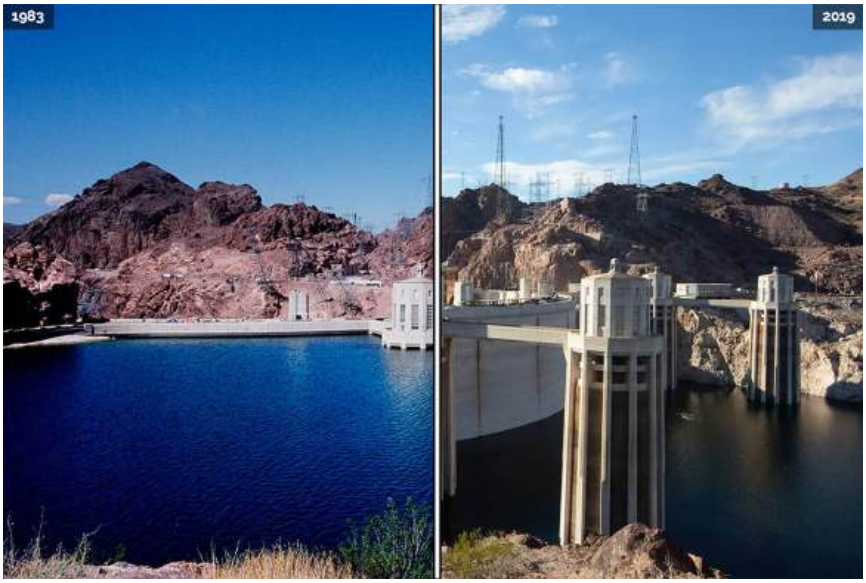


Figure 1. Hoover Dam and Lake Mead water levels

The figure above illustrates the stark difference in water capacity from 1983 to 2019. The amount of water lost is clear. With the major drought the Colorado River Basin region has experienced over the past few decades (an unprecedented, prolonged event that hints at the influence of a changing climate), it is clear that the dam has struggled in recent years (“Colorado River Basin”). Such systems, although once incredibly beneficial to the United States, have now aged and become less efficient to manage. However, simply abandoning all effort toward these projects is not an optimal path forward either, as stable energy output remains necessary in a way that does not further harm the environment.

To address this predicament, many nations have set long-term, climate-focused renewable energy objectives that include greater shifts toward renewable energy systems. Each method of power generation carries unique risks and economic and environmental sensitivities that affect cost-benefit forecasts for long-term power generation. Environmental tolerance and general output elasticities vary across systems, and many require relatively stable or predictable climate conditions to be effective. However, as the global climate continues to evolve, power generation systems that are more environmentally sensitive will face greater challenges with output variability, asset depreciation, maintenance, and other factors that could depress or upend their cost-benefit metrics. Forms of clean energy that do not rely on the surrounding environment as

heavily or at all do exist, such as nuclear and geothermal energy, yet it is quite likely that their research, development, and investment are lacking relative to energy systems that depend on the environment.

With numerous nations across the globe setting targets to reach certain carbon emission levels, such as the Paris Agreement and the net-zero emissions goal by 2050, it is absolutely imperative to determine whether the energy systems being invested in and promoted in the future are truly reliable. To move forward with our research analysis, some key terms must first be defined so that all information introduced in this paper is clearly understood and addressed. Firstly, an energy system is defined as a single site, project, or structure that generates electricity for the surrounding region through a known or conventional method. Environment-dependent energy systems, or EDES, rely heavily on specific aspects of the local climate, weather, and environment to produce reliable electricity, including solar, wind, hydroelectric, and tidal power. Environment-independent energy systems, or IES, are systems that do not rely on local climate or weather to produce reliable electricity, including nuclear, fossil fuel, biomass, and geothermal power. Although geothermal power is viewed by some individuals as environment-dependent, it will not be categorized as such for the purposes of this study, given that once a location with the required heat capacity is identified, geothermal power sites largely produce energy quite stably. To determine whether these energy systems are truly effective at providing power, they will be analyzed for their reliability, that is, their ability to produce electricity consistently over the expected period without requiring extensive monitoring beyond conventional servicing. Reliability will be measured using various metrics, which will be introduced and defined later in the methodology. With these definitions in mind, this research paper seeks to answer the question: are environment-dependent or environment-independent energy systems more reliable for the United States of America in the future, considering the extreme weather associated with climate change?

To anticipate potential questions and concerns in the United States of America regarding reliance on environment-dependent energy systems, we must first understand the extent to which such systems are currently used in the country. In 2023, 60% of the United States' electricity generation relied on fossil fuels, while 21.4% was derived from renewable energy sources (“What Is U.S. Electricity Generation”). This is understandable, as sources such as natural gas, coal, and petroleum are already established, trusted, and heavily utilized by Americans today. It can also be said that these sources are more cost-efficient, as they are cheaper to obtain and implement, making them more desirable from an economic perspective. Among the 21.4% of energy generated via renewables, wind, hydropower, and solar ranked at the top of the list, producing 10.2%, 5.7%, and 3.9% of the 21.3% share, respectively (“What Is U.S. Electricity Generation”). It is crucial to note that all three of these sources rely heavily on the environment, the implications of which we will discuss later. On paper, these statistics check out, as the United States has long relied on these sources to meet its renewable energy needs, using large dams, wind turbines, and solar panels. Yet, with increasingly variable conditions, these systems are slowly becoming outdated and less efficient. The previous example of Lake Mead and the

Hoover Dam is a perfect illustration of this truth, as the Colorado River System that feeds both has lost approximately 10 trillion gallons of water since 2000 (Canon). This has resulted in greater inefficiency at the Hoover Dam, as it struggles to generate the same amount of electrical power it once did due to the water loss. Not only does this display dangerous signs of climate dependency, but it is also becoming much more costly to maintain with fewer results. Such a growing predicament illustrates just one of the many increasing questions that come with environment-dependent energy systems that rely on stable climates to remain efficient, and as global weather patterns continue to fluctuate in more erratic manners, further consideration must be conducted in order to ensure whether or not such energy sources are truly viable for the future. As the United States continually pushes its renewable targets, it cannot be mindless and ignore the quiet yet growing flaws of hydropower, wind, and solar power, among others, signaling the necessity of this research.

**Methodology**

To best answer this question, a quantitative analytical research method was used to assess the reliability of a particular energy system by breaking it down into three major factors. The first factor is its cost-effectiveness, which will be analyzed using two major metrics: the Levelized Cost of Energy (LCOE) and the Internal Rate of Return (IRR). The LCOE will be particularly useful in determining whether the electricity produced over a particular source's lifetime is truly of value, given the initial investment and maintenance required to keep it operating. The IRR helps indicate the financial risk of investing in a particular energy source; if the value is low, it is likely not worth investing heavily in that source. The second factor in determining the reliability of energy sources is their stability, which will be assessed using the Effective Lifespan Ratio. This metric will help determine whether specific energy sources truly last for their initially advertised duration and will be crucial, given that the United States will continue to require constant power. The third factor is viability, and it will be analyzed using the System Average Interruption Duration Index (SAIDI as well as the Intermittency Index. The SAIDI will be crucial in determining whether an energy source can truly support the United States' power demands, as frequent energy cuts can quickly become catastrophic if additional measures are not taken. The Intermittency Index will be used to determine the variability of a source's output, as frequent variability can be quite damaging to the source.

Overall, the quantitative measures used to methodically analyze potential energy sources are intended to support more concrete, data-driven decisions about which sources may be best for the United States of America. Given that the rush for renewables has generally been on an increasingly quickening rise, it is imperative to take a step back and truly determine what will provide the greatest value for the inevitably large investment that each will require, so that the country will be receiving not an alternative in power, but a more optimal form of it.

**Results**

***U.S. Grid Analysis***

Given the sheer diversity and expanse of the United States of America, it encompasses numerous regions with unique climates that shape how energy systems are developed to best meet local needs. Therefore, it is crucial to understand these variabilities within the nation's grid so that more accurate conclusions can be made about what is truly best for the United States of America moving forward. The United States power grid is organized into three major regions: the Western Interconnection, the Eastern Interconnection, and the Texas Interconnected System. Each regional system has its own unique resource endowments, infrastructure challenges, and strategic factors, limiting the efficacy of a "one-size-fits-all" approach to managing and developing energy infrastructure. Driven by goals for a net-zero future for American energy production, plans across these regions show substantial growth in the use of renewable energy systems (clean energy), largely composed of environment-dependent systems like solar- and wind-based generation infrastructure. For instance, the Western Interconnection plans to add approximately 172 gigawatts of new capacity, with a large share coming from solar and wind power. This trend toward high EDES penetration is occurring as the overall demand for electricity continues to rise (U.S. electricity consumption has increased 3.8% since 1990) ("Use of Electricity"). However, this reliance on EDES infrastructure introduces significant, somewhat novel systemic complexity and cost to the North American energy grid. Using System-Level Cost Adders, we show that EDES-heavy energy portfolios incur high system integration costs that environment-independent energy systems (IES) largely avoid. The Achilles' heel of the North American energy grid is its ailing, antiquated transmission network, and while we are by no means the first to highlight this flaw, the optimal plan to address this issue remains elusive. Within North America, Canada also has its own power grid system. As for Canada, its government has approved 71 deployment projects, 26 of which are commissioned, with 496 megawatts of capacity planned for installation ("Western Assessment").

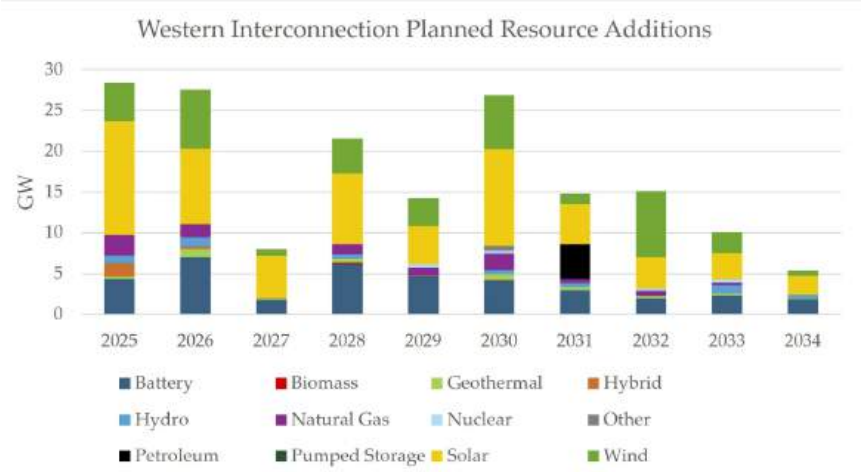



Figure 2. Western Interconnection planned resource additions

It is vital to understand which renewable energy sources will provide the most viable and stable output given North America's continually increasing energy demands. Within just five years, all three energy sectors, commercial, industrial, and residential, increased sales by 2.6%, 2.1%, and 0.7%, respectively ("Use of Electricity"). This is simply a five-year difference, and it clearly highlights an ever-growing, widespread demand for energy in the United States of America. Although 2020 had comparatively low energy demand, this does not mean that energy demand has not continued to rise throughout history; electricity consumption has increased by 3.8% since 1990 ("Use of Electricity"). The number may seem small in isolation, but given that its units are in billions of kilowatt-hours and it provides energy for a first-world country, the percentage is quite substantial. Yet the country has provided relatively stable energy; however, the distribution of available electricity is severely underfunded and experiencing significant degradation. To transport electricity generated for citizen use, the power must be distributed through extremely bulky, clunky transmission lines built decades ago and receiving minimal, if any, updates. This has been met with considerable backlash, as these lines have been seen as an eyesore by residents, yet they are still crucial for producing the energy they require. This issue becomes even more prominent when one realizes that energy-dependent sources are often location-dependent as well. Dams or long lines of solar panels cannot simply be built and placed anywhere, so the energy produced at these sites would need to be transported to urban areas for effective use. Yet, if production of these energy sources only increases, then the production voltage transmission lines would also have to be increased, further aggravating the issue of them being an eyesore.

***Energy System Analysis***

In February of 2021, Winter Storm Uri hit the state of Texas and exposed one of the most integral factors to an energy system's success: viability. When the severe weather hit the state, numerous blackouts occurred, with 69% of Texans reporting power outages at some point during the storm, and 49% reporting issues with water service. The storm also caused the deaths of a minimum of 210 individuals, with a lack of stable energy and water being a substantial cause. From a financial standpoint, the state lost about $80–130 billion due to the storm's impact, with key economic sectors such as supply chains and agricultural sites struck and damaged. This year, natural gas accounted for 51% of Texas's energy-generating capacity, with other renewable sources, such as wind, accounting for 24.8% (see Figure 6). Clearly, wind and other environment-dependent sources were especially ineffective during this period, revealing their key dangers. As climate-related variables continue to rise, relying on such sources would present an increasingly significant risk that must be analyzed.

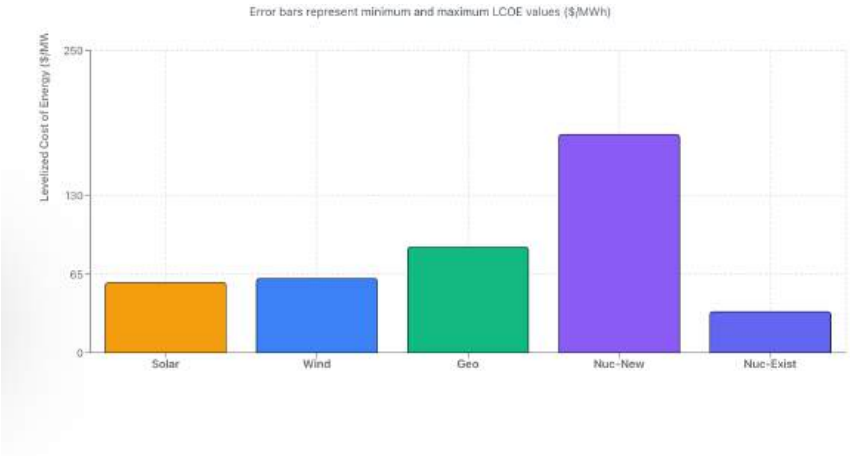



Figure 3. Levelized cost of energy by technology

First, to determine whether a particular energy source is economically efficient to implement at such large scales across the United States, the LCOE, or Levelized Cost of Energy, metric is used, measured in dollars per megawatt-hour. To serve as a benchmark for comparison with the following energy sources, the LCOE for gas and coal in America ranged from 149–251 and 71–173, respectively.

The LCOE of Solar PV utility was 38–78, clearly indicating a lower range and overall interval than those of gas and coal. This, however, is to be expected, as there is a clear reason why governments, private companies, and homeowners are choosing to rely on carbon-intensive sources that are more cost-effective and can generate more power at the same price. Additionally, solar production simply cannot match the continuous electricity generation that fossil fuel sources can provide, given its reliance on sunlight. In regions where the weather is warm and sunny for most of the day, solar power can be an effective way to provide clean, reliable electricity in the surrounding area. However, given that solar panels are deployed not only in sunny regions but also in regions with widely varying climates, this investment can become quite risky and ineffective, as solar's ability to produce electricity will inadvertently become more costly. Even so, cities are continually rushing to invest heavily in this source, with New York City planning to install 6 gigawatts of solar panels. As previously mentioned, however, the city's precipitation levels have risen to unusually high levels over the past decade or so, so this investment does not seem particularly shrewd.

This is where another key issue arises, seemingly benevolent on the surface, but potentially incredibly dangerous in the future. Nearly all states within the U.S. continue to roll out plans and enact legislation at various summits to assert that they will meet certain carbon-intensity targets in the coming years. Again, New York is a clear example of this, as it outlines plans to reach net-zero carbon-intensity rates as soon as possible in its Scoping Plan. These targets are often politically effective in encouraging citizens that governments are doing their best to clean up the environment, yet in reality, they create a rushed process that leads governments to focus less on the actions they are actually taking and more on simply

demonstrating that they are utilizing renewable technologies. With solar PV's lower LCOE, however, greater consideration must be given to how it is invested and developed. This is not to say that solar power is entirely inefficient and unreliable; however, it is to say that governments should take greater time and effort to analyze where it would truly be most efficient, given the fact that the citizen's tax money is being used to pay for significant portions of these projects, so it is only fair that the energy they provide is worth it in terms of value.

When analyzing the LCOE rates of these five forms of energy, it is clear that the new nuclear energy plants are much more expensive than the other four, with an LCOE of approximately 190.

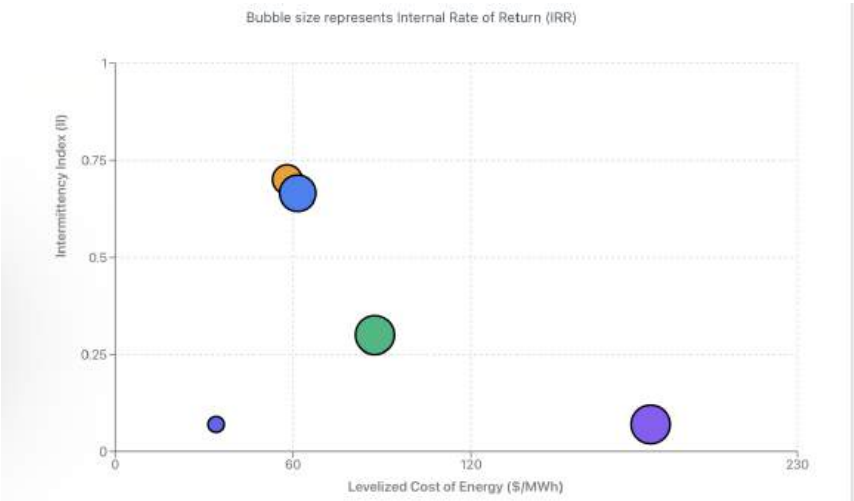



Figure 4. Cost-intermittency trade-off analysis

However, when the nuclear plant is pre-existing and the initial cost of developing the plant itself is not considered, it is, in reality, the cheapest of the five sources. Therefore, compared to what numerous politicians may proclaim about the economic prowess of solar—or wind-powered energy sources, nuclear power, often cited as a popular and well-known path to net-zero carbon emissions, may in reality provide a more efficient alternative to cleaner energy sources. This statistic often illustrates the goal of economists who advocate for more intensive use of nuclear energy. They point out that, in the long run, although nuclear power plants require higher upfront capital investment, they can recover those initial construction costs over time because they generate electricity at a lower long-term cost per kilowatt-hour than solar and wind power.

This graph provides yet another illustration of the cost, LCOE, on the x-axis and the intermittency index on the y-axis, showing a particular energy source's ability to produce reliable power. To put it simply, the higher a particular energy source is on the y-axis, the more variability exists in a source's ability to provide functional power. As the graph shows, both new and existing nuclear plants are much more efficient at maintaining reliable power output than

solar and wind power plants. This is often attributed to the fact that these sources are highly dependent on the conditions of their surrounding environment. More specifically, wind power, of course, depends on ample wind to drive the turbines and produce electrical energy, while solar panels depend on the reliable availability of sunlight. However, nuclear power and, surprisingly, geothermal power are often more reliable than those that are more heavily marketed politically, suggesting a more economically viable approach.

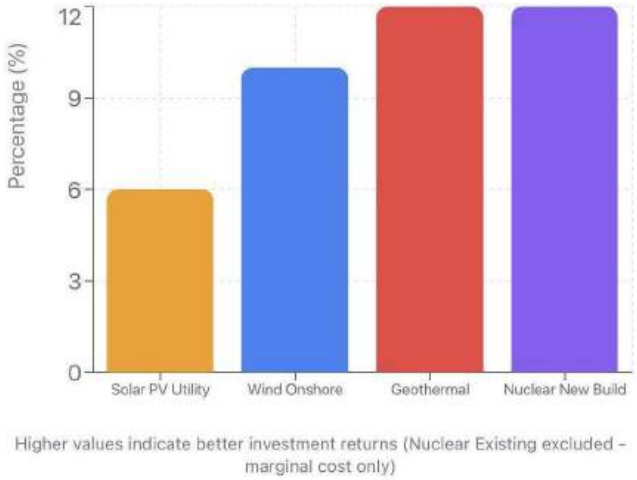



Figure 5. Internal Rate of Return (IRR) by technology. Higher values indicate better investment returns (Nuclear Existing excluded; marginal cost only).

The next metric used to determine the economic value of a particular energy source is the IRR, or Internal Rate of Return. Based on a previous study, an IRR of 8–10% was determined to be required for an investment in a particular energy source to be worthwhile, following the traditional IRR for coal plants. However, our analysis finds that the IRR for solar power ranges from 4–8%, with only its upper bound meeting the generally accepted minimum threshold for investment viability. This metric further illustrates that a heavy reliance on solar utility can be economically inefficient. If it were used as an auxiliary or secondary source, there may be a stronger case to utilize solar. However, given that the annual return rate of solar is less than half that of fossil fuels, it is clear that it is not effective enough to serve as a primary energy source. The IRR metric further supports this pattern. A prior study found that an IRR of 8–10%, based on the traditional IRR of coal plants, is generally needed for an energy investment to be considered worthwhile. This analysis finds that solar power's IRR ranges from 4–8%, with only the upper bound meeting this threshold. This suggests that heavy reliance on solar carries meaningful economic risk; solar may be better suited as a secondary or auxiliary power source rather than a primary one, particularly given that its annual return can fall to less than half that of fossil fuel investments.

**Table 1: Comprehensive Energy Technology Performance Metrics**

| Technology | LCOE ($/MWh) | IRR (%) | ELR (years) | SAIDI* (min/yr) | II |
|---|---|---|---|---|---|
| Solar PV | 38–78 | 4–8 | 1–1.1 | 118.4 | 0.7 |
| Wind Onshore | 37–86 | 8–12 | 1.5 | 123.8 | 0.665 |
| Geothermal | 66–109 | 12–12 | 1 | 123.9 | 0.3 |
| Nuclear New | 141–220 | 12–12 | 1.75 | 123.9 | 0.07 |
| Nuclear Existing | 30–38 | N/A† | 1.5–2 | 123.9 | 0.07 |

**Abbreviations:** LCOE = Levelized Cost of Energy; IRR = Internal Rate of Return; ELR = Energy Load Ratio; SAIDI = System Average Interruption Duration Index; II = Intermittency Index

* SAIDI values exclude major events

† Nuclear Existing represents marginal operating costs; IRR not applicable

Table 1. Comprehensive energy technology performance metrics.

To determine the reliability of a particular energy source, the ELR (Effective Lifespan Ratio) metric can be used to estimate its expected lifespan. This can be found by dividing the actual effective operational lifespan of a source by its designed lifespan. Given this, it is crucial that any energy source has an ELR of at least 1.0 to ensure it operates as expected. Generally, for natural gas and coal plants across the United States, the average ELR is around 1.1–1.75, indicating they tend to perform slightly longer than expected. For Solar PV, the average ELR is 1.0–1.1 per year, indicating that large arrays of solar panels often perform as expected or slightly better. Although this is slightly less efficient than fossil fuels, solar panels still operate as expected. Thus, in terms of reliability, solar panels can remain useful throughout their lifespan, making them strong in this regard.

Another metric used to assess the reliability of energy sources is the SAIDI (System Average Interruption Duration Index). This index calculates the sum of all customer interruption durations, i.e., the total minutes of electricity supply interruption, and divides it by the total number of customers served. The index is measured in minutes per year, i.e., the number of minutes energy systems were disrupted over the course of a year. However, for the pAIDI metric in particular, a clear distinction must be made to account for outliers. SAIDI is calculated in two ways: with and without major events. For example, when Winter Storm Uri hit Texas in 2021, all 254 counties were affected to some degree, with 69% of Texans experiencing a power outage at some point during the storm. Of course, a large-scale event such as this would skew the SAIDI for that region, so separate calculations must be made for it. Solar PV Utility in the U.S. experienced roughly 366.6 minutes of disruption per customer per year when major events were included, and approximately 118.4 minutes without. Clearly, solar power is much more vulnerable to major weather events; however, in the absence of such events, its reliability is on par with that of other energy sources, such as natural gas.

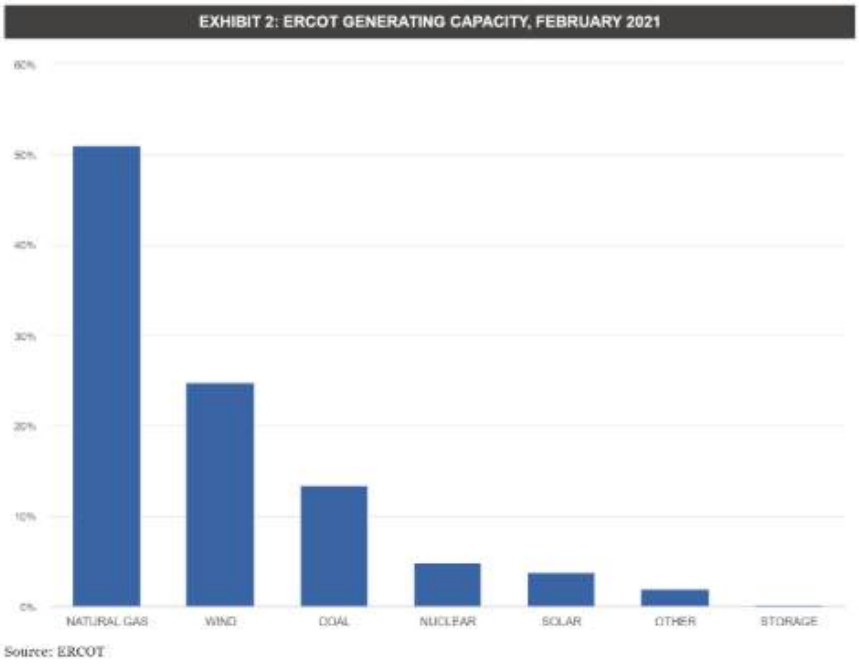



Figure 6. ERCOT generating capacity, February 2021

**Future Direction**

Ultimately, it has become clear that there are stark differences between the economic reliability and efficiency of environment-dependent and environment-independent energy sources. Power sources such as wind and solar are often too dependent on the environment and the stable availability of wind and sunlight. Although both of these sources have been researched and marketed to a greater extent, this study has shown that they are often less cost-effective at producing sufficient electrical energy. Additionally, other indices show that solar and wind power are both more prone to power cuts, which affect their ability to provide continuous electricity. The analysis above has illustrated the overall trend that environment-independent energy systems are more stable and cost-effective at generating reliable electricity. However, this is often counteracted by the fact that most politicians are pushing for increased use and development of more environmentally friendly energy sources.

Although with the arrival of President Donald Trump's second term, most of the plans and incentives to further push the usage of environment-dependent sources of energy have ended, under former President Joe Biden, several investments were made in order to increase the overall amount of renewable energy systems utilized in the United States, particularly illustrating a clear trend toward environment-dependent energy systems. However, as we have previously discussed, these forms of energy are often easier to market to the public because they have a lower upfront cost that taxpayers are more likely to pay. Yet, as we have analyzed, in the long run, environment-independent energy sources, such as nuclear power, are more effective and stable for producing continuous energy.

**Works Cited**


Anke. “New York Will Turn On 1,000,000 Solar Panels — 500,000 kW Project to Forget About Coal and Gas.” Energies Media, 8 Jan. 2026, energiesmedia.com/new-york-will-turn-on-1000000-solar-panels/.

Bernecker, J. “Levelized Production Cost Savings.” ERCOT, 19 Jan. 2021, www.ercot.com/files/docs/2021/01/14/Levelized_Production_Cost_Savings_RPG_01192021.pdf.

Canon, Gabrielle. “Western US States Fail to Agree on Plan to Manage Colorado River Before Federal Deadline.” The Guardian, 12 Nov. 2025, www.theguardian.com/environment/2025/nov/12/colorado-river-agreement-state-negotiators/.

“Colorado River Basin.” Drought.gov, U.S. Department of Agriculture, www.drought.gov/watersheds/colorado.

Edwards, Jess. “Where Lake Mead Water Levels Stand Amid Heavy California Rain.” Newsweek, 5 Jan. 2026, www.newsweek.com/where-lake-mead-water-levels-stand-amid-heavy-california-rain-11308356.

“Global Renewable Energy Investment Still Reaches New Record as Investors Reassess Risks.” BloombergNEF, 26 Aug. 2025, about.bnef.com/insights/clean-energy/global-renewable-energy-investment-reaches-new-record-as-investors-reassess-risks/.

“Precipitation.” New York State Climate Impacts Assessment, 26 Dec. 2023, nysclimateimpacts.org/explore-the-assessment/new-york-states-changing-climate/nysc-precipitation/.

United States, Energy Information Administration. “Use of Electricity.” EIA, 18 Dec. 2023, www.eia.gov/energyexplained/electricity/use-of-electricity.php.

---. “What Is U.S. Electricity Generation by Energy Source?” EIA, 29 Feb. 2024, www.eia.gov/tools/faqs/faq.php?id=427&t=3.

“Western Assessment of Resource Adequacy 2024: Planned Resource Additions by Subregion.” Electric Reliability and Security for the West, WECC, www.wecc.org/sites/default/files/documents/products/2025/WARA%202024%20Planned%20Resource%20Additions%20by%20Subregion%20v2.pdf.

**How Do Economic Systems Exploit Predictable Irrationality? A Game-Theoretic and Behavioral Analysis By Teddy Hsu**

**Keywords**

Game Theory, Psychology, Economic Systems, Irrationality, Psychological Manipulation, Loss Aversion, Sunk Cost Fallacy, Lottery, Penny Auction, Illusion of Control, Psychological Biases, Gambler's Fallacy, Behavioral Analysis

**Abstract**

We live in a world where trust is not so easily given as it once was before. Corporations are willing to do anything to get your money. Through all the tactics and manipulation, it becomes hard to discern what constitutes a rational choice.

This paper attempts to explain the rational analysis of economic choices using game theory and psychology. First, it will explain game theory through the classic and famous prisoner's dilemma example and describe its application in the fields of economics by connecting the prisoner's dilemma to the concept of price wars. Then, it will analyze different economic systems through the lens of game theory and psychology, including penny auctions, gacha games, and the lottery. Through this, we conclude that people may be drawn to make seemingly irrational decisions that lose money due to the exploitation of psychological principles in an economic system.

Overall, we show how human behavior can be predictable, and thus can be shaped and exploited for people's money. Through these examples, this paper argues that irrational behavior is not always random. Sometimes people are influenced by predictable psychological biases, while in other cases the structure of the system itself makes seemingly harmful behavior rational for the individual. Game theory and psychology help explain how human behavior can be predictable, and how economic systems can shape the choices people make.

**What is Game Theory?**

Have you ever wondered whether there is a way to fully rationalize human behavior into some type of quantifiable logical system? Whether we could create an equation that fully explains every single human behavior if we plug in the variables? Well, while I wish game theory was that, unfortunately it is not. But, it is one of the closest things we have to it!

Game theory uses logical and mathematical thinking to study the rational decision making of entities and how they interact with each other. By taking a "game" with multiple "players," which can be any interacting people or parties with choices that can give rewards or punishments, game theory allows us to study the results that come from these choices and analyze what the most rational choice would be for each entity. These games show up everywhere in the world, and prove that game theory is crucial to analyzing our decisions.

**Prisoner's Dilemma**

The most famous example of game theory by far is the prisoner's dilemma. I'm sure whether you know of game theory or not, you've been exposed to some version of it. The prisoner's dilemma, and variations of it, can be found in numerous situations today that all can be analyzed with game theory.

The "typical contemporary version" of the game is described in William Poundstone's 1993 book Prisoner's Dilemma, whose setup of the problem I will borrow:

"Two members of a criminal gang are arrested and imprisoned. Each prisoner is in solitary confinement with no means of speaking to or exchanging messages with the other. The police admit they don't have enough evidence to convict the pair on the principal charge. They plan to sentence both to a year in prison on a lesser charge. Simultaneously, the police offer each prisoner a Faustian bargain. If he testifies against his partner, he will go free while the partner will get three years in prison on the main charge. Oh, yes, there is a catch... If both prisoners testify against each other, both will be sentenced to two years in jail. The prisoners are given a little time to think this over, but in no case may either learn what the other has decided until he has irrevocably made his decision. Each is informed that the other prisoner is being offered the very same deal." (Poundstone, 1993) [(1)]

So, what would the best decision be? What about a decision where you are fully selfish or fully loyal to your partner? There are essentially four outcomes to this situation. You can either testify or stay silent, and your partner can also testify or stay silent leading to 2 * 2 = 4 outcomes. The outcomes can be viewed in the following grid in Table 1.

Table 1. Game table for the prisoner's dilemma

| **Prisoner's Dilemma** | **Partner Stays Silent** | **Partner Testifies** |
|---|---|---|
| **You Stay Silent** | You: 1 \ Partner: 1 | You: 3 \ Partner: 0 |
| **You Testify** | You: 0 \ Partner: 3 | You: 2 \ Partner: 2 |

Assuming you are looking out for your own benefit, if your partner stays silent, you'd rather testify as you'd be free instead of having to serve one year. However, if your partner testifies, it is also better to testify against him as it lessens your sentence from three years to two years. In this way, no matter what your partner does, the best decision for you is to testify. Now, if your partner is rational, they will also realize this, and thus both of you will testify and serve two years. By being rational, both you and your partner ended up choosing the option with the worst collective outcome. If both you and your partner had stayed silent, you both would have collectively got a shorter sentence. But, by being rational, you both end up worse off. The analysis of using rational reasoning to make your decision and to study the outcomes of each decision to help you make your choice is part of what game theory is all about.

This paper argues that irrational behavior is often not random. In many scenarios within economics, people make seemingly irrational choices because the system is designed and created

to exploit predictable irrational behaviors. Game theory helps explain the incentive structure of these situations: what a perfectly rational person, company, or country might do. Psychology helps explain why real humans often act differently from the rational model. When applied together, game theory and psychology show how institutions, companies, and governments can shape behavior by making certain choices feel urgent, rewarding, or necessary, even when those choices may irrationally harm the individual or society in the long run.

**Strategic Traps: Price Wars**

The prisoner's dilemma is useful because it shows that choices that are rational for individual benefit may be harmful for the collective benefit. In these cases, the problem is not necessarily that people are irrational. Instead, the issue is that the structure of the issue pushes rational actors toward escalation, competition, or distrust, rather than cooperation.

**Game Theory's Application in Economics (Price wars)**

Let's say there are two companies that sell similar products. Currently, both companies sell the product at $10 each and as such, they both have similar revenue from that product. However, each company has a choice to lower the price of the product to become cheaper and become the favorable product to buy for consumers. So, let's say lowering the price to $8 is an option for the company to take all the revenue from the two companies. Their profits would approximately double while the other company loses their profits as their revenue has shifted to the company with a lower price. In this case, "lowering the price" incurs the same effect as "testifying" in the prisoner's dilemma. The chart given in Table 2 explains this.

Table 2. Game Theory Table for Price Wars

| Price Wars | Company B keeps price | Company B lowers price |
|---|---|---|
| **Company A keeps price** | A: $10 \ B: $10 | A: $0 \ B: $16 |
| **Company A lowers price** | A: $16 \ B: $0 | A: $8 \ B: $8 |

From a game theory perspective, what to do in this scenario is identical in principle to the prisoner's dilemma. If company B keeps their price, company A would rather lower their price so they become more competitive than company B and take their customers. Their profits would increase from 10 to 16. (They have a lower price, but double the customers.) If company B lowers their price, company A would also rather lower their price so company B doesn't take all the profits and to stay competitive in the industry, bring their profits from $0 to $8. So, no matter what company B does, the optimal choice for company A to make as much money as possible would be to lower their price. Company B would have the same situation so they would also want to lower their prices. So, similar to the prisoner's dilemma, both companies sell the product

at $8 each where they could've instead each kept the same price and sold the product at $10 each. Both companies lose out where both could've profited more instead by cooperating.

In practice, this is not entirely bad. Competition between companies in price wars prevents monopolies from forming and helps keep prices fair. Companies want to keep a balanced price to stay competitive and make a good amount of money. Price wars show one major lesson from game theory: seemingly rational behavior can still produce harmful outcomes when cooperation is difficult. Each side has a selfish reason to protect itself by competing, even though both sides could mutually benefit through cooperation. These examples are different from later examples such as penny auctions, gacha games, and lotteries. In price wars and arms races, the main issue is strategic conflict. In the later economic examples, the problem is psychological manipulation.

**Economic Manipulation**

Economic systems are especially useful for studying irrationality because money gives us a clear qualitative measure to compare payoffs. If a person repeatedly spends more money than they are likely to gain, the behavior appears irrational from a financial perspective. However, these choices become easier to understand when we look at the design of the system. Penny auctions, gacha games, and lotteries all use small costs, rare rewards, and emotional pressure to make people continue participating despite the seemingly irrational outcomes.

While this is cool, the usefulness and practicality of such analysis may be questioned. I refute any claim of its apparent lack of use, and instead aim to show all the many applications of game theory in the world. Just from the example above, the prisoner's dilemma, the most famous form of game theory, is used in price wars, two very different scenarios. Game theory finds its application whenever there are at least two rational entities that want to maximize their benefits, which can be seen in pretty much every area of life.

However, what happens when humans seemingly make irrational decisions? What happens when they are manipulated into making the option that makes them lose more? In this case, one could turn to psychology as an explanation for these events. By studying human behavior, numerous psychological phenomena have been named that defy rational choice, instead leading to people making irrational choices. Analysis of both psychological effects and why such irrationality occurs in humans is a main topic of this paper. Through game theory and psychology, such irrationality will be analyzed and explained, and it will be shown how some practices and psychological vulnerabilities are exploited to manipulate humans into making these irrational decisions.

**What are Penny Auctions?**

Penny auctions are an interesting type of auction where you have to pay pennies to place a bid. Instead of getting to bid for free, you pay an incredibly small price to place a bid, at the same time restarting the timer of the auction, which is usually kept to be very short (around 10 seconds). Even if you lose the auction, you have to pay the price it cost you to place a bid, and

when you win the auction, you have to pay the price of the product as well as the prices of the bids you placed. So, while the seller usually will sell their product for less, they still get money from each individual bid placed to compensate.

On DealDash, one of the most popular penny auction sites, each bid raises the price by $0.01. The price of bids is 13 cents and each new bid brings the timer to 9 seconds. While buyers can sometimes snag good deals, sellers usually receive the profits and beneficial results. Every time the price goes up $1, that means 100 new bids have been placed, equaling $13 for the seller. For example, on DealDash, a Macbook and Briefcase product has a value of around $4,200. For the most recent winner (at the time of writing), the winning bid was $1,314.18. Seems good, right? They make a $3,000 profit! Well, when you take in the fact that the bidder had to place 26,010 bids to win that auction, the price they paid turns out to be $3,901.51. And then when you take into account that other people also paid money through many bids, as the auction had 131,418 bids in total, then you start to realize where the huge profit margins come from. Every penny added to the "price" is essentially 14 cents added to the seller's revenue. This means that this product would have to sell for $300 or less for the seller to lose money in the auction. From this one product alone, there are 24 sales saved in history, and only 7/24 sales came under that $300 price point. This was usually by a small margin too, while final prices that came above $300 sometimes stretched into the thousands. A study in 2015 found that: "As a result of bidder behavior, auctioneer profit margins from the original penny auction site empirically exceeded 50% across more than 166,000 auctions" (Augenblick, 2015)[(2)]. All this goes to show that while penny auctions can have deceptively low sticker prices, they prove highly profitable for sellers.

***Why Bid?***

So, if it has been proven how penny auctions generally always benefit the seller, then why still bid? Why would perfectly rational people know that they will most likely lose money? It's not that hard to see that unless you actually win the whole auction, you will always lose money. Even if you win, there are still high chances that you will overpay. To illustrate the rationale of bidding compared to abstaining, Table 3 provides a game theory analysis on whether a rational agent would bid in a penny auction.

Table 3. Game Theory Table for Penny Auctions

| **Penny Auctions** | **You bid** | **You don't bid** |
|---|---|---|
| **No one else bids** | +99 | 0 |
| **Someone else bids** | -1 | 0 |

As you can see, the game theory analysis shows how one might rationally decide to bid. When you bid, there's a chance for a huge payout, represented by +99, and if someone else bids after you, you only get a small loss, represented by -1. However, one issue with this

representation of the situation is that it inaccurately equates the chances of someone bidding vs. someone not bidding. There are many bidders for each auction, but only one winner, so you are more likely to lose small than win big. In this way, penny auctions are similar to a lottery system. As a 2009 study states, “On one hand, the auction appears to resemble a lottery, with large numbers of participants losing relatively little, one participant winning a significant prize, and the auctioneer making large profits” (Augenblick, 2009)[3]. Expected turnout is impossible to calculate, as it relies on the exact chances of someone else bidding. But, as the auction goes on, people are less likely to bid, being deterred from the higher price. While there may be the occasional win, most buyers lose money overall. As such, penny auctions have lost popularity over time. A quick search on Google trends[4] shows penny auction searches peaked in 2010, but by 2014, it died out and has remained stagnantly low ever since. Overall, the public has realized that it is generally a losing game and have stopped using it.

**So, It's Irrational?**

With such obvious metrics showing how bidders lose money on average, it should be clear why people wouldn’t want to bid on it. However, despite this seeming irrationality, people still bid on it anyway. The nature of penny auctions takes advantage of many psychological principles in order to get players.

First, penny auctions give players an “illusion” of control, making them believe that they can cheat the system to make a profit. According to Wang 2024, “The empirical analysis of DealDash auctions reaffirms that users tend to strategically time their bids to maximize their chances of winning. The optimization of bids, especially observed in the last seconds before timer expiration, underscores the influence of the pay-per-bid auction structure on participants' decision-making” (Wang, 2024)[5]. Players often believe that by bidding at the last second, they can strategically win the auction. Another factor that plays into this is the Fear of Missing Out, or “FOMO” that comes with penny auctions. As the little time dwindles out, the player doesn’t have the time to rationally process what to properly do. They develop a fear that they are missing out on a great deal by not bidding, and with the very short timer, many choose to bid on a split second decision. Once the first bid is made, you have now contributed money for the sunk cost fallacy to start. The sunk cost fallacy is the principle that states that if you spend time, effort, or money on something, you will be more prone to seeing it through, even if it hurts you more than it helps. As Arkes and Blumer put it, the sunk-cost effect refers to the tendency “to continue an endeavor once an investment in money, effort, or time has been made” (Arkes, 1985)[6]. Most people see the situation in an irrational manner, arguing that if they already spent money on a bid, quitting now would guarantee a loss, while continuing to bid may result in a win. A sense of pride kicks in, rationalizing that to quit now would only admit defeat. While human persistence is generally beneficial in staying motivated, here it is exploited to maximize company profits, and to take as much money as it can from the player.

One final principle that works to keep people bidding is loss aversion. The principle of loss aversion shows how people are especially averse to losing money. In cases where there is an

equivalent gain or loss, the equivalent loss hurts more than the equivalent gain makes someone feel good. When looking for good deals, people will be more conscious and analytic on whether a deal loses them money. Not losing money is more important than gaining money to them. When people stop bidding in a penny auction, they are admitting defeat. They know they will lose their money, instead of having a chance to snag a good deal. Because people are so adamantly against losing, it makes it that much harder for people to stop bidding once they start.

**Okay so..?**

Despite penny auctions seemingly being active and profitable for all, many players tend to realize the harms that come with continuously playing penny auctions. In fact, there are very few players of penny auctions who are considered "veterans" and a large majority of players seem to be new. According to Wang 2012, "A central finding of this paper is that BigDeal profits from a revolving door of new bidders, but loses significant amount of money to experienced bidders as a group" (Wang, 2012)[(7)]. Newer players are drawn into the system by the seemingly good deals, but after they lose, most of them realize the deceptive nature of the auctions and tap out, while the few experienced players are able to make as much money as they can.

Wang 2012 concludes that such auction sites are unsustainable, and it's not hard to see why. The system relies on the ignorance of players and their overconfidence in believing that they can fool the system. Penny auctions have essentially lost all forms of relevance in current times, and the penny auction site: "BigDeal.com" mentioned in Wang 2012, has had its domain sold to another entity.

But, why is this important? The nature of penny auctions teaches us that human rationality can and has been exploited before. While humans tend to believe they are rational, systems like penny auctions operating under the guise of rationality end up deceiving many people. In today's capitalist society, penny auctions aren't the only place where this is relevant. Many more of these systems operate in order to take the money of unassuming customers, and it is important to understand and analyze these predatory practices.

**Gacha Games: What are they?**

Unlike the prisoner's dilemma or price wars, gacha games are not a perfect example of classical game theory because the player is not competing strategically against another equal player. Instead, gacha games are better understood as decision-making under risk. The company designs the probabilities, rewards, visuals, and currencies, while the player decides whether to keep spending. This still fits the larger argument of the paper because the system uses predictable psychological biases to make continued spending feel rational or exciting.

Gacha games are games that rely on a "gacha" system for engagement and money as a main structure of their gameplay. "Gacha" derives from Japanese "gachapon", little vending machines that spit out a randomized toy from a certain selection. Usually, these games try to get players to spend in-game currencies on these gachas for a randomized chance to get an item of a

character, not unlike a loot box system. Some of the currencies can be obtained through free gameplay, but the ultimate goal is to monetize the currency.

This randomized design has been proven to be addictive, and it attracts many players to spend copious amounts of money. Despite being free to start playing, these games hide their true cost in microtransactions. The addictive feature has given rise to comparisons to gambling, and there has been pushback and resistance to regulate gacha games more like gambling is regulated. Despite this resistance, gacha games still prove to be extremely profitable. In December of 2025, the highest revenue grossing gacha games all made over 19 millions dollars, with the top game: Love and Deepspace, making over 80 million.[8] This allows gacha games to continue to make their games of higher quality, to keep player retention as high as possible and in turn, make even more money.

**The Gacha Illusion of low cost: Why do people buy?**

Gacha games have an inherent dishonesty in how they truly present the cost of items, often making it seem like it's lower than it actually is. Instead of buying characters or items outright, you instead have to go through a whole sequence of currencies to get the chance of the item you desire. By hiding the true cost of each character in a gacha mechanic and in numerous currencies, the true cost of each character becomes hidden. In the game “Genshin Impact” getting a limited character requires up to $200. And that doesn’t even factor in the “50/50” systems implemented by many of these games, giving you only a 50% chance of getting your desired character. Many professional popular games are only between $50-100, and those are the games with the highest quality. These gacha games are incredibly deceitful in hiding their absurd prices with the apparent “free-to-play” characteristics. The deceit of the true cost is one key reason why people continue to spend copious and excessive amounts of money on gacha games that other games couldn’t even come close to getting.

However, a study suggests that people spend more money on randomized systems because of design as opposed to the chance for rewards. According to a 2024 study on this, “Our results show that the design of loot boxes, rather than the random rewards they provide, encourages players to overspend” (Cordes, 2024)[9]. They continue, “For example, if gamers get distracted by fancy visual effects, they might become more prone to make statistical errors. A nice design of loot boxes might also result in a more favorable view of the game developer and the odds it offers.” As Cordes speculates, designs are more important than the actual rewards themselves. When it comes to gacha games, this tactic is abused heavily. Gacha games will build hype around certain characters and show the character extensively so people feel a more personal connection to them. They may also make newer characters more powerful and more needed to do well in the game, a concept which is known as “power creep.” All this is done to make people rationalize spending money for a character, by building a character’s personality to be likable and relatable (pathos) and by making them the character that has the best abilities in the game (logos). This along with flashy animations, and feelings of FOMO, and it can become easy to get sucked into the gacha game sphere.

**Game Theory And Psychological Analysis**

The game theory analysis for gacha games also gives us some insight into how people may rationalize their desire to participate in gacha games. This involves a similar setup to penny auctions where a small sum is paid for the potential to win a huge prize. While you lose most of the time, there is always that chance for you to hit that jackpot and get the item or character that you wanted. So, when you win, you win big, but when you lose, you lose a small amount, which, in essence, is a system of gambling.

However, while penny auctions give you an illusion of control that makes you want to bid, gacha games don't rely on this factor. People know that gacha games are random, but they are still able to motivate people to spend money. An article by Milijana Komad dives into the psychology behind gacha games, and explains many of the psychological factors that go into them. Firstly, they talk about how the principles of operant conditioning are used to foster as much engagement as possible. As Komad says, "Gacha mechanics can be analysed through the lens of operant conditioning, specifically Skinner's variable-ratio reinforcement schedule. This schedule, characterized by delivering rewards after an unpredictable number of responses, results in a high response rate and resistance to extinction" (Komad, 2023)[(10)] By delivering the jackpot after a random amount of money spent, it creates the most stable pattern of reinforcement, leading to a habit that is extremely hard to break.

Gacha games also employ the near miss effect to motivate more player engagement. When players feel that they have come so close to getting what they want, it distorts their expectations of rationale. Winning feels more likely than the odds suggest. This is the reason why the "50/50" system exists, as even when you get a rare item or character, you could still lose out based on a 50/50 chance, which activates the near-miss effect for players. Komad gives further insight on other psychological principles that boost player engagement, citing the the Zeigarnik effect, where people remember uncompleted tasks better than completed ones. As he says, "Incomplete character collections or unachieved rarities compel players to persist." Perceived progress also motivates people to continue playing, as players nonetheless make slow progress toward winning even when they lose.

All of these psychological effects build up to create the engagement that gacha games receive today. But, perhaps this particular situation is less about the irrationality of humanity. Instead, people engage with the gacha system because it's fun. No one goes into a gacha game expecting a profit, unlike penny auctions, but they know that they can pay to enjoy a game. Despite the psychological manipulation used to make the gacha system as addicting as possible, many people continue to play these games. However, the classification of gacha games remains ambiguous on whether it should be classified as gambling or a harmless game. While there is luck, most people know that they will lose most of the time, and yet they still chase the elusive win, because well, they find the games fun!

**Lottery**

The lottery is similar to gacha games because it is not mainly a strategic game between two players. It is a system of risk, probability, and expectation. A rational financial analysis would usually discourage people from participating because the expected return is generally negative for most players. You almost never win and the expected payout is usually under the cost. However, psychology explains why people continue to buy tickets anyway. While only a small population engages in penny auctions and uses the gacha system, the lottery is extremely popular and widespread where you can find it in most convenience stores around the world.

For this analysis, I will be focusing on the big national lottery systems for the United States, named Powerball and Mega Millions, both of which operate under the same system. The player buys a ticket for a small price ($2), in which they choose six numbers. Five of them are between 1 and 69, and the last one is between 1 and 26. The jackpot can be won by matching all 6 numbers to the winning ones, but there are also smaller prizes for matching less numbers.

Most of the time, buying the lottery isn't worth it. The odds of winning a jackpot can be done with simple arithmetic, being around 1 in 292 million. With a cost of $2 per ticket, the jackpot would have to be at least 584 million to consider that the expected payout could be higher than the cost spent, assuming risk-neutrality. This is not uncommon for the lottery to reach this high, so would this be a financially sound decision?

The answer, while seemingly yes, still does not hold up. When the jackpots rise to amounts that high, the lottery starts receiving more and more participants. With more participants, multiple people can win the jackpot. When this happens, the prize is split, so it becomes more unlikely that someone will be the sole winner of a huge jackpot.

Another thing that makes the number seem deceivingly high is taxes. When you win big in the lottery, your prize is taxed heavily. 24% of the prize is automatically withheld, and the winnings could lead to you having to pay more taxes from a higher tax bracket. All this is to say that very few people tend to make a profit from participating in the lottery. However, the presence of these few people are one of the main reasons why people continue to play in the lottery. An article by the University of Texas Permian Basin talks about how publicizing these winners affects participation. "By publicizing winners, lotteries promote the idea that wins are commonplace, when that couldn't be farther from the truth" (Why Do We Play the Lottery?, 2023)[11] The availability bias, or availability heuristic, is a phenomenon where people tend to disproportionately judge the likelihood of situations based on the examples that they've witnessed. With so many success stories of people who win, people develop an irrational belief that winning the lottery is quite common and happens to many people. The publication of certain events to create different perceptions of real world events is a tactic that the media uses every day, from wars to politics, and it can even affect lottery players too.

The same article also discusses the effect of "illusion of control" and how many players are misled to continue playing the lottery: "Illusion of control occurs when a person overestimates the influence of their choices on outcomes, even when those outcomes are left to chance. Anyone who's experienced a near miss—perhaps one number off from a winning ticket—and felt like they were a hair's breadth from a payout has been influenced by this

self-serving bias." When you are presented with a blank lottery ticket, you know that you have the possibility of winning, that if you choose the right numbers, you will win. Even though you have no control over whether you win, the ability to control the numbers you choose creates the illusion of control that makes people believe that they can win the jackpot through skill.

Prospect theory also works to add to the overconfidence. Prospect theory describes how people's beliefs on gains and losses can be shifted from the rational conclusion. With the lottery, it explains how smaller probabilities tend to be overweighted by players. This further adds to make the odds seemingly feel not as bad as they actually are.

Finally, the last psychological principle that affects lottery players is the gambler's fallacy. Some players believe that if a number showed up recently they will not show up again and as such, they can decode the system by interpreting the previous winning numbers. This phenomenon is known as the gambler's fallacy, and it's the same reason why we think the next coin flip will be tails if the first one is heads. Despite each individual lottery number choosing being completely independent of each other, this still affects players, and adds to the illusion of control that players feel where they think they can increase their odds of winning by choosing skillful numbers.

***Counterargument: Are These Choices Really Irrational?***

One possible objection to this paper is that many of these choices are not truly irrational. A person who buys a lottery ticket may not do so believing they are making a good investment. They may simply enjoy the feeling of imagining a different life. Similarly, a person who spends money on a gacha game may not expect to make a profit. They may value the entertainment, the characters, or the excitement of the random reward and find that spending money to be worth it for them. Outcomes can be, and often are, measured by emotions and feelings rather than the objective payoffs in money and rewards. In this broader sense, some decisions that look irrational under a narrow financial or strategic model may be rational when you factor in emotional, social, or moral values into the analysis.

However, this does not eliminate the concern about manipulation. The key question is whether people freely choose these actions with full understanding, or whether the system is designed to hide costs, exaggerate rewards, create pressure, or exploit bias. The problem is not that humans have emotions or make non-financial choices. The problem is that institutions can design systems that use those emotions against them in order to exploit these biases and principles for their own benefit.

**Conclusion**

Game theory and psychology explain two different but connected parts of human decision-making. Game theory shows the structure of a situation: the players, the choices, the incentives, and the possible outcomes. Psychology shows how real people respond to that structure when they are influenced by pride, social pressure, or limited information. Both are

important in analysing the behaviors and choices that people make and whether they are rational or irrational.

Across the examples in this paper, irrationality is not random. Rational competition can lead to harmful collective outcomes. Penny auctions, gacha games, and lotteries show how economic systems can use small costs, rare rewards, and psychological biases to encourage people to keep spending. These examples are not the only examples of irrationality, but they help show how psychology and game theory are extremely useful tools for behavioral analysis in so many areas. Recently, we have seen a new digital age in which technology becomes more and more powerful. Our behaviors and habits are often exploited and manipulated by companies for profit and control over people. For example, engagement "streaks," social-media notifications, and limited-time offers, are used to give users powerful feelings of FOMO. Logging off of social media for one day can lose you a streak or a good deal; Notifications constantly pester users to go back online. Social media understands the psychology of people extremely well and also understands how to abuse it for as much engagement and profit as possible. Understanding the intersection between game theory and psychology helps us understand why we do what we do and fight the exploitation of it.

Ultimately, I do not argue that people are simply irrational or foolish. Instead, I argue that human decision-making is predictable, and because it is predictable, it can be shaped. Sometimes that shaping can be harmless or even useful. But when systems are designed to hide costs, exploit bias, or pressure people into choices that benefit institutions more than individuals, irrationality becomes a tool of manipulation. Understanding both game theory and psychology helps us recognize these systems and make better decisions within them.

**Works Cited**


1.Poundstone, William. Prisoner's Dilemma. Anchor Books, 1992.

2.Augenblick, Ned. "Consumer and producer behavior in the market for penny auctions: A theoretical and empirical analysis." Unpublished manuscript. Available at www. stanford. edu/~ ned789 (2009).

3.Augenblick, Ned, et al. Consumer and Producer Behavior in the Market for Penny Auctions: A Theoretical and Empirical Analysis. 2009, https://faculty.haas.berkeley.edu/ned/Penny_Auction_Original.pdf.

4."Explore - Google Trends." Google.Com, Trends, 2026, https://trends.google.com/explore?q=penny%20auctions&date=all&geo=US.

5.Wang, Han. "Profit Mechanisms in Penny Auctions: Unraveling the Impact of Sunk Costs." SHS Web of Conferences, edited by Azlina Md Yassin, vol. 188, 2024, p. 03016, https://doi.org/10.1051/shsconf/202418803016.

6.Arkes, Hal R, and Catherine Blumer. "The Psychology of Sunk Cost." Organizational Behavior and Human Decision Processes, vol. 35, no. 1, Feb. 1985, pp. 124–40, https://doi.org/10.1016/0749-5978(85)90049-4.

7.Wang, Zhongmin, and Minbo Xu. Learning and Strategic Sophistication in Games:, Mar. 2012, gwern.net/doc/psychology/cognitive-bias/sunk-cost/2012-wang.pdf.

8."Revenue Data." Game Revenue Tracker, 2024, https://revenue.ennead.cc/revenue.

9.Cordes, Simon, et al. "What Drives Demand for Loot Boxes? An Experimental Study." Journal of Economic Behavior & Organization, vol. 228, Nov. 2024, p. 106755, https://doi.org/10.1016/j.jebo.2024.106755.

10.Milijana Komad. "Gacha Systems and Variable Reinforcement in Games." Medium, Bootcamp, 28 Aug. 2023, https://medium.com/design-bootcamp/product-design-and-psychology-exploring-gacha-mechanics-in-video-game-design-1015511cf00c.

11."The Psychology of the Lottery: Gambling With Our Mental Limits | UT Permian Basin Online." Online.Utpb.Edu, 19 Apr. 2023, https://online.utpb.edu/about-us/articles/psychology/the-psychology-of-the-lottery-gambling-with-our-mental-limits/.

**Corporate Disclosure and Information Asymmetry in the Intangible Economy**
**By Daniyal Kamruddin**


## Abstract

As firms derive a greater share of their value from intangible resources such as software, technological capabilities, and intellectual property, the extent to which traditional financial reporting captures this economic value has become an important issue. Because many economically significant intangible resources are not fully represented in recognised accounting information, this may result in an information gap between managers and investors. Software, AI capabilities, and proprietary technologies that add value to technology-intensive businesses may not be adequately represented by current accounting frameworks. The main concern this paper will address is the extent to which corporate disclosure can supplement financial reporting and reduce information asymmetry if financial statements alone are insufficient to fully capture intangible value. Existing research has primarily used theoretical modelling, archival analysis of financial and accounting data, and reviews of empirical literature to investigate the relationship between disclosure, information asymmetry, and intangible assets. Building on this literature, this paper combines a review of academic research with a qualitative comparative case study of corporate disclosures from NVIDIA, Microsoft, and Oracle to evaluate how disclosure communicates intangible sources of value and can reduce information asymmetry. This paper argues that corporate disclosure can reduce information asymmetry by making intangible sources of value more visible to investors; however, it cannot fully resolve the limitations created by the current accounting treatment of internally generated intangible resources. The paper applies this problem to technology and AI-intensive firms, illustrating how corporate disclosure communicates intangible sources of value in an economy where intangible resources play an increasingly important role.


## Introduction

Advances in modern technology as well as artificial intelligence (AI) have shifted the way in which firms derive value. An increasing share of this value is generated by intangible resources, which are non-physical resources such as software, intellectual property, and organisational capabilities that can generate future economic benefits. Historically, firms relied more heavily on tangible forms of capital, whereas intangible investment has become of greater importance to business activity and firm value. Corrado et al. demonstrate how intangibles and knowledge capital have become important drivers of economic growth (683). This trend has become further visible with the expansion of digital business models in recent years. Technological and digital services have emerged as the profit drivers for some of the world's largest firms, and consequently, intangible-driven firms have become many of the highest-valued companies in the economy. NVIDIA, for instance, has experienced a substantial increase in valuation alongside growing investor expectations surrounding AI and software, illustrating how investors may place significant value on the future earnings potential associated with intangible

resources. However, a central challenge for these firms is that conventional financial reporting does not fully capture the economic value of many internally generated intangible resources.

Many of the intangible assets held by firms are internally generated rather than acquired. Traditional accounting frameworks (the accounting rules governing which resources are recognised and measured in financial statements) face difficulties in reliably measuring many internally generated intangible resources, leaving a systematic gap between what firms report and their underlying economic value (Wyatt 8). This limitation becomes more significant as intangible investment accounts for a growing share of economic activity (Corrado et al. 673). As a result, investors may receive incomplete information about the resources underlying firms whose value is heavily dependent on intangible assets.

This incomplete representation can contribute to information asymmetry between managers and investors. Managers are likely to understand their intangible resources and future strategies better than external investors, meaning investors rely on disclosed and public information to evaluate value. Akerlof demonstrates that this kind of uncertainty can lead to adverse selection, where the inability to distinguish quality can cause weaker firms to appear just as attractive as stronger ones, ultimately reducing the efficiency of the market as a whole (490). Furthermore, as Easley and O'Hara show, the quality and quantity of information can affect asset prices, highlighting the importance of the information environment to investor decision-making (1554). Given these limitations, an important question arises: if financial reporting alone cannot fully communicate intangible value, to what extent can corporate disclosure reduce this information gap?

Corporate disclosure is one of the primary methods used to address this information gap. Firms provide mandatory and voluntary disclosures through financial reports, management discussion and analysis, risk disclosures, and other filings (Healy and Palepu 1). These disclosures provide investors with information not fully captured in financial statements and can therefore reduce information asymmetry. Verrecchia argues that greater disclosure reduces the private information advantage of managers by narrowing the gap between what managers and investors know, thereby improving the efficiency of the market process (105). However, the extent to which disclosure is sufficient for intangible-driven firms remains the main issue this paper will evaluate.

Although corporate disclosure can improve transparency and reduce information asymmetry, it cannot eliminate the underlying uncertainty surrounding the measurement and valuation of many intangible resources. Even relevant disclosures on intangibles can fail to convey reliable information to investors, as aggregate figures such as Research and Development (R&D) expenditure combine many different value-creating activities without revealing which components are actually expected to generate future benefits (Wyatt 16).

Therefore, this research addresses the gap at the intersection of disclosure, information asymmetry, and intangible assets. Though each area has been thoroughly studied individually, this paper examines how effectively corporate disclosure can reduce information asymmetry by communicating the sources of value in firms whose competitive advantages are primarily driven

by intangible resources, particularly technology and AI-related businesses. The increasing reliance on intangible resources as drivers of firm value makes this issue especially relevant to accounting and finance research.

This paper evaluates the extent to which corporate disclosure reduces information asymmetry in financial markets where intangible assets drive firm value. Though it can considerably reduce information asymmetry, this paper contends that disclosure cannot completely overcome the measurement constraints related to intangible assets. This argument is evaluated through theoretical analysis and comparative case studies of NVIDIA, Microsoft, and Oracle, focusing on large US-listed technology firms and their corporate disclosures, primarily through Form 10-K filings. The analysis therefore focuses on the information environment of large technology firms operating in the US market and does not attempt to generalise the findings across all firms, jurisdictions, or forms of intangible capital. The paper then concludes with a discussion of the implications for corporate disclosure and financial reporting.

**Information Asymmetry, Corporate Disclosure, and the Intangible Economy**

Information asymmetry becomes particularly important when investors must value firms whose key sources of economic value are not fully observable in financial statements. Information asymmetry arises when one party in an economic transaction holds information relevant to the transaction's value that the other party lacks, creating an imbalance that can distort pricing and decision-making. Applying this to financial markets, Easley and O'Hara demonstrate that information relevant to a firm's value is divided between private and public sources, with managers possessing information that is not immediately available to external investors (1557).

This imbalance is consequential for intangible-intensive firms because investors may have limited evidence of the resources and strategic investments expected to generate future economic benefits. As a result, adverse selection may arise, whereby investors may struggle to distinguish between high- and low-quality firms and therefore price both according to their average expected quality. Consequently, sellers of genuinely high-quality assets are unable to obtain a fair return and may withdraw from the market altogether, leaving the market to be dominated by lower-quality offerings (Akerlof 490).

Beyond these effects on market efficiency, information asymmetry also creates information risk for uninformed investors. According to Easley and O'Hara, stocks with a higher proportion of private information compared to public information command higher expected returns because uninformed investors need more compensation to bear the associated uncertainty, as they are unable to fully infer this information from market prices (1563). This implies that the quality of a company's information environment, in addition to the firm's underlying economic fundamentals, can affect investors' required returns and, in turn, its cost of capital. Since the quality of information can affect investment decisions and market outcomes, reducing information asymmetry becomes crucial, creating the rationale for corporate disclosure.

Corporate disclosure provides one mechanism through which firms can narrow this information gap by making information held internally more accessible to external investors. It

works through mandatory disclosure that takes place through regulated financial reports, including financial statements, footnotes, and management discussion and analysis, and voluntary disclosure such as management forecasts, analyst presentations, and press releases (Healy and Palepu 1). Verrecchia's description of disclosure theory holds that firms weigh the costs and benefits of releasing information (3). Diamond and Verrecchia further show that greater disclosure can improve the liquidity of a firm's securities, which then lowers its cost of capital, giving firms a financial incentive to disclose (1326). Although voluntary disclosure can further narrow the information gap, its effectiveness can be influenced by managerial incentives, as managers may selectively disclose information or present the firm's prospects in an overly favourable manner. This is why mechanisms such as mandatory reporting requirements and third-party auditors exist to reinforce the credibility of what managers disclose (Healy and Palepu 1).

The problem of information asymmetry becomes more pronounced in an economy where firms increasingly derive value from intangible resources. Corrado et al. demonstrate that intangible investment had already become a major component of business investment before the more recent expansion of AI and digital platforms (662). Intangible assets such as software, patents, and firm-specific capabilities can generate substantial future economic benefits, yet their value is often difficult to observe and measure reliably (Wyatt 15). This is particularly relevant to technology-intensive firms, where competitive advantage may depend on proprietary software and innovation capabilities that are only partially reflected through accounting. Therefore, this further accentuates the need for information beyond financial statements in order for investors to assess the resources underlying future earnings.

Despite their growing economic importance, many intangible assets remain difficult to recognise within conventional financial reporting frameworks. Wyatt suggests that because many intangible assets involve uncertain future benefits and are difficult to measure reliably, they create challenges for accurate financial reporting (7). These reliability concerns make accounting frameworks conservative, generally recognising acquired intangible assets, such as those obtained through an acquisition, while placing stricter restrictions on the recognition of internally generated intangible assets. This distinction remains under current standards. Ibrahim et al. find that recognised intangible assets average 32% of total assets in UK firms that grow through acquisitions, compared with only 9% in firms that grow organically, with these firms partly compensating by disclosing more information about the internally generated intangibles they cannot recognise (857). This difference illustrates the consequences of the recognition restrictions surrounding internally generated intangible assets and the resulting reliance on disclosure to communicate information not reflected in financial statements.

The implications of this recognition gap can extend further. Resources firms build through R&D, software development, etc. can create genuine future economic benefits without appearing as recognised assets on the balance sheet. As Barth et al. demonstrate, differences between market value and book value can reflect value-relevant information not captured in financial statements, including the effects of unrecognised intangible assets (28). Dugar and

Pozharny show that this gap has widened over time, finding that the value relevance of book value and earnings declined for high-intangible-intensity industries in the United States and internationally between 1994 and 2018 (3). Corrado et al. further demonstrate the scale of this omission, estimating that including intangible assets substantially increases measured business capital stock by approximately $3.6 trillion (682). Because of this, investors could have limited information about the resources generating a firm's future earnings, increasing uncertainty about firm value. Wyatt notes that this lack of complete information can lead investors to rely on additional sources, such as notes to annual reports, to adjust financial statements before applying valuation models (80). This suggests that investors draw on information beyond recognised financial measures in order to better assess intangible-driven firms.

However, disclosure still depends on managerial judgment and cannot compensate for assets that are not readily quantified. A further problem is the lack of a uniform framework for disclosing intangible assets voluntarily, which renders the information provided even more challenging to interpret for investors (Wyatt 80). These limitations provide the basis for examining whether the narrative disclosures of intangible-intensive firms give investors meaningful information about the resources, strategies, and future prospects that are not fully captured by recognised accounting measures.

**Evaluating Corporate Disclosure in the Intangible Economy: Comparative Case Study Evidence**

In order to better examine how intangible-intensive firms communicate sources of value that may not be fully represented in financial statements, this paper uses a qualitative comparative case study approach. This involves examining the annual reports and Form 10-K filings of NVIDIA, Microsoft, and Oracle to evaluate whether corporate disclosure provides investors with additional information that can reduce information asymmetry (Microsoft Corporation; NVIDIA Corporation; Oracle Corporation). Form 10-K filings provide standardised disclosures regarding a firm's financial performance, business operations, risk factors, and management's discussion of future prospects. The analysis focuses on the financial statements, which provide recognised accounting information, alongside the Management Discussion and Analysis (MD&A) and Business sections, which provide additional explanations of strategic resources and value drivers. These SEC filings are appropriate because they combine mandatory financial reporting with management discussion of strategic resources, risks, investments, and future growth opportunities, giving insight into the technological capabilities, innovation activities, and business ecosystems discussed above. The selection of technology firms is justified on the basis that their competitive advantage increasingly depends on intangible resources rather than physical assets, including software, AI capabilities, and other knowledge-based resources, making them a relevant setting for examining the accounting challenges discussed previously.

NVIDIA was selected as it represents a firm where valuation is strongly connected to technological innovation and expectations surrounding AI development, with its Form 10-K

providing discussion of its AI computing platforms, technological ecosystem, and strategic growth areas (NVIDIA Corporation). Microsoft was included as it represents a mature technology firm whose value increasingly depends on software platforms, cloud infrastructure, and AI capabilities, with its latest Form 10-K providing disclosures regarding these intangible resources and related strategic investments (Microsoft Corporation). Oracle provides a useful comparison because it represents a more established software-focused firm transitioning towards cloud and AI-related services, with its Form 10-K providing information regarding its software platforms, cloud operations, and evolving business strategy (Oracle Corporation).

For the purpose of a systematic comparison, each case is evaluated against the same three criteria. First, which specific intangible resources the firm identifies as sources of competitive advantage. Second, whether disclosures are forward-looking, offering insight into future strategy and investment instead of only current performance. Third, whether disclosures acknowledge the risks or limitations associated with these resources, rather than presenting them only in favourable terms. These criteria are applied to each firm in turn before being compared directly in the Comparative Evaluation section.

**NVIDIA**

The first firm analysed is NVIDIA, an appropriate case study due to its competitive position being heavily dependent on intangible resources, including software capabilities, technological knowledge, and ecosystem effects. Drawing from the "Business" section of its 10-K annual report, NVIDIA describes itself as "a data center scale AI infrastructure company reshaping all industries" (NVIDIA Corporation 4). The company's disclosures highlight how these intangible resources contribute to its competitive advantage, especially through its software ecosystem, developer community, and GPU technologies. NVIDIA describes its technology stack as centred on the CUDA development platform that runs across its GPUs, supported by hundreds of domain-specific software libraries, frameworks, and APIs (NVIDIA Corporation 4). This disclosure suggests that NVIDIA's competitive advantage rests as much on this underlying software layer as on the hardware itself, a distinction investors could not infer from the balance sheet alone. Although these software platforms and development tools represent important sources of economic value, they are not fully recognised within financial statements. These disclosures make NVIDIA's technological capabilities more visible to investors, particularly where those capabilities are not reflected in accounting measures.

NVIDIA also provides forward-looking disclosures that assist investors in evaluating future expectations rather than only reporting current performance. For example, NVIDIA discloses that all major cloud service providers use its data center-scale infrastructure to accelerate the services they deliver to billions of end users (NVIDIA Corporation 4). This information about the utilisation of its data centres at scale helps the company’s investors understand future demand for its services long before such information would appear in the financial statements. By communicating who is adopting its technology and where demand is

emerging, management provides investors with insight into potential future growth and earnings sustainability.

Furthermore, NVIDIA provides detailed disclosure regarding its innovation strategy, allowing investors to assess future value creation. NVIDIA discloses that it has invested over $76.7 billion in research and development since its inception, and frames this spending as building its technology roadmap for the broader AI ecosystem (NVIDIA Corporation 4). The figure itself says little about future returns, but naming the scale of investment gives investors a basis for judging NVIDIA's commitment to sustaining its technological lead. Although R&D investment itself does not directly demonstrate future value creation, these disclosures provide context regarding long-term innovation, engineering capabilities, and technological development.

Overall, NVIDIA's disclosures provide insight into competitive advantages, future growth drivers, and strategic priorities, enabling investors to make more informed assessments of long-term value. In this sense, disclosure enhances transparency and reduces uncertainty. However, these disclosures also demonstrate the limitations of corporate reporting in the intangible economy. While management explains why resources such as CUDA, its developer ecosystem, and engineering expertise are strategically valuable, the disclosures do not quantify their economic value. NVIDIA therefore illustrates a key distinction between communicating intangible value and measuring it, as its disclosures make the strategic importance of CUDA and continued R&D investment more visible, but do not provide a monetary measure of the economic value generated by these resources. This supports Verrecchia's argument that disclosure can reduce information asymmetry by communicating information unavailable through recognised financial statements (105). Despite disclosure communicating information about internally generated intangible assets, it cannot fully overcome the accounting limitations associated with the intangible economy.

**Microsoft**

Microsoft provides a contrasting case where value is distributed across interconnected intangible resources, including software platforms, cloud infrastructure, customer relationships, and AI capabilities. "We develop and support a broad portfolio of technology solutions... focusing on secure, trusted, and innovative platforms and applications that meet evolving customer needs across AI, cloud computing, productivity and collaboration, and personal computing" (Microsoft Corporation 3). Microsoft's value, depending on the integration of its resources rather than individual technologies, strengthens the reason for its selection.

A prime example of how Microsoft discloses the strategic resources supporting its cloud ecosystem is through its "Business" section of the 10-K report, which describes Azure as a comprehensive set of cloud services that give developers, IT professionals, and enterprises the freedom to build, deploy, and manage applications (Microsoft Corporation 7). As well as disclosing the existence of Azure, Microsoft also discloses its role as a platform that enables customers to develop and adopt digital technologies.

Microsoft's forward-looking disclosures additionally provide insight into the relationship between current investment and future value creation. Microsoft discloses that its investments in cloud and AI infrastructure will continue to raise operating costs and could reduce operating margins, framing this spending as necessary to meet growing demand for AI services (Microsoft Corporation 35). Instead of only highlighting expected AI growth, Microsoft communicates the trade-off between short-term financial performance and the development of future intangible resources, increasing transparency with investors.

Microsoft's communication of intangible resources beyond products and platforms also extends to human capital. The company reveals that 77,000 of its staff are engaged in product research and development (Microsoft Corporation 9), which is a specific figure that provides a better understanding of the organisational capabilities that its other disclosures describe only in general terms. While NVIDIA's human capital disclosure is closely connected to advancing specialised computing technologies, Microsoft's disclosure demonstrates the organisational capabilities required to continuously develop and integrate resources across multiple business areas.

A further distinctive aspect of Microsoft’s disclosure relates to responsible AI. Microsoft states that it is committed to developing and implementing AI in an ethical manner, incorporating security and safety from the outset (Microsoft Corporation 3). This conveys the idea that Microsoft's intangible resources go beyond technological innovation to include the capacity for risk management and governance. This differs from NVIDIA’s disclosure focus on technological advancement and market demand, as Microsoft demonstrates how governance and ethical responsibility contribute to sustaining the value of AI-related resources. Similarly to NVIDIA’s case, the described resources and capabilities are only partially reflected by accounting-based indicators, meaning that the provided information is useful but insufficient for an accurate assessment.

**Oracle**

Oracle provides a third perspective, illustrating how an established enterprise software firm communicates the value of long-developed customer relationships, software capabilities, and an evolving cloud platform. In its 2026 Form 10-K, Oracle Corporation discloses that its enterprise applications and infrastructure offerings are increasingly enhanced by AI technologies (1). This illustrates how Oracle combines enterprise software, cloud infrastructure and AI-enabled services to support organisations' IT operations. Oracle's disclosures primarily focus on its integrated cloud ecosystem, describing Oracle Cloud Applications and Oracle Cloud Infrastructure as providing customers the flexibility to choose when and how they deploy Oracle's technologies (Oracle Corporation 4). Rather than simply describing individual products, Oracle explains how its integrated cloud ecosystem and flexible deployment model support long-term customer value. Since these organisational capabilities are not directly observable through financial statements, such disclosures provide investors with greater insight into the firm's sources of competitive advantage.

Oracle also discloses that it expects to be in the early stages of a material migration of its existing customer base from on-premise products to the Oracle Cloud, and notes that the proportion of cloud revenues relative to total revenues has already increased and is expected to keep rising (Oracle Corporation 2). These disclosures aid investors in understanding the company's strategic transformation, anticipated customer migration, and changing revenue model instead of only reporting financial performance. Because it links management's expectations for the company's future revenue model and cash flows to an observable strategic transition (from customer migration to the cloud), this disclosure is especially applicable to valuation.

Oracle's disclosures also demonstrate how the firm continues to develop and strengthen its existing intangible resources. The company states that it invested $10.3 billion in research and development, directed at improving the performance, reliability, and cost-effectiveness of its offerings (Oracle Corporation 2). The focus of this expenditure is on refinement rather than reinvention, which is appropriate for a company that manages an existing platform rather than creating entirely new technology categories. Oracle's R&D disclosures concentrate on improving and integrating an established enterprise platform, in contrast to NVIDIA, whose disclosures mainly highlight technological innovation. This provides investors with context for evaluating the company's competitiveness in the future without suggesting that R&D spending is the only source of value.

Oracle thus adds value to the comparison, as its disclosures render the potential economic implications of customer migration and transformation toward the cloud more apparent, though the relationships, software and organisational capabilities that underpin the transformation remain largely unrecognised in financial statements.

**Comparative Evaluation**

Across the three cases, corporate disclosure makes intangible sources of value more visible to investors. NVIDIA provides the clearest emphasis on technological capabilities and the AI ecosystem, Microsoft demonstrates how value can arise from interconnected software and governance capabilities, and Oracle illustrates how disclosure can communicate an ongoing transition in customer relationships and business models. Despite these variations, all three companies offer information that cannot be fully seen in financial statements regarding strategic investment, technological capabilities, and anticipated sources of future growth. The cases therefore support the argument that disclosure can improve the information environment by narrowing the information gap between managers and investors (Verrecchia 105).

However, management disclosure involves strategic choices about what information to provide. Managers may present information selectively or in an unduly positive light, as noted by Healy and Palepu (3). As a result, certain disclosures could serve promotional purposes that may not be as useful for investors. Both Microsoft's portrayal of its AI investments as essential to secure computing and NVIDIA's claim that it is transforming every industry are examples of language that implies strategic intent but is challenging for investors to independently confirm. Therefore, it is difficult to distinguish between selectively optimistic disclosure and informative

disclosure, especially when claims are made in qualitative terms without external verification. Nevertheless, the existence of specific numbers, like Microsoft's disclosed research headcount or NVIDIA's disclosed R&D spending, provides investors with a partial check against purely promotional language because these numbers can be verified and have repercussions if they turn out to be false. Therefore, rather than being read in isolation, disclosure is most helpful to investors when qualitative claims about strategic advantage are supported by quantifiable data.

However, the cases also reveal a common limitation. None of the three firms provides a reliable monetary measure of the economic value of internally generated software ecosystems, AI capabilities, customer relationships, or organisational resources. Instead, their disclosures explain how these resources contribute to competitive advantage/expected growth, without resolving the measurement problem that prevents many internally generated intangibles from being recognised as assets. This supports the distinction between measuring and communicating intangible value. While disclosure significantly reduces information asymmetry regarding the nature and strategic significance of intangible resources, it is unable to remove the uncertainty surrounding their potential economic value.

The broader implication of these findings is not that traditional financial reporting is inadequate or no longer useful, but that it faces increasing limitations when representing an economy in which firm value depends heavily on internally generated knowledge, software, and technological capabilities. The case evidence indicates that developing complementary disclosure practices that increase the transparency and utility of intangible resources is likely a better course of action than replacing conventional accounting recognition principles. These disclosures can enhance the information environment in which investors assess companies whose most significant sources of value are not fully reflected in their financial statements, but they cannot completely remove the uncertainty resulting from the challenge of measuring intangible value.

**Conclusion**

To conclude, corporate disclosure can reduce information asymmetry in the intangible economy by providing investors with information about resources, strategies, and expected sources of future value that are not fully captured by financial statements. Evidence from the case studies shows that NVIDIA, Microsoft, and Oracle all use narrative disclosure to make important intangible resources and sources of value more visible to investors. Although each firm emphasises different sources of value, the cases demonstrate how narrative disclosure can communicate the strategic importance of intangible resources. However, disclosure cannot resolve the underlying measurement uncertainty surrounding many internally generated intangible resources, and thus cannot fully overcome the limitations of financial reporting. It therefore represents an important complement to financial reporting rather than a complete solution to information asymmetry in the intangible economy.

The continued growth of AI, software, and digital business models will increase pressure on financial reporting systems. The findings also highlight areas that are likely to become

increasingly important as the intangible economy continues to develop. More standardised disclosure of intangible resources could provide investors with more decision-useful information, while avoiding a false sense of precision around resources that remain difficult to measure reliably. At the same time, accounting regulators may face more pressure to balance comparability, reliability, and relevance as firms derive more of their value from internally generated AI and software capabilities that current standards were not designed to fully capture. Future financial reporting must therefore provide investors with more meaningful information while recognising the measurement challenges associated with intangible resources. As the intangible economy continues to expand, corporate disclosure will remain an important mechanism through which firms communicate value creation and investors assess increasingly complex business models.

**Acknowledgments**

I would like to thank my mentors, Dr Aivaliotis and Ms Howland, for their support, guidance, and patience throughout this project. This research would not have been possible without them.

**Works Cited**


Akerlof, George A. ‘The Market for “Lemons”: Quality Uncertainty and the Market Mechanism’. The Quarterly Journal of Economics, vol. 84, no. 3, Aug. 1970, p. 488–500. https://doi.org/10.2307/1879431.

Barth, Mary E., et al. ‘Analyst Coverage and Intangible Assets’. SSRN Scholarly Paper no. 263981, Social Science Research Network, 1 Mar. 2000. papers.ssrn.com, https://doi.org/10.2139/ssrn.263981.

Corrado, Carol, et al. ‘Intangible Capital And U.S. Economic Growth’. Review of Income and Wealth, vol. 55, no. 3, Sept. 2009, pp. 661–85. https://doi.org/10.1111/j.1475-4991.2009.00343.x.

Diamond, Douglas W., and Robert E. Verrecchia. ‘Disclosure, Liquidity, and the Cost of Capital’. The Journal of Finance, vol. 46, no. 4, 1991, pp. 1325–59. https://doi.org/10.1111/j.1540-6261.1991.tb04620.x.

Dugar, Amitabh, and Jacob Pozharny. ‘Equity Investing in the Age of Intangibles’. SSRN Scholarly Paper no. 3770088, Social Science Research Network, 30 Dec. 2020. papers.ssrn.com, https://doi.org/10.2139/ssrn.3770088.

Easley, David, and Maureen O’Hara. ‘Information and the Cost of Capital’. The Journal of Finance, vol. 59, no. 4, 2004, pp. 1553–83. https://doi.org/10.1111/j.1540-6261.2004.00672.x.

Healy, Paul M., and Krishna Palepu. ‘Information Asymmetry, Corporate Disclosure and the Capital Markets: A Review of the Empirical Disclosure Literature’. SSRN Scholarly Paper no. 258514, Social Science Research Network, 1 Dec. 2000. papers.ssrn.com, https://doi.org/10.2139/ssrn.258514.

Ibrahim, Salma, et al. ‘Recognition and Disclosure of Intangibles Under International Financial Reporting Standards’. Abacus, vol. 61, no. 4, 2025, pp. 857–96. https://doi.org/10.1111/abac.70015.

Microsoft Corporation. Annual Report 2026 (Form 10-K). Microsoft Investor Relations, 2026. https://microsoft.gcs-web.com/static-files/8a9beeed-0e9e-48e6-b4fb-593d558e2693.

NVIDIA Corporation. Annual Report 2026 (Form 10-K). NVIDIA Investor Relations, 2026. https://d18rn0p25nwr6d.cloudfront.net/CIK-0001045810/e361e58a-7483-44f5-bc62-a9080ae6ec72.pdf.

Oracle Corporation. Annual Report 2026 (Form 10-K). Oracle Investor Relations, 2026. https://d18rn0p25nwr6d.cloudfront.net/CIK-0001341439/0723dfa7-be5f-4227-9da6-eff3cce376a8.pdf.

Verrecchia, Robert E. ‘Essays on Disclosure’. SSRN Scholarly Paper no. 276699, Social Science Research Network, 1 June 2001. papers.ssrn.com, https://doi.org/10.2139/ssrn.276699.

Wyatt, Anne. ‘What Financial and Non-Financial Information on Intangibles Is Value Relevant? A Review of the Evidence’. SSRN Scholarly Paper no. 1103443, Social Science Research Network, 10 Mar. 2008. papers.ssrn.com, https://doi.org/10.2139/ssrn.1103443.

**AI Exposure and Wage Inequality Amongst Skilled and Unskilled Workers: Evidence from Brazil By Henrique Gomes**


## Abstract

This paper investigates whether AI exposure widens the wage gap between skilled and unskilled (tertiary v non-tertiary educated) workers in Brazil, and whether informality moderates this effect. Using RAIS microdata from 2017 and 2022 for São Paulo (SP), Minas Gerais (MG) and Rio de Janeiro (RJ), and PNAD Contínua household survey data (Q4 2017, Q4 2022) covering the formal and informal labour market across the same three states, this study estimates repeated cross-sectional regressions interacting tertiary education, occupational exposure score and time period. The RAIS analysis found a modest increase in tertiary-educated workers' wage premium (~3%, $p = 0.002$) that was unstable, varying when controls were removed. Additionally, there was no statistically significant evidence that the wage premium had a sharper increase in more exposed occupations. The PNAD Contínua analysis found no significant changes within the informal market, and no changes when evaluating both formal and informal altogether. More importantly, when reexamining the formal sector data wage premium, three different definitions of the formal labour market sector were used; although all three point estimates were negative, they disagreed with each other in magnitude and statistical significance, and also disagreed in sign with the positive RAIS result. That instability, rather than a confirmed widening of the gap, is this study's main finding. A supplementary case study of a large retail bank in Brazil was used to give a firm-specific picture of how AI is actually being used in the country, though it doesn't directly test or support the wage-premium results. Overall, this paper finds no stable evidence that AI exposure has widened the wage gap between skilled and unskilled workers in Brazil.



**Keywords:** AI exposure, wage inequality, informality, Brazil, skill premium


## 1. Introduction

Artificial intelligence is the most consequential labour-displacing technology in decades, yet its distributional effects in developing economies remain poorly understood, particularly where a large share of workers operate in the informal sector of the labour market. The integration of AI creates a displacement effect in which demand for routine and low skill workers drastically decreases, and also creates a productivity effect which benefits the higher end of the labour market, skilled workers whose jobs are complemented by the technology (Acemoglu and Restrepo). Acemoglu and Restrepo's main prediction is that the wage gap will continuously widen if AI displaces jobs faster than it complements and creates new ones, especially in developing countries that contain fewer resources to help workers adapt.

Evidence from the International Labour Organization (Gmyrek et al.) indicates that in Latin America, AI exposure is heavily concentrated in the formal sector of the labour market. It shows how AI has been complementary to skilled workers, leading to wage gains, but has created displacement risk for routine/low skilled employees. The IDB working paper on the distributive impacts of AI in the Latin labour market (Ciaschi et al.) complements this, projecting

widening skill premiums for skilled workers, with Brazil listed as one of the most exposed countries.

From a domestic perspective, Brazil specifically has followed a path in which its wage inequality has declined from the 2000s to the early 2010s due to a minimum wage policy and education expansion, but has stagnated in recent years (Firpo and Portella). Another ILO paper by Adamczyk et al. shows the skill transition in the country with official RAIS data (Brazil's formal sector employer registry), determining which occupational groups face the most automation exposure.

Brazil stands out in Latin America due to its massive informal labour market. Approximately 40% of the country's workers operate in the informal sector (PNAD Contínua), and as Firpo and Portella show, the formal-informal wage gap is a major driver of inequality. This means that the standard automation driven wage inequality seen in US and Europe are very different from what is seen in Brazil: informal workers are less exposed to AI, but are also less likely to benefit from the productivity gains it could bring like skilled workers do (Gmyrek et al.).

This paper investigates the extent to which AI exposure widens the wage gap between skilled and unskilled workers in Brazil. This question matters for two main reasons that remain unresolved among existing literature.

Firstly, all existing literature contains evidence and empirical data on how automation and the adoption of AI has affected high-income economies in well developed countries with very high formal employment. For instance, Acemoglu and Restrepo and Autor and Salomons use US and European data, while the IMF's 2024 discussion of generative AI - which argues that its productivity gains will mostly benefit high-skilled occupations and that labour market outcomes will depend on how quickly complementary skills are formed - though it covers advanced, emerging, and developing economies more broadly still concentrates its detailed analysis on formal-sector, specialized occupations (Cazzaniga et al.). Whether existing literature generalizes AI's effect without major emerging economies consideration is an unresolved and open question.

Second, informality has been acknowledged in these studies, but not as rigorously tested. Gmyrek et al. identify informality as a potential buffer against displacement, but end up treating it more as a superficial observation and not as an active mechanism to be measured. Similarly, the Inter-American Development Bank paper on Latin America (Ciaschi et al.) acknowledges informality as a potential factor but doesn't include it in its analysis. This paper attempts to address this issue by treating informality as a major moderating variable of the evaluation. This is considerably relevant given that Brazil's informal sector tends to expand during economic shocks like the COVID-19 pandemic, which means that the size and composition of the sector - and consequently AI exposure within it - is unlikely to have been static over the period studied.

If AI exposure leads to a widening formal-sector skill premium while leaving informal employees behind, then Brazil risks seeing inequality grow two ways: inequality will increase both within the formal sector and within the formal and informal sector simultaneously. That

would complicate standard responses like minimum wage increases, new wage policies, which are mainly applied through the formal sector.

This paper hypothesizes that AI exposure widens the wage gap between tertiary-educated and non-tertiary-educated workers in the formal sector, and that high levels of informality moderate this effect, as informal workers are less affected by both job displacement and the benefits of working alongside AI. (Gmyrek et al.; Acemoglu and Restrepo). If confirmed, this would imply the inequality effect is less intense in high-informality sectors than in low-informality ones. Concretely, in a state or occupation in which informal employment is high, the average skilled/unskilled wage gap would widen by less than in another dominantly formal state or occupation group, given that a smaller share of that workforce is exposed to either AI's displacement risk or its productivity complementarities in the first place.

The remainder of the paper proceeds as follows. Section 2 describes the data sources and empirical methodology, including the construction of the occupational scores and the supplementary case study. Section 3 displays the results of the RAIS and PNAD Contínua investigation, as well as the supplementary case study's findings. Section 4 discusses these findings regarding the study's hypothesis and cited literature. Section 5 discusses the findings and outlines directions for future research in this area.

**2. Materials and Methods**

The study used secondary quantitative publicly available government microdata. It did not involve any type of direct primary data collection. Two terms are distinguished throughout: AI adoption refers to when firms actually introduce AI into their business, while AI exposure (or potential exposure/automation susceptibility) refers to the extent to which an occupation's task(s) can be impacted by AI. Formal sector employment data was collected from RAIS, Brazil's official ministry of labour, as employers and employees are legally required to be registered in this system. Informal-sector data came from PNAD Contínua, a quarterly household survey conducted by Brazil's Institute of Geography and Statistics (IBGE). It is important to understand a few differences between PNADC and RAIS, as it matters for the results interpretation: RAIS covers the full population of registered formal employees, while PNADC only samples people and treats that sample as representative of the whole population. So PNADC's numbers should be used as comparisons within its own sample, not as representative estimates for Southeast Brazil as a whole. The analysis focuses on São Paulo, Minas Gerais and Rio de Janeiro because they are the country's most populous states and its second and third largest GDP and by formal employment (RAIS registered). Combined, these three states are accountable for a substantial share of Brazil's economic activity and were selected as a computational suitable starting point, as processing RAIS' full national microdata was not feasible given the study's available resources (see below). The PNAD Contínua analysis uses the same three states for consistency when comparing it with RAIS in the analysis.

Each RAIS file was read using 6 main variables (out of 60 total): occupation, education, hours, wage, municipality, and active employment status. The states were identified from their

municipality code. The sample only included employee relationships that were active on December 31st, with positive wages/hours and a classified education code. That means this analysis can only speak to wages, not job displacement. The full files were too large for the study's computer to handle, so a smaller subsample was used instead (N = 89,490). This subsample came from randomly picking eligible employment records - after applying the filters described above - using a fixed random seed so the sample could be reproduced later. It isn't limited to any specific occupation or firm, just the three states already chosen. Running the full file just wasn't possible given the computing power available.

A binary tertiary indicator (completed undergraduate studies or above) and a binary period indicator (2017 vs 2022) were built. An approximate ten category occupational exposure score was assigned based on each occupation's CBO (Brazil's official system for coding occupations, similar to the U.S.A.'s SOC): the first digit of Brazil's six digit code, which sorts every occupation into one of ten broad categories (managers, production workers, clerical workers etc.). These exposure values were based on the exposure-to-computerization concept introduced by Frey and Osborne, who asked machine-learning professionals to subjectively rate a set of benchmark U.S. occupations on task "bottlenecks" to automation - such as the need for manual dexterity, creativity or social intelligence for instance - and then used a classifier to extend these ratings across the full O*NET occupation list. As this paper dates itself back to 2013, it captures general susceptibility to computerization and robotics rather than exposure to artificial intelligence specifically. More recent sources cited in the introduction, such as Gmyrek or Adamczyk et al., construct these measures directly around AI's capabilities, with Adamczyk et al.'s built directly from RAIS microdata. This paper's coarse ten-category measure was used instead of newer indices primarily due to the data-access and computational resources available for this study, and this is a major limitation of the analysis, discussed further below: every worker in the same occupation group was given the same exposure value. There are only ten exposure values in the analysis, no matter how many workers were in each group. So when testing whether exposure affected the results, the analysis really only had ten distinct values to work with, not the full number of workers in the sample - a real limit on the statistical power of this particular test.

Two OLS regressions were estimated. The first is:

$$log(wage) = \beta_0 + \beta_1 \cdot Tertiary + \beta_2 \cdot Post + \beta_3 \cdot (Tertiary \times Post) + State\ FE + Occupation\ FE + \beta_4 \cdot Hours + \varepsilon$$

where the coefficient of interest, $\beta_3$, captures how the tertiary-education wage gap changed from 2017 to 2022. The second equation adds a three-way interaction with the occupational exposure score:

$$log(wage) = \beta_0 + \beta_1 \cdot Tertiary + \beta_2 \cdot Post + \beta_3 \cdot (Tertiary \times Post) + \beta_4 \cdot (Tertiary \times Post \times Exposure) + State\ FE + Occupation\ FE + \beta_5 \cdot Hours + \varepsilon$$

With $\beta_4$ as the coefficient of interest.

Although the models compare changes over time, they are not an actual difference-in-differences model because the education groups are not a clear treatment and control group, and with only one pre- and one post-period, the parallel trends assumption cannot be tested. Fixed-effects OLS of this kind is a standard baseline approach in the cited literature (e.g., Acemoglu and Restrepo, Adamczyk et al.) and was chosen here as a transparent, easily interpretable starting point. More sophisticated approaches - things like quantile regressions tracing the effect across the wage distribution, or event-study designs with additional pre-periods for example - were not used due to the computational and data resources available for this study, so the results shown below should be interpreted as preliminary, reduced-form evidence and not a definitive causal finding. Heteroskedasticity-robust (HC1) standard errors were used. Whether these errors should instead be clustered by occupation remains uncertain because there are only ten exposure groups. 4 figures were produced using ggplot2 and are analyzed and interpreted in the results section.

A supplementary PNAD Contínua analysis (Q4 2017, Q4 2022, SP, MG and RJ) tested whether informality has any impact on the skilled vs unskilled wage gap. More specifically, it tests whether informality moderates the tertiary-education premium. Workers were classified as formal if they had a signed work card (Carteira de Trabalho e Previdência Social - CTPS), were statutory or military public servants, or were self-employed or employers who contribute to social security - an operational definition consistent with IBGE's own official criteria for PNAD Contínua. The rest was classified as informal, following the complementary operational definition of informality proposed by Ulyssea's (2018) definition. Thus, the analysis is only partly comparable to the RAIS analysis. It also uses an unweighted sample and relies on self-reported income rather than recorded wages. PNADC groups occupations into its own eleven categories instead of using RAIS' CBO codes. Because of this, the exposure score built from PNADC's classification is conceptually similar to the RAIS-based measure, but the numbers don't line up. So the exposure interaction terms from the two datasets can't be compared directly. On top of that, tertiary education is defined differently, and the data is quarterly rather than based on December 31st employment like RAIS.

A case study from a large retail bank in Brazil was utilized, using anonymized internal data, covering the bank's debt collection operations from November 2025 to January 2026. Unlike RAIS and PNADC, which are the central sources for the empirical investigation from this paper, this study directly observes a firm that has adopted AI and asks whether that adoption is associated with different operational outcomes - conversion and discount rates - between an AI-assisted channel, and the regular fully human operated channels. It is included to provide a concrete exemplification of how AI is being adopted in Brazil. It is complementary to, but not a direct test of, the wage premium hypothesis examined in the study.

Two tables from the bank's debt-collection operations were used: an eligibility table with 664,144 delinquent accounts, split into "test" and "control", including each account's balance, and an outcomes table with 270,726 negotiated agreements and their negotiated debt, discount and the channel used, including the bank's AI conversational agent/channel. The two tables were

linked using an internal agreement ID. The combined data was used to compare conversion rates between the test and control group to compare discounts at the different channels (human vs AI agent). As the files were too large to be processed directly, they were summarized and analyzed using R script. It is important to note that the groups are slightly unbalanced in size, so the results should be read as firm-specific evidence rather than causal effect.

## 3. Results

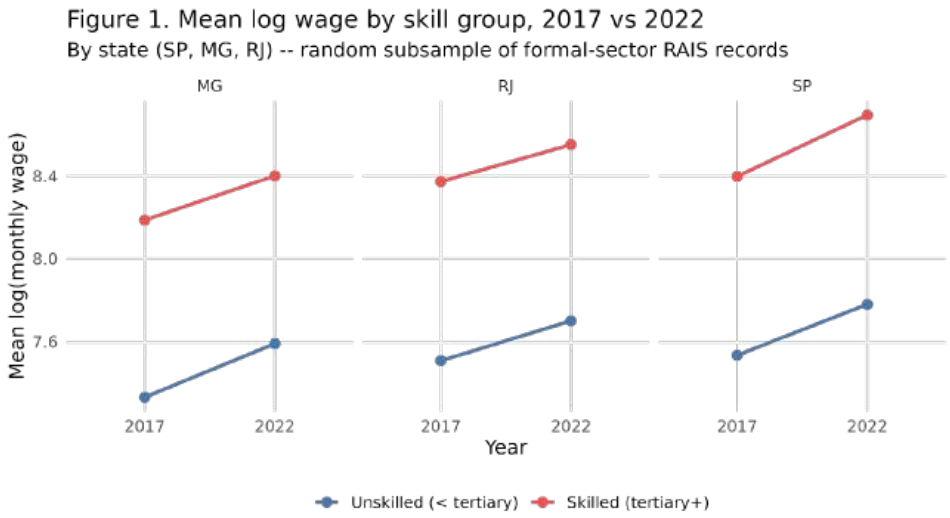
### 3.1 RAIS analysis (formal jobs, SP, MG, RJ)

Two repeated regressions were modelled across the formal sector, evaluating how the wage difference associated with tertiary education changed from 2017 to 2022, and whether that change differed by the respective assigned occupational exposure score. Both regressions used data.table (an R package for data manipulation) for data preparation, fixest (an R package for fixed-effects estimation), and heteroskedasticity (HC1) for standard errors. The first regression tests specifically if the wage gap between tertiary-educated workers and non-tertiary-educated workers changed over the time period studied. As shown in Table 1 and Figure 1, the gap grew by 3%. More specifically: coefficient = 0.031, SE = 0.010, t = 3.10, p = 0.002, based on 89,490 workers. Another less complex model without controls produced a smaller coefficient (0.018, SE = 0.013) whose 95% confidence interval included zero. This sensitivity shows that the result depends on the controls included and requires further investigation into whether changes in occupation, state, or hours worked between the two education groups are responsible for the difference, rather than assuming that the fully controlled model is automatically correct.

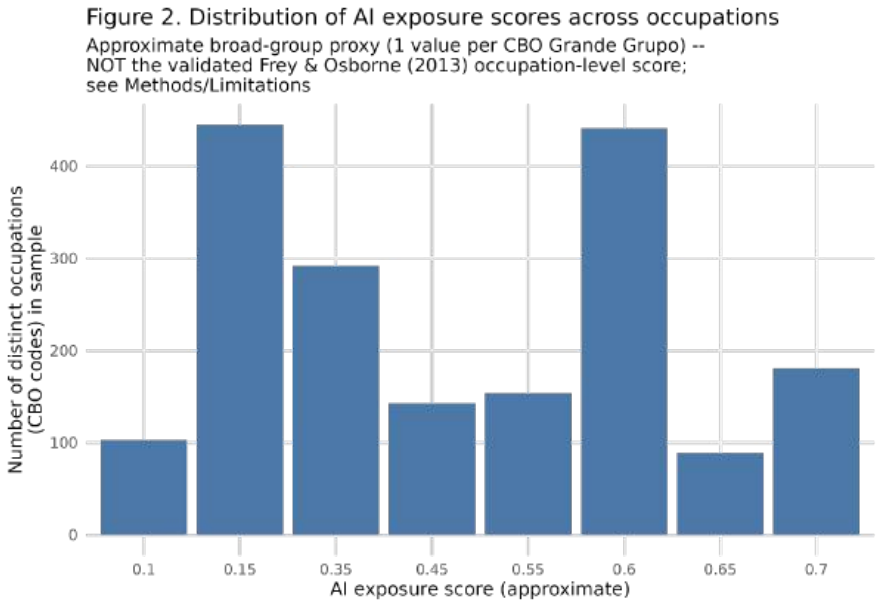
**Table 1. RAIS Regression Results (tertiary x period interaction)**

| Specification | Coefficient | SE | t | p | N |
|---|---|---|---|---|---|
| **(a) Naive (no controls / FE)** | 0.018 | 0.013 | 1.40 | 0.163 | 89,490 |
| **(b) Baseline (hours + state + occupation FE)** | 0.031 | 0.010 | 3.10 | 0.002 | 89,490 |
| **(c) Exposure-interaction (triple interaction)** | -0.052 | 0.059 | -0.88 | 0.38 | 89,490 |

Note: Rows (a) and (b) report the coefficient on tertiary education × period. Row (c) reports a different quantity: the triple interaction of tertiary education × period × exposure score, and is not directly comparable in magnitude to rows (a) and (b). All standard errors are heteroskedasticity-robust (HC1)




The second regression focused on testing whether that gap grew more in jobs that are more exposed to automation. The results yielded the following: coefficient = -0.052, SE = 0.059, t = -0.88, p = 0.38. That is small, and is not statistically significant. As the exposure score is only composed of ten values no matter the sample size - illustrated in Figure 2, which shows how the RAIS samples occupation cluster into only a few distinct exposure levels - this null result should be interpreted cautiously, since the actual variation used to identify the effect is much smaller than suggested by the sample size. The exposure score's own main effect could not be separately estimated, as it is constant within, and therefore absorbed by the occupation fixed effect.




### 3.2 PNAD Contínua Analysis (formal + informal-sector workers)

The PNAD Contínua sample is composed of 56,323 people in Q4-2017 and 48,787 in Q4-2022 (total of 105,110), and the same empirical approach as the RAIS analysis was used here, including the same SE specification and controls for state and detailed occupation.

This section of the analysis tests the informality aspect of the hypothesis using the same approach. As summarized in Table 2, in the informal sector, the tertiary-education-by-period interaction was -0.011 (SE = 0.024, t = -0.47, p = 0.64, N = 37,876). Looking at formal and informal employment, the interaction was -0.026 (SE = 0.026, t = -0.99, p = 0.32, N = 105,110). Neither result is statistically significant. That doesn't mean nothing changed though - PNADC's sample is smaller, unweighted, and only covers one quarter per year, so there just isn't enough evidence here to say either way.

Re-examining formal employment using PNADC produced an important result, also summarized in Table 2. Three definitions of formal employment yielded very different results. A broad definition (signed-card employees plus contributing self-employed workers and employers, N = 67,234) produced -0.023 (SE = 0.011, t = -2.12, p = 0.034). Under a stricter definition (only signed-card employees, the closest match to RAIS's classification and to IBGE's own primary formal-employment criterion, with N = 53,987) produced -0.006 (SE = 0.011, t = -0.50, p = 0.62). A definition covering only contributing self-employed workers and employers (N = 13,247) yielded -0.094 (SE = 0.030, t = -3.15, p = 0.002). All three point estimates are negative, but they disagree sharply in magnitude and statistical significance: two are statistically significant and one is not, and all three disagree in sign with the positive RAIS estimate (0.031, p = 0.002). This disagreement across different configurations will be treated as a finding rather than using the one most consistent with the study's hypothesis. It also shows that the paper's central instability finding is, in part, a direct function of which formal-sector definition is used.

**Table 2. PNAD Contínua regression results (tertiary × period interaction)**

| Specification | Coefficient | SE | t | p | N |
|---|---|---|---|---|---|
| **Informal sector** | -0.011 | 0.024 | -0.47 | 0.64 | 37,876 |
| **Formal + informal (pooled)** | -0.026 | 0.026 | -0.99 | 0.32 | 105,110 |
| **Formal - broad definition** | -0.023 | 0.011 | -2.12 | 0.034 | 67,234 |
| **Formal - strict definition (signed-card only)** | -0.006 | 0.011 | -0.50 | 0.62 | 53,987 |
| **Formal - self-employed only** | -0.094 | 0.030 | -3.15 | 0.002 | 13,247 |

Note: Every row shows the tertiary education x period effect, except "Formal + informal (pooled)", which shows a three-way interaction (tertiary education x period x formal status) instead. SE's use a correction (HC1) that stays accurate even when variability differs across the data.

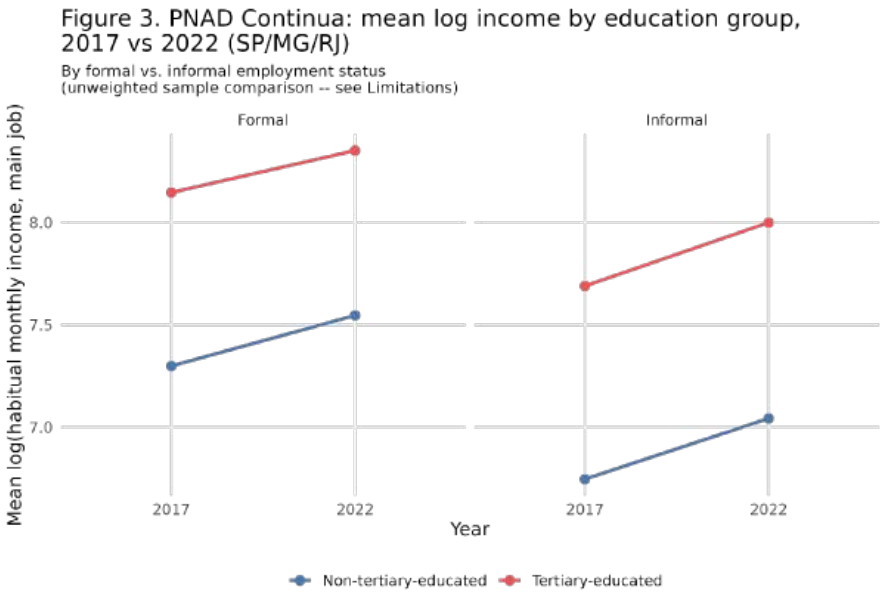



Figure 3 illustrates these patterns, plotting mean log income by education group and formality status across the two periods.

### 3.3 Limitations of the Analysis

First, the ten-category exposure indicator is a broad measure inspired by Frey and Osborne, not a specifically detailed modified version to match Brazilian occupations, which limits the amount of useful variation in the analysis. As discussed in Section 2, this is a major limitation of the exposure-interaction tests specifically, since they rest on only ten distinct exposure values. Second, the analysis cannot be fully treated as a difference-in-differences study, as there is no clear control and test/treatment group, no direct measure of AI adoption of employers, and only one period before and after 2022, making it very difficult to test whether the groups followed similar trends before the change.

Additionally, the comparison between 2017 and 2022 does not prove that 2022 represents a specific point of AI exposure; if anything, firm-level adoption of AI in Brazil appears to have remained limited for much of this window and to have accelerated mainly afterward, based on IBGE's Pintec Semestral technology-adoption surveys (IBGE, "PINTEC Semestral"), so the 2017-2022 window may in part precede the period in which AI exposure could plausibly move wages. Other events like COVID-19 and the economic recovery could have had an impact on the results, including a temporary rise in informal employment, and no structural-break test around the pandemic was performed. Wages were not inflation-adjusted - since the analysis focuses on the relative gap between education groups within each cross-section rather than on real wage levels across years - though this is nonetheless noted as a limitation, particularly for any reading of Figure 1's wage levels across years, and weren't adjusted by working hours. The PNADC analysis is also not perfectly compatible with the RAIS evaluation: it uses a self reported and top-coded income measure that considers only the respondent's main job (self-reported income

of this kind is commonly under-reported relative to administrative payroll data such as RAIS), and uses a separate, non-identical AI exposure measure built from PNADC's own occupation classification.

### 3.4 Supplementary Case-Study: AI-Assisted Debt Collection at Brazilian Retail Bank

To evaluate AI adoption within an actual Brazilian firm, rather than occupation based estimates, this analysis uses internally generated data from a large Brazilian retail bank's debt collection operations, covering November 2025 to January 2026. The bank had assigned eligible delinquent accounts (N = 664,144) to a test group, which was 79.5% of total eligible balance (R$ 1.9998 billion, 528,622 accounts) or a control group (R$513.3 million, 135,522 accounts), and outcomes were compared using the share of eligible balance converted into a negotiated agreement. The test group converted 2.2009% of its eligible balance, compared with 1.8482% for control, a difference of 0.35 percentage points that represents around R$ 7.05 million in additional recovered balance if applied to the test group's balance. Due to the difference in size between the two groups, this estimate should be read as associational rather than proof that AI caused the higher recovery rate.

Because the test group was about four times as large as the control group, this paper tested whether the conversion rate gap was caused by the difference in group sizes. One thousand random subsamples of the test group, each equal in size as the control group, was drawn independently, and the conversion rate gap was recalculated for each sample. The average gap was 0.352 percentage points (SD = 0.059; 95% range: 0.237-0.473), almost the same as the 0.353 gap from the previously calculated unbalanced raw groups. So the size difference between the two groups doesn't seem to be driving the conversion rate gap. The balanced groups' recovered balance is much lower (R$1.80 million, SD = R$301,715) than the earlier R$7.05 million, but that's just because it's scaled to a smaller group. What actually matters here is the recovery rate, not the total amount recovered.

A secondary analysis examined the test group's composition. Among accounts that converted, 10.6% of the test's conversions used the bank's AI channel, compared with only 1.1% in the control group. Still, AI was a minority channel within the test group - most conversions came through the regular collection channels. So the test/control split looks more like a broader operational difference than a clean AI-vs-human comparison.

Another analysis evaluated whether the AI channels were offering less discounts (more revenue for the bank). Agreements closed through the AI channel had a discount rate of 40.6%, compared with 43.1% across all other channels combined. That gap wasn't consistent across individual channels though - it was mostly driven by the largest human-staffed channel, which had a 57.1% discount rate, while the other two human channels actually offered lower discount rates than AI. So the AI channel did offer a lower discount rate overall, just not lower than every single channel.

Internal customer service transcripts also provided an example that supposedly illustrates how the AI channel deals better with customers. In the specific example, a customer going

through a phase of financial hardship mentioned that they were considering suicide. The AI model handled it very well, consoling the person and providing them with the national suicide hotline. This example should not be used as a systematic pattern but as an example of how well the model can operate in diverse situations and still manage to continue pursuing a debt resolution. These findings should be interpreted as a firm-specific detailed example of how AI adoption is occurring in Brazil, not part of the paper's central empirical testing.

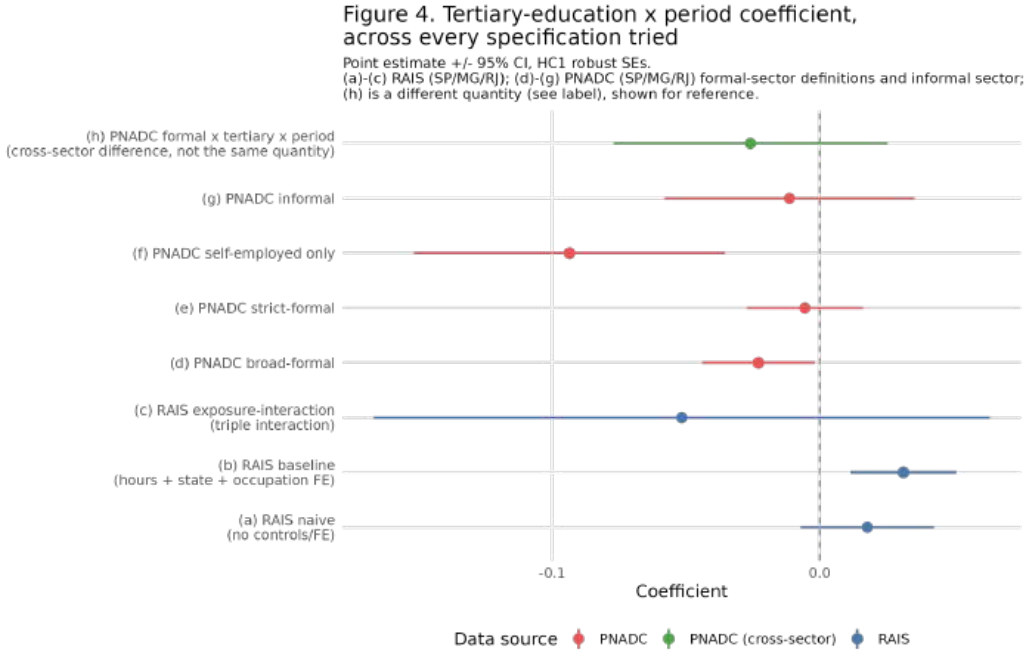
**4. Discussion**

The hypothesis of the study has three claims: within formal employment, workers whose skills complement AI-exposed tasks may receive a larger wage premium; in the informal market, this effect may be weaker because AI exposure is lower; and across formal and informal employment, inequality may still increase if automation complements formal, tertiary-educated workers without reaching informal workers. RAIS (covers only formal employment) could directly test the first claim. The second and third claim, which require informal sector data, were tested separately using PNADC.

Regarding the first claim, the estimated tertiary-education wage premium was approximately 3% larger in 2022 than in 2017 in the fully controlled specification ($p = 0.002$). Once the controls were removed though, the estimate shrank and stopped being significant. So the increase might just come down to differences in occupation, state, or hours worked between the two groups, not education itself. The exposure interaction regarding this claim was small, negative and statistically insignificant ($p = 0.38$). There are several possible reasons for this result. Maybe the wage gap just didn't widen the way the hypothesis said it would. Or maybe the exposure measure was too broad and missed real differences between occupations in Brazil. The 2017-2022 window might not have been long enough either. AI adoption among Brazilian firms stayed pretty low for most of that stretch anyway. Other things unrelated to AI exposure could have gotten in the way too, like how COVID-19 disrupted the labour market.

The second and third claims, tested using formal-informal data from PNADC, produced no statistically significant results ($p = 0.64$ and $p = 0.32$). However, this result should be seen as a matter of lack of significant evidence rather than evidence that no effect exists, given the limitations already discussed regarding PNADC (single-quarter, income capped design etc.). The most important finding of this phase, however, is in relation to the first claim, not the second or third one. Reexamining the formal sector employment premium using PNAD Contínua showed that its estimated direction and significance are unstable across data sources and definitions of formal employment. The estimate varied significantly from positive and significant in RAIS ($+0.031$, $p = 0.002$), to negative and statistically significant under a broad PNADC definition ($-0.023$, $p = 0.034$), to a small and statistically insignificant negative estimate under the definition most consistent with RAIS' coverage ($-0.006$, $p = 0.62$), and to a larger negative and statistically significant estimate when restricted to self-employed workers and employers ($-0.094$, $p = 0.002$). Possible explanations for such differences may be due to RAIS' employer-reported wages and PNADC's self-reported income, or differences between census and household survey

with sampling error across subgroups. These unstable outcomes are treated as central findings to the study rather than manipulated in favor of a specific estimate that would be the best match to confirm the hypothesis. Figure 4 plots every specification's point estimate and 95% confidence interval side by side. None of them stay both stable in sign and statistically significant once you change the data source or how formality is defined.




This pattern kind of matches what Acemoglu and Restrepo describe: wage inequality only widens when a technology's displacement effect outweighs its task-creation effect. It also lines up with Firpo and Portella's finding that Brazilian wage inequality has recently stagnated instead of sharply increasing. Maybe there just aren't real differences across groups, or maybe the exposure measure is too broad to pick them up. Either way, that could explain the lack of significant exposure and cross-sector effects. Gmyrek, Winkler and Garganta and Adamczyk, Ehrl and Monasterio both suggest that more detailed occupation-level measures would be more likely to catch these differences, if they're really there. The sensitivity of the informality results backs this up too: the classification followed Ulyssea's approach, but the results still shifted depending on how formal employment was defined. All three estimates came out negative, but they varied a lot in size and significance. So the definition itself seems to be shaping the estimated effect, not just how precisely it's measured.

The supplementary case study provides a firm-level illustration of how AI adoption currently looks like in Brazil, showing a modest conversion-rate increase and, overall but not consistently across all channels, smaller discounts when operated by the AI agent (more revenue for the bank). Other data sets, like IBGE's Pintec Semestral survey (IBGE, "PINTEC Semestral"), covering the 2022 and 2024 rounds, directly measure firm-level adoption of AI and other advanced digital technologies across more Brazilian firms, and could complement or

extend that single-firm illustration and contextualization in future works. An additional sensitivity check confirmed that the conversion-rate gap was not caused by the substantial size difference between the test and control groups: averaging the gap across 1000 randomly drawn, size-matched subsamples produced a result of 0.352 percentage points, statistically indistinguishable from the original 0.353-point estimate based on the unbalanced groups. As stated in Section 2, this case-study speaks to AI adoption and its firm-level operational effects - a different objective from the central sources PNADC and RAIS in the empirical investigation above - so it should be read as a contextual illustration rather than strengthening evidence for the wage-premium findings discussed above.

**5. Conclusion**

This study utilized RAIS employer-employee microdata from 2017 and 2022 to evaluate, within the formal labour market of São Paulo, Rio de Janeiro and Minas Gerais, whether the tertiary-education wage premium changed between the period studied, and whether that effect differed depending on occupational exposure to AI automation. This section of the analysis did not directly measure firm-level AI adoption, and does not contain a full traditional difference-in-differences model, since there is no clear treatment/control group, and only 2 years being tested. The analysis found a slight increase in tertiary-educated workers' skill premium - sensitive to specifications, though -, but no statistically significant evidence that this increase was caused by automation exposure, or that this change was more intense in more exposed occupations.

A supplementary analysis conducted using PNAD Contínua formal and informal labour market data tested the remaining two parts of the hypothesis. It found no significant statistical evidence of change in the skill premium within the informal sector, neither between formal and informal workers. Furthermore, as seen when running the calculations, it is revealed that the direction and significance of this estimate are not stable across different specifications, data source or sector definitions. This instability is reported as a finding instead of filtering for the results most consistent with the hypothesis. Given that RAIS and PNADC don’t fully line up with each other, this analysis should be treated as a solid first attempt at understanding the informality factor, not a definitive answer.

A supplementary case-study covering how a large Brazilian retail bank has implemented AI into its debt-collection operations illustrated how automation has been occurring in Brazil in practice. Analyzing internal-data, a modest increase in conversion-rates were confirmed, and smaller discounts associated with the AI channel (though not completely uniform). A consistency check was conducted to ensure equal sample sizes, and the results were the same, showing that the increase is not attributable to the significant size imbalance between the reported comparison groups. This descriptive single-firm case-study serves as a clear contextualization of how automation is taking place in Brazil, and should not be treated as additional support to the study’s main empirical wage-premium findings reported above.

Several extensions could sharpen this study's findings. Firstly, because firm-level AI adoption in Brazil appears to have accelerated mainly after 2022, future work should extend the time period used for analysis to more recent RAIS and PNADC data, supplementing the 2017-2022 comparison covered in this paper with a later period in which AI's labour market effects are more likely to be detected. Secondly, IBGE's Pintec Semestral survey (IBGE, "PINTEC Semestral"), which directly measures firm-level adoption of AI or any other advanced technologies, could be either used as a case-study or to construct a second exposure measure complementing the occupation-level exposure scores used in this study. Third, future work should explicitly adjust and test for COVID-19 effects on the labour market - for example, a pandemic-related structural break in the wage and informality trends studied, given the pandemic's documented effect on Brazilian informal employment. Lastly, replacing the ten-category Frey and Osborne based exposure score with a more detailed, Brazil-tailored, AI index - possibly a measure somewhat similar to those constructed by Gmyrek et al. or Adamczyk et al. - would directly address this study's most significant measure limitation, and could reveal exposure effects that were undetected by the current coarse measure.

**Works Cited**


Acemoglu, D., and P. Restrepo. "Tasks, Automation, and the Rise in U.S. Wage Inequality." Econometrica, vol. 90, no. 5, 2022, pp. 1973-2016, doi.org/10.3982/ECTA19815.

Adamczyk, W., et al. "Skills and Employment Transitions in Brazil." International Labour Review, vol. 162, 2023, pp. 1-30, doi.org/10.1111/ilr.12412.

Autor, D., and A. Salomons. "Is Automation Labor-Displacing? Productivity Growth, Employment, and the Labor Share." Brookings Papers on Economic Activity, vol. 2018, no. 1, 2018, pp. 1-63, doi.org/10.3386/w24871.

Cazzaniga, M., et al. "Gen-AI: Artificial Intelligence and the Future of Work." IMF Staff Discussion Note No. 2024/001, International Monetary Fund, 2024, doi.org/10.5089/9798400262548.006.

Ciaschi, M., et al. "The Potential Distributive Impact of AI-Driven Labor Changes in Latin America." IDB Working Paper, Inter-American Development Bank, 2025, doi.org/10.18235/0013677.

Firpo, S., and A. Portella. "Declining Wage Inequality in Developing Countries: The Case of Brazil." Journal of Economic Literature, vol. 63, no. 3, 2025, pp. 875-915, doi.org/10.1257/jel.20241697.

Frey, C. B., and M. A. Osborne. "The Future of Employment: How Susceptible Are Jobs to Computerisation?" Oxford Martin School, University of Oxford, 2013, www.oxfordmartin.ox.ac.uk/downloads/academic/The_Future_of_Employment.pdf.

Gmyrek, P., et al. "Buffer or Bottleneck? Employment Exposure to Generative AI and the Digital Divide in Latin America." ILO Working Paper 121, International Labour Office and World Bank, 2024, doi.org/10.54394/TFZY7681.

Instituto Brasileiro de Geografia e Estatística. "Pesquisa Nacional por Amostra de Domicílios Contínua (PNAD Contínua)." IBGE, 2023, www.ibge.gov.br/estatisticas/sociais/trabalho/17270-pnad-continua.html.

Instituto Brasileiro de Geografia e Estatística. "Pesquisa de Inovação Semestral (PINTEC Semestral): Indicadores de adoção de tecnologia pelas empresas." IBGE, in partnership with ABDI and UFRJ, 2024, www.ibge.gov.br/estatisticas/investigacoes-experimentais/estatisticas-experimentais/35867-pesquisa-de-inovacao-semestral.html.

Ministério do Trabalho e Emprego. "Relação Anual de Informações Sociais (RAIS): Microdados vínculos." Governo Federal do Brasil, 2022, www.gov.br/trabalho-e-emprego/pt-br/assuntos/estatisticas-trabalho/rais.

R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, 2026, doi.org/10.32614/R.manuals.

Ulyssea, G. "Firms, Informality, and Development: Theory and Evidence from Brazil." American Economic Review, vol. 108, no. 8, 2018, pp. 2015-2047, ora.ox.ac.uk/objects/uuid:553b76be-e2f7-4db0-b5a5-d5342b48e453.

# The Twenty-Five Dollar Rule: Emergence of Decision Heuristics in a Ten-Year Longitudinal Single-Subject Study of 644 Micro-Lending Choices (2016–2026)

By Aaliyah Chaudhary

## ABSTRACT

Behavioral economics has studied many of the shortcuts people use when making decisions, but it is much harder to see how one of those shortcuts actually develops. Doing that requires a long record of the same person's real choices over time. This study uses a record of 644 micro-loans made by one lender, who is also the author, on Kiva.org from age 7 to 17 (July 2016 to February 2026; $22,720 across 52 countries). Five analyses compared an early period (2016–2019, ages 7–10) with a later period (2020–2026, ages 11–17). Three main patterns suggest that decision rules became more stable over time. Loans made at exactly $25, Kiva's minimum contribution, increased from 8.0% to 74.6% ($\chi^2 = 93.2$, $p < 0.001$), while the number of loans made also increased sharply. Loans to sectors connected with women's economic activity increased from 28.0% to 49.7% ($p = 0.005$). The percentage of female borrowers did not significantly change (76.0% vs. 78.1%, $p = 0.87$), but the year-to-year variation became much smaller (SD fell from 15.0 to 6.6 percentage points). Two original hypotheses were not supported: sector diversity increased instead of decreasing (Shannon entropy 4.83 to 5.10 bits), and the overall level of the gender preference stayed about the same. A repayment check also found no advantage for loans to female borrowers. Overall, the results suggest that some parts of the lender's decision-making became more automatic while other parts stayed flexible.

**KEYWORDS** Behavioral and Social Sciences; Behavioral Economics; Heuristics; Prosocial Development; Micro-lending.

## INTRODUCTION

How do decision-making rules form? Behavioral economics has described many shortcuts that adults use when making choices. These include mental accounting, where people separate money into different categories, fast-and-frugal rules that reduce the amount of information needed for a decision, and warm-glow giving, where people get satisfaction from giving even when they receive no financial benefit. Most studies, however, look at people after these habits are already in place. Lab studies usually last one session, and even longer studies often rely on self-reports instead of a continuous record of real choices. Because of this, it is difficult to observe the point where repeated decisions start turning into a stable rule.

This study looks at a record where that change may be visible. The dataset includes 644 micro-loans made on Kiva.org between July 2016 and February 2026, from age 7 to age 17, totaling $22,720 across 52 countries. Kiva is a nonprofit lending platform where users choose individual borrowers and can contribute starting at $25. Since repaid money can be lent again, regular users can build a long record of repeated choices. In this case, the lender, who is also the

author, started using Kiva in childhood and continued through adolescence. This creates an unusual record of real financial decisions made by the same person over almost ten years.

Previous research on Kiva has mostly looked at large groups of lenders. Galak, Small, and Stephen found that lenders often favor borrowers who are similar to them in areas such as gender and occupation.

Developmental research has also shown that prosocial preferences start forming early. Children around ages 7–8 often show stronger preferences for equal sharing, while ideas about fairness continue to develop through adolescence. What these two areas of research do not show very well is how one person's giving choices change across the same developmental period.

This study asks whether the lending record shows a shift from more variable choices to more consistent, rule-like choices. Five analyses examine contribution amount, sector choice, borrower gender, geographic and institutional breadth, and repayment outcomes. The clearest change was the use of exactly $25: this increased from 8.0% of loans in the early period to 74.6% later ($\chi^2 = 93.2$, $p < 0.001$). Loans to sectors connected with women's economic activity also increased from 28.0% to 49.7% ($p = 0.005$). The percentage of female borrowers did not significantly increase, but it became more consistent from year to year. Two original hypotheses were not supported: sector diversity increased instead of narrowing, and the overall gender-preference level stayed statistically unchanged. The repayment analysis also did not show a financial advantage for lending to women, which fits better with a values-based explanation than a returns-based one.

What this study can offer is narrow, but it is specific. It follows one developing decision-maker closely enough to show a possible shift from choosing each case separately to using more stable rules. The biggest break in the record appears around age 11. That timing may reflect development. It may equally reflect the fact that an eleven-year-old has more money to lend, and more say over the account, than a seven-year-old does. Because this is a single-subject study, it cannot be used to estimate how common the same pattern is in other people. Instead, it provides one detailed example that could later be compared with larger samples.

## LITERATURE REVIEW

### DEVELOPMENT OF PROSOCIAL BEHAVIOR IN CHILDHOOD AND ADOLESCENCE

Prosocial behavior begins to take shape early in childhood. In a set of incentivized choice experiments, Fehr, Bernhard, and Rockenbach found that children became much more likely to prefer equal sharing between ages 3–4 and 7–8, even when equality came at a personal cost. Fairness-related behavior continues changing after that point. McAuliffe and colleagues describe fairness norms as developing through late childhood and adolescence.

Most of this research is based on cross-sectional studies and laboratory tasks, such as sharing games, sticker divisions, or hypothetical situations. These studies are useful, but they usually observe each child only once and use choices created by the researcher. Longitudinal

studies exist, but they often collect data at large time intervals and depend on parent or self-report. As a result, there is still very little continuous evidence of how one child's real prosocial decisions change over several years. The lending record in this study provides one example of that kind of data.

## BEHAVIORAL ECONOMICS OF GIVING AND HEURISTIC DECISION-MAKING

A second area of research focuses on the thinking behind adult giving and decision-making. Andreoni's theory of impure altruism argues that people can receive a personal sense of satisfaction, or a "warm glow," from giving even when they do not receive a material reward. Kahneman and Tversky showed that people often use simpler judgments instead of fully analyzing every option when making decisions under uncertainty. Gigerenzer and Goldstein similarly argued that fast-and-frugal rules can be useful because they reduce effort without always reducing decision quality. Thaler's work on mental accounting is especially relevant here because it explains how people divide money into separate categories and use informal rules for how each category is spent. Thaler and Sunstein also showed how defaults and other features of choice architecture can shape people's decisions.

Most of this literature studies heuristics after they already exist. Researchers usually ask whether a shortcut is present, how strong it is, or when it is used. Much less research follows one person closely enough to see a specific shortcut develop over time. To observe that process, the same person would need to make many comparable choices in a relatively stable setting. The Kiva record gives a chance to look at that kind of transition.

## MICRO-FINANCE PLATFORMS, LENDER CHOICE, AND GENDER

A third area of research looks at which borrowers lenders choose to support. In a large Kiva study, Galak, Small, and Stephen found that lenders were more likely to support borrowers who were socially similar to themselves, including similarities in gender and occupation. The wider microfinance literature also shows that the industry often focuses on women borrowers and that women generally have strong repayment records. Kiva has historically reported repayment rates close to 97%, which means repayment risk may not be a major factor separating many loan options. This leaves more room for personal preferences and values to influence the decision.

These studies usually look at many lenders at one point in time or over a limited period. They can show that certain preferences exist, but they do not show clearly how those preferences began or became more consistent. To the author's knowledge, no published study follows one lender's full Kiva portfolio across childhood and adolescence. This study uses the opposite approach: one lender, 644 decisions, and almost ten years of data. The goal is to see whether preferences that appear in large-group studies can also be seen developing within one person.

## METHODS

## SUBJECT AND SETTING

The subject is a United States-based lender who made her first Kiva loan in July 2016 at age 7. The record continued through February 2026, when she was 17. The subject is also the author of this study, which creates possible bias and is discussed later under Ethical Considerations and Limitations. Lending took place through a standard individual Kiva account, and repaid principal was usually lent again, creating activity across the full time period.

Kiva.org is a nonprofit person-to-person micro-lending platform founded in 2005. Users browse profiles of individual borrowers or borrowing groups. The profiles include information such as the borrower's name, photo, country, activity, reason for the loan, and partner institution. Lenders can contribute starting at a minimum of $25. The lender does not earn interest, and repaid money returns to the account so it can be withdrawn or lent again. Each funded loan can therefore be treated as a separate, timestamped decision involving real money.

## DATASET

The dataset is the subject's complete Kiva loan record, exported from the account on April 3, 2026. It contains 644 loans totaling $22,720 across 52 countries and 97 field-partner institutions. The variables used were the amount the subject contributed ("You Loaned"), purchase date, borrower gender as listed by Kiva (502 female and 142 male in this dataset), country, activity, field partner, and loan status. There were 82 different activity labels in the record. Repayment-related fields included amount paid back, currency loss, and amount defaulted. No loans were removed from the dataset.

Two parts of the dataset need clarification. First, borrower gender is based on Kiva's label. For group loans, which make up 4.8% of the records, the gender category applies to the group as Kiva classified it and does not necessarily describe every member. Second, the status "ended with loss" mostly reflects small currency-exchange losses rather than borrowers failing to repay. Where a currency loss occurred, the mean loss was $0.60. There were only 8 true defaults out of 644 loans (1.24%), totaling $72.68. Because these are different types of loss, the repayment analysis treats them separately.

## PERIOD DEFINITION

The analysis compares an early period (2016–2019, ages 7–10; n = 50 loans) with a later period (2020–2026, ages 11–17; n = 594 loans). The split was chosen because a clear change in loan amount, volume, and pace appears between 2019 and 2020. It also falls around the subject's eleventh year. However, the split was partly chosen after looking at the same data that were later tested, so these comparisons should be treated as exploratory rather than fully confirmatory. Year-by-year values are included when possible so the pattern can also be evaluated without relying only on the two-period comparison.

## ANALYSES

Five analyses were completed in Python using pandas and SciPy. Statistical significance was evaluated at $\alpha = 0.05$. Chi-square tests of independence were used for comparisons of proportions between the early and later periods, and binomial proportions are reported with Wilson 95% confidence intervals.

Analysis 1 – Loan denomination. For each year, I calculated the percentage of loans made at exactly $25, which is the platform minimum, and reported Wilson confidence intervals. The early and later periods were compared using a chi-square test. The percentage of $25 loans was used as a measure of whether a fixed contribution rule had developed, consistent with the idea of mental accounting.

Analysis 2 – Sector allocation. I compared the percentage of loans in a researcher-defined group of sectors commonly connected with women's economic activity in the microfinance literature. These included farming, food production or sales, retail, general store, clothing sales, fruits and vegetables, and grocery stores. The early and later periods were compared using a chi-square test. Since this grouping was created by the researcher rather than Kiva, the full activity distribution was also measured using Shannon entropy in Analysis 3.

Analysis 3 – Sector diversity. Shannon entropy, measured in bits, was calculated for the activity distribution by year and for each pooled period. The original hypothesis predicted that entropy would decrease as sector choices became narrower. The result was reported even if the pattern went in the opposite direction.

Analysis 4 – Borrower gender. The percentage of loans to female borrowers was calculated by year with Wilson confidence intervals. The early and later periods were compared using a chi-square test. I also compared the standard deviation of the yearly female-borrower percentages to see whether the preference became more consistent over time, even if the overall level did not increase.

Analysis 5 – Breadth and robustness. I measured geographic and institutional breadth using the number of unique and cumulative countries, the number of field partners, and within-year Gini concentration across countries. I also compared lender-loss rates by borrower gender. Currency losses and true defaults were kept separate to test whether the preference for female borrowers could be explained by better financial outcomes.

## ETHICAL CONSIDERATIONS

This study is a retrospective analysis of the researcher's own transaction history. No outside participants were recruited, surveyed, or interviewed. The borrower information used in the dataset is limited to information Kiva already publishes publicly, and no borrower is identified in the paper beyond grouped statistics. Institutional review was not sought for these reasons. Since the researcher is also the subject, there is still a risk of interpreting past behavior in a favorable way. I tried to reduce that risk by stating hypotheses in advance, reporting results that did not support them, and limiting the main claims to patterns that can be calculated from the transaction record rather than relying on memory. All statistics were calculated from the exported

file without changing the records, and the analysis code and dataset are available from the author on request.

## RESULTS

### OVERVIEW OF THE RECORD

Table 1 summarizes the lending record by year. Two changes stand out. First, the number of loans increased sharply around the period boundary. There were 50 loans from 2016–2019 (ages 7–10) and 594 from 2020–2026 (ages 11–17). The approximate number of loans per active month also increased from around 2–6 to around 15–33. Second, several other measures change around the same point, which is described in the analyses below.

**Table 1.** Lending record by year, 2016–2026.

| Year | n | $25 share (%) | Mean loan ($) | WEE share (%) | Entropy (bits) | Female share (%) | 95% CI |
|---|---|---|---|---|---|---|---|
| 2016 | 9 | 0.0 | 86 | 33.3 | 2.95 | 100.0 | [70, 100] |
| 2017 | 12 | 16.7 | 67 | 33.3 | 3.42 | 75.0 | [47, 91] |
| 2018 | 17 | 11.8 | 72 | 17.6 | 3.73 | 70.6 | [47, 87] |
| 2019 | 12 | 0.0 | 75 | 33.3 | 3.25 | 66.7 | [39, 86] |
| 2020 | 61 | 85.2 | 33 | 54.1 | 4.19 | 65.6 | [53, 76] |
| 2021 | 69 | 88.4 | 28 | 53.6 | 4.41 | 76.8 | [66, 85] |
| 2022 | 136 | 88.2 | 29 | 53.7 | 4.54 | 80.1 | [73, 86] |
| 2023 | 131 | 48.1 | 37 | 53.4 | 4.76 | 79.4 | [72, 85] |
| 2024 | 119 | 72.3 | 33 | 48.7 | 4.38 | 82.4 | [75, 88] |
| 2025 | 49 | 67.3 | 33 | 22.4 | 4.61 | 81.6 | [69, 90] |
| 2026 | 29 | 96.6 | 26 | 44.8 | 4.05 | 69.0 | [51, 83] |

### ANALYSIS 1: ADOPTION OF A FIXED CONTRIBUTION UNIT

During the early period, exactly $25 was not a common contribution amount. Only 4 of 50 loans (8.0%) were made at $25, and the mean contribution was $74.00, with amounts ranging from $25 to $150. Starting in 2020, $25 became the most common amount. In the later period, 443 of 594 loans (74.6%) were exactly $25, and the mean contribution fell to $32.02 ($\chi^2 = 93.2$, $df = 1$, $p < 0.001$; exact $p = 4.8 \times 10^{-22}$). The change was sudden rather than gradual (Figure 1): the $25 share went from 0% in 2019 to 85.2% in 2020. It stayed at 48% or higher in every later year and reached 96.6% in the partial year 2026. At the same time, the total number of loans increased. The smaller average contribution therefore did not mean less giving. The same money was reaching more borrowers.

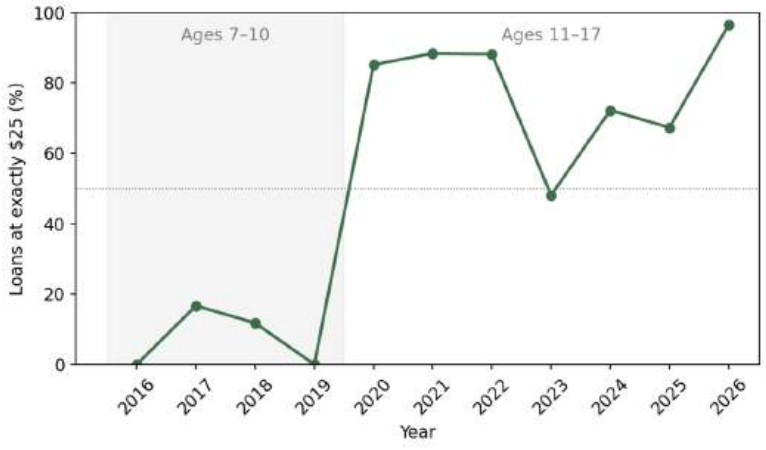



**Figure 1.** Percentage of loans made at exactly $25, Kiva's minimum contribution, by year from 2016–2026. The shaded area shows the early period (2016–2019, ages 7–10). The percentage rises sharply at the 2019/2020 boundary. The lower value in 2023 (48.1%) happened during a temporary increase in $50 contributions.

The lower $25 share in 2023 (48.1%; mean $37.33) came from a temporary increase in $50 contributions rather than a full return to the more varied early-period pattern. Together, $25 and $50 contributions made up 96.9% of loans in 2023. If fixed platform denominations are defined as either $25 or $50, the later-period share is 98.1% compared with 44.0% in the early period. Across the full dataset, 638 of 644 loans (99.1%) were exact multiples of $25, including the early years. This shows that the platform's denomination system influenced contribution amounts throughout the record. What changed after 2019 was the much stronger concentration on the smallest amount.

## ANALYSIS 2: REALLOCATION TOWARD WOMEN'S ECONOMIC-ACTIVITY SECTORS

The percentage of loans in the researcher-defined WEE sector group increased from 28.0% (14/50) in the early period to 49.7% (295/594) in the later period ($\chi^2 = 7.83$, df = 1, $p = 0.005$). In every early year, the percentage was between 17.6% and 33.3%. In later years it was between 44.8% and 54.1%, with 2025 as the main exception at 22.4% because that year included a cluster of sanitation-system loans. Since the increase appears across several years, it does not seem to be caused by just one unusual year.

## ANALYSIS 3: SECTOR DIVERSITY (PRE-SPECIFIED HYPOTHESIS NOT SUPPORTED)

The original hypothesis predicted that sector choices would become narrower as a heuristic formed. The data did not support this. Pooled Shannon entropy increased from 4.83 bits in the early period to 5.10 bits in the later period. Yearly entropy also generally increased, from 2.95 bits with 8 distinct activities in 2016 to a high of 4.76 bits with 45 activities in 2023 (Table 1; Figure 3). Some of this increase is expected because later years had more loans, which makes

it easier for more categories to appear. Even so, the overall direction is clear: activity choices became more diverse, even while lending became more concentrated in the WEE group and at the $25 amount. This means the pattern of heuristic formation was not the same across every type of decision.

## ANALYSIS 4: BORROWER GENDER – STABLE LEVEL, COLLAPSING VARIABILITY

The percentage of loans to female borrowers was almost the same across the two periods: 76.0% (38/50) early and 78.1% (464/594) later ($\chi^2 = 0.03$, df = 1, p = 0.87). Therefore, the hypothesis that the gender preference would increase was not supported. The more noticeable change was consistency. In the early period, yearly female-borrower percentages ranged from 66.7% to 100%, with a standard deviation of 15.0 percentage points. In the later period, they ranged from 65.6% to 82.4%, with a standard deviation of 6.6 points. This is a reduction of more than half (Figure 2). The early yearly samples are small, so some of the higher variability is expected from sample size alone. Still, the overall pattern fits a preference that was already present, becoming more consistently applied rather than becoming stronger. Across all 644 loans, 78.0% went to female borrowers.

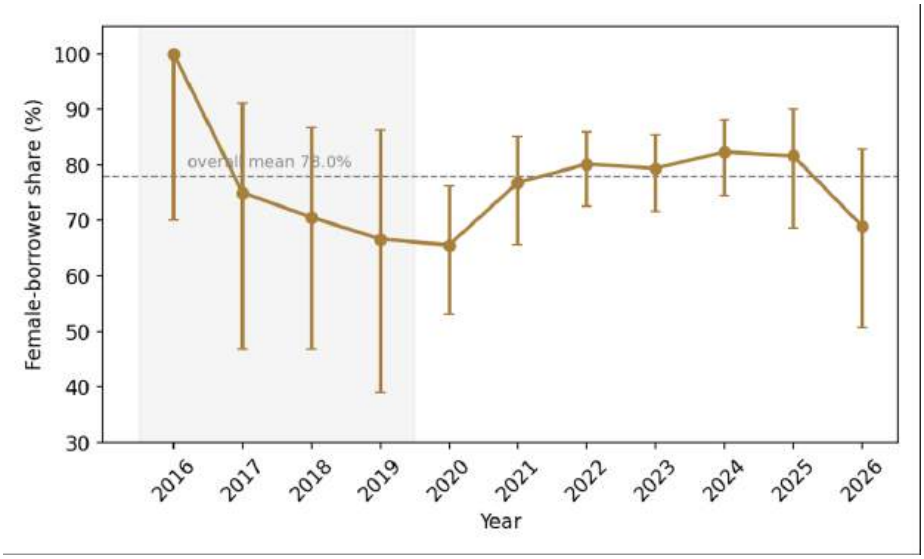



**Figure 2.** Female-borrowers share by year with Wilson 95% confidence intervals. The dashed line shows the overall mean of 78.0%, and the shaded area shows the early period. After 2020, yearly percentages stay closer to the overall mean. The standard deviation falls from 15.0 percentage points in 2016–2019 to 6.6 in 2020–2026, while the pooled percentage does not significantly change (p = 0.87).

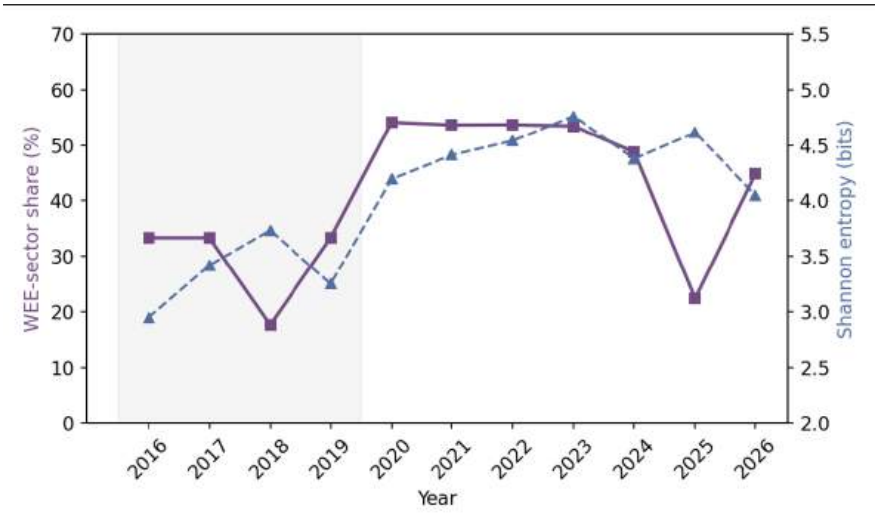



**Figure 3.** The left axis shows the percentage of loans in women's-economic-activity (WEE) sectors by year, and the right axis shows Shannon entropy of the full sector distribution. The WEE percentage increased from 28.0% to 49.7% across the pooled periods ($p = 0.005$), while overall sector diversity also increased from 4.83 to 5.10 bits. This goes against the original prediction that sector choices would narrow.

## ANALYSIS 5: BREADTH – AND A ROBUSTNESS CHECK ON RETURNS-SEEKING

Geographic and institutional breadth increased over time. The total number of countries reached 21 by the end of 2019 and 52 by 2026. Unique field partners increased from 28 in the early period to 81 in the later period, while the largest single partner's share of loans fell from 18.0% to 9.9%. At the same time, within-year country concentration, measured with Gini, increased from 0.13 -- 0.39 in the early years to 0.30 -- 0.61 later. Looking at the monthly data helps explain this. Lending often happened in country-specific bursts. For example, Ecuador received 24 of its 54 loans (44%) in July 2023. Bolivia received 21 of 33 loans (64%) and the Democratic Republic of the Congo received 17 of 27 (63%) in June 2022. Paraguay received 21 of 69 loans in January 2022. This suggests that certain countries became a temporary focus for a group of loans before attention moved elsewhere.

To test whether the preference for female borrowers could be explained by better financial outcomes, lender-loss rates were compared by borrower gender. Loans to female borrowers ended with some type of loss, either currency loss or default, in 39.6% of cases compared with 31.0% for male borrowers. The difference was not statistically significant ($\chi^2 = 3.17$, $df = 1$, $p = 0.075$). Most of these losses were very small currency-exchange losses, with a mean of $0.60, rather than true non-repayment. There were only 8 defaults out of 644 loans (1.24%), totaling $72.68 out of $22,720 in principal. The analysis therefore shows no financial advantage to lending to women, which leaves a returns-based explanation of the 78% allocation with little to stand on.

## DISCUSSION

The lending record shows that some parts of one person's decision-making became much more consistent between childhood and adolescence, while other parts stayed open and varied. Three parts of the pattern are especially important to discuss.

THE $25 RULE AS MENTAL ACCOUNTING – AND AS CHOICE ARCHITECTURE

The clearest result is the strong shift toward Kiva's $25 minimum contribution. In terms of Thaler's mental accounting framework, the $25 amount may have become a simple unit for lending rather than something that had to be decided separately for every borrower. Instead of asking "how much should I lend to this person?", the decision could become closer to "how many borrowers should I fund?" This would reduce one decision for every loan. The change also happened alongside a large increase in the number of loans and a wider geographic reach, which is what a fast-and-frugal rule is supposed to buy: more decisions, less effort spent on each one.

There are at least two possible explanations for where the $25 rule came from. One possibility is that it developed through experience, as the subject gradually learned that contributing more than the minimum was not necessary for each loan. Another possibility is that the platform itself encouraged the rule because $25 is Kiva's minimum and an obvious default amount. Research on choice architecture shows that people often adopt options provided by a platform or institution as their own routine. The data cannot fully separate these explanations. However, the subject did not use the minimum consistently during the first four years, when the average contribution was much higher. The shift happened suddenly, from 0% at $25 in 2019 to 85.2% in 2020. The platform default was available the whole time. Something in 2020 made it stick.

**DIMENSION-SPECIFIC RULE FORMATION**

The sector-diversity result is useful because it shows that not every part of the lending behavior became narrower. If contribution amounts, sectors, gender, and countries had all become more repetitive at the same time, the pattern could simply reflect habit or less attention. Instead, the amount choice became much more fixed, the gender preference became more stable, and sector and country choices became broader. In 2023 alone, the subject funded 45 different activity types. This fits Gigerenzer's view that heuristics get applied selectively, in the places where simplifying saves the most effort rather than everywhere at once.

At the same time, sector choices did show a clear direction. The WEE share increased from 28.0% to 49.7% ($p = 0.005$), even though total sector diversity increased. In other words, the subject did not stop exploring different business types. Instead, more of those choices came from sectors connected with women's economic activity. The preference became more visible without making the whole portfolio narrow. The country-level bursts may show another effort-saving pattern. Instead of choosing a completely new country for every loan, the subject sometimes appeared to focus on one country for several loans in a row. This is consistent with ideas such as status quo bias or anchoring on a recent choice. However, the platform may also have contributed to the pattern if Kiva showed several borrowers from the same partner or country together. The record cannot tell exactly how much came from the lender and how much

came from the platform. Either way, batching could reduce the number of separate country decisions made during one lending session while still producing a portfolio that covered 52 countries overall.

## THE GENDER PREFERENCE: STABLE IN LEVEL, AUTOMATIC IN APPLICATION

The gender result needs to be interpreted carefully. The preference for female borrowers did not get stronger over time. It was already present in the earliest loans: 100% of loans in 2016 were to female borrowers, and the pooled early-period share was 76.0%. The later-period share was 78.1%, which was not significantly different ($p = 0.87$). What changed more clearly was the year-to-year consistency. The preference appears to have become more regularly applied rather than larger.

Galak, Small, and Stephen offer another possible explanation: Kiva lenders often favor borrowers who are similar to themselves, including by gender, and the subject in this study is female. This study cannot fully separate a preference for supporting women's economic activity from similarity-based sympathy, because both could produce the same pattern. Two results still help narrow the interpretation. First, loans to female borrowers did not provide better financial returns, so the preference does not appear to be explained by returns-seeking. This is more consistent with a giving motive such as warm glow. Second, lending also shifted toward sectors connected with women's economic activity, which is not fully explained by gender similarity alone. The preference looks values-based rather than strategic, and it became more consistent in how it was applied. Similarity to the borrower may reinforce it, but similarity alone does not account for the sector shift.

## THE TIMING: AGE ELEVEN, AND ITS RIVALS

The biggest change in the record falls between ages 10 and 11. One possible explanation is developmental. Classical theories often place the move from concrete-operational to formal-operational thinking around age 11. The later lending pattern, which uses fixed amounts and more stable allocation rules, could fit with the development of more abstract ways of organizing decisions. However, this study only follows one person, so it cannot show that development caused the change. Other explanations line up with the same period. The subject likely had more money available as she got older, which would make a many-small-loans strategy easier. She may also have had more control over the Kiva account than she did at age 7. Finally, the change happened in 2020, when the pandemic changed daily routines and time spent at home. The rules become visible at this boundary. Why they become visible here is a question this record cannot settle.

## LIMITATIONS

The design limits what these numbers can support. One person made all 644 of these decisions, so nothing here estimates a rate in any population; the record shows what is visible in a single detailed case. The early and later split was also chosen partly after I had already seen a

major break in the data, which makes the period comparisons exploratory. The yearly samples from 2016 to 2019 are small (n = 9 to 17), and small samples vary more on their own, which inflates some of the early-period variability sitting underneath the gender result. I am also the subject. Analyzing my own transaction history creates an obvious risk of reading the past in a flattering direction. Stating hypotheses in advance, reporting the two that failed, and restricting claims to what the exported file can compute all reduce that risk. None of them removes it.

Several things change with age at once: how much money I had available to lend, how much control I had over the account, whatever Kiva altered on its own side, and the arrival of the pandemic in 2020. Several measures also rest on Kiva's labels rather than mine, since borrower gender and activity type are the platform's classifications while the WEE grouping is my own. The repayment analysis is thin for a different reason: almost all of the recorded loss was small currency movement, and only 8 loans actually defaulted. And the file records choices, not reasons. Every motivational reading in this paper is inference laid over a transaction log.

## FUTURE IMPLICATIONS

The obvious next step is more lenders. A cohort of young Kiva users followed across the same span would show whether fixed contribution amounts and stabilizing preferences arrive at similar ages in other people, or whether age eleven is a feature of this record alone.

Interviews would answer a different question. The log cannot tell me whether "I always lend $25" was ever a decision or whether it only looks like one in hindsight. Asking lenders while they are still lending would. Platform experiments would go further still: vary the minimum, vary the default, and see whether the rule follows. If design can set a young user's giving habits for a decade, that matters well beyond this dataset.

## CONCLUSION

Two of the pre-specified hypotheses were not supported, and they failed in a direction worth reporting. Sector diversity was predicted to narrow as rules took hold. It widened instead, with Shannon entropy rising from 4.83 to 5.10 bits and 45 different activity types funded in 2023 alone. The share of female borrowers was predicted to climb. It did not move (76.0% to 78.1%, $p = 0.87$). What changed there was consistency rather than level, as year-to-year variation fell from 15.0 to 6.6 percentage points.

Set against those two failures, the prediction that did hold is sharper than it would have been on its own. Across 644 loans made from ages 7 to 17, contributions at exactly $25 rose from 8.0% of loans to 74.6% ($p < 0.001$), and lending shifted toward sectors connected with women's economic activity, from 28.0% to 49.7% ($p = 0.005$). The amount became a rule. The choice of borrower did not.

A single-subject record cannot show that the same process runs in anyone else, and it cannot establish why any of it happened. What it can show is that rule formation was uneven within a single repeated decision: one dimension hardened into a default while the others stayed open.

**Works Cited**

Andreoni, J. Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving. Economic Journal, 1990, 100(401), 464–477.

Fehr, E.; Bernhard, H.; Rockenbach, B. Egalitarianism in Young Children. Nature, 2008, 454(7208), 1079–1083.

Galak, J.; Small, D.; Stephen, A. T. Micro-Finance Decision Making: A Field Study of Prosocial Lending. Journal of Marketing Research, 2011, 48(SPL), S130–S137.

Gigerenzer, G.; Goldstein, D. G. Reasoning the Fast and Frugal Way: Models of Bounded Rationality. Psychological Review, 1996, 103(4), 650–669.

Kahneman, D.; Tversky, A. Judgment under Uncertainty: Heuristics and Biases. Science, 1974, 185(4157), 1124–1131.

McAuliffe, K.; Blake, P. R.; Steinbeis, N.; Warneken, F. The Developmental Foundations of Human Fairness. Nature Human Behaviour, 2017, 1(2), 0042.

Raghubir, P.; Srivastava, J. The Denomination Effect. Journal of Consumer Research, 2009, 36(4), 701–713.

Thaler, R. H. Mental Accounting Matters. Journal of Behavioral Decision Making, 1999, 12(3), 183–206.

Thaler, R. H.; Sunstein, C. R. Nudge: Improving Decisions About Health, Wealth, and Happiness. Yale University Press, 2008.

**Examining the Role of Socio-Economic Factors in Women Entrepreneurs' Access to Formal Business Credit By Anika Vivek**


## Abstract

This paper, "The Role of Socio-Economic Factors in Women Entrepreneurs' Access to Formal Business Credit" examines how women entrepreneurs navigate formal lending institutions to fund and grow their businesses. Specifically, it presents a comparative case study of six women entrepreneurs in India, grouped by the size of the formal business loan they obtained: below ₹10 lakhs, between ₹10–50 lakhs, and ₹50 lakhs and above, and an aspect from the ex-State Bank of India's chief credit officer. By comparing women across these tiers, this case study can uncover how class, prior financial exposure, and access to networks or guarantors differently shape the credit-seeking experience, and determine whether the barriers women face at the bottom of the credit ladder are the same barriers faced by those higher up, or whether they are qualitatively different. This study adopts a qualitative, primary research design structured around case study narrative methodology, supplemented by a secondary literature review. Semi-formal interviews were conducted with each participant using a series of questions developed from an in-depth review. The interview questions were designed to capture each participant's credit-access journey from beginning to end, allowing the ability to understand not only the outcome of the borrowing process but also the experiences and challenges encountered at different stages. The interview with the ex-SBI banker provided an institutional expert perspective and personal opinion on women entrepreneurs' access to credit across different loan brackets. In conclusion, this study identifies that socio-economic factors do change a woman entrepreneur's experience in accessing formal business credit. However, the extent to which it does so differs between different loan sizes and backgrounds of women.


## 1. Introduction

Access to formal business credit remains one of the most persistent barriers to women's economic participation in India, despite a decade of policy initiatives aimed at closing the gender credit gap. This study examines how socio-economic circumstances shape the experiences of women entrepreneurs as they navigate formal lending institutions to fund and grow their businesses. Specifically, it presents a comparative case study of six women entrepreneurs in India, grouped by the size of the formal business loan they obtained: below ₹10 lakhs, between ₹10–50 lakhs, and ₹50 lakhs and above, and an account from the ex-State Bank of India's chief credit officer. The three loan-size tiers were constructed as an analytical framework to examine whether increasing credit requirements are associated with differences in collateral, documentation, and institutional scrutiny. By comparing women across these tiers rather than treating 'women entrepreneurs' as a single homogeneous group, this case study can discover how class, prior financial exposure, and access to networks or guarantors differently shape the credit-seeking experience, and determine whether the barriers women face at the bottom of the

credit ladder are the same barriers faced by those higher up, or whether they are qualitatively different.

This topic is important to investigate because access to credit is not merely a financial constraint but a structural gatekeeper to entrepreneurship itself. In India, women-led enterprises are frequently characterised as under-financed, and formal lenders have historically shown documented hesitancy in extending credit to women applicants, often citing perceived risk, lack of collateral in the applicant's own name, or informal business histories that do not fit standard underwriting criteria. This hesitancy is compounded by a widely observed pattern in which women entrepreneurs themselves report lower confidence in approaching banks, negotiating loan terms, or managing debt once acquired. A case study design is well suited to this problem because institutional bias and self-confidence are contextual experiences that are difficult to capture through large-scale survey data alone. In-depth interviews allow these dynamics to surface in the entrepreneurs' own words, and a comparative structure across income groups allows patterns to be identified without collapsing very different lived experiences into a single average.

Prior research on women's access to finance in India has tended to focus on microfinance and self-help groups at the lowest end of the credit spectrum, or on aggregate statistics documenting the gender gap in loan approval rates, interest rates, and collateral requirements. These works have been valuable in establishing that disparities exist, but have said comparatively little about how these disparities are experienced differently as women move into higher loan brackets and engage with mainstream commercial banking rather than microfinance channels. Before undertaking this research, my own assumption was that the central barrier for women entrepreneurs would be a lack of awareness, that women were either unaware that formal credit products existed or unaware of which schemes and loan categories they were eligible for. This study uses loan size as a lens onto socio-economic differences, not just business scale. Comparing entrepreneurs across three loan brackets tests that assumption: does a loan officer's doubt feel the same at ₹8 lakhs as at ₹75 lakhs? Does confidence grow with wealth, or come from family support and prior exposure to banking instead? This comparative design shows where socio-economic circumstance changes the experience, and where gender alone still drives the same outcome.

At the aggregate level, the evidence for a gender credit gap in India is strong and consistent. Using Census data on over a million manufacturing firms, Chaudhuri, Sasidharan, and Raj found that male-owned firms had 10–12% higher odds of loan approval than female-owned firms even after controlling for firm characteristics, and that this gap was driven by genuine discrimination rather than differences in firm quality (Chaudhuri et al.) Chiplunkar and Goldberg arrived at a related conclusion through a different route, showing that women face no excess cost entering informal entrepreneurship but pay roughly double the cost of formalising a business, the very step that formal lenders require before extending credit (Chiplunkar et al.). Earlier, Menon and Rodgers used decades of NSSO survey data to show that formal loans have a far stronger effect on women's self-employment than men's, precisely because women have so

few alternative routes into paid work (Menon and Van Der Meulen Rodgers). Together these studies establish that the gap is real and large, but because they rely on survey data, they say little about how individual women experience the lending process itself, or whether the size of the loan being sought changes that experience.

A woman's main obstacle is a lack of awareness of loan products and eligibility. Kumar and Singh's Delhi-based study found that half of women entrepreneurs who received no state support cited lack of policy awareness as the reason, alongside bureaucratic hurdles and insufficient collateral (Kumar and Singh). A survey-based study of rural women in Punjab reached a related conclusion; it grouped barriers to financial inclusion into psychological, economic, physical, socio-cultural, and technical factors, and found that a psychological factor, low self-confidence and reluctance to approach unfamiliar institutions, were the biggest barriers. Overall, the study found that women's lack of confidence and hesitation to approach financial institutions were bigger barriers to accessing loans than simply not knowing how to apply for them. (Sharma et al.) This suggests that although a lack of awareness may contribute to difficulties in accessing finance, it may be a consequence of deeper confidence-related barriers rather than the underlying cause. However, on the other hand, a Nagaland-based study of microcredit access found that women were, if anything, more aware of self-help group resources than men, yet still significantly less likely to obtain credit, indicating that awareness alone cannot explain the gender gap (Rehman).

**2. Research Methods**

This study adopts a qualitative, primary research design structured around case study narrative methodology, supplemented by a secondary literature review. The central question here focuses on women's personal experiences, the challenges they face when making financial decisions, and their interactions with formal lending institutions. These dimensions are best understood through the perspectives and experiences of the women themselves, rather than through numerical measures alone. A qualitative approach therefore allows the research to explore how factors such as income, education, social networks, and financial knowledge influence women's experiences of seeking credit. It also provides insight into the challenges, perceptions, and strategies that may remain uncovered within purely quantitative data. The qualitative data allows personal insight into the credit process, and interviewees can express additional information a quantitative study may not capture. Thus, the strength of this study lies in capturing the nuances and granular perspectives that may otherwise be missed.

A statistic showing that women receive smaller loans on average does not explain why and for what reasons, whether it is lack of assets, relational access to bank officials, confidence in self-presentation, documentation readiness, or some combination of these. It was this gap between the quantitative "what" and the qualitative "why" that made a case study narrative design the more appropriate methodological choice for this research question. Case study methodology allows each participant's credit access journey to be treated as a distinct independent journey, capturing its own sequence of events, decision points, and relationships,

while still permitting cross-case comparison to identify patterns that recur across different socio-economic situations.

Semi-formal interviews were conducted with each of the participants using a series of questions developed from an in-depth review of existing literature on women's entrepreneurship, microfinance, and gendered access to credit in India. The questionnaire was designed to be semi-structured and open-ended to account for the flexibility required in capturing the nuances of individual experiences in non-homogeneous settings. The interview questions were designed to capture each participant's credit-access journey from beginning to end, allowing the ability to understand not only the outcome of the borrowing process but also the experiences and challenges encountered at different stages. The questions covered participants' personal and business backgrounds, motivations for seeking credit, awareness of available loan options, interactions with financial institutions, application and documentation requirements, approval or rejection experiences, and the process of managing and repaying the loan.

Participants were also encouraged to reflect on their emotional experiences throughout the process and identify factors that could have made accessing credit easier. Although the same core questions were used across all interviews to maintain consistency, follow-up questions and probes were introduced based on participants' responses. Each interview lasted approximately 30–45 minutes and was conducted either in person or on the phone. This semi-formal structure provided sufficient consistency for comparing experiences across participants while allowing individuals to describe their experiences, perceptions, and challenges in their own words.
The sample was constructed using a snowball sampling technique. An initial set of contacts, identified through personal and professional networks, were interviewed and subsequently asked to refer to other women entrepreneurs who had experience seeking formal business credit. This word-of-mouth approach was selected because women entrepreneurs, particularly those with limited business visibility, may be difficult to identify through formal sources.

Seven in-depth interviews were conducted, six with women entrepreneurs across a range of business types and credit histories, and one with an ex-senior banking professional (a retired Chief Credit Officer at State Bank of India) who provided an institutional expert perspective against which participant accounts could be evaluated and judged. To enable structured cross-case comparison and understand an in-depth perspective of different experiences of women, the six women entrepreneurs were divided into three tiers based on the size of the largest formal loan they had sought. This naturally gave insight into the size of the business and the capital invested in the business.

All participant identities have been anonymised in this paper. Entrepreneurs are referred to as Low-Tier Participant, Mid-Tier Participant, or High-Tier Participant, rather than by name or business, and personal information is omitted where it was not analytically necessary. The three tiers are broken up, and participant descriptions have been given below.

- **Low-tier (₹10 lakhs and below):** two participants. One runs a physiotherapy clinic, and another runs a home-based food production business. Both accessed collateral-free,

government-backed schemes (Mudra/MSME) as first-time borrowers with limited prior credit history.

- **Mid-tier (₹10–50 lakhs):** two participants, one running a salon franchise and the other a clothing retail business, respectively. Both had progressed through multiple rounds of borrowing
- **High-tier (₹50 lakhs and above):** two participants, one running a services company and the other a manufacturing business, respectively. Both entered the credit process with pre-existing property, established professional networks, or prior business and credit history.

The similarities between participants in the same tier were not pre-decided but analysed after the interviews. Each participant was independent of the others and did not influence or have any involvement in the decision-making processes of other entrepreneurs. This tiering was intentionally adopted as an analytical framework to examine the role of capital position in shaping participants' experiences. It allowed the analysis to assess whether the barriers, relationships, and concerns reported by participants varied systematically as the capital obtained increased, independent of the specific industry each woman was operating in.

Each transcript was first read in full to understand the participant's narrative as a whole and the shape of her journey from the beginning of approaching the bank to the final loan collection. Before being reread with the research question held explicitly in mind, it is important to consider which aspects may be influenced by the participant's socio-economic circumstances. The analysis involved several readings of the interviews. The first reading focused on identifying repeated words, ideas, and experiences. The second reading followed the chronological order of each interview to identify at which stages of the loan process barriers were most common. The final reading compared participants across the three capital tiers to identify which experiences were specific to certain tiers and which were common across all participants.

Then themes were evaluated against the research question directly: does this pattern help explain how socio-economic circumstances shape credit access experience? Other interesting patterns that were not directly related to the research question, such as differences in personal motivation to become an entrepreneur, were noted but given less importance in the final analysis. The themes were then compared with the ideas of the ex-senior banker, as she had an alternative perspective, one from the inside, that could draw parallels to the interviewee responses

**3. Key Findings: Role of Socio-economic Circumstances in Credit Access**

Across nearly every tier, participants secured their loan through a person they already knew rather than through a cold application process. Low-Tier Participant 1 obtained her Mudra loan specifically because of an existing relationship with a bank manager, stating plainly, "If you know the manager and are friendly with them, she can help, or he can help. Otherwise, it's difficult. They won't give the loan." Mid-Tier Participant 1 approached the same manager who had handled her earlier preschool business account, noting, "My school account was there

earlier, so I somehow know that manager... then also I approached him." High-Tier Participant 2 similarly reported, "I went with the banker I'm most familiar with," and High-Tier Participant 1 relied on the bank that already managed her family's personal finances, explaining she went to them because she "knew they'd already be open to knowing our background."

Low-Tier Participant 2 is an important exception. She lacked any prior banking relationship; she went to a third-party non-profit organisation dedicated to empowering disadvantaged Indians by transforming business ideas into sustainable enterprises. The organisation helped in arranging introductions, giving her advice, and connecting her directly to a bank already familiar with their referrals, suggesting the organisation functioned to manufacture the trust that a personal relationship would otherwise have supplied. This suggests that personal and professional connections can play an important role in facilitating access to credit, regardless of an individual's income. Women who do not have access to such informal networks may consequently rely more on formal institutions to provide the guidance, support, and connections needed to navigate the credit process.

Separately from relational access, several participants identified the unfamiliarity of their business category to the lender as an independent source of problems. Mid-Tier Participant 1 recounted that "no bank is ready to offer a loan" for a salon franchise at the time she sought hers, because lenders "don't know what will be the turnover." High-Tier Participant 1, describing her company's founding in an entirely new industry category, reflected that "the bank at that time would have found it too new... we were pioneers in that industry" and was "not sure they would have even given us a loan... if we had asked for it." This suggests that the relative newness of a business can create a separate barrier that is not directly related to gender. This is also supported by the ex-banker, as she mentions, "Generally, you have a sense of what will work. Somewhere there is no water, obviously, you can't be lending for agriculture." This suggests that bankers have a fine distinction of what businesses are viable in that area, and prefer to lend to what will work. However, it also reflects internal biases that may be based on outdated notions or systems and that resultantly may not allow new businesses to test the market.

Low-Tier Participant 2 talked about a huge challenge which she had to face early on in the credit process- she described being repeatedly questioned by the bankers handling her application about why her academic degree (zoology) did not match her food-business venture. She was questioned about why she wanted to enter the entrepreneurial journey and felt discouraged at first. To her, it seemed like she was being mistreated, as it was her first time obtaining a loan. It felt like the bank assumed she did not know what she was doing and further interpreted this as gendered, "especially like when you're a married woman, when you're in your 30s or 40s, then they think that you're here to get your hobbies started." Through this statement, there was a gender disparity in obtaining a loan, which was made evident to the participant through the rigour of questioning.

The ex-banker explicitly distinguishes lending experiences by class, describing women from a "middle class and upper" background as arriving "totally brilliant... capable... independent... they have a clear vision... they are convincing," implicitly contrasting this with

women from lower socioeconomic strata, who she says are typically routed through NGO or self-help group-mediated channels rather than approaching a bank directly, thus hinting towards the fact that these women generally fear approaching the bank individually and feel more comfortable when they have a group or an organization to support them throughout the credit process. However, many of the participants mentioned they had a passion for being an entrepreneur and truly believed in their business, but still got rejected from the bank multiple times, describing the process as tedious. Suggesting that they were not confident in explaining their vision and convincing their bank; thus, business passion did not necessarily translate into the confidence or ability to communicate their vision convincingly to lenders. This emphasis on being able to communicate one's business confidently is also reflected in the advice of a low Tier 2 participant, who advises first-time women entrepreneurs seeking credit that "you have to be sure of the product or service you are going to give. You should not second-guess or have doubts if it will work. So if you go to the bank and they ask questions, you should be able to tackle it smartly," hence suggesting that the display of confidence is essential to convince the lender of your abilities. This directly corroborates the ex-banker's response.

Anxiety about loan repayment was universal across the sample, but the resources available to manage that anxiety were not. High-Tier Participant 1 described a family member who was prepared to "chip in" if needed, and who ultimately "just helped pay back the final piece of the loan." Mid-Tier Participant 2's father repeatedly closed out her existing loans so she could qualify for the next tier of borrowing. By contrast, Mid-Tier Participant 1 described managing slow months by "rotating" payments and delaying royalty obligations to her franchisor, while Low-Tier Participant 2 described a period of trial-and-error production losses with no equivalent family support, although this was her choice to avoid familial resources. Mid-Tier Participant 2's account of pursuing a larger loan illustrates the resistance built into this system from the lender's side: "you'll hear like 100 no's before you can get somebody to say yes you are eligible." Ex-banker's own account explains this resistance structurally rather than as arbitrary gatekeeping. Mentioning that individual loan officers are personally penalised when a loan they approve turns bad: "the banker is put to hardships... the bank penalizes the official who has lent," which she says causes officers to become risk-averse over time: "he gets hardened and does not lend." Read together, the "100 no's" Mid-Tier Participant 2 experienced is not simply institutional indifference but the accumulated caution of individual officers each independently protecting themselves from personal consequences for a bad lending decision. A dynamic that disproportionately disadvantages applicants without the collateral or credentials to make a loan officer's decision feel personally "safe."

Another connection was drawn when the ex-banker was asked about repayment differences by gender, stating women are "a tad more sincere in repaying, and a tad more committed... as they are committed to a family, they are also committed to this." Low-Tier Participant 2's own stated reasoning for wanting a formal loan rather than informal help independently echoes this framing. She described not wanting to "rely on anybody else for financial dependence" and not wanting to "make things ugly" with people close to her, explicitly

formalizing the arrangement to protect those relationships. The banker's generalization about women's family-rooted commitment and Low-Tier Participant 2's individual account of family-protective motivation for formalizing her debt speak to the same underlying pattern, which strengthens the claim. Women in rural India have very limited access to regular paid jobs. As a result, once a loan becomes available, starting a small business becomes the most practical and logical way for women to earn an income. Whereas men, who already have more job options, are less influenced by the loan in the same way. Suggesting that a woman's outlook on a loan itself is different, and that once they obtain a loan, it is upon them to repay it without incurring any debt (Menon and Van Der Meulen Rodgers)

**4. Discussion: Implications for Women Entrepreneurs' Financial Inclusion**

Low-Tier Participant 2's two-month delay in securing her loan was spent almost entirely building the legal identity a bank could recognise (GST registration, food safety certification, firm registration) rather than waiting on the bank itself. Chiplunkar & Goldberg's structural analysis of India's Economic Census data explains why this stage carries such weight. Women face "substantially higher costs of formalizing their businesses," twice men's cost in 1998 and still 25% higher by 2005, and the paper frames formalization explicitly as "the gateway to institutional finance," since "informal firms are effectively locked out of formal credit." Read against Low-Tier Participant's account, this suggests that her documentation delay was not simply an isolated difficulty, but may reflect a broader structural challenge faced by women formalizing a first-generation business. (Chiplunkar et al.)

Mid-Tier Participant 2's comment that her Mudra loan of roughly five lakhs "doesn't do much" and Mid-Tier Participant 1's cap at ₹20 lakhs against a stated need of ₹40 lakhs. Both point to loan size as a meaningful constraint. The SHG/Community Investment Fund study makes the same argument at the household level. Typical self-help-group loans are "too small to meaningfully shift household dynamics", whereas larger Community Investment Fund grants "substantially increase women's decision-making power". This suggests that the adequacy of the loan amount should be considered separately from simply whether or not credit was obtained. (Kochar et al.)

Awareness of available credit options and institutional procedure emerged as a recurring barrier among the study's lower-tier participants, consistent with broader findings in the literature on financial inclusion among Indian women. The GNAU rural financial inclusion study, surveying 600 rural women, found that despite 91% holding a bank account, 77% demonstrated only a low level of financial inclusion, meaning they rarely engaged with services beyond basic banking. The study identified psychological factors, including low self-esteem, limited confidence in financial decision-making, and hesitation to approach "individuals outside their immediate social network, such as bank officials or financial advisors", as the single largest contributor to this gap (Sharma et al.). This finding is directly reflected in Low-Tier Participant's account of learning about her Mudra loan only because a bank manager who already knew her personally volunteered the information, rather than through any independent search or awareness

of the scheme on her part. When asked if she was aware of any other banking scheme, she responded saying no, and was only aware of her options due to the bank manager Low-Tier Participant 2's account extends this pattern into the documentation and compliance stage of the credit process specifically. She described having no prior awareness of the Know Your Customer (KYC) requirements or the broader formal documentation a bank would expect until an external support organisation intervened to guide her through each requirement: "they had this bunch of books with them. They'll tell you to get everything before the bank asks you. They'll tell you from A to Z what are the documents needed". Without this intervention, she indicated she would not have known what to prepare or in what order, underscoring that her barrier was not simply awareness of the loan scheme itself but awareness of the procedural requirements.

This procedural dimension of the awareness gap is corroborated at scale by the Delhi COVID-era study, which solely looks at the awareness of schemes as a distinct measured category; 50% of respondents cited "lack of policy awareness" as a barrier to accessing state support during the pandemic, alongside 21.4% citing bureaucratic hurdles more broadly (Kumar and Singh)

In relation to confidence and awareness barriers, the ex-banker explains that women from lower socioeconomic backgrounds often face "confidence issues, awareness issues" as well as social and cultural restrictions that may discourage them from leaving the home or approaching formal institutions ("the milieu that they are in, and the kind of forbidden charm that you cannot step over the house and all those rules and regulations"). Women are around 20% less likely to apply for formal financing due to concerns about high interest rates, complex application procedures, extensive documentation requirements, and the pressure of loan repayment (Peter et al.). However, she also notes that there has been "a bit of a change from the erstwhile fear of coming to the branch," with more women now approaching banks due to increased education and greater exposure to media. This suggests that although confidence and awareness continue to act as barriers to accessing formal credit, increased education and exposure to information through media and social platforms may be gradually reducing the fear and hesitation women experience when approaching banks.

Even in other countries, the World Bank surveys in Latin America and the Caribbean reveal that household obligations and domestic work reduce women's participation in formal business training programs (Valque et al.). Thus, women in general have higher household obligations, limiting their access to gain awareness of credit facilities and general bank processes. Hence, it highly depends on the household norms of the women and their freedom to gain an education.

The ex-banker consistently reiterates that "sex does not matter" and women are treated as any other candidate. In addition, she mentions that females are better at repayment than men and can handle a business more productively. The banker also suggests that the willingness to lend to female enterprises is marginally higher, as they feel responsible for handling the loan appropriately. However, this contradicts a paper (Chaudhuri et al.) that suggests women-owned

firms significantly underperform male-owned firms in output, employment, labor productivity, and total factor productivity, and that male-owned firms have 10–12% higher odds of obtaining a loan than women-owned firms.

As the ex-SBI banker states, the only way sex matters is when a woman is getting a concession in the interest rates or an extended repayment duration. She further emphasizes that a lower interest rate is offered to women to encourage financial independence and upliftment of society. Hence, the overall motive of the government bank is directed towards the encouragement of women and giving them an opportunity to become financially sufficient. The public bank tends not to approach the candidate with bias. The underlying root of women not availing bank loans or the opportunities given to them is their awareness of the banking system and loans.

The main obstacle to the effectiveness of government schemes is not generating awareness but facilitating the transition from awareness to actual utilization. The loan process is slowed down by excessive bureaucracy, poor outreach to local communities, and a lack of support after loans are issued. People may have awareness of the government schemes but may not want to use them. (Arora) As the ex-banker also mentions, governments must go further to conduct workshops to ignite entrepreneurship in people, and "spoon-feed them" with the information

**4.1 Scope and Limitations of the Study**

Several participants across all three tiers, including those with the highest loan tier, described the loan approval stage as more stressful and time-consuming than the loan process itself. Once the loan was approved, disbursement and repayment were generally described as straightforward. This suggests that the main difficulties in accessing credit occur at the initial stage when women are searching for and convincing lenders to provide credit, rather than throughout the entire borrowing process.

Second, gender was not identified as a factor by five of the six entrepreneur participants. Only the participant at the low-tier requirement and least professional qualifications described experiencing gender-based scrutiny directly and repeatedly. When considered alongside the ex-banker's comments about differences across socio-economic groups, this suggests that gender bias in accessing credit may not affect all women in the same way. Instead, it may become more noticeable among women from lower socio-economic backgrounds (less credit history, less awareness, and education). This suggests that socio-economic circumstances may influence how strongly gender becomes a barrier to accessing credit. It is also difficult to directly classify an action as gendered bias, as women may feel a decision is made based on their gender; however, it could be due to another issue as well. Therefore, it is difficult to directly understand the true nature of a decision

This study has a few limitations. First, the relatively small sample size may not be able to determine a generalisable conclusion, as the sample is not large enough to represent a large population. Next, as all the samples were collected in one specific city, the results may not be

applicable to other cities in India, therefore limiting the reliability of the results. Lastly, the data is based on participants' recollection of their credit process. This can be influenced by perception and false remembering or misinterpretation of their experiences.

**5. Conclusion**

This study used the approach of three loan-size tiers as an analytical framework to examine whether increasing credit requirements changed the experiences faced by women in different loan brackets. By comparing women across these tiers, this case study was able to uncover how class, prior financial exposure, and access to networks differently shape the credit-seeking experience, and to determine whether the barriers women face at the bottom of the credit ladder are the same barriers faced by those higher up, or whether they are qualitatively different. The study grouped the size of the formal business loan they obtained into three categories: below ₹10 lakhs, between ₹10–50 lakhs, and ₹50 lakhs and above.

Firstly, wealth and assets play a major role in the credit process, as bankers are more willing to lend to individuals with assets, as it provides an extra sense of security. The fear of whether the loans will be paid back is gone, as the bankers are assured they will receive their investments back regardless. Secondly, social networks, such as knowing or being related to a bank employee, drastically change the credit application, as an individual is then able to seek extra guidance, is trusted, or can get easier access to the system, which is otherwise difficult to navigate. Therefore, social class and the related networks play an important role in access to credit. Although education may not appear to be a major factor, it can act as a catalyst in the loan process, as it reflects the applicant's understanding of financial obligations and their ability to manage and repay the loan effectively. Confidence and knowledge about the loan certainly depict a sense of dominance as the participant approaches with a clear focus that remains unwavered even when questioned by the bank.

However, gender as such may not play as much of a factor, as banks have made it clear they approach each client without any bias. Only one interviewee out of the six had explicitly stated gendered bias and was treated unfairly, as well as the ex-banker repeatedly stating women are not treated differently to men but may even be supported better through lower interest rates. Yet existing data states that male-owned firms have 10–12% higher odds of obtaining a loan than women-owned firms. Therefore contradicting the results of the study. It may be difficult to determine whether a result is gender-based because the underlying gender differences in a patriarchal society like ours may not be readily apparent. It goes on to show how deeply rooted the internalised gender based biases are.

In conclusion, this study identifies that socio-economic factors do change a woman entrepreneur's experience in accessing formal business credit. However, the extent to which it does so will differ between different loan sizes and backgrounds of women.

**Works Cited**


Arora, Ishaan. "Entrepreneurship in India – a Discourse to Evaluating the Government's Mandates on Spurring Entrepreneurship." Journal of Asia Entrepreneurship and Sustainability, vol. 22, no. 4S, July 2026, pp. 132–37. https://doi.org/10.66635/tzqraw80.

Baby, Ann, et al. "The Knowledge Ecosystem of Women Entrepreneurs in the Era of Digital Finance Mapping: Global Scholarly Evidence." Rajagiri Management Journal, vol. 20, no. 2, June 2026, pp. 177–94. https://doi.org/10.1108/ramj-02-2026-0029.

Chaudhuri, Kausik, et al. "Gender, Small Firm Ownership, and Credit Access: Some Insights From India." Small Business Economics, vol. 54, no. 4, Dec. 2018, pp. 1165–81. https://doi.org/10.1007/s11187-018-0124-3.

Chiplunkar, Gaurav, et al. AGGREGATE IMPLICATIONS OF BARRIERS TO FEMALE ENTREPRENEURSHIP. Working Paper 28486, National Bureau of Economic Research, 2021, www.nber.org/system/files/working_papers/w28486/w28486.pdf.

Deshpande, Ashwini, and Smriti Sharma. "Entrepreneurship or Survival? Caste and Gender of Small Business in India." Working Paper No. 228, by Centre for Development Economics and Department of Economics, Delhi School of Economics, 2013.

Guerin, Isabelle, et al. "The Gender of Debt and Credit: Insights From Rural Tamil Nadu." 2020. IZA Discussion Paper Series, by IZA Institute of Labor Economics, report, 13891, IZA Institute of Labor Economics, Nov. 2020, docs.iza.org/dp13891.pdf.

Kochar, Anjini, et al. "Financial Access and Women's Role in Household Decisions: Empirical Evidence From India's National Rural Livelihoods Project." Journal of Development Economics, vol. 155, Jan. 2022, p. 102821. https://doi.org/10.1016/j.jdeveco.2022.102821.

Kumar, Sanjeev, and Neha Singh. "Entrepreneurial Prospects and Challenges for Women Amidst COVID-19: A Case Study of Delhi, India." Fulbright Review of Economics and Policy, vol. 1, no. 2, Nov. 2021, pp. 205–26. https://doi.org/10.1108/frep-09-2021-0057.

Menon, Nidhiya, and Yana Van Der Meulen Rodgers. "How Access to Credit Affects Self-employment: Differences by Gender During India's Rural Banking Reform." The Journal of Development Studies, vol. 47, no. 1, Jan. 2011, pp. 48–69. https://doi.org/10.1080/00220381003706486.

Peter, Serin, et al. "Unveiling the Nexus: Financial Inclusion, Financial Literacy, and Financial Performance as Catalyst for Women-owned Enterprises in India." Journal of the International Council for Small Business, vol. 6, no. 4, Jan. 2025, pp. 721–51. https://doi.org/10.1080/26437015.2024.2432385.

Rehman, A. "Does Gender Define Access to Microcredit? Evidence From India." Review of Business and Economics Studies, vol. 12, no. 1, June 2024, pp. 37–50. https://doi.org/10.26794/2308-944x-2024-12-1-37-50.

Research Portal. scholarship.libraries.rutgers.edu/esploro/outputs/bookChapter/Womens-labor-market-status-and-economic/991031558847804646.

Sharma, Manu, et al. “OBSTACLES FACED BY RURAL WOMEN IN GAINING COMPLETE FINANCIAL INCLUSION.” Journal of Management & Technology, vol. XIV–XIV, no. I, June 2023, www.gnauniversity.edu.in/assets/pdf/journal/7.pdf.

Valque, Rosa Ysabel Bazán, et al. “Influence of Socioeconomic Factors on Female Entrepreneurship: An Analysis Using Structural Equation Modeling (PLS-SEM).” Frontiers in Sociology, vol. 10, Oct. 2025, p. 1684697. https://doi.org/10.3389/fsoc.2025.1684697.

# Understanding Retirement Planning and Savings Behavior of Working People in Pakistan

By Amaan Siddiqui

## Abstract

Due to the increasing cost of living and the need for financial security after retirement, retirement planning has become increasingly important. Financial knowledge, saving behavior, financial planning, and future financial expectations may influence retirement readiness in Pakistan. This study aimed to examine retirement planning and saving behavior among working people in urban areas of Pakistan and explore the relationships between saving behavior, future expectations, financial planning, financial knowledge, and retirement preparedness. A quantitative method was adopted, with data collected from 36 respondents through an online survey using convenience sampling. The survey assessed demographics, saving behavior, future expectations, financial planning, financial knowledge, and retirement preparedness. Data were analyzed using descriptive statistics, Pearson correlation analysis, and the Mann–Whitney U test. Retirement planning had the lowest mean score, while future expectations had the highest. Saving behavior showed a weak relationship with retirement planning. Future expectations, financial planning, and financial knowledge were significantly and positively related to retirement planning. Overall, the findings suggest that retirement preparedness involves more than saving alone. Improving financial knowledge, active financial planning, and clear future financial goals may strengthen retirement preparedness among working people in Pakistan.

## Keywords

Retirement planning, saving behavior, Pakistan, working people, and financial planning

## Introduction

Given increasing life expectancy and the high cost of living and healthcare, planning for retirement has become vital. Financial resources are critical after retirement to sustain one's standard of living; therefore, saving and financial planning are important during the working years. Retirement planning may involve saving money, making investment decisions, budgeting for retirement expenses, and planning for long-term financial objectives. Research indicates that those who take an active role in personal financial planning exhibit better retirement planning behavior (Mohidin, 2013).

But a variety of financial and behavioral factors affect retirement savings. Financial awareness, spending patterns, self-control, income, family obligations, and future planning can all influence people's financial readiness for later life. Financial literacy is especially crucial, as people must be able to manage their income, savings, investments, and other financial resources effectively. In addition, self-control and financial knowledge have been shown to influence savings among younger adults, suggesting that being financially savvy does not necessarily translate into saving if one lacks developed financial habits (Hartono & Isbanah, 2022).

These factors are particularly relevant in Pakistan, where households' saving behavior is closely linked to their income and demographic characteristics. Based on the HIES data for 2018-19, Khalid (2023) found that household income is positively associated with saving, and that a higher dependency ratio adversely affects households' capacity to save. The study also found that saving rates differ between rural and urban areas and that income does not influence saving. Likewise, Ahmad (2015) found that in the case of Pakistan, GDP per capita, financial development, inflation, dependency ratio, and fiscal variables are significant contributors to private savings during 1972-2012. The results indicate that both personal financial choices and economic conditions affect retirement preparedness in Pakistan.

Other factors, such as financial literacy and behavior, also impact whether an individual uses available income to help prepare for the future. Liu et al. (2021) concluded that financial literacy and behavioral factors, such as financial self-control, optimism, herding, and loss aversion, affect financial inclusion and financial decision-making in Pakistan. Further, a study conducted among female government officials in Pakistan revealed that financial risk-taking, future time perspective, and retirement goal clarity positively influenced retirement planning, with financial literacy in between (Hussain et al., 2024). This implies that retirement readiness is not just about financial resources and preparedness, but also about financial awareness, retirement planning, and setting clear goals.

Nevertheless, the importance of retirement planning in Pakistan has not been explored as extensively as in other countries, especially among younger adults and the working population. Existing research has mainly focused on household saving behavior, financial literacy, financial inclusion, or specific groups, but has not directly investigated how people prepare for and save for retirement. This also requires a better understanding of the relationship among saving behavior, financial knowledge, future expectations, and retirement preparedness. Thus, this study examines saving behavior, financial planning, financial knowledge, and retirement preparedness among working people in urban areas of Pakistan. It also explores whether sociodemographic characteristics such as age, gender, income, educational attainment, and occupational status influence retirement planning and financial preparedness.

## Methodology

### Tools used

A quantitative survey was carried out in the present study. The survey consists of six sections. Section 1 collected demographic information like age, gender, occupation, education level, annual income, place of residence, family employment background, and marital status. These variables enabled comparisons across different demographic groups. Section 2 assessed saving behavior using six statements rated on a 5-point Likert Scale (1 = Strongly Disagree to 5 = Strongly Agree). The next section examined future financial expectations through nine Likert-scale statements related to financial security and retirement outcomes. Section 4 measured financial planning behavior using four Likert-scale statements focused on retirement planning

and goal setting. Section 5 assessed financial knowledge using three Likert-scale statements that evaluated participants' understanding of financial and retirement-planning concepts. Lastly, Section 6 measured perceived financial preparedness using four Likert-scale statements regarding retirement savings, financial readiness, and retirement planning.

**Recruitment strategy and sample characteristics**

This study used convenience sampling to collect data through an online survey from 39 participants. Participants represented various demographic backgrounds, including different age groups, genders, occupations, education levels, income levels, residential locations, family employment backgrounds, and marital statuses. Table 1 presents the frequencies and percentages of respondents' socio-demographic characteristics.

**Table 1:** Sociodemographic characteristics of respondents (n= 36)

| **Socio-demographic characteristics** | **n** | **%** |
|---|---|---|
| Age | | |
| *18–25* | 7 | 19.4 |
| *26–35* | 5 | 13.9 |
| *36–45* | 8 | 22.2 |
| *46–55* | 14 | 38.9 |
| *56+* | 2 | 5.6 |
| Gender | | |
| *Male* | 17 | 47.2 |
| *Female* | 19 | 52.8 |
| Education | | |
| *High School* | 3 | 8.3 |
| *Undergraduate* | 8 | 22.2 |
| *Graduate* | 8 | 22.2 |
| *Post graduate* | 16 | 44.4 |
| *Chartered Accountant* | 1 | 2.8 |
| Annual Income | | |
| *Below PKR 500,000* | 7 | 19.4 |
| *PKR 500,000 – 1,000,000* | 3 | 8.3 |
| *PKR 1,000,001 – 2,000,000* | 5 | 13.9 |
| *PKR 2,000,001 –5,000,000* | 7 | 19.4 |
| *Above PKR 5,000,000* | 14 | 38.9 |
| Type of Employment | | |
| *Service* | 6 | 16.7 |
| *Business* | 30 | 83.3 |
| Relationship Status | | |
| *Single* | 9 | 25 |

| | | |
|---|---|---|
| *Married*<br>*Divorced* | 25<br>2 | 69.4<br>5.6 |
| Children<br>*Yes*<br>*No* | <br>22<br>14 | <br>61.1<br>38.9 |

**Ethical considerations**

To protect participants' privacy, this study ensured anonymity throughout the data collection and analysis process. No personally identifiable information, such as phone numbers, email addresses, or IP addresses, was collected. Responses were gathered through Google Forms and recorded without linking them to individual identities. Before completing the survey, participants were informed of the study's purpose and the voluntary nature of participation. All data were used solely for academic research purposes, stored securely, and reported only in aggregate form to ensure that no individual participant could be identified.

**Data collection procedure**

The data for this study were collected using an online survey created through Google Forms. The survey link was distributed via email, WhatsApp, and social media to reach a diverse range of respondents. This method provided participants with a convenient and accessible way to complete the survey while ensuring anonymous data collection. All responses were automatically recorded and securely stored within Google Forms for subsequent analysis.

**Data analysis strategy**

Data from the Google Form were used to examine different statistical methods to understand retirement planning and savings behavior among 36 participants. The information collected was analyzed for basic demographic data, including age, gender, education, and occupation, using frequency counts and percentages. To clearly present the responses and their patterns, tables and graphs were used. The relationship between retirement planning and saving behavior was analyzed using Pearson correlation, and the strength and direction of the correlation between variables were determined. The Mann-Whitney U test was also used to compare responses across groups and determine whether statistically significant differences existed. In most cases, these methods helped identify patterns, relationships, and differences in the collected data on participants' retirement planning and saving behavior.

**Results and Discussion**

**Table 2:** Descriptive statistics of different sub-scales of the respondents

| **Variables** | **Mean (SD)** | **Minimum** | **Maximum** |
|---|---|---|---|

| | | | |
|---|---|---|---|
| **Saving Behavior** | 18.26 (6.042) | 10 | 30 |
| **Future expectation** | 21 (6.68) | 16 | 40 |
| **Financial planning** | 10.73 (3.52) | 4 | 20 |
| **Financial knowledge** | 10.36 (2.57) | 5 | 14 |
| **Retirement planning** | 8.56 (3.93) | 4 | 20 |

The descriptive statistics for the variables used in the study are shown in Table 2. The highest mean score with SD is for those who believe they have a high expectation of the future (21.00 ± 6.68), followed by saving behavior with a mean score of (18.26 ± 6.04). The mean score for financial planning is 10.73 (SD = 3.52), and for financial knowledge is 10.36 (SD = 2.57). The lowest mean score is for retirement planning (8.56, SD = 3.93). The minimum scores for saving behavior, future expectations, financial planning, financial knowledge, and retirement planning are 10, 16, 4, 5, and 4, respectively. The overall results indicate that the respondents scored highest on future expectations and saving behavior, and lowest on retirement planning.

**Table 3**: Pearson's correlation analysis with retirement planning and saving behavior, future expectation, financial planning, financial knowledge, and retirement planning (n= 36)

| **Variables** | **Retirement planning** | **p** |
|---|---|---|
| **Saving Behaviour** | 0.14 | 0.47 |
| **Future Expectation** | 0.43 | 0.02* |
| **Financial Planning** | 0.53 | 0.003* |
| **Financial knowledge** | 0.48 | 0.01* |

*Note: $p<0.05$**

Table 3 presents the correlation analysis between retirement planning and saving behavior, forward-looking plan, financial planning, and financial knowledge, based on Pearson's values. Saving behavior is poorly correlated with retirement planning ($r = 0.14$; $p = 0.47$) and is not statistically significant. In contrast, there is a moderate, positive, and significant association between future expectations and retirement planning ($r = 0.43$, $p = 0.02$). There is a moderate, positive, and significant relationship between financial planning and retirement planning ($r = 0.53$, $p = 0.003$), and between financial knowledge and retirement planning ($r = 0.48$, $p = 0.01$). Overall, the results indicate a significant association between higher future expectations, financial planning, knowledge of finance, and increased retirement planning, whereas non-financial savings do not show a significant association.

**Table 4:** Mann-Whitney U analysis of different variables by gender (n = 36)

| Variables | Groups | Rank Sum | U | p |
|---|---|---|---|---|
| **Saving Behaviour** | **Male** | 173.50 | 82.50 | 0.49 |
| | **Female** | 232.50 | | |
| **Future Expectation** | **Male** | 190.50 | 95.50 | 0.93 |
| | **Female** | 215.50 | | |
| **Financial Planning** | **Male** | 211.00 | 75.00 | 0.30 |
| | **Female** | 195.00 | | |
| **Financial Knowledge** | **Male** | 202.50 | 83.50 | 0.51 |
| | **Female** | 203.50 | | |
| **Retirement Planning** | **Male** | 180.50 | 89.50 | 0.71 |
| | **Female** | 225.50 | | |

Table 4 compares the savings habits, future outlook, financial planning, financial knowledge, and retirement planning in men and women. The results indicate no statistically significant differences between boys and girls across any of the variables, with all P values greater than 0.05. There was no significant difference in the saving behavior of males and females (U = 82.50, p = 0.49). Similarly, no significant differences were found for future expectation (U = 95.50, p = 0.93), financial planning (U = 75.00, p = 0.30), and financial knowledge (U = 83.50, p = 0.51). The results indicate that gender does not have a significant effect on saving behavior, future expectation, financial planning, and financial knowledge among the respondents.

## Discussion

The results of this research study will shed light on the retirement planning and saving behavior of Pakistan's working population. The descriptive statistics indicated that future expectations had the maximum average score among the variables, followed by saving behavior. The mean score for financial planning was compared with the mean scores for financial knowledge and retirement planning and was found to be of a similar level. The mean scores for financial knowledge and retirement planning were also found to be of similar levels. These findings indicate that when people consider their financial future and engage in saving behaviors, this does not necessarily imply that they are actively saving in a specific way to prepare for retirement. This is vital because saving in general and mentally planning for retirement aren't the

same as developing explicit financial plans and goals for retirement (Mohidin, 2013; Hartono & Isbanah, 2022).

A Pearson correlation analysis was also used to clarify factors related to retirement planning. There was a relatively weak positive correlation between saving behavior and retirement planning, but it was not statistically significant. This indicates that higher savings rates were not necessarily associated with improved retirement-planning outcomes among participants in this study. Future expectations regarding the statistically significant relationship. Those with higher expectations and thoughts about their financial future reported higher levels of retirement planning overall (Hussain et al., 2024).

Retirement planning and financial planning were moderately and significantly positively correlated, and financial knowledge and financial planning were also positively correlated. The results here show that, beyond saving patterns, some financial knowledge and conscious financial planning might be more closely linked to retirement preparation. This indicates that, in line with the concept presented at the beginning of the study, retirement readiness comprises financial resources, financial awareness, financial planning ability, and financial aims (Mohidin, 2013; Liu et al., 2021; Hussain et al., 2024).

Furthermore, the Mann-Whitney U test was used in this study to assess differences between male and female participants for each variable. The results indicated no statistically significant differences in future expectations, financial planning, financial knowledge, retirement planning, or saving behavior between males and females, as all p-values were above 0.5. This indicated that, in this study sample, there was no significant gender difference in financial behaviors and retirement preparation across items. The results should be interpreted with caution as the number of respondents was just 36.

The results of this study can also be contrasted with established prior studies on financial and retirement planning. Mohidin (2013) concluded that there was a significant positive relationship between personal financial planning, in terms of money management and investment planning, and retirement planning behavior. This finding corroborates the current study's results that financial planning is more strongly positively related to retirement planning. Similarly, a previous study on working women in Pakistan found that a future time perspective and clarity of retirement goals were important predictors of retirement planning, and that financial literacy was also a key predictor of the relationship between retirement planning and these factors. Financial literacy, too, is significant in the decision-making process and financial habits of people in Pakistan, consistent with findings in more advanced nations. In Pakistan, research has also shown that financial literacy plays a crucial role in shaping financial decisions and behaviors.

This is important because it implies that increases in saving cannot be shown to have a significant impact on retirement readiness without corresponding increases in other behaviors. More emphasis could be placed on individuals' financial literacy, providing them with clear retirement goals and motivating them to take an active role in their financial future. However, there are some limitations to this study. This study was conducted with 36 participants using convenience sampling, so the results may not generalize to the wider working population of

Pakistan. Future studies might require larger and more diverse samples from varied geographic areas, income levels, ages, and occupations to better understand retirement planning and saving behavior in Pakistan.

**Conclusion**

The main objective of the study was to examine the retirement planning and savings habits of Pakistan's working population and to identify the major factors contributing to savings and retirement preparation. In this study, people from various backgrounds responded to the survey that was sent out. It investigated the relationship between various factors, including age, work status, marital status, the number of children, life position, and saving and retirement activities of working individuals. The participants were drawn from a diverse range of occupations and age groups, reflecting different perspectives and experiences in saving and planning for retirement. The main findings revealed that saving behavior was positively associated with retirement planning, though this association was not statistically significant. Future expectations, financial planning, and financial knowledge, however, were all positively correlated with retirement planning. Financial planning had the strongest correlation, indicating that being proactive about it is especially significant for retirement planning. The Mann-Whitney U analysis also showed no statistically significant differences between male and female participants in their saving behavior, future expectations, financial planning, financial knowledge, or retirement planning. These results provide an indication of retirement planning and saving, but the study participants numbered only 36 and were convenience-sampled. These results may not be representative of the general working population of Pakistan; therefore, future studies might involve a larger or more diverse sample from other regions of Pakistan, as well as additional research on the relationship between age, income, occupation, education, and preparation for retirement and family responsibilities. In general, the results show that financial planning, financial knowledge, and planning for future financial needs are key factors when making retirement plans.

**Works Cited**


Afthanorhan, Asyraf. “Framing the Retirement Planning Behavior Model towards Sustainable Wellbeing among Youth: The Moderating Effect of Public Profiles.” Framing the Retirement Planning Behavior Model towards Sustainable Wellbeing among Youth: The Moderating Effect of Public Profiles, MDPI, 02 09 2020, https://www.mdpi.com/2071-1050/12/21/8879. Accessed 13 07 2026.

Ahmad, Fawad. “Determinants of Savings Behavior in Pakistan: Long Term-Short Term Association and Causality.” Timisoara Journal of Economics and Business, vol. 8, no. 1, 2015. IDEAS/RePEc

Aslam, Maria, et al. “Assessing the Household Saving Behavior of Urban and Rural Households in Pakistan: Evidence from PSLM/HIES 2018-19.” Pakistan Journal of Humanities and Social Sciences, vol. 10, no. 4, 2022, pp. 1439–1446. doi:10.52131/pjhss.2022.1004.0299. iRASD Journals

Hussain, Rana Tanveer, Muhammad Aslam, and Muhammad Asif. “Determinants of Financial Planning for Retirement: A Study on Working Women in Government of Pakistan.” Journal of Asian Development Studies, vol. 13, no. 3, 2024, pp. 1703–1715. doi:10.62345/jads.2024.13.3.137. ResearchGate

Isbanah, Yuyun. “Students' saving behavior: what are the motives that influence them to save?” Students' saving behavior: what are the motives that influence them to save?, Jurnal Ekonomi Bisnis dan Kewirausahaan (JEBIK), 03 11 2022, https://pdfs.semanticscholar.org/342d/648c0c8c12bc7189c0bd0d04a9dfe27a8bb9.pdf. Accessed 15 07 2026.

Jackson, Richard. “The Global Retirement Crisis.” The Global Retirement Crisis | The Geneva Papers on Risk and Insurance - Issues and Practice | Springer Nature Link, The Geneva Papers on Risk and Insurance - Issues and Practice, 01 10 2002. Accessed 10 07 2026.

Liu, Siming, et al. “The Behavioral Role of Digital Economy Adaptation in Sustainable Financial Literacy and Financial Inclusion.” Frontiers in Psychology, vol. 12, 2021, article 742118. doi:10.3389/fpsyg.2021.742118. Frontiers

Mathers, Colin D. “Causes of international increases in older age life expectancy.” Causes of international increases in older age life expectancy - The Lancet, The Lancet, 07 02 2015. Accessed 10 07 2026.

Mohidin, Rosle, et al. “Revisiting the Relationship Between Attitudes and Retirement Planning Behavior: A Study on Personal Financial Planning.” International Journal of Multidisciplinary Thought, vol. 3, no. 2, 2013, pp. 449–461. UMS Institutional Repository

**What Portfolio Characteristics Predict Retail Investor Outperformance? A Systematic Review and Quantitative Synthesis of Trading Activity, Diversification, and Investment Horizon By Bhargava Janga**


## Abstract

Retail investors are often characterized as underperforming market benchmarks, yet investment outcomes vary substantially across individuals. This variation raises a more useful question than whether retail investors underperform on average: what characteristics distinguish stronger-performing retail investors from weaker-performing ones? This study systematically reviews empirical research examining the relationship between retail investor performance and three characteristics of portfolio management: trading activity, diversification and portfolio concentration, and investment horizon and experience. Peer-reviewed studies using account-level or transaction-level retail investor data were identified from major financial journals and academic research repositories. Because differences in outcome measurement, portfolio definitions, time horizons, and overlapping investor datasets make a single pooled effect estimate potentially misleading, findings are synthesized quantitatively where comparable and narratively where methodological differences prevent direct aggregation. The evidence indicates that frequent trading is the most consistently negative characteristic. Across large datasets from the United States, Taiwan, Sweden, Germany, and other markets, highly active investors generally experience weaker net or risk-adjusted performance, although investor experience and information advantages create important exceptions. Evidence on diversification is more conditional. Under-diversification generally increases idiosyncratic risk and is associated with poorer risk-adjusted outcomes for typical retail investors, but concentrated portfolios can outperform when held by experienced or better-informed investors. Evidence supporting longer investment horizons is largely indirect because few studies isolate holding duration from turnover and investor experience. Overall, the literature suggests that retail outperformance is less associated with constant portfolio activity than with disciplined trading, adequate diversification, and the ability to avoid persistent behavioral errors.


## 1. Introduction

Retail participation in financial markets has expanded substantially as brokerage commissions have fallen, mobile trading has become widespread, and financial information has become easier to access. Individual investors can now construct and alter portfolios with far less friction than earlier generations of market participants. These changes have made direct investing more accessible, but they have not eliminated the behavioral and portfolio-management problems documented in the academic literature.

A large body of research finds that the average individual investor does not consistently outperform broad market benchmarks after costs. The most influential evidence comes from Barber and Odean (2000), who studied tens of thousands of U.S. brokerage households and found that investors who traded the most earned substantially weaker net returns than investors

who traded less. Subsequent research across other countries and market structures has repeatedly identified similar patterns. Excessive turnover, poor timing, attention-driven purchases, and behavioral biases can reduce the returns that investors actually earn even when the underlying securities perform reasonably well.

However, describing retail investors as a single group hides substantial variation. Some individuals trade almost every day, while others rarely change their portfolios. Some hold only a few stocks, while others maintain broad diversification. Some investors repeatedly make the same mistakes, while others appear to learn from prior outcomes. The existence of these differences suggests that the more interesting question is not simply whether retail investors underperform, but which observable characteristics are associated with stronger outcomes.

This paper focuses on three related portfolio characteristics that have received substantial empirical attention: trading activity, diversification and concentration, and investment horizon and experience. Trading activity captures how frequently investors alter their positions. Diversification describes how broadly investors distribute capital across securities and how much firm-specific risk they retain. Investment horizon and experience capture the extent to which investors maintain positions, learn from previous decisions, and avoid reacting to short-term fluctuations.

These characteristics are related but not identical. A concentrated investor can trade infrequently, while a diversified investor can turn over an entire portfolio rapidly. An investor can also have a long average holding period while retaining losing positions for behavioral reasons. For that reason, the evidence must be considered jointly rather than interpreted as a set of isolated rules.

The central research question is:

**Which portfolio characteristics are most consistently associated with retail investor outperformance in the empirical literature?**

The paper evaluates three hypotheses. First, higher trading activity is expected to be associated with weaker net and risk-adjusted returns. Second, severe under-diversification is expected to be associated with weaker risk-adjusted outcomes for the typical retail investor, although informed concentration may represent an important exception. Third, longer investment horizons and greater investor experience are expected to be associated with improved outcomes when they reduce unnecessary trading and repeated behavioral mistakes.

Rather than constructing a new investor-level dataset, this study synthesizes findings from existing empirical research. This design allows evidence from multiple countries, brokerage datasets, and time periods to be compared. At the same time, it avoids creating a misleading pooled estimate when studies use different performance measures or repeatedly analyze the same underlying investor samples. The goal is therefore a systematic review with quantitative evidence synthesis rather than a single conventional meta-analytic coefficient.

## 2. Theoretical Background

Modern portfolio theory and behavioral finance offer complementary explanations for why retail investors experience different outcomes.

Markowitz (1952) established the basic logic of diversification. When asset returns are not perfectly correlated, combining securities can reduce idiosyncratic risk without proportionally reducing expected return. From this perspective, a severely concentrated portfolio exposes an investor to company-specific outcomes that could have been diversified away. Diversification does not guarantee superior raw returns, but it generally improves the relationship between expected return and avoidable risk.

Behavioral finance adds a second perspective. Investors do not always trade according to fully rational expectations. They may become overconfident, overweight recent information, focus on attention-grabbing securities, sell winning investments too early, or retain losing investments too long. These behaviors can influence both portfolio construction and realized performance.

Overconfidence is particularly relevant to turnover. If investors overestimate the precision of their information or their ability to identify mispriced securities, they may trade more frequently than is justified by expected gains. Each additional trade may create direct costs, but even when explicit commissions are small, frequent trading can still expose investors to bid ask spreads, taxes, poor timing, and adverse selection. Barber and Odean's research is consistent with this mechanism because the investors who traded most tended to experience the weakest net results.

The disposition effect provides a useful counterpoint to the idea that longer holding periods are automatically beneficial. Odean (1998) showed that investors display a tendency to sell profitable investments while holding losing investments. This behavior demonstrates that patience and inertia are not the same thing. A long holding period can reflect disciplined long term investing, but it can also reflect reluctance to recognize an unsuccessful decision.

Diversification also interacts with investor skill. In theory, an investor with valuable private information or exceptional security-selection ability may rationally concentrate capital in a small number of positions. For the typical retail investor, however, concentrated portfolios often appear alongside overconfidence, familiarity bias, local bias, and limited financial sophistication. The literature therefore suggests a conditional relationship: concentration may be beneficial for unusually informed investors but costly for the average investor.

A central theme of this review is consequently that portfolio characteristics should not be interpreted mechanically. Low turnover is not valuable because trading is always harmful. Diversification is not valuable because owning more stocks automatically raises returns. Long holding periods are not valuable simply because time passes. These characteristics matter because they can reflect or influence the quality of investor decision-making.

## 3. Methodology

This study uses a structured systematic-review approach to identify and synthesize empirical research on retail investor portfolio characteristics and performance. The objective is

not to estimate a single universal effect size, but to compare the direction, magnitude, consistency, and limitations of findings across studies.

### 3.1 Scope and eligibility

Studies were considered relevant when they examined individual or retail investors using actual brokerage records, account-level holdings, or transaction-level data. Studies based exclusively on professional mutual funds, institutional asset managers, or hypothetical experimental portfolios were not treated as primary evidence because their decision environments differ materially from those of retail investors.

Eligible studies also needed to examine at least one of the following relationships: trading activity and investment performance, diversification or concentration and performance or risk, holding behavior and realized outcomes, or investor experience and subsequent trading performance. Classic theoretical papers were used to establish background but were distinguished from empirical retail-investor evidence.

### 3.2 Search strategy

The review focused on peer-reviewed finance and economics journals and well established academic working-paper repositories. Search terms included combinations of retail investor, individual investor, portfolio turnover, trading frequency, diversification, portfolio concentration, holding period, disposition effect, investor learning, brokerage accounts, and investment performance. Reference lists from foundational papers were also used to identify related studies.

### 3.3 Avoiding double counting

An important methodological problem is that multiple influential papers analyze the same or overlapping brokerage datasets. For example, several Barber and Odean studies use data from a large U.S. discount brokerage during the 1990s. Treating each publication as an independent investor sample would artificially increase the apparent amount of evidence.

The synthesis therefore distinguishes between independent datasets and multiple analyses of the same underlying sample. Repeated analyses remain informative because they test different mechanisms, but they are not counted as independent replications of the same result.

### 3.4 Outcome classification

Performance measures vary substantially across the literature. Some studies report raw portfolio returns, others report returns net of trading costs, and others calculate market-adjusted or risk-adjusted abnormal returns. Studies also differ in whether performance is measured before or after transaction costs.

To preserve these distinctions, results were classified into three broad categories: net realized performance, benchmark-adjusted performance, and risk-adjusted performance. Conclusions are based most heavily on studies that account for trading costs or systematic risk

because raw returns alone can reward investors for simply assuming more risk.

### 3.5 Evidence synthesis

A conventional meta-analysis generally requires sufficiently comparable effect-size estimates and statistically independent observations. The retail-investor literature does not consistently meet these conditions. Studies use different definitions of turnover and concentration, different return horizons, different currencies and markets, and different statistical models. Several studies also rely on overlapping investor samples.

For these reasons, forcing all results into a single pooled coefficient could create false precision. This paper instead uses quantitative synthesis where directly reported magnitudes can be meaningfully compared, while relying on structured narrative synthesis for heterogeneous results. Greater weight is placed on large samples, direct account-level evidence, risk-adjusted outcomes, and findings replicated across independent markets.

## 4. Results

### 4.1 Trading activity and portfolio performance

The strongest and most consistent result in the reviewed literature concerns trading frequency. Highly active retail investors generally earn lower net returns than less-active investors.

Barber and Odean (2000) provide the foundational U.S. evidence. Using 66,465 household accounts from a large discount brokerage, they found that the most active traders substantially underperformed the market on a net basis. The average household also lagged a passive market benchmark after costs, but the shortfall was considerably larger among accounts with high turnover. Their central conclusion was not simply that transaction costs matter. The pattern was consistent with overconfidence causing investors to trade even when the expected benefit of doing so was insufficient.

Barber and Odean (2001) examined trading differences between men and women using related brokerage data. Men traded more frequently, particularly unmarried men, and the additional trading was associated with a larger reduction in net performance. Because this study uses overlapping data with Barber and Odean (2000), it should not be considered an independent replication. It is nevertheless important because it supports the proposed behavioral mechanism linking overconfidence to excessive turnover.

Evidence from Taiwan provides a major independent test. Barber, Lee, Liu, and Odean (2009) analyzed the complete trading history of investors in the Taiwan Stock Exchange and estimated that individual trading generated economically meaningful performance losses in aggregate. The breadth of the dataset is important because it is not restricted to clients of a single U.S. broker. The results suggest that the turnover-performance relationship is not solely a feature of one brokerage population or one institutional setting.

Research from Germany also identifies a relationship between trading intensity and

investor characteristics. Dorn and Huberman (2005) found that more risk-tolerant and overconfident investors tended to trade more aggressively. Although the paper focuses heavily on the determinants of trading and portfolio choice rather than a single return ranking, it reinforces the interpretation that high turnover often reflects behavioral tendencies rather than superior information.

The overall evidence therefore strongly supports the first hypothesis. Frequent trading is the portfolio characteristic with the clearest negative association with retail investor outcomes. The exact magnitude differs across datasets because transaction costs, tax systems, market structures, and performance measures differ, but the direction of the relationship is unusually consistent.

There are, however, important exceptions. Not every trade destroys value, and not every active investor lacks skill. Some studies find evidence that subsets of investors possess information or learn over time. The correct conclusion is therefore not that trading should always be minimized. Rather, unusually high turnover is a strong warning sign when it is not supported by persistent informational advantage.

**4.2 Diversification and concentration**

The evidence on diversification is more nuanced than the evidence on turnover. Goetzmann and Kumar (2008) analyzed more than 40,000 individual-investor accounts and documented widespread under-diversification. Many investors held portfolios with substantial idiosyncratic risk that could have been reduced through broader diversification. Under-diversification was especially common among investors displaying characteristics associated with limited sophistication and behavioral bias.

Calvet, Campbell, and Sodini (2007) reached a related conclusion using detailed Swedish household financial data. Their analysis showed that households frequently made investment mistakes involving participation, diversification, and risk allocation. Importantly, financial sophistication was associated with better portfolio decisions. This supports the idea that diversification is not merely a mechanical feature of portfolio size. It is also connected to investor knowledge and financial decision quality.

At the same time, concentration cannot automatically be equated with poor stock selection. Ivkovic, Sialm, and Weisbenner (2008) found evidence that concentrated individual investors can display superior performance in the stocks to which they allocate unusually large weights. One interpretation is that some households concentrate precisely where they possess better information. This finding provides an important limitation to simple diversification rules.

Taken together, the studies indicate that severe under-diversification is generally undesirable for the typical retail investor because it increases uncompensated firm-specific risk. However, the relationship between concentration and raw returns is not uniformly negative. Informed investors may concentrate in securities where they possess an advantage, while less informed investors may concentrate because of familiarity, overconfidence, or limited attention.

This distinction also explains why risk-adjusted performance is particularly important. A

concentrated portfolio can outperform dramatically if one position succeeds, but that outcome does not necessarily imply a superior risk-return tradeoff. Concentration increases the dispersion of possible outcomes. For most investors, diversification remains the more reliable method for controlling idiosyncratic risk.

The second hypothesis is therefore supported with qualification. Diversification generally improves portfolio efficiency and reduces avoidable risk, but the literature does not establish that the most diversified investor will always earn the highest raw return.

### 4.3 Holding behavior, investment horizon, and the disposition effect

Evidence on holding period is less direct because relatively few studies isolate average holding duration as an independent predictor of performance. Most of the relevant evidence appears through turnover, the disposition effect, and investor learning.

Odean (1998) provides one of the most influential findings. Investors were more likely to realize gains than losses, meaning that they tended to sell winning positions and continue holding losing positions. More importantly, the winners that investors sold subsequently performed better than the losing securities they retained. This pattern contradicts the idea that a longer holding period is automatically beneficial.

The finding changes how the investment horizon should be interpreted. A long holding period is potentially valuable when it reflects conviction in a sound investment and resistance to unnecessary short-term trading. It is harmful when it reflects reluctance to recognize losses or update beliefs.

Research on investor learning provides evidence that experience can improve decisions. Nicolosi, Peng, and Zhu (2009) found that individual investors can learn from their trading experience and that past performance affects subsequent behavior. Their findings suggest that investor ability and behavior are not necessarily fixed. Some individuals become better at investment decisions as they accumulate experience.

Seru, Shumway, and Stoffman (2010), however, demonstrate why observed improvement must be interpreted carefully. Apparent learning among active investors can partly result from selection because unsuccessful investors may stop trading while successful investors remain in the sample. Their work suggests that cross-sectional comparisons between experienced and inexperienced investors can overstate true learning if investor exit is ignored.

These studies support a more precise version of the third hypothesis. Longer horizons appear beneficial when they are associated with lower unnecessary turnover, thoughtful security selection, and learning. Holding duration by itself is not sufficient. The empirical literature therefore provides moderate rather than definitive support for treating investment horizon as an independent predictor of outperformance.

### 4.4 Investor experience and heterogeneity

A recurring result across the literature is that retail investors differ substantially in skill. This matters because average effects can obscure persistent differences between investors. Some

investors repeatedly trade poorly, while others demonstrate evidence of information advantages or improved decisions. Experienced investors may be better at identifying securities, controlling risk, or recognizing when trading is justified. At the same time, experience does not automatically eliminate behavioral biases.

This heterogeneity helps reconcile apparently conflicting results. The typical highly active investor underperforms, yet a minority of informed active investors may trade profitably. The typical concentrated investor bears unnecessary idiosyncratic risk, yet some investors appear to concentrate where they possess information. The typical investor may exhibit a disposition effect, yet some individuals learn to reduce repeated mistakes.

The relevant characteristics therefore operate probabilistically rather than deterministically. High turnover raises the likelihood of underperformance, but does not guarantee it. Concentration raises idiosyncratic risk, but can be rational for an informed investor. Longer holding periods reduce activity, but can also preserve losing positions. The most successful retail behavior appears to involve matching activity to genuine informational advantage rather than maximizing or minimizing any single variable.

## 5. Comparative Evidence Summary

Table 1 summarizes several of the most influential empirical studies included in the synthesis. The studies are not treated as statistically independent when they use overlapping datasets.

| Study | Market / Sample | Primary characteristic | Main finding |
|---|---|---|---|
| Barber & Odean (2000) | U.S. brokerage households | Turnover | Highest-turnover households earned markedly weaker net returns. |
| Barber & Odean (2001) | U.S. brokerage households, overlapping sample | Trading / overconfidence | Higher trading among men was associated with larger performance reductions. |
| Barber et al. (2009) | Taiwan Stock Exchange investors | Trading activity | Individual trading produced substantial aggregate performance losses. |
| Goetzmann & Kumar (2008) | U.S. individual accounts | Diversification | Under-diversification was widespread and associated with investor characteristics linked to lower sophistication. |
| Calvet et al. (2007) | Swedish households | Diversification / sophistication | Households made diversification and risk allocation mistakes; sophistication improved decisions. |
| Ivkovic et al. (2008) | U.S. individual accounts | Concentration | Some concentrated investors showed evidence consistent with superior information. |
| Odean (1998) | U.S. brokerage accounts | Holding behavior | Investors disproportionately sold winners and held losers, consistent with the disposition effect. |
| Nicolosi et al. (2009) | Individual brokerage investors | Experience / learning | Trading experience was associated with changes in behavior and evidence of learning. |

| | | | |
|---|---|---|---|
| Seru et al. (2010) | Individual investors | Learning / selection | Apparent learning partly reflected unsuccessful investors exiting trading. |

## 6. Discussion

The literature produces a clear hierarchy in the strength of evidence across portfolio characteristics.

Trading activity has the strongest and most consistent relationship with performance. Large studies using different markets repeatedly find that investors who trade heavily tend to earn weaker net results. The consistency of this finding is particularly notable because the studies span different transaction-cost environments and institutional settings. Excessive activity appears to harm investors through a combination of explicit costs and behavioral mistakes.

This result remains relevant even as commissions approach zero. Eliminating brokerage commissions removes only one component of trading cost. Investors still face spreads, potential tax consequences, poor timing, and the opportunity cost of replacing a successful position with an inferior one. More importantly, commission-free trading does not remove overconfidence or attention-driven decision-making. Lower friction may even make it easier for behavioral impulses to translate into transactions.

Diversification produces a more conditional conclusion. The traditional argument for diversification is well supported when the objective is to reduce idiosyncratic risk. Retail households frequently hold portfolios that are less diversified than standard portfolio theory would recommend. This can leave investors exposed to risks for which markets do not necessarily provide additional expected compensation.

However, the evidence also shows why concentration should not be treated as inherently irrational. Investors with genuine information may rationally place greater weight on securities where they possess an advantage. The important distinction is therefore between informed concentration and accidental or behaviorally driven concentration. Unfortunately, an investor usually cannot know with certainty which category describes his or her own portfolio.

Investment horizon is the least cleanly measured of the three characteristics. The literature strongly suggests that unnecessary short-term trading is harmful, but it does not follow that holding every investment for a long period is optimal. The disposition effect demonstrates that investors can hold positions too long for the wrong reason. A useful interpretation is that long-term orientation is beneficial when combined with willingness to update beliefs when evidence changes.

These findings also highlight the importance of distinguishing portfolio characteristics from investor skill. A successful investor may trade more because new information justifies trading. Another investor may trade more because of overconfidence. Their turnover statistics can look similar even though the underlying process is different. Likewise, two concentrated portfolios can reflect either superior information or inadequate diversification.

For that reason, portfolio characteristics are best interpreted as indicators of likely behavior rather than universal causal rules. Among typical retail investors, high turnover and

severe concentration are warning signs. But the strongest conclusion from the literature is not simply to trade less or own more stocks. It is that performance improves when portfolio decisions become more deliberate, evidence-based, and resistant to behavioral error.

### 6.1 Implications for behavioral finance

The review supports several core ideas in behavioral finance. First, overconfidence appears capable of translating directly into costly portfolio activity. Second, investors often depart from theoretically efficient diversification. Third, realized gains and losses affect selling behavior in ways that are difficult to reconcile with purely rational tax or information explanations.

The persistence of these effects across datasets suggests that they are not isolated anomalies. Instead, they represent recurring features of household financial decision-making.

### 6.2 Implications for retail investors

The evidence has practical implications, although the observational nature of the studies means these implications should not be treated as guaranteed investment rules. The strongest lesson concerns unnecessary trading. Before replacing an investment, an investor should have a reason strong enough to overcome not only explicit costs but also the possibility that the decision is being driven by attention, emotion, or overconfidence. A second implication concerns diversification. Investors without a clear informational advantage generally benefit from avoiding portfolios dominated by a small number of individual companies. Broad diversification cannot prevent market losses, but it can reduce losses arising from company-specific events.

Finally, investors should distinguish patience from refusal to update. A long-term strategy can reduce behavioral reactions to short-term noise, but long holding periods should not become an excuse to ignore deteriorating fundamentals or evidence that an original investment thesis was wrong.

## 7. Limitations

This review has several limitations.

First, the underlying literature is heterogeneous. Studies use different countries, investor populations, return measures, market benchmarks, and definitions of trading activity. This variation limits the validity of calculating one pooled effect size across all studies.

Second, influential papers sometimes analyze overlapping datasets. Multiple publications from the same brokerage sample provide important evidence about different mechanisms, but they cannot be treated as independent replications. This review explicitly accounts for that limitation, but it reduces the number of truly independent large-scale samples.

Third, public academic datasets often come from earlier periods in market history. Commission structures, mobile brokerage interfaces, social media, and fractional trading have changed dramatically. Behavioral mechanisms documented in earlier datasets may persist, but their magnitudes may differ in modern markets.

Fourth, observational studies cannot perfectly separate investor skill from investor behavior. An investor may trade frequently because of overconfidence or because of genuine information. A concentrated portfolio may reflect poor diversification or superior knowledge. Statistical controls reduce this problem but do not eliminate it.

Fifth, this review focuses primarily on directly held equities. Results may not generalize to investors whose portfolios consist mainly of index funds, retirement accounts, options, cryptocurrencies, or other assets.

Finally, the evidence on the holding period is less direct than the evidence on turnover and diversification. Future research using modern account-level datasets could explicitly measure average holding duration while controlling for turnover, investor experience, portfolio concentration, and security characteristics.

## 8. Conclusion

Retail investors do not experience uniform outcomes. The empirical literature shows substantial differences in trading behavior, portfolio concentration, experience, and performance across individuals. This review asked which observable characteristics are most consistently associated with stronger retail investment outcomes.

The clearest answer concerns trading activity. Across large investor datasets and multiple markets, unusually frequent trading is repeatedly associated with weaker net or risk-adjusted performance. The relationship is consistent with transaction costs, overconfidence, attention driven investing, and poor timing.

Diversification also matters, but the relationship is more conditional. Severe under diversification exposes typical investors to avoidable company-specific risk and is frequently associated with lower portfolio efficiency. Nevertheless, concentration can be rational and even profitable when investors possess meaningful information advantages.

Longer investment horizons appear useful primarily when they reduce unnecessary activity and allow disciplined investment strategies to develop. Holding an investment for a long period is not independently evidence of skill, as the disposition effect demonstrates that investors sometimes retain unsuccessful positions for behavioral reasons.

Taken together, the literature suggests that retail outperformance is not characterized by constant activity. Stronger outcomes are more often associated with disciplined trading, adequate diversification, experience, and the ability to avoid repeated behavioral errors. The evidence therefore shifts attention away from finding a single ideal portfolio structure and toward understanding the decision processes that produce portfolio choices.

The central distinction is not simply between active and passive investors, concentrated and diversified investors, or short-term and long-term investors. It is between portfolio decisions supported by information and disciplined reasoning and decisions driven by behavioral bias. That distinction provides the most consistent explanation for why some retail investors perform better than others.

**Works Cited**


Barber, B. M., Lee, Y.-T., Liu, Y.-J., & Odean, T. (2009). Just how much do individual investors lose by trading? The Review of Financial Studies, 22(2), 609-632.

Barber, B. M., & Odean, T. (2000). Trading is hazardous to your wealth: The common stock investment performance of individual investors. The Journal of Finance, 55(2), 773-806. https://doi.org/10.1111/0022-1082.00226

Barber, B. M., & Odean, T. (2001). Boys will be boys: Gender, overconfidence, and common stock investment. The Quarterly Journal of Economics, 116(1), 261-292. https://doi.org/10.1162/003355301556400

Calvet, L. E., Campbell, J. Y., & Sodini, P. (2007). Down or out: Assessing the welfare costs of household investment mistakes. Journal of Political Economy, 115(5), 707-747. Dorn, D., & Huberman, G. (2005). Talk and action: What individual investors say and what they do. Review of Finance, 9(4), 437-481.

Goetzmann, W. N., & Kumar, A. (2008). Equity portfolio diversification. Review of Finance, 12(3), 433-463. https://doi.org/10.1093/rof/rfn005

Ivkovic, Z., Sialm, C., & Weisbenner, S. (2008). Portfolio concentration and the performance of individual investors. Journal of Financial and Quantitative Analysis, 43(3), 613-655.

Markowitz, H. (1952). Portfolio selection. The Journal of Finance, 7(1), 77-91. Nicolosi, G.,

Peng, L., & Zhu, N. (2009). Do individual investors learn from their trading experience? Journal of Financial Markets, 12(2), 317-336.

Odean, T. (1998). Are investors reluctant to realize their losses? The Journal of Finance, 53(5), 1775-1798.

Seru, A., Shumway, T., & Stoffman, N. (2010). Learning by trading. The Review of Financial Studies, 23(2), 705-739.

**The Effect of Bankruptcy-Law Reform on Entrepreneurship Varies According to National Culture By James Nichols**

**Abstract**

Bankruptcy law is widely used as a tool by governments to influence entrepreneurship. The logic behind it makes sense and research in the past expresses it well: in a creditor-friendly environment the already large downside of a bankruptcy is expanded for the debtor when the repercussions are more harsh, therefore fewer entrepreneurs will enter the market due to fear of bankruptcy, and vice versa for a debtor-friendly environment (Puchakayala and Veluchamy, 2023).

We have seen these methods used in the past successfully. For example, Lee et al. (2011) found that across 29 countries over 19 years, countries with more lenient bankruptcy law had higher rates of startups and entrepreneurship. But this pattern isn't consistent.

When studies have looked at standalone countries and how the implementation of new bankruptcy laws affected entrepreneurship, inconsistent results were found. Finland's Bankruptcy Act was introduced in 2004. It is considered a debtor-friendly reform of Finland's previous laws, but the results did not bear out what previous research had posited — there was hardly any change at all in Finland's level of entrepreneurship (Sadeghi and Kibler, 2022). Similarly, in Japan the case for bankruptcy law having its intended effects was not wholly backed up by the evidence. Reforming its bankruptcy laws to a more debtor-friendly environment, Japan only saw an increase in entrepreneurship among the "elite" — those from top universities or over 40 years of age (Eberhart et al 2017.). This shows that the effects of reforming bankruptcy law are not consistent.

Although vague patterns may emerge when a large data set is used, when a magnifying glass is placed over one case study, that pattern seems to fall apart and become unreliable. In the past this inconsistency has been attributed to variables that were not considered, such as the speed of the legal process or the type of entrepreneur in that system (Fu et al. 2020). So it is known that bankruptcy law on its own is unlikely to fully explain how enterprising and risk-taking a nation is.

One variable that has received comparatively little attention in this specific context is national culture. Existing research has already established that cultural dimensions — measured, for instance, through Geert Hofstede's (Hofstede, Hofstede and Minkov, 2010) six-criteria framework — correlate with a nation's level of entrepreneurship (Dheer, 2017; Çelikkol, Kitapçı and Döven, 2019). What hasn't yet been done is connecting that established relationship back to the specific puzzle above: why the same kind of legal reform produces such inconsistent outcomes across countries, and even across different groups of entrepreneurs within one country. That is the purpose of this study — to draw together the existing evidence on culture and entrepreneurship and use it to explain the inconsistency documented in the research discussed above.

Therefore the hypothesis of this research paper is that national culture has been overlooked when assessing how bankruptcy law affects entrepreneurship. It is not a denial that bankruptcy law has an effect - it has been proven to have one, nor is it to say that culture is the most important determinant of entrepreneurship. It is to add to the list of contributing factors that affect entrepreneurship. To help form a conclusion I interviewed experts from Xero as well as representatives from The Entrepreneurial Refugee Network (TERN) in order to find the most conclusive evidence.

**Context**

There is significant research, both modern and dated, which shows compelling evidence to support the view that bankruptcy law is the most significant factor in controlling entrepreneurship. Entrepreneurship has become a more and more significant factor of every economy. In terms of scale the Global Entrepreneurship Monitor's 2024/2025 Global Report found that one in eight working aged people were involved in entrepreneurial activity. The World Bank's Entrepreneurship Database also shows something similar: newly registered limited liability companies per 1,000 working-age adults rose about 50%, from 2.9 in 2006 to 4.3 in 2024. This growth is not some short run temporary phenomenon - the rebound from COVID 19 shows this.The 2024–2025 GEM US Report found Total Entrepreneurial Activity at a historic high of 19%. This new economics chapter is called “entrepreneurial economy” (Audretsch, Carree, and Thurik, 2001) characterized also by large firms breaking down the devolving into smaller more flexible firms. With ever increasing technological advances barriers to entry are lower for nearly all business structures (Etro, 2009).

As entrepreneurship has become an essential factor of any economy trying to grow, governments have resorted to bankruptcy law to guide the extent and magnitude of entrepreneurial activity (Lee et al., 2011). But despite extensive initial research about which policies were most effective, the distinct aspects of each economy mean that there is no straightforward path to controlling entrepreneurial activity (Davidsson and Wiklund, 1997; Frederking, 2004). It is hard to argue against it but it is not to say that implementation of bankruptcy law has not had success in the past. One example is Germany's Insolvenzordung in 1999. Unifying the bankruptcy procedure, it introduced a US Chapter 11–style reorganization plan (Insolvenzplan), and created personal consumer insolvency with debt discharge (Restschuldbefreiung). The reforms brought a much needed change to a system that was very harsh against debtors. Introducing a ‘fresh start’ mechanism which meant entrepreneurs were not personally liable for their business if it went bankrupt allowing risk taking people to have multiple goes at starting a business in order to improve growth in the German economy.

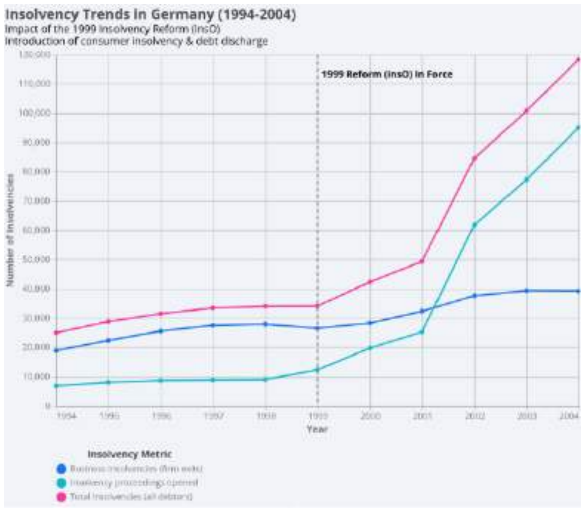



Figure 1.

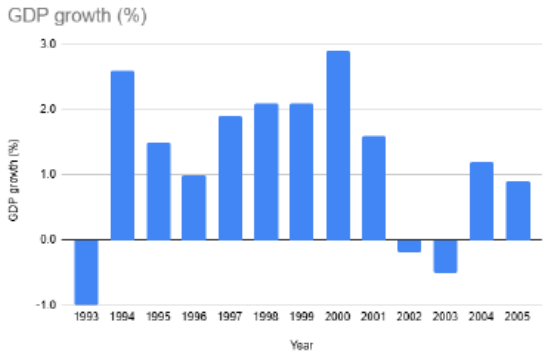



Figure 2

Figure 1 shows that the reform had the intended direct effects. Because of a greater insurance effect (Fan and White, 2003), business owners felt more confident in taking risks, hence the higher rate of insolvencies going from ~32,000 in 1999 to ~84,000 in just 3 years. This increased risk-taking position business owners took is a signal for a higher rate of entrepreneurship (Kihlstrom and Laffont, 1979) as risk appetite is an essential factor of entrepreneurship. Stuart Miller (Miller, 2026), director of public policy and technology research at Xero agrees with this saying that risk appetite is a key evaluation point for levels of entrepreneurship. But insolvencies don't ultimately translate to higher growth which is the goal for higher entrepreneurship (Radha Krishna and Harika, 2025). Debtor friendly bankruptcy law follows the chain of reading which starts with entrepreneurs being less risk-adverse, evidence of this can be found in the number of failures as shown by figure 1. Therefore it can be concluded that the risk appetites of entrepreneurs in Germany did increase as a result of the Insolvenzordung. This does point towards bankruptcy having an effect on entrepreneurship but it does not account for the indirect effects that may have contributed to the indicator of higher insolvency rate. When the insurance for starting a business is so low, it attracts people who wouldn't naturally be entrepreneurs because of the lower barriers to entry (Landier, 2005). This means that the policy of Insolvenzordung may have had the direct effect of encouraging entrepreneurs to not be afraid of failure but at the same time encouraged those not from the

‘elite’ background (Sorensen and Chang, 2006) which dilutes the growth of the economy as these entrepreneurs are more likely to keep failing instead of providing innovation and growth to the economy. As seen in Figure 2, which shows no consistent growth after 1999, Germany experienced no definitive results from its implementation. So although the signifier of a successful policy was there in the form of higher insolvencies, the real ultimate intended effect was not evident.

Sometimes we see new bankruptcy laws have little or no effect towards the intended outcome. Finland introduced reforms to its bankruptcy system in 2004. Finland's Bankruptcy act brought into law the No-Discharge provision which allowed creditors to further pursue the collection of their debt. A change that would bring massive benefit to creditors and in theory would take away confidence from entrepreneurs. To evaluate if the reform had any effect Sadeghi & Kibler (2022) created a ‘synthetic Finland’ by compiling aspects of countries which most closely resembled Finland’s economy prior to the act and compared the changes seen in the synthetic Finland with the real Finland. They found that there was no detectable change in levels of entrepreneurship.

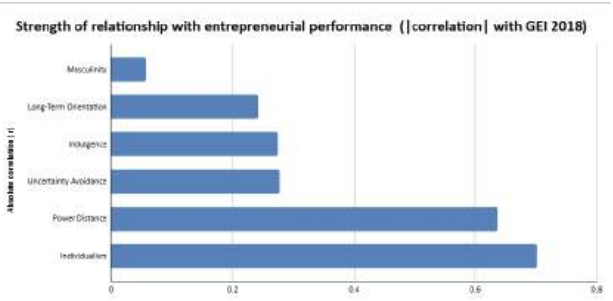
But we also see examples of successful implementation of new bankruptcy law. India introduced its Insolvency and Bankruptcy code in 2016 in order to solve the issue of high insolvency rates. The reform allowed control of insolvencies to creditors rather than to the court. This creditor friendly policy was brought into effect in order to reduce the risk appetite of potential entrepreneurs which in turn would lower the number of businesses started and therefore lower the number of insolvencies the India economy had to deal with (Puchakayala and Veluchamy, 2023). Following it,the creditor recovery rate in India rose from 26 cents on the dollar to 71.6 cents on the dollar; average resolution time fell from 4.3 years to 1.6 years (Doing Business report, 2020). Therefore we have an example of bankruptcy law being introduced to create a creditor friendly environment (La Porta et al., 1998).

The evidence from these three examples suggests that while using legal reform as a lever to manage entrepreneurship is theoretically sound (Lee et al., 2011), the actual outcomes remain notably unpredictable. The research indicates a spectrum of efficacy, spanning from the specific successes documented in Japan (Eberhart, Eesley and Eisenhardt, 2017) to the negligible impact observed in Finland (Sadeghi and Kibler, 2022). This reinforces the argument that a uniform policy framework rarely functions consistently when applied to differing national environments (Davidsson and Wiklund, 1997; Frederking, 2004).

**Analysis: The Interplay Between Cultural Dimensions and Entrepreneurial Performance**

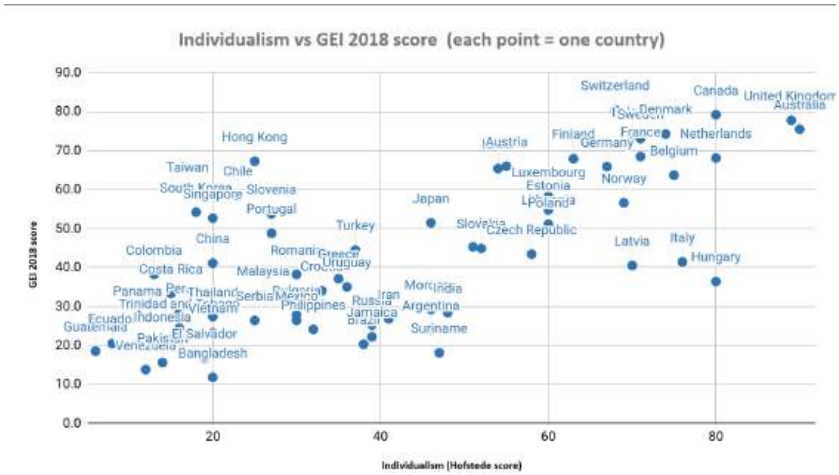
Analysis of cross-national datasets reveals persistent patterns between the cultural attributes of a nation and its entrepreneurial efficacy (Kleinhempel, Klasing and Beugelsdijk, 2023). In this study, entrepreneurial performance metrics (GEDI 2018) were integrated with Hofstede's cultural dimensions across 95 countries. Statistical analysis—including Pearson correlation, multivariate OLS regression, and income-controlled sensitivity tests—was employed to isolate the independent contribution of culture to variations in entrepreneurial activity.

Although the case studies discussed in Section 2 vary temporally, Hofstede's assertion that culture is a stable construct transmitted intergenerationally justifies the application of these findings to contemporary contexts. The fundamental cultural traits of Germany, Finland, and Japan remain relatively constant, providing a reliable basis for interpreting the outcomes of legislative reform.

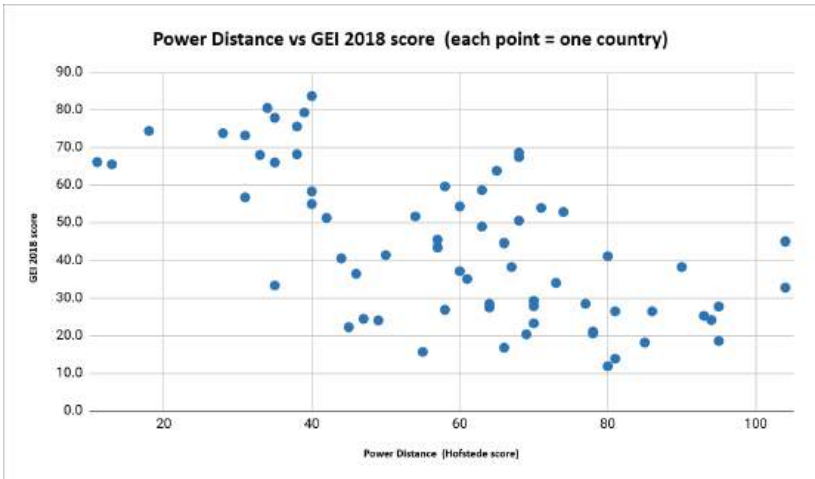



| Rank | Cultural dimension | Direction that helps | Correlation r | Strength of \|r\| | p value | Significant at 5%? | Hypothesis test verdict | Countries (n) |
|---|---|---|---|---|---|---|---|---|
| 1 | **Individualism** | Higher individualism | 0.702 | 0.702 | 0.00000 | **Yes** | Reject H0 | 67 |
| 2 | **Power Distance** | Lower power distance | -0.637 | 0.637 | 0.00000 | **Yes** | Reject H0 | 67 |
| 3 | **Uncertainty Avoidance** | Lower uncertainty avoidance | -0.278 | 0.278 | 0.02281 | **Yes** | Reject H0 | 67 |
| 4 | **Indulgence** | Higher indulgence | 0.275 | 0.275 | 0.00996 | **Yes** | Reject H0 | 87 |
| 5 | **Long-Term Orientation** | Higher long-term orientation | 0.242 | 0.242 | 0.02374 | **Yes** | Reject H0 | 87 |
| 6 | **Masculinity** | No clear direction | -0.057 | 0.057 | 0.64919 | **No** | Accept H0 | 67 |

By plotting a country's overall GEDI score against their score for each aspect of Hofstede's model of culture we can see a clear correlation between Individualism and Power distance when it comes to levels of entrepreneurship in a country. With an |r| value 0.702 and a p value of 0.0000 high individualism is a strong indicator of high levels of entrepreneurship. Using data from 67 countries which had published data on both the Geert Hofstede website and the GEDI website, it passed the hypothesis test at 5% by a large margin. Individualism is defined by Hofstede as “the extent to which people feel independent, as opposed to being interdependent as members of larger wholes.”. Logic checks out here as well, independent is required to start one's own business, being self employed requires one to set their own goals and plan processes for themselves (Vantilborgh, Joly and Pepermans, 2015). Examples of countries that follow the expected trend are Malaysia, Belgium and The Netherlands. Outliers include Hong Kong, Taiwan and Chile - this is because of a strong history of aggressive economic liberalization because of state-led or export-led economies and so their individualism score is proportionally much higher than their GEDI score.




Similarly, power distance is an important factor which plays into entrepreneurship according to this model. Geert Hofstede defines power distance as “Power Distance is the extent to which the less powerful members of organizations and institutions (like the family) accept and expect that power is distributed unequally.” and the regression gives an |r| value of 0.637 with a p value of 0.0000 as well. This shows that the more distant the power in a country/culture, the less likely entrepreneurism is going to be a strong suit of that country/culture. Once again this checks out logically, the less willing to accept powerlessness a country is, the more likely they are to start their own business. Self-efficacy and belief in their own ability to succeed is significantly higher in entrepreneurs (Chen, Greene and Crick, 1998). Some countries that follow the trend closely are Luxemburg, France and the Philippines. Countries that are clear outliers are the United States, France and Pakistan. Here The US and France lie above the line of best fit because of abnormally high rates of entrepreneurship and Pakistan lies below it because of a lack of entrepreneurship which is expected in such a power distant country.



But not all aspects of national culture have an effect, 4 of the 6 had a weak correlation. All 4 had |r| values below 0.3, but only masculinity failed the hypothesis test with a p value of 0.64919. So although it could be interpreted that 3 more of the aspects of culture could be interpreted as having an effect on entrepreneurism, for the sake of reliability I will treat that as variable without any correlation and only use individualism and power distance to explain the gaps in the previous studies mentioned in section 1. However Long term orientation is an exception. Although there is no direct link between long-term orientation and entrepreneurship levels, it does have a relationship with receptiveness to new policy and is therefore relevant when talking about bankruptcy law because the three case studies are all examples of policy being implemented.

**Discussion**

Germany's Insolvenzordung was clearly successful in promoting startup culture in Germany but wasn't successful because it created an environment in which Entrepreneurs were unafraid to fail rather than one where the benefits of starting a business outweighed the risk (Eberhart, Eesley and Eisenhardt, 2017). The outcome can be best explained when compared to the national culture of Finland as both countries have very similar profiles in terms of Individualism as well as power distance. In terms of individualism Germany has a score of 67 while Finland has a score of 63 putting them both around 1 standard deviation above the mean. For power distance Germany has a score of 35 and Finland 33 putting them both over 1 standard deviation below the mean. So both countries are comparable in these two metrics but both within the most extreme quartile for these metrics. So the similarity between the cultures of these two countries explains why they had similar outcomes when reforms were implemented - they both ultimately saw no real change. But the difference in the subtleties is that Germany had a rapidly increasing rate of insolvencies shortly after reforms were introduced. To explain this slight difference in outcome we can refer to long-term orientation, which the two countries vary massively. Germany has the sixth highest score at 83, meaning its national culture makes its residents (and therefore entrepreneurs) very malleable to law implementation, this explains why there was such a large reaction in 1999. Finland on the other hand has a score of 38 meaning its culture leads people to be more suspicious of change - leading to virtually no change following bankruptcy law reforms.

Finland presents an interesting case. Its relatively high level of individualism and low power distance suggest a cultural environment that is, in broad terms, supportive of entrepreneurial activity. Finland scores 63 on individualism and 33 on power distance, while its Global Entrepreneurship Index score is 67.9. However, Finland's comparatively low long-term-orientation score of 38, especially when contrasted with Germany's score of 83, may help explain why the reform produced only a limited behavioural response. This interpretation should be treated cautiously: long-term orientation is not a direct measure of receptiveness to policy change, and the available evidence is correlational rather than causal.
Germany provides a useful contrast. Its broadly similar levels of individualism and power

distance were accompanied by a much higher long-term-orientation score. This may have contributed to a more pronounced initial response to bankruptcy-law reform, reflected in the subsequent increase in insolvency activity. However, the evidence does not establish that the reform encouraged excessive numbers of unskilled or unprepared entrepreneurs. A more defensible interpretation is that the reform increased risk-taking and business entry, but did not produce sustained economic growth. Higher insolvency rates should therefore not be treated as equivalent to higher levels of successful entrepreneurship.

Taken together, the Finnish and German cases suggest that bankruptcy-law reform does not operate uniformly across national contexts. Finland's reform generated little detectable change in entrepreneurial activity, whereas Germany experienced a stronger initial response without clear evidence of sustained growth. National culture may offer a plausible explanation for this difference, but it should be presented as a moderating factor rather than as a complete or proven explanation.

India however follows a different chain of reasoning to explain its success when implementing reforms. India has a much higher power distance than the previous two countries at 77. This means those with the least power in India are relatively willing to accept that they are less powerful than others. This can be interpreted as more law obedient when it comes to government policy, so it makes sense that India's reforms had the intended effects. Its individualism score is above average at 48 and its long-term orientation is 56 - just above average.

It is reasonable to assume therefore that because of India's high power distance, reforms that the powerful government put into place are likely to have the intended effect.
What this discussion provides is a somewhat reasonable explanation as to why these three case studies all reached different conclusions having all studied similar reforms on bankruptcy law - with the main difference being the country in which the reforms were based. It shows how the unique national culture of each of these countries can be used to provide tailored explanations for why that nation had an outcome in terms of entrepreneurship to the implementation of its bankruptcy laws.

**Conclusion**

This paper has addressed the inconsistent results seen in previous case studies which have found varied entrepreneurial results following implementations of bankruptcy law across different countries. Individually the existing literature provides a basis for the claim that bankruptcy law affects the level of entrepreneurship in an economy: by changing the risk levels when entrepreneurs enter the market, policy makers can discourage or encourage new entrants from starting in the first place or from many attempts at starting a business. However the evidence that has been reviewed shows that the effectiveness of bankruptcy law reforms is very inconsistent, working in some countries but not others and sometimes working but not as intended. Finland saw no change following reforms, Japan saw benefits but only in certain groups which made the overall intended changes incomplete, Germany saw some intended

signals of higher entrepreneurship but ultimately it amounted to nothing. When we look at all these cases together it suggests that there is a missing factor that has not been taken into account across all the studies.

This cross-national analysis provides support for the hypothesis of this paper. It shows that individualism is positively associated with entrepreneurship and that power distance is negatively associated. These relationships reinforce the point that national culture acts as a filter for legal incentives through expectation on personal agency, authority and risk acceptance. A more individualistic economy and one with less power distance is one where entrepreneurs view business ownership as a legitimate and plausible career, making them more reactive when debtor-friendly policy is introduced. Long term orientation is also significant because of the nature of this argument, although there is not conclusive evidence to say that there is a direct correlation with entrepreneurship, a country's willingness to accept reforms balances on if they are susceptible to change in the long term. A country who is not adaptive will not embrace reforms well.

The case studies used demonstrate this argument effectively but do not provide a system of diagnosis case by case. Germany and Finland have broadly similar cultures when it comes to the criteria that affect entrepreneurship but different long-term orientation which is a plausible explanation for why they had slightly different results when reforms were introduced. India presented a contrasting case in which reforms were introduced in a high power setting economy. Although the reforms in each of the cases were generalised by what their intended effects were in the long term, this study is not to give exact answers and details to why each case happened, it only shows that national culture is an appreciated criteria for evaluating bankruptcy law. The analysis should therefore present culture as a moderating factor rather than as a complete explanation or a substitute for institutional analysis.

The data provides some support for a cautiously stated conclusion: bankruptcy law reforms are an effective tool for policymaking; however, their impact on entrepreneurship will be determined to some extent by the cultural and social context in which they exist. This research's most compelling finding is not that individualism or power-distance determine entrepreneurship but rather that they can contribute to explaining why the same legislative action has created varying forms of business formation (i.e., new ventures), business failure, business recovery and venture growth. Researchers should now focus on testing the relationship using panel data, interaction terms between country culture and bankruptcy law regimes, and indicators that differentiate between new business formation vs. business survival, repeat entrepreneurship, innovation and job creation. A within-country migration design combined with information regarding community support and stigmas associated with bankruptcy could potentially offer a promising approach to moving beyond anecdotal and correlational accounts to developing a more complete explanation of how law and culture interact to influence entrepreneurial behavior.

**Works Cited**


Ács, Z.J., Szerb, L., Lafuente, E. and Lloyd, A. (2018) Global entrepreneurship and development index 2018. Cham: Springer. doi: 10.1007/978-3-030-03279-1.

Audretsch, D.B. and Thurik, A.R. (2001) 'What's new about the new economy? Sources of growth in the managed and entrepreneurial economies', Industrial and Corporate Change, 10(1), pp. 267–315. doi: 10.1093/icc/10.1.267.

Çelikkol, M., Kitapçı, H. and Döven, G. (2019) 'Culture's impact on entrepreneurship and interaction effect of economic development level: an 81 country study', Journal of Business Economics and Management, 20(4), pp. 777–797. doi: 10.3846/jbem.2019.10180.

Chen, C.C., Greene, P.G. and Crick, A. (1998) 'Does entrepreneurial self-efficacy distinguish entrepreneurs from managers?', Journal of Business Venturing, 13(4), pp. 295–316. doi: 10.1016/S0883-9026(97)00029-3.

Davidsson, P. and Wiklund, J. (1997) 'Values, beliefs and regional variations in new firm formation rates', Journal of Economic Psychology, 18(2–3), pp. 179–199. doi: 10.1016/S0167-4870(97)00004-4.

Dheer, R.J.S. (2017) 'Cross-national differences in entrepreneurial activity: role of culture and institutional factors', Small Business Economics, 48(4), pp. 813–842. doi: 10.1007/s11187-016-9816-8.

Eberhart, R.N., Eesley, C.E. and Eisenhardt, K.M. (2017) 'Failure is an option: institutional change, entrepreneurial risk, and new firm growth', Organization Science, 28(1), pp. 93–112. doi: 10.1287/orsc.2017.1110.

Etro, F. (2007) Competition, innovation, and antitrust: a theory of market leaders and its policy implications. Berlin: Springer. doi: 10.1007/978-3-540-49601-4.

Fan, W. and White, M.J. (2003) 'Personal bankruptcy and the level of entrepreneurial activity', The Journal of Law and Economics, 46(2), pp. 543–567. doi: 10.1086/382602.

Frederking, L.C. (2004) 'A cross-national study of culture, organization and entrepreneurship in three neighbourhoods', Entrepreneurship and Regional Development, 16(3), pp. 197–215. doi: 10.1080/0898562042000197126.

Fu, K., Wennberg, K. and Falkenhall, B. (2020) 'Productive entrepreneurship and the effectiveness of insolvency legislation: a cross-country study', Small Business Economics, 54(2), pp. 383–404. doi: 10.1007/s11187-018-0040-6.

Global Entrepreneurship Monitor (2025a) Global Entrepreneurship Monitor 2024/2025 global report: entrepreneurship reality check. London: Global Entrepreneurship Research Association. Available at: https://gemconsortium.org/report/gem-20242025-global-report-entrepreneurship-reality-check-4 (Accessed: 1 September 2026).

Global Entrepreneurship Monitor (2025b) 2024/2025 USA national report: national entrepreneurship assessment for the United States. Babson Park, MA: Babson College. Available at:

https://gemconsortium.org/report/20242025-usa-national-report-national-entrepreneurship-assessment-for-the-united-states-2 (Accessed: 1 September 2026).

Hofstede, G. (no date) The 6D model of national culture. Available at: https://geerthofstede.com/culture-geert-hofstede-gert-jan-hofstede/6d-model-of-national-culture/ (Accessed: 1 September 2026).

Hofstede, G., Hofstede, G.J. and Minkov, M. (2010) Cultures and organizations: software of the mind. 3rd edn. New York: McGraw-Hill.

Insolvenzordnung (Insolvency Code) 1994, BGBl. I p. 2866 (in force 1 January 1999). Available at: https://www.gesetze-im-internet.de/englisch_inso/index.html (Accessed: 1 September 2026).

Kihlstrom, R.E. and Laffont, J.-J. (1979) 'A general equilibrium entrepreneurial theory of firm formation based on risk aversion', Journal of Political Economy, 87(4), pp. 719–748. doi: 10.1086/260790.

Kleinhempel, J., Klasing, M.J. and Beugelsdijk, S. (2023) 'Cultural roots of entrepreneurship: evidence from second-generation immigrants', Organization Science, 34(5), pp. 1800–1819. doi: 10.1287/orsc.2022.1645.

Landier, A. (2005) Entrepreneurship and the stigma of failure. Working paper. Available at: https://ssrn.com/abstract=850446 (Accessed: 1 September 2026).

La Porta, R., Lopez-de-Silanes, F., Shleifer, A. and Vishny, R.W. (1998) 'Law and finance', Journal of Political Economy, 106(6), pp. 1113–1155. doi: 10.1086/250042.

Lee, S.-H., Yamakawa, Y., Peng, M.W. and Barney, J.B. (2011) 'How do bankruptcy laws affect entrepreneurship development around the world?', Journal of Business Venturing, 26(5), pp. 505–520. doi: 10.1016/j.jbusvent.2010.05.001.

Miller, S. (2026) Interview with the author. [Add interview date and format, e.g. 'Interviewed by James Nichols, 14 August 2026, video call'.]

Puchakayala, N.V.V.S. and Veluchamy, R. (2023) 'Impact of resolving insolvency on entrepreneurial activity in India', SEDME (Small Enterprises Development, Management and Extension Journal), 50(2), pp. 169–181. doi: 10.1177/09708464221150680.

Radha Krishna, T. and Harika, K. (2025) 'The role of entrepreneurship in economic growth and development in India', IOSR Journal of Economics and Finance, 16(3), pp. 74–77. doi: 10.9790/5933-1603047477.

Sadeghi, A. and Kibler, E. (2022) 'Do bankruptcy laws matter for entrepreneurship? A synthetic control method analysis of a bankruptcy reform in Finland', Journal of Business Venturing Insights, 18, e00346. doi: 10.1016/j.jbvi.2022.e00346.

Sorensen, J.B. and Chang, P.M.Y. (2006) Determinants of successful entrepreneurship: a review of the recent literature. Working paper. Available at: https://ssrn.com/abstract=1244663 (Accessed: 1 September 2026).

Statistisches Bundesamt (2026) Insolvency Proceedings: long-term series (table lrs 01). Wiesbaden: Statistisches Bundesamt. Available at:

https://www.destatis.de/EN/Themes/Economic-Sectors-Enterprises/Enterprises/Business-Notifications-Insolvencies/Tables/lrins01.html (Accessed: 1 September 2026).
Trading Economics (2026) Germany bankruptcies. Available at: https://tradingeconomics.com/germany/bankruptcies (Accessed: 1 September 2026).
Vantilborgh, T., Joly, J. and Pepermans, R. (2015) 'Explaining entrepreneurial status and success from personality: an individual-level application of the entrepreneurial orientation framework', Psychologica Belgica, 55(1), pp. 32–56. doi: 10.5334/pb.be.
World Bank (2020) Doing business 2020: comparing business regulation in 190 economies. Washington, DC: World Bank. Available at: http://hdl.handle.net/10986/32436 (Accessed: 1 September 2026).
World Bank (2025) New business density (new registrations per 1,000 people ages 15–64), World Bank Entrepreneurship Database. Available at: https://data.worldbank.org/indicator/IC.BUS.NDNS.ZS (Accessed: 1 September 2026).
World Economics (2026) Germany GDP annual growth rate. Available at: https://www.worldeconomics.com/GrossDomesticProduct/GDP-Annual-Growth-Rate/Germany.aspx (Accessed: 1 September 2026).

# Consumer Preferences Towards Emerging Digital Tools such as AR, VR & 3D in Real Estate Sales & Marketing By Manya Juvvadi [1], Dr. T P Singh[2]

## Introduction

The real estate business has experienced tremendous change in the course of the last decade which can be attributed to the booming growth and application of emerging technologies. The sector has traditionally relied on face-to-face interactions, visits to physical properties, and manual records, but it is currently being transformed by the increasing use of digital platforms, data-driven applications, and intelligent systems that change the nature of property marketing and consumption. Such technological innovations have not only made the operational processes much easier, but have also completely transformed the consumer expectations, tastes and preferences as well as the decision-making patterns in the real estate market.

The advent of Property Technology (PropTech) has brought some new technology designs like virtual tours or online listing websites, artificial intelligence (AI), big data analytics, blockchains transactions, and smart buildings. These instruments will allow accessing information about property faster, increase the correctness of prices, transparency, and safety in transactions. Consequently, consumers are ceasing to be passive recipients of selective information as delivered by the agents, i.e., consumers can now read and compare alternatives in real time and make more informed buying decisions as they read through the digital resources. Access to information about existing properties, foresight about the market, and immersive experiences of viewing have made consumer interests shift towards convenience, speed, personalization, and transparency.

Digital communication mediums have altered the communication between buyers, sellers, and real estate professionals. Using online platforms and mobile applications, information asymmetry is minimized and enhances trust in transactions of property. Smart technologies and data-driven management systems have an additional impact on the preferences of tenants in commercial real estate by improving the efficiency of operations, sustainability, and the overall experience of the user. This has proved that technology is not just a supportive tool but a strategic driver that is changing the market forces and consumer behaviour.

As the trend towards technology use in real estate marketing is progressively increasing, there is a need to study the impact of the developments on consumer preferences. This is because knowing how to navigate this relationship is essential to real estate firms with the aim of staying competitive in a growing digital world. The purpose of this research paper is to discuss the role of new technologies and their influence on consumer preferences in the real estate marketing industry and how digital transformation affects the expectations of buyers, their patterns of engagement, and the decision-making process in modern property markets.

As real estate has traditionally been a business built on personal visits, face-to-face talks,

[1] Class 12, Oakridge International School, Gachibowli, Hyderabad
[1] manya.juvvadi@gmail.com
[2] Research Mentor: Editor-in-Chief, Applied Researches,
[2] tpsingh@live.in

and trust between the buyer and the seller, a buyer would usually visit many properties before choosing one, and this process would often take a long time and cost a lot of effort. In the last few years, this old way of buying and selling property is slowly changing because of new digital tools. Among these tools, Augmented Reality (AR), Virtual Reality (VR) and 3D visualization have started to play an important part in how properties are shown and understood by buyers. These tools allow a buyer to see a property without actually being there in person. A VR tour can give a buyer a feeling of walking through a house from their own home. An AR tool can let a buyer place furniture inside a room to see how it would look. A 3D model can show the full layout and size of a property much better than flat photographs. Real estate companies and agents are now using these tools to attract buyers, save time, and make the property search process easier and more interesting.

However, even though these tools are becoming more common, it is not yet fully clear how much buyers actually trust and accept them. Some buyers may feel comfortable and excited to use AR, VR and 3D tools, while others may still prefer the old way of visiting a property in person. There may also be worries about data privacy, since these tools often need access to a device's camera or personal information. It becomes important, then, to study how consumers actually feel about these tools before real estate companies invest further money and effort into them.

This paper studies consumer preference towards AR, VR and 3D technology in real estate sales. It looks closely at how comfortable people are with this technology, how useful they find it, and how much they trust it while making a real property decision. The findings of this paper can help real estate companies, developers and technology providers understand where the market currently stands, and what steps can be taken to build stronger trust and wider use of this growing technology.

## Literature Review

### *Growth of PropTech and Digital Change in Real Estate*

Although the Real Estate industry traditionally believed in face-to-face interactions, some of the recent literature indicates that awareness of digital real estate technologies is increasing, however, the actual adoption still remains uneven. A review of the Indian real estate context found that developers have begun using AR and VR to provide immersive property-viewing experiences, but consumer adoption remains limited and is concentrated primarily in higher-end residential and commercial segments. The study identified technological literacy, implementation cost, cultural perceptions, and consumer readiness as major barriers to broader adoption (Ramesh & Singh, 2024). This suggests that the availability of digital tools alone does not guarantee their acceptance by end users.

Similarly, research on VR-based real estate applications suggests that the technology can improve property presentation and customer engagement, particularly where buyers need to view properties remotely or before construction is completed. VR enables users to explore

computer-generated environments at a realistic scale and may permit them to observe design alternatives, layouts, furniture arrangements, and interior elements before making a purchase decision (Yin et al., 2017). Such capabilities are particularly valuable in off-plan housing sales, where buyers cannot physically inspect the finished property.

The broader PropTech literature also indicates that adoption is shaped by the interaction of technological, organisational, and consumer-level factors. From a technology acceptance perspective, users are more likely to adopt a system when they perceive it as useful, easy to use, enjoyable, reliable, and compatible with their needs. In immersive environments, these conventional acceptance variables are supplemented by experiential dimensions, including presence, immersion, visual realism, perceived control, and emotional engagement. Recent research on AR acceptance similarly notes that enjoyment, immersion, aesthetics, and social interactivity can be as influential as traditional variables such as perceived usefulness and effort expectancy (Carrión et al., 2021; Castillo & Bigné, 2021; Jung et al., 2020).

Trust is also becoming increasingly important in digital property transactions. AR and VR systems may capture movement data, visual data, location information, facial features, and behavioural interaction patterns. Therefore, prospective users evaluate not only the benefits of immersive property tools but also the accuracy, reliability, privacy, and ethical use of the information provided. Trust in the platform, developer, agent, and technological system can consequently influence satisfaction and continued usage intention (Alam et al., 2021).

***Prior Exposure and Comfort with Technology***

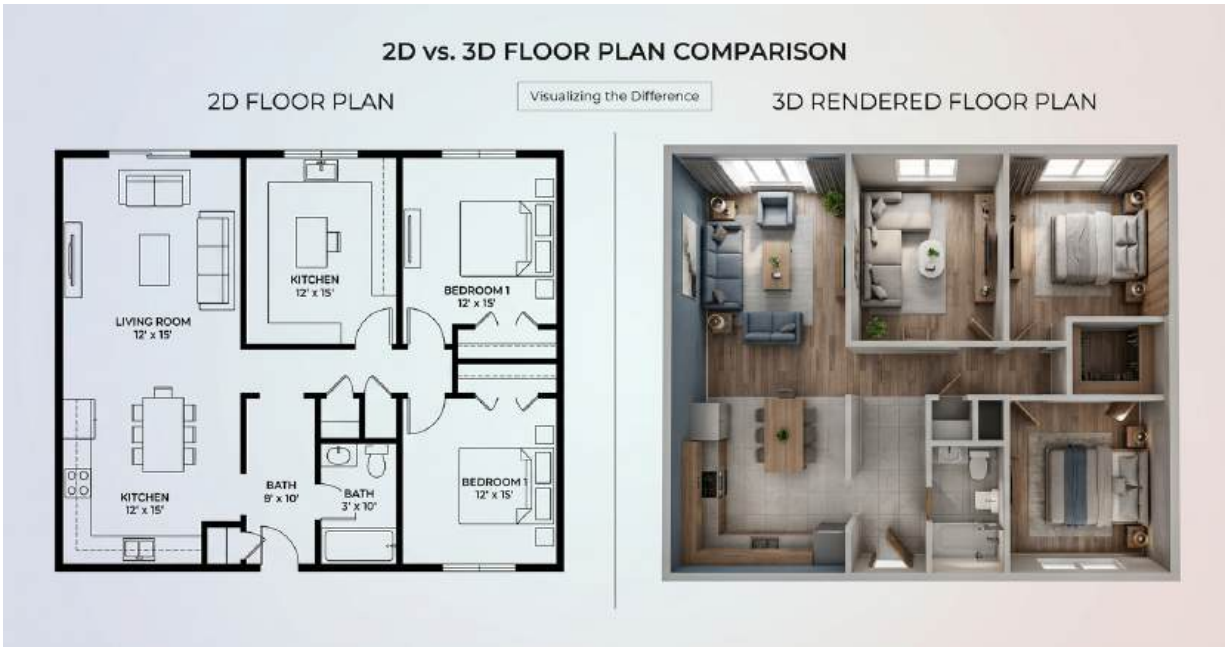
Prior exposure may be particularly important in the Indian real estate context, where access to smartphones is widespread but familiarity with dedicated VR headsets and advanced AR interfaces may vary substantially across age groups, income levels, educational backgrounds, and locations. A recent review of AR and VR adoption in Indian real estate noted that consumer readiness and technological literacy remain important barriers, despite increasing developer interest in immersive property marketing (Ramesh & Singh, 2024).

Studies on consumer intention to use VR in the real estate industry similarly identifies perceived ease of use as an important determinant of adoption. Consumers are more likely to use VR in the home-buying process when they perceive the technology as useful, easy to operate, and enjoyable. In particular, ease of use can reduce the cognitive effort needed to explore a virtual property and can make the experience more accessible to individuals with limited prior technical exposure (Malik & Annuar, 2019).

The importance of technological comfort is also supported by research on VR applications in housing and design. In a study of potential house buyers, participants generally viewed VR systems as useful and expressed positive attitudes and behavioural intentions toward the technology. Nevertheless, the study concluded that perceived ease of use remained an area requiring improvement, indicating that users may recognise the benefits of VR while still experiencing operational difficulties (Yusof et al., 2022).

***Perceived Usefulness and Immersive Experience***

Perceived usefulness is a foundational construct in technology acceptance research. It refers to the extent to which users believe that a technology improves their task performance, increases efficiency, saves time, or assists in making better decisions. Three-dimensional property visualisation appears particularly valuable because conventional two-dimensional floor plans require buyers to mentally translate symbols and measurements into a realistic understanding of space.




Source: www.le-studio3d.com, www.biorev.com & weaverbird.co.in

Many buyers may find this difficult, especially when evaluating properties under construction. A study conducted at the University of Twente found that 3D floor plans produced higher purchase intention than 2D floor plans. The mean purchase intention for 2D plans was

3.38, compared with 3.76 for 3D plans. The highest purchase intention was observed for a 3D floor plan with additional design elements, which recorded a mean score of 4.00 (Chrześcijańska, 2024). These findings suggest that 3D presentation can enhance the attractiveness and perceived comprehensibility of a property.

Similarly, Wang et al. (2018) found that VR technology improved prospective buyers' understanding of pre-sale housing projects. Participants perceived VR as useful because it helped them understand the project more clearly and increased their intention to purchase. This is particularly relevant in property markets where buyers often make decisions based on incomplete construction, brochures, model apartments, or conceptual designs.

A study of VR property-viewing preferences found that potential buyers preferred VR show-house experiences over brochures, social media, and 3D mock-up models because VR was perceived as more visually rich and lifelike (Rahman et al., 2023). This supports the proposition that immersive visualisation can generate stronger engagement than traditional real estate marketing materials.

***Trust, Adoption Intention, and Value Perception***

Trust is particularly important in real estate because property transactions are high-involvement decisions involving substantial financial, legal, and emotional consequences. Buyers may use AR, VR, and 3D tools to form early impressions, compare alternatives, and shortlist properties. However, they may hesitate to rely on immersive visuals if they believe that room dimensions, finishes, views, lighting, neighbourhood conditions, or other property features are being exaggerated or manipulated.

Evidence suggests that immersive technologies can improve buyer confidence when they are perceived as credible and informative. Tan's (2018) study of homebuyers found that VR had a positive influence on purchase experience and purchase intention. The study identified confidence and trust toward the property developer, assistance in evaluating alternatives, and improved purchase experience as the three main mechanisms through which VR accelerated property decision-making.

A recent study examining AR and VR in relation to real estate investment decision-making identified four latent constructs: Confident Value Perception, Innovative Investment Appeal, Trusted Property Transactions, and Effortless Property Engagement. Among these, Trusted Property Transactions represented a distinct factor influencing investor decision-making and property purchase experience (Kaur et al., 2025). This finding is closely aligned with the present study's trust-and-adoption factor, which includes confidence in agents, virtual property information, and the broader use of immersive tools in property transactions.

Most of the notified empirical studies have been conducted in markets outside India, including Singapore, Malaysia, Taiwan, the Netherlands, Nigeria, and other developed or emerging economies. While these studies provide useful insights, consumer attitudes toward property technology may differ substantially across national and regional contexts. Differences in digital literacy, income, housing affordability, real estate regulation, urban development

patterns, cultural attitudes, and trust in developers may affect the adoption of AR, VR, and 3D property tools.

Thus, the present study therefore seeks to contribute to the emerging PropTech literature by examining AR, VR, and 3D technologies together within the Indian real estate context. It focuses on prospective users' technological comfort, perceived usefulness, immersive experience, trust, privacy concerns, and intention to adopt these tools. The resulting three-factor structure provides an empirically grounded framework for understanding how consumers evaluate immersive property technologies in a developing and rapidly digitising real estate market.

***Industrial applications and outcomes***

The industrial use cases of AR, VR and 3D technologies from Indian Real Estate suggest that these techs have emerged as a disruptive catalyst in transforming traditional property marketing and asset transactional frameworks.

A primary example of this paradigm shift is observed in the strategies deployed by the Prestige Group for their high-volume residential developments in Hyderabad. The firm faced the critical challenge of securing substantial inventory bookings during the high-risk pre-construction phase, long before physical site infrastructure could establish buyer confidence. To mitigate this issue, the developer deployed virtual reality (VR) property showrooms within centralized sales galleries, enabling local and remote buyers to walk through photorealistic digital twins of the properties. The deployment of these immersive environments yielded a profound commercial impact, allowing the Prestige Group to successfully sell 35% of the total project inventory prior to the commencement of physical construction, a milestone that traditionally required eight to nine months of active construction progress to achieve (banyantreerealty.in).

Similarly, Mahindra Lifespaces addressed structural inefficiencies within both the pre-sales and post-sales consumer journeys by migrating toward digital-first engagement platforms. The developer sought a mechanism to streamline online buyer conversions and mitigate post-sale friction or customer dissatisfaction regarding space perception during the final apartment handovers. Mahindra Lifespaces' embedded mobile-responsive augmented reality (AR) interactive walkthroughs resolved spatial convenience bottlenecks, resulting in sustained remote sales, improved customer retention, and a significantly shortened overall transaction cycle (theintechgroup.com).

Furthermore, the suburban market dynamics encountered by the Rutu Group during their Rutu City project highlight the utility of spatial tech in product differentiation. Operating in a highly competitive residential corridor, the developer faced the challenge of establishing an emotional connection with first-time home buyers and nuclear families who struggled to visualize lifestyle experiences from sterile technical plans. The group utilized a tailored suite of photorealistic 3D exterior renderings and narrative-driven 4-minute cinematic 3D interior walkthroughs that embedded active lifestyle directly into the visual data. Deployed across digital

marketing channels and interactive experience centers, this spatial asset strategy substantially increased consumer click-through rates and directly converted 8% of the entire project inventory during the pre-launch phase alone (weaverbird.co.in).

Collectively, these empirical use cases demonstrate that technologies such as AR, VR, 3D and other spatial computing tools serve as a core driver of financial efficiency and market resilience within contemporary Indian real estate operations.

## Research Methodology

This study uses a quantitative research design and is descriptive and exploratory in nature. It tries to describe the consumer preferences towards emerging digital tools, and it is exploratory because it uses Exploratory Factor Analysis to find out if there are any hidden groups or patterns sitting behind the individual questions.

A structured questionnaire was prepared with twenty questions, each written as a simple statement that a respondent could agree or disagree with. Every question was answered on a five-point scale, where 1 means "Strongly Disagree" and 5 means "Strongly Agree." Data for this study was collected from 62 respondents through an online survey form. Respondents were reached through convenience sampling.

After collection of data, it was cleaned and checked for errors, and no missing values were found. The following statistical steps were then carried out on this data:

- Descriptive statistics- mean, median, mode and standard deviation, were calculated for each question, to understand the average opinion and how much that opinion varied among respondents.
- The Kaiser-Meyer-Olkin (KMO) test and Bartlett's Test of Sphericity- to check if the data was fit for factor analysis. Both tests must give a good result before moving further, otherwise the factor analysis result cannot be trusted.
- Exploratory Factor Analysis (EFA) was then carried out using the Principal Component Method with Varimax rotation, to find out if the twenty questions could be grouped into a smaller number of hidden factors.

## Result Analysis and Interpretation

### *Descriptive Statistics*

| Code | Question (abridged) | Mean | Median | Mode | Std. Dev. |
|---|---|---|---|---|---|
| Q1 | I have previously experienced a virtual property tour (VR/360° video/3D walkthrough) before visiting a property in person. | 2.34 | 2.0 | 1 | 1.51 |
| Q2 | I am comfortable using digital devices (smartphone, VR headset, laptop) to explore property listings with immersive technology. | 1.73 | 1.0 | 1 | 1.18 |

| | | | | | |
|---|---|---|---|---|---|
| Q3 | Using VR/3D virtual tours helps me understand the spatial layout and size of a property more clearly than photographs. | 1.69 | 1.0 | 1 | 1.11 |
| Q4 | AR tools that allow me to visualize furniture and interior décor in a property help me assess whether the space suits my needs. | 1.74 | 1.0 | 1 | 1.04 |
| Q5 | Using immersive technology (AR/VR/3D) saves me significant time by reducing the number of physical property visits required. | 1.95 | 1.5 | 1 | 1.18 |
| Q6 | I find AR/VR/3D property tools easy to navigate and operate without technical expertise. | 2.26 | 2.0 | 2 | 1.12 |
| Q7 | Real estate platforms that offer VR/3D virtual tours are easier to use than traditional property listing websites. | 2.02 | 2.0 | 1 | 1.03 |
| Q8 | Experiencing a property through VR/3D technology increases my confidence in making a purchase decision. | 2.23 | 2.0 | 1 | 1.22 |
| Q9 | Real estate agents who use AR/VR/3D tools appear more professional and trustworthy to me. | 2.50 | 2.5 | 3 | 1.18 |
| Q10 | Virtual staging (digitally furnished interiors shown via AR/3D) positively influences my interest in purchasing a property. | 2.31 | 2.0 | 1 | 1.15 |
| Q11 | A VR/3D virtual property tour gives me a sense of "being there" that photographs or videos cannot replicate. | 2.07 | 2.0 | 1 | 1.16 |
| Q12 | I feel more connected to a property when I experience it through an immersive VR tour compared to static images. | 2.11 | 2.0 | 1 | 1.20 |
| Q13 | AR-enabled features (e.g., overlaying property details, neighbourhood data, or pricing on a live camera view) enhance my property search experience. | 1.92 | 2.0 | 1 | 1.06 |
| Q14 | Given the option, I prefer real estate agents or developers who offer VR/3D tools over those who do not. | 2.36 | 2.0 | 1 | 1.22 |
| Q15 | I would actively seek out properties listed with virtual tours or 3D walkthroughs over those with only photographs. | 2.23 | 2.0 | 1 | 1.26 |
| Q16 | I believe AR/VR/3D technology will become a standard and essential part of real estate sales in the future. | 1.81 | 1.0 | 1 | 1.21 |
| Q17 | I trust that virtual tours and 3D models provided by real estate companies accurately and honestly represent the actual property. | 2.52 | 2.5 | 3 | 1.28 |
| Q18 | I have concerns about data privacy when using AR/VR platforms provided by real estate companies (e.g., data collected through a VR headset or AR app). | 2.66 | 3.0 | 3 | 1.29 |
| Q19 | The availability of AR/VR/3D tools by a real estate developer or agent positively influences my overall brand perception of that company. | 2.08 | 2.0 | 1 | 1.21 |
| Q20 | Overall, I am satisfied with the use of AR/VR/3D technology as a medium for exploring and evaluating real estate properties before purchase | 1.89 | 2.0 | 1 | 1.01 |

The above table presents descriptive statistics for 20 questionnaire items measuring respondents' perceptions of augmented reality (AR), virtual reality (VR), and 3D technology in real estate. The statistics include mean, median, mode, standard deviation, minimum, and maximum scores. All items have responses ranging from 1 to 5, indicating that the full response scale was used. However, most mean scores fall between 1.69 and 2.66, and the mode for most items is 1. This pattern indicates generally low endorsement of AR/VR/3D technology across the sample, assuming that higher values represent stronger agreement with the statements.

The first set of items relates to prior exposure and technological comfort. Q1, concerning previous experience with virtual property tours, has a mean of 2.34 and the highest standard deviation of 1.51. This suggests limited overall exposure but substantial variation among respondents: while many have not used virtual tours, some have experience with them. Q2 records a low mean of 1.73, indicating limited comfort with devices such as smartphones and VR headsets. Similarly, Q6, which measures ease of navigating AR/VR/3D property tools without assistance, has a mean of 2.26. These findings suggest that technological familiarity may be a major barrier to adoption. Respondents may need simple interfaces, demonstrations, and guided assistance before they can use immersive property technologies confidently.

The second group concerns perceived usefulness and immersive experience. Q3, which assesses whether VR/3D tours improve understanding of property layout, has the lowest mean in the table, at 1.69. Q4, related to using AR to visualise furniture and interior décor, also has a low mean of 1.74. These results suggest that respondents do not yet strongly perceive immersive tools as useful for understanding spatial design or visualising interiors. Q5, measuring whether AR/VR/3D saves time, has a mean of 1.95, while Q7, measuring whether immersive property platforms are easier than conventional listing websites, has a mean of 2.02. Thus, respondents do not clearly perceive AR/VR/3D technology as more convenient or efficient than traditional online property-search platforms. The immersion-related items show similarly modest results. Q11, measuring the feeling of "being there" during a virtual property tour, records a mean of 2.07. Q12, measuring connection with a property after an immersive experience, has a mean of 2.11. Q13, which concerns AR-enabled information overlays such as property and neighbourhood details, has a mean of 1.92. Together, these results suggest that respondents may not yet experience AR/VR/3D tools as sufficiently realistic, engaging, or useful to create strong emotional connection with a property.

The third group concerns trust, adoption intention, privacy, and satisfaction. Q9, regarding whether agents using AR/VR/3D appear more professional, has a relatively higher mean of 2.50 and a mode of 3. This indicates a comparatively neutral-to-positive perception of technology-enabled agents. Q17, measuring trust in the accuracy of virtual tours and 3D models, has a mean of 2.52 and a mode of 3, suggesting cautious neutrality rather than strong trust or distrust. Privacy concerns are the strongest issue in the table. Q18 has the highest mean of 2.66, with both median and mode equal to 3. This indicates that respondents are relatively more concerned about data privacy than about other aspects of AR/VR/3D adoption. Overall satisfaction with immersive technology is low, as shown by Q20's mean of 1.89. Moreover, Q16

indicates that respondents do not strongly believe AR/VR/3D will become a standard feature of real estate, with a mean of 1.81.

Overall, the findings suggest that AR/VR/3D adoption in real estate is still at an early stage among the respondents. Greater awareness, user-friendly design, trustworthy information, transparent virtual staging, improved privacy protection, and integration with physical property visits may be necessary to improve acceptance and satisfaction

***KMO and Bartlett's Test***

| **KMO Measure of Sampling Adequacy & Bartlett's Test of Sphericity** | |
|---|---|
| Overall KMO Measure of Sampling Adequacy | 0.909 |
| Bartlett's Test — Approx. Chi-Square | 1022.77 |
| Bartlett's Test — Degrees of Freedom (df) | 190 |
| Bartlett's Test — Sig. (p-value) | < 0.001 |

Before running the factor analysis, it is important to check if the data is fit for this kind of test. The Kaiser-Meyer-Olkin (KMO) value came out at 0.909. Any value above 0.90 is seen as "excellent" for factor analysis, so this data is very well suited for finding hidden factors. The Bartlett's Test of Sphericity also gave a strong result (chi-square value of 1022.77, with 190 degrees of freedom, and p value less than 0.001). This means the 20 questions are related to each other in a meaningful way, and they are not just random numbers. Together, these two tests confirm that it was correct and safe to go ahead with the Exploratory Factor Analysis.

***Exploratory Factor Analysis — Total Variance Explained (Eigenvalues, PCA extraction)***

| **Component** | **Eigenvalue** | **% of Variance** | **Cumulative %** |
|---|---|---|---|
| 1 | 11.048 | 55.24 | 55.24 |
| 2 | 1.683 | 8.41 | 63.65 |
| 3 | 1.186 | 5.93 | 69.58 |

***Rotated Factor Loadings (Varimax Rotation, 3-Factor Solution, PCA Extraction)***

| **Code** | **Factor 1** | **Factor 2** | **Factor 3** |
|---|---|---|---|
| **Q1** | | | **0.813** |
| **Q2** | | | **0.704** |
| **Q3** | | **0.738** | |

| | | | |
|---|---|---|---|
| **Q4** | | **0.806** | |
| **Q5** | | **0.591** | |
| **Q6** | | | **0.616** |
| **Q7** | | **0.587** | |
| **Q8** | **0.632** | | |
| **Q9** | **0.788** | | |
| **Q10** | **0.643** | | |
| **Q11** | | **0.819** | |
| **Q12** | | **0.649** | |
| **Q13** | | **0.782** | |
| **Q14** | **0.634** | | |
| **Q15** | **0.694** | | |
| **Q16** | **0.572** | | |
| **Q17** | **0.765** | | |
| **Q18** | **0.700** | | |
| **Q19** | **0.728** | | |
| **Q20** | **0.512** | | |

Principal Component Analysis with Varimax rotation extracted 3 factors with eigenvalues greater than 1 (11.05, 1.68, and 1.19 respectively), consistent with the scree plot 'elbow' after the first component, and together these 3 factors explain approximately 69.6% of total variance — a strong result for social-science survey data.

Factor 1 (Trust, Adoption Intention & Value Perception) loads most strongly on items Q8-Q10 and Q14-Q20 (confidence, agent trustworthiness, virtual staging influence, preference for VR-enabled agents, future standardisation, trust in accuracy, privacy concern, and satisfaction). Factor 2 (Perceived Usefulness & Immersive Experience) loads on Q3-Q5, Q7, and Q11-Q13 (spatial clarity, AR furniture visualization, time savings, ease-of-use vs. traditional listings, sense of presence, emotional connection, and AR-overlay features). Factor 3 (Prior Exposure & Technology Comfort) loads on Q1, Q2 and Q6 (past VR experience, device comfort, and ease of navigation). This three-factor structure aligns conceptually with an extended Technology Acceptance Model (TAM) framing: prior exposure/comfort as an antecedent, perceived usefulness/experience as a mediating belief, and trust/adoption intention as the outcome construct.

**Naming of the factors**

**Factor 1 – Trust, Adoption Intention and Value Perception:** This factor brought together questions about confidence in buying, trust in real estate agents who use these tools, virtual staging, preference for agents who offer VR tools, belief that this technology will become common in future, trust in the accuracy of virtual tours, privacy worries, and overall satisfaction. In simple words, this factor is about whether buyers trust the technology enough to actually use it in their buying decision.

**Factor 2 – Perceived Usefulness and Immersive Experience:** This factor grouped questions about understanding the size and layout of a property, seeing furniture through AR, saving time, ease of use compared to normal listing websites, the feeling of "being there," feeling connected to the property, and AR features that show extra details. This factor is about whether the technology is actually useful and gives a good, real feeling experience.

**Factor 3 – Prior Exposure and Technology Comfort:** This factor is made up of only three questions, about earlier experience with virtual tours, comfort with using digital devices, and ease of moving around inside the AR/VR tool. This factor is small but important, as it shows the starting point of a buyer, that is, how much they already know and how comfortable they already are with such tools.

**Conclusion**

This study set out to understand how consumers feel about the use of AR, VR and 3D technology in real estate sales, and to check if these feelings can be grouped into clear factors. Based on the responses of 62 participants, it is clear that most buyers are still cautious about this technology. The mean scores for almost every question stayed below the middle of the scale, showing that trust and full acceptance have not yet been built among the wider public, even though the technology itself is available and growing in the market.

At the same time, the statistical tests give confidence that the data is sound. The KMO value of 0.909 and the significant Bartlett's Test show that the 20 questions used in this study measure something real and connected, not random noise. The Exploratory Factor Analysis further confirms that consumer opinion on this topic is not scattered but can be explained through three clear and meaningful factors: prior exposure and comfort with technology, perceived usefulness of the immersive experience, and trust with intention to adopt the technology for real buying decisions.

Thus, the study concludes that AR, VR and 3D adoption in real estate is not simply about giving buyers a new gadget to use. It is a step-by-step journey that starts with comfort, moves through usefulness, and ends with trust. Real estate companies that want buyers to actually use and value this technology need to work on all three stages together, and cannot expect trust to grow only by improving the technical features of the tool.

**Recommendations**

Based on the findings of this study, the following recommendations are put forward for real estate companies, developers, technology providers and future researchers:

- Real estate firms should first focus on building basic comfort among buyers, such as offering simple training or short demo videos, before pushing buyers straight into full AR/VR experiences, since many buyers still show low comfort with the underlying technology.

- Developers of AR/VR/3D property tools should make the tools very easy to use, with clear buttons and simple steps, so that even a first-time user with no technical background can move through a virtual tour without confusion.
- Companies should be open and honest about how virtual tours and 3D models are made, and should avoid over-decorating or misrepresenting a property through virtual staging, since trust in the accuracy of these tools directly affects whether buyers will use them for real decisions.
- Firms should clearly explain to buyers how their personal data is collected, stored and used when they use AR/VR platforms, as data privacy concerns are closely linked to overall trust in the technology.
- Real estate agents should be trained to use AR/VR/3D tools confidently, since agents who use this technology are seen by buyers as more professional and trustworthy.
- Marketing teams should target early adopters first, using them as reference points and word-of-mouth promoters, while slowly building comfort among the more cautious majority of buyers.
- Future studies should use Confirmatory Factor Analysis (CFA) and reliability tests such as Cronbach's Alpha to check and confirm the three-factor structure found in this study before it is used in larger models such as Structural Equation Modeling (SEM).

**Limitations**

Like any research study, this study also has some limits that should be kept in mind while reading the results.

- The sample size of 62 respondents is small, so the results may not fully represent the views of all property buyers in the wider market.
- The data was collected from a limited group of respondents, and the study did not separately study different age groups, income levels or locations, which could have shown different patterns of technology acceptance.
- The study is based only on self-reported answers from a questionnaire, so it depends on how honestly and correctly each respondent understood and answered each question.
- The data was collected at one point in time, so it only shows a picture of consumer opinion at that particular time and cannot show how opinions may change over a longer period.
- Since AR, VR and 3D technology is still new and fast-changing, the views collected in this study may shift quickly as more people start using such tools and as the tools themselves improve.
- The study used Exploratory Factor Analysis, which is good for finding hidden patterns for the first time, but it does not confirm or test these patterns statistically; that would require a Confirmatory Factor Analysis in future work.

**Works Cited**


Abdullahi, M., Yusuf, A., Ibrahim, S., & Bello, M. (2025). Impact of immersive technologies on consumer confidence and trust in real estate transactions. Bayero Journal of Administrative and Environmental Studies, 16, 1–15.

Abu Bakar, N. A., Rahman, N. A., & Ismail, S. (2023). The influence of augmented reality on purchase intention in the housing industry. Human Behavior and Emerging Technologies, 2023, Article 5728088. https://doi.org/10.1155/2023/5728088

Alam, S. S., Susmit, S., Lin, C. Y., & Ahsan, M. N. (2021). Does augmented reality influence customer satisfaction and behavioural intention? The role of trust in augmented reality applications. Journal of Hospitality and Tourism Technology, 12(4), 1–18.

Carrión, G. C., Noguera, J. M., & Minguillón, J. (2021). Augmented reality in consumer contexts: A review of adoption drivers and behavioural outcomes. International Journal of Human–Computer Interaction, 37(12), 1121–1138.

Castillo, S., & Bigné, E. (2021). A model of adoption of augmented reality in the retail environment. Journal of Retailing and Consumer Services, 58, Article 102222.

Chrześcijańska, A. (2024). Properly about property floor plans: The effect of 2D and 3D floor plans on purchase intention [Bachelor's thesis, University of Twente]. University of Twente Student Theses Repository.

Jung, T., tom Dieck, M. C., Lee, H., & Chung, N. (2020). Effects of virtual reality and augmented reality on visitor experiences in tourism: A review. Tourism Management Perspectives, 33, Article 100623.

Kaur, P., Singh, R., & Sharma, V. (2025). Influence of augmented reality and virtual reality on real estate investment decision-making and property purchase experience. International Journal of Housing Markets and Analysis. Advance online publication.

Malik, M. I., & Annuar, S. N. S. (2019). Factors influencing consumer intention to use virtual reality in the real estate industry. International Journal of Academic Research in Business and Social Sciences, 9(11), 1220–1234.

Rahman, M. A., Ahmad, N., & Rahim, M. A. (2023). Consumer preference for virtual reality show houses in residential property marketing. Journal of Advanced Research in Applied Sciences and Engineering Technology, 29(2), 126–134.

Ramesh, P., & Singh, A. (2024). The impact of augmented reality and virtual reality on the Indian real estate sector: A review of adoption, consumer behaviour, and market implications. International Research Journal of Modernization in Engineering Technology and Science, 6(2), 1–8.

Tan, K. H. (2018). The impact of virtual reality on homebuyers' purchase experience and purchase intention [Master's thesis, National University of Singapore]. NUS ScholarBank.

Wang, C. H., Chiang, Y. H., & Hsu, C. Y. (2018). Developing and evaluating a virtual reality-based navigation system for pre-sale housing sales. Applied Sciences, 8(6), Article 952. https://doi.org/10.3390/app8060952

Yin, C., Zhang, X., & Wang, J. (2017). A low-cost and lightweight 3D interactive real estate-purposed indoor virtual reality application. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, IV-4/W4, 307–314. https://doi.org/10.5194/isprs-annals-IV-4-W4-307-2017

Yusof, N. A., Mahmud, M., & Yusof, S. A. M. (2022). Virtual reality in residential property marketing: Perceived usefulness, ease of use, attitude, and behavioural intention. International Journal of Real Estate Studies, 16(2), 75–85.

**Knowledge Alone Is Not Enough: Behavioral and Motivational Predictors of Physical Activity Among Chinese Adolescents and Young Adults By Ziyue Ren**


## Abstract

This study asks what predicts Chinese young adults' weekly hours of moderate-to-vigorous physical activity (MVPA), and specifically whether knowledge of physical activity's health benefits translates into greater participation. I use an anonymous online survey of 93 high school and young adult respondents conducted via Tencent. I found that knowledge of physical activity's benefits for general health, cardiovascular disease, or sleep has no statistically significant relationship with activity levels. Instead, the strongest positive predictors are satisfaction with one's current activity level and intention to improve health in the next three months, both significant at the 5% level. Demographic factors (age, gender), boarding status, varsity participation, and parental support are all statistically insignificant. This suggests that physical inactivity among Chinese youth is not an information problem but a motivational and time-allocation problem, consistent with present bias: students know exercise is beneficial but prioritize immediate rewards. Rather than health education alone, policymakers and schools should focus on making exercise intrinsically enjoyable and building sustained motivation. This would be vital for China's future labor productivity as a country with a rapidly aging population.



**Keywords:** Physical Activity, Moderate-to-Vigorous Physical Activity (MVPA), Adolescents, Behavioral Economics, Chinese Youth, Decision-making


## Introduction

A healthy labor market depends not only on how many people are willing to work, but also on how productively they can work. Physical activity, often viewed as an indicator of personal health, can therefore have broader economic consequences by influencing workers' productivity, participation, and ultimately the supply of labour. However, physical activity levels do not meet standards. Around 27.5% of adults and 81% of adolescents worldwide did not meet the recommended level of physical activity in 2016 (World Health Organization [WHO]). Low physical activity may lead to health issues, such as obesity and heart disease (MedlinePlus). As a result, physical inactivity is one of the leading contributors to global mortality, with an estimated 4-5 million deaths that could be prevented if the global population were more physically active (WHO).

The issue of shortage in labour supply is critical in aging societies, especially those in East Asia, such as Japan, South Korea, and China (Gaudiaut). As productivity in an economy decreases due to a shortage of labour caused by population aging, there would be less economic activity, followed by higher taxes to raise government revenue and lower pensions due to a restricted budget. As the aging population is slowly becoming a major issue in China, significant economic and social challenges are emerging before the country reaches the development level of many high-income countries. China is currently classified as a developing nation, with GDP

per capita ranked 73rd worldwide. Thus, labour productivity must be high for China to become a high-income country before the aging population takes over productivity. As fertility rates decline and life expectancy rises due to the previous one-child policy China has set, the proportion of elderly citizens increases while the relevant size of the working-age population shrinks. This creates challenges for maintaining labour supply and supporting the social protection system (Maags et al.). More money will be spent on pensions and healthcare by the government, yet less economic growth will happen due to a shrinking labour force. As a result, in the face of growing challenges due to a declining number of people of working age, the health and efficiency of China's future labour force are becoming ever more important, emphasising the importance of studying variables affecting health behaviour.

A more robust labor supply could be supported by higher levels of physical activity amongst labor market participants. A good state of health contributes to people's ability to engage in study and labour and to their productivity in the future, whereas poor health limits people's ability to work and implies certain costs due to absenteeism and loss of productivity as well as higher healthcare spending (World Bank Group). Thus, physical activity can be seen as an investment in future health and productivity. The transition from adolescence to early adulthood becomes particularly relevant for studying these decisions due to the growing number of academic and social activities of individuals (Hayes et al.). Simultaneously, physical activity patterns formed in this period may influence health and behavior patterns in the future (Miller et al.). Hence, the study of the factors influencing decision-making towards engaging in physical activities becomes important not only from a public health point of view, but also from an economic point of view.

The literature has found that physical inactivity among adolescents is a significant global concern, influenced by a complex interaction of individual knowledge, social pressures, and environmental barriers. According to the Global Status Report on Physical Activity (WHO), over 80% of adolescents across the globe fail to meet the recommended levels of moderate-to-vigorous physical activity (MVPA), which is 60 minutes per day. Moreover, even after being aware of the benefits of physical activity, the majority of adolescents are unable to engage in a sufficient amount of MVPA. Engagement of adolescents in physical activities depends upon their decision-making environment. Even though individuals understand the benefits of physical activities, barriers such as lack of facilities or other social pressures may influence their decision (WHO). According to "The cost of inaction on physical inactivity to public health-care systems: a population-attributable fraction analysis," an attempt to increase physical activity among young people needs to be multifaceted in order to address all these influences on their behavior. The authors also identify limitations in existing research concerning the broader economic consequences of physical inactivity, noting that productivity losses and mortality costs were excluded because of methodological challenges (Santos et al.). Therefore, understanding these behavioral and environmental influences is essential for developing effective strategies to increase adolescents' MVPA engagement.

Individuals place disproportionately greater value on immediate costs and benefits than on those that occur in the future (behavioralecon). For example, individuals may prefer watching short videos on phones over going to the gym since they value the utility derived from the current state over utility that will be gained in the future. In the context of physical activity, the cost of effort and fatigue are experienced immediately, whereas many of its positive health benefits occur in the future. Present bias may therefore cause individuals to prioritise alternatives that bring immediate rewards rather than physical activity, regardless of recognising the long-term health benefits of exercise. Yet learning to do the hard things at a young age is important in the future, since adults would have competing objectives, leading to higher opportunity costs of working out. Thus, it is important to establish a routine of regular physical activity habits from a young age. Understanding those behavioural factors can provide insight into how and why individuals allocate their limited time to physical activity. The Physical Activity Profile 2022 for China shows that amongst adolescents aged 11 to 17, 80% of boys and 89% of girls were classified as physically inactive. This supports the idea that Chinese adolescents are experiencing insufficient MVPA, making it necessary to investigate why they do not engage in enough physical activity (WHO). Furthermore, higher-income individuals are more likely to engage in physical activity, suggesting that participation in exercise is influenced not only by individual preferences but also by economic constraints and available resources (Kari et al.). Therefore, there are several control factors to be explored when studying physical activity behaviours.

Research on adolescents' and young adults' physical health exists globally, yet data on physical health from adolescence to young adulthood in China remain limited due to strict government regulations. This creates a gap that may limit evidence-based policy. Thus, this research aims to identify what influences Chinese young adolescents' decisions to engage in moderate-to-vigorous physical activity (MVPA) and what beliefs and knowledge they have about physical activity. I made an anonymous survey link and shared it with students aged 14 to 20 in China, who may be from higher economic backgrounds. Based on the literature and behavioral economic theory discussed above, I hypothesized that individuals with greater knowledge of the benefits of physical activity will engage in higher levels of MVPA. Understanding this relationship may help identify how knowledge influences adolescents' physical activity decisions and potential interventions to promote healthier behaviors. Adolescence is a period where health behaviors and decision-making habits are established, which would potentially affect physical activity levels in adulthood. In Section 1, I present my method of research. In Section 2, I present my results from the data, including two tables generated by Gretl based on my survey results. In Section 3, I conclude and summarize the key findings, discuss their implications, and identify directions for future research.

**Methods**

I conducted an anonymous survey on Tencent from August 17 to 28, 2026. The link was shared via WeChat, and a total of 19 questions were asked. Participants were students aged

14–20 enrolled in public and private schools in China, recruited through convenience sampling via personal WeChat networks using a Tencent survey link. The survey asked for the participants' consent and assured them of their anonymity. Data were cleaned in Google Sheets and then analyzed in Gretl.

**Equation 1**

$$\begin{aligned}\text{pa_hours_week}_i &= \beta_0 + \beta_1\,\text{age}_i + \beta_2\,\text{male}_i + \beta_3\,\text{boarding}_i + \beta_4\,\text{varsity}_i + \beta_5\,\text{no_school_sport}_i \\ &+ \beta_6\,\text{sch_team}_i + \beta_7\,\text{sch_indiv}_i + \beta_8\,\text{sch_oppon}_i + \beta_9\,\text{out_team}_i \\ &+ \beta_{10}\,\text{out_indiv}_i + \beta_{11}\,\text{study_hours_wknd}_i + \beta_{12}\,\text{sleep_hours_wknd}_i \\ &+ \beta_{13}\,\text{par_support_index}_i + \beta_{14}\,\text{par_instilled}_i + \beta_{15}\,\text{know_exercise}_i \\ &+ \beta_{16}\,\text{know_disease}_i + \beta_{17}\,\text{know_sleep}_i + \beta_{18}\,\text{pa_satisfied}_i \\ &+ \beta_{19}\,\text{intent_improve}_i + \varepsilon_i\end{aligned}$$

I estimated the determinants of weekly physical activity hours ($pa_hours_week_i$) using an ordinary least squares (OLS) regression, where my primary variables of interest are health knowledge indices across three domains—exercise ($know_exercise_i$), disease prevention ( $know_disease_i$), and sleep hygiene ($know_sleep_i$), each constructed from three survey questions—alongside psychological drivers measuring satisfaction with current physical activity ($pa_satisfied_i$) and intention to improve engagement ($intent_improve_i$).

To isolate these effects from potential confounding factors, I incorporated an extensive set of control variables spanning student demographics and living conditions (age, gender, and boarding status), participation profiles in school and extracurriculars, weekend time allocations (study and sleep hours), and home environment characteristics such as parental support and instilled exercise habits, with $\varepsilon_i$ as the error term.

**Results**

My primary data include 93 responses from adolescents and young adults in China, ranging in age from 14 to 20. This includes all genders, with my data consisting of 57% female and 43% male. 50.5% are boarding students, and 49.5% are commuter students. 61.3% participate in at least one varsity sport at school, 21.5% do not participate in any varsity sport at school, and 17.2% claim they do not participate in any sporting activities at school. On average, the individuals participate in 2.98 days of moderate-intensity activity per school week and 2.03 days of vigorous-intensity activity per school week. In a normal school day, the individuals participate in 1.76 hours of moderate-intensity activity and 1.33 hours of vigorous-intensity activity. In total, the individuals participate in 10.87 hours of physical activity per week.

It is worth noting that 57% of responses are from the city of Shanghai, 11% from Jiangsu province, and 32% from other provinces. This indicates that many respondents may be in a large city with comparably easier access to physical activity facilities. Additionally, the

boarding/commuting split creates further variation: boarders live on campus with built-in access to school gym facilities, while commuters navigate travel time and home environments that may either support or hinder physical activity. The responses consist of students who go to both public and private schools, creating a balance between the environments for physical activity.

Most of the responses are from individuals aged 17, accounting for 23.7% of the responses, and the fewest responses are from individuals aged 19, consisting of only 4.3%. It is also worth noting that the mean age in this research is 16.73, and only 10.8% are aged 19 or 20, meaning that it is likely that a random respondent is in high school. Meanwhile, 37% play team sports outside school, 42% play individual sports outside school, and 40% play opponent sports outside school, suggesting that extracurricular activity is substantial but not universal. While 87% know that regular physical activity maintains health and 95% know sleep aids recovery, the average student reports only 2.98 days per week of moderate activity and 2.03 days of vigorous activity, which is well below recommended levels.

The mix implies that most respondents participate in sporting activities in school, as there are only 17.2% who do not. This is relevant since many survey questions are about how much an individual participates in physical activities. The gender balance also helps ensure fair representation, avoiding single-gender samples that may reduce response accuracy. The students have high health knowledge (mean health knowledge index of 2.47 out of 3) and moderate satisfaction with their current activity levels (mean 3.45 out of 5). They also express a reasonably strong intention to improve their health in the next three months (mean 3.78 out of 5). However, they face real barriers: 51% cite academic workload, 48% prefer other activities (phone or friends), 47% are too tired, and 27% say exercise takes time away from studying.

However, there are also limitations in my data. Individuals from small towns in China may be underrepresented, since 57% of responses were from Shanghai, one of the biggest cities in China. There may also be an overrepresentation of varsity members; there may not be enough opportunities for everyone to participate in a varsity sport in public and private schools in China. Additionally, the study relies on self-reported data, which may introduce social desirability bias: students may overstate their activity levels or underreport barriers that feel personal. Thus, the data collected may show a deeper knowledge regarding the external benefits physical activity can bring to individuals.

**Table 1. Descriptive results of the Survey**

| Variables | Min | Mean | Std. dev. | Max |
|---|---|---|---|---|
| Age in years | 14.00 | 16.73 | 1.62 | 20.00 |
| Biological gender | 0.00 | 0.43 | 0.50 | 1.00 |
| Boarding vs commuting student last semester | 0.00 | 0.51 | 0.50 | 1.00 |
| Participates in a varsity sport at school | 0.00 | 0.61 | 0.49 | 1.00 |
| Does no sporting activity at school at all | 0.00 | 0.17 | 0.38 | 1.00 |
| Plays team sports at school | 0.00 | 0.53 | 0.50 | 1.00 |
| Plays individual sports at school | 0.00 | 0.29 | 0.46 | 1.00 |

| | | | | |
|---|---|---|---|---|
| Plays opponent sports at school | 0.00 | 0.33 | 0.47 | 1.00 |
| Days per school week of moderate-intensity activity | 0.00 | 2.98 | 1.79 | 7.00 |
| Days per school week of vigorous-intensity activity | 0.00 | 2.03 | 1.75 | 7.00 |
| Hours per school day of moderate-intensity activity | 0.00 | 1.76 | 1.26 | 6.00 |
| Hours per school day of vigorous-intensity activity | 0.00 | 1.33 | 1.34 | 6.00 |
| Weekly moderate activity hours (constructed) | 0.00 | 6.54 | 7.89 | 42.00 |
| Weekly vigorous activity hours (constructed) | 0.00 | 4.33 | 6.65 | 42.00 |
| Total weekly physical activity hours (constructed) | 0.00 | 10.87 | 13.00 | 84.00 |
| Hours studying outside school on a typical weekend | 0.00 | 4.43 | 2.99 | 10.00 |
| Hours slept on a typical weekend | 0.00 | 7.86 | 1.74 | 10.00 |
| Plays team sports outside school | 0.00 | 0.37 | 0.48 | 1.00 |
| Plays individual sports outside school | 0.00 | 0.42 | 0.50 | 1.00 |
| Plays opponent sports outside school | 0.00 | 0.40 | 0.49 | 1.00 |
| Does no sport outside school | 0.00 | 0.18 | 0.39 | 1.00 |
| How important physical health is to the respondent | 1.00 | 2.71 | 0.67 | 4.00 |
| Knows regular physical activity maintains health | 0.00 | 0.87 | 0.34 | 1.00 |
| Knows activity reduces cardiovascular disease risk | 0.00 | 0.66 | 0.48 | 1.00 |
| Knows sleep aids recovery and function | 0.00 | 0.95 | 0.23 | 1.00 |
| Health knowledge index (constructed) | 0.00 | 2.47 | 0.75 | 3.00 |
| Satisfaction with current physical activity level, reverse-coded so higher = more satisfied | 1.00 | 3.45 | 0.99 | 5.00 |
| Barrier: exercise takes time away from studying | 0.00 | 0.27 | 0.45 | 1.00 |
| Barrier: academic workload | 0.00 | 0.51 | 0.50 | 1.00 |
| Barrier: long commute / occupied | 0.00 | 0.16 | 0.37 | 1.00 |
| Barrier: prefers phone or friends | 0.00 | 0.48 | 0.50 | 1.00 |
| Barrier: does not know how to work out | 0.00 | 0.15 | 0.36 | 1.00 |
| Barrier: too tired | 0.00 | 0.47 | 0.50 | 1.00 |
| Barrier: no friends to work out with | 0.00 | 0.16 | 0.37 | 1.00 |
| Barrier: does not think working out is important | 0.00 | 0.08 | 0.27 | 1.00 |
| Number of barriers selected (constructed) | 1.00 | 2.28 | 0.77 | 3.00 |
| Likelihood of taking active steps to improve health in the next 3 months | 1.00 | 3.78 | 1.03 | 5.00 |
| Parents encourage joining school sports club | 0.00 | 0.71 | 0.46 | 1.00 |
| Parents do not care/do not stop respondent | 0.00 | 0.24 | 0.43 | 1.00 |
| Parents pay for gym or other memberships | 0.00 | 0.39 | 0.49 | 1.00 |
| Parents hire a personal trainer | 0.00 | 0.16 | 0.37 | 1.00 |
| Parents support activity only if grades are good | 0.00 | 0.04 | 0.20 | 1.00 |
| Parents discourage or prevent sporting activity | 0.00 | 0.03 | 0.18 | 1.00 |
| Active parental support index (constructed) | 0.00 | 1.26 | 0.81 | 3.00 |
| Parents tried to instil the value of physical activity | 0.00 | 0.82 | 0.39 | 1.00 |

**Table 2: OLS Regression Results for Weekly Physical Activity Hours**

| Independent Variable | Coefficient (Std. Error) | t-statistic (p-value) |
|---|---|---|
| **Intercept** | | |
| Constant (const) | −40.156** (16.986) | −2.364 (0.021) |
| **Demographics & Living** | | |
| Age (age) | 1.216 (0.899) | 1.353 (0.180) |
| Male (male) | 1.527 (3.201) | 0.477 (0.635) |
| Boarding Student (boarding) | 1.170 (3.011) | 0.389 (0.699) |
| **School Athletics & Extracurriculars** | | |
| Varsity Athlete (varsity) | 2.430 (3.562) | 0.682 (0.497) |
| No School Sports (no_school_sport) | −1.648 (4.883) | −0.338 (0.737) |
| School Team Sports (sch_team) | −0.170 (3.763) | −0.045 (0.964) |
| School Individual Sports (sch_indiv) | −2.038 (3.639) | −0.560 (0.577) |
| School Opponent Sports (sch_oppon) | −1.546 (3.106) | −0.498 (0.620) |
| Outside Team Sports (out_team) | −1.437 (3.850) | −0.373 (0.710) |
| Outside Individual Sports (out_indiv) | −0.173 (3.394) | −0.051 (0.960) |
| **Time Allocation** | | |
| Weekend Study Hours (study_hours_wknd) | 0.685 (0.461) | 1.486 (0.142) |
| Weekend Sleep Hours (sleep_hours_wknd) | −0.513 (0.821) | −0.625 (0.534) |
| **Parental Factors** | | |
| Parental Support Index (par_support_index) | 0.605 (1.699) | 0.356 (0.723) |
| Parental Habits Instilled (par_instilled) | −4.043 (3.628) | −1.115 (0.269) |
| **Health Knowledge & Perceptions** | | |
| Knowledge: Exercise (know_exercise) | −0.819 (5.089) | −0.161 (0.873) |
| Knowledge: Disease (know_disease) | −3.455 (3.024) | −1.143 (0.257) |
| Knowledge: Sleep (know_sleep) | 13.263 (8.276) | 1.603 (0.113) |
| PA Satisfaction (pa_satisfied) | 3.659** (1.670) | 2.191 (0.032) |
| Intent to Improve (intent_improve) | 3.002** (1.383) | 2.171 (0.033) |
| **Model Fit Statistics** | | |
| Observations (N) | 93 | — |
| Dependent Variable Mean (SD) | 10.871 (12.998) | — |
| Standard Error of Regression | 12.072 | — |
| R^2 | 0.315 | — |
| Adjusted R^2 | 0.137 | — |
| F-statistic (19, 73) | 1.771* | p = 0.043 |

*Note: Dependent variable is pa_hours_week. Standard errors are reported in parentheses.*

I hypothesized that there would be a positive correlation between an individual's knowledge of physical health and their participation in physical activities. This is because greater knowledge of physical health may increase the awareness of its significance and, thus, the motivation to participate in physical activities. However, all three knowledge indices (knowledge

on exercise, disease, and sleep) were statistically insignificant. On the other hand, physical activity satisfaction led to 3.659 additional physical activity hours/week, while the intent to improve led to 3.002 additional PA hours/week, each statistically significant at the 5% level.

A significant finding in my results is that an individual knowing about the benefits of physical activity does not directly translate their knowledge into behaviour; thus, simply providing individuals with information about the health benefits of exercise may have limited behavioural impact if students do not enjoy physical activity or intend to change their behaviour. In this dataset, where students have high knowledge of the benefits of physical activity, it seems that more knowledge does not nudge them to be more active. This suggests that there is more of a motivational barrier and competing time-use objectives. Consequently, sustaining student motivation appears to be the central factor for maintaining long-term physical activity across diverse student groups.

Several other control variables, including demographic factors such as age and sex, and participation in school or out-of-school sports clubs, did not show statistically significant relationships with weekly physical activity time. Instead, activity was strongly linked to satisfaction and intention to improve, suggesting that personal motivation and engagement drive participation far more than institutional settings or demographic traits. The lack of statistical significance of control variables is, however, also noteworthy. Age and sex do not have a statistically significant relationship with weekly physical activity time, suggesting some equality in physical activity, which may help the labor force. Similarly, none of the coefficients in relation to sports activities in schools or from out-of-school teams, either group or individual, are statistically significant. This indicates that formal club participation alone does not appear to explain substantial differences in overall weekly physical activity in this sample. This might imply that physical activity is broadly accessible to different types of students, and keeping students' motivation towards physical activity would be important for sustaining participation.

**Conclusion**

Physical activity is important since good physical health can reduce healthcare costs associated with preventable diseases and improve productivity. In this research, I conducted a survey regarding the correlation between Chinese individuals' knowledge of physical health and their participation using Tencent. The findings show strong correlations between physical activity satisfaction and intention to improve, which suggests that students' motivation and engagement may be related to their level of physical activity rather than their demographic backgrounds and institutional settings for involvement. Thus, instead of purely teaching students theoretical knowledge on physical health, it may be beneficial for the government to provide support in making students more motivated to do sports and build more physical activity sites such as gyms and fields. This would allow a more productive labor force in the future. Organizations and employers would also benefit from implementing gym sites beside offices to increase accessibility and motivation. In a broader context, China is facing an aging population, subpar replacement rates, and a looming contraction of the labor force that needs intervention.

This increased labor force may come from increasing retirement age, accepting or welcoming immigrant workers, and advancement in robotics. In the context of increasing retirement age and maintaining productivity levels, physical activity policies such as accessibility and education would be needed.

Several limitations need to be considered when discussing the findings. First, the study is based on self-reported questionnaires that may be affected by personal bias, since people tend to overstate their levels of physical activity or give answers that might appear desirable to others even though the study is anonymous. Second, when individuals are required to recall their activities over an extended period, participants may have difficulty accurately remembering the duration and frequency of their activities. In the survey, I asked about time spent on an average school week, while the survey was conducted during summer break. Thus, recall bias may also affect the reliability of reported weekly data. Third, the study has 93 responses, which may not adequately represent the wider population of Chinese adolescents. Differences in school environment, socioeconomic background, access to sports facilities, and academic demands could influence individuals' physical activity levels. Lastly, I found that strong correlations exist between physical activity satisfaction and intention to improve, yet it is unclear whether individuals participate in physical activity with the intent to be satisfied with their physical health, or if they participate in physical activity and naturally become satisfied with their physical health. Thus, further research is needed to clarify the direction of this relationship and determine whether satisfaction drives physical activity intentions or results from participation.

Future research should address this study's main limitations while testing the behavioral mechanism more directly. A larger, more geographically diverse sample, including rural and small-town students, would test whether the finding that knowledge does not translate into activity holds where baseline knowledge is lower and facilities scarcer, while objective measurement such as accelerometers and a longitudinal design could reduce self-report bias and establish whether satisfaction and intention genuinely cause activity. Most importantly, incorporating validated delay-discounting tasks and intrinsic motivation scales would reveal whether students with steeper discount rates show a weaker knowledge–behavior link, reducing present bias to an extent. Finally, linking adolescent activity patterns to later academic, health, and labor-market outcomes would directly test the human-capital argument motivating this study, offering stronger guidance for policy in aging economies such as China.

**Works Cited**


behavioralecon. Present bias. In BehavioralEconomics.com | The BE Hub. https://www.behavioraleconomics.com/resources/mini-encyclopedia-of-be/present-bias/. Accessed July 22, 2026.

Bonthuis, Boele, et al. "Population Aging and Pension Reforms in China." IMF Working Papers, vol. 2026, no. 027, Feb. 2026, p. 1, https://doi.org/10.5089/9798229038522.001. Accessed August 26, 2026.

Ding, Ding, et al. "The Economic Burden of Physical Inactivity: A Global Analysis of Major Non-Communicable Diseases." The Lancet, vol. 388, no. 10051, Sept. 2016, pp. 1311–24, https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(16)30383-X/abstract. Accessed August 26, 2026.

Gaudiaut, Tristan. "How Fast Is the Asian Population Ageing?" Statista Daily Data, Statista, 26 Jan. 2026, https://www.statista.com/chart/35738/ageing-trends-in-asia-change-in-share-of-people-65-and-above-by-country/. Accessed July 29, 2026.

Hayes, Gráinne, et al. "Tracking of Physical Activity and Sedentary Behavior From Adolescence to Young Adulthood: A Systematic Literature Review." Journal of Adolescent Health, vol. 65, no. 4, June 2019, https://doi.org/10.1016/j.jadohealth.2019.03.013. Accessed July 29, 2026.

Kari, Jaana T., et al. "The Individual-Level Productivity Costs of Physical Inactivity." Medicine & Science in Sports & Exercise, vol. 55, no. 2, Sept. 2022, pp. 255–63, https://doi.org/10.1249/MSS.0000000000003037. Accessed August 26, 2026.

Maags, Christina, and Jingwen Zhang. "Active Aging in the People's Republic of China: A Case Study of Working After Retirement." Asian Development Review, vol. 42, no. 02, June 2025, pp. 111–46, https://www.adb.org/sites/default/files/publication/1063031/adr-vol42no2-6-aging-prc-working-after-retirement.pdf. Accessed August 26, 2026.

MedlinePlus. "Health Risks of an Inactive Lifestyle." MedlinePlus, National Library of Medicine, 1 Sept. 2017, https://medlineplus.gov/healthrisksofaninactivelifestyle.html. Accessed August 8, 2026.

Miller, Kelcie, et al. "Types of Leisure-Time Physical Activity Participation in Childhood and Adolescence, and Physical Activity Behaviours and Health Outcomes in Adulthood: A Systematic Review." BMC Public Health, vol. 24, no. 1, July 2024, https://doi.org/10.1186/s12889-024-19050-3. Accessed August 8, 2026.

Santos, Andreia Costa, et al. "The Cost of Inaction on Physical Inactivity to Public Health-Care Systems: A Population-Attributable Fraction Analysis." The Lancet. Global Health, vol. 11, no. 1, Dec. 2022, pp. S2214–109X(22)00464-8, https://doi.org/10.1016/S2214-109X(22)00464-8. Accessed August 26, 2026.

WHO guidelines on physical activity and sedentary behaviour. World Health Organization, 2020, https://www.ncbi.nlm.nih.gov/books/NBK566039/. Accessed August 26, 2026.

Witynski, Max. “Behavioral Economics, Explained.” News.UChicago.Edu, The University of Chicago, 2024, https://news.uchicago.edu/explainer/what-is-behavioral-economics

World Bank Group. “From Data to Opportunity: Putting Human Capital in People’s Hands.” Worldbank.Org, World Bank Group, 3 June 2026, https://www.worldbank.org/en/news/feature/2026/05/05/from-data-to-opportunity-putting-human-capital-in-people-s-hands. Accessed September 8, 2026.

World Health Organization. “Physical Activity Profile 2022”. Www.Who.Int, 2022, https://cdn.who.int/media/docs/default-source/country-profiles/physical-activity/physical-activity-chn-2022-country-profile.pdf?sfvrsn=64471272_6&download=true. Accessed August 8, 2026.

World Health Organization. “WHO guidelines on physical activity and sedentary behaviour”. In World Health Organisation. World Health Organization, 2020, https://www.who.int/publications/i/item/9789240015128. Accessed August 8, 2026.

World Health Organization. “The Global Status Report on Physical Activity 2022.” Www.Who.Int, 19 Oct. 2022, https://www.who.int/teams/health-promotion/physical-activity/global-status-report-on-physical-activity-2022. Accessed August 8, 2026.

World Health Organization. “Global status report on physical activity 2022: Country profiles.” Www.Who.Int, 31 Jan. 2023, https://www.who.int/publications/i/item/9789240064119. Accessed August 8, 2026.

Xiao, Jing Jian, and Nilton Porto. “Present Bias and Financial Behavior.” Financial Planning Review, vol. 2, no. 2, June 2019, https://doi.org/10.1002/cfp2.1048. Accessed August 26, 2026.

**Market opportunity analysis of creator-led sports broadcasting in Southeast Asia**
**By Dia Dsilva**

**Abstract**

This paper investigates the potential market opportunity for creator-led sports broadcasting in Southeast Asia. The study uses a mixed-method approach, with data collected through a survey of consumers across Southeast Asia and interviews with three industry professionals working in sports media and sports rights. Some of the main findings include an interest in creator-led sports broadcasting, a shift in the way sports have traditionally been consumed, and a growing preference for creator-led and short-form content. However, the results also showed weaker levels of trust and willingness to spend money on products or services promoted by creators. Industry professionals saw potential in this model of sports distribution, but emphasized the importance of scale, professionalism, rights management, and monetization. Overall, the study concludes that creator-led sports broadcasting in Southeast Asia has potential, but it is more likely to work alongside traditional sports media rather than fully replace it.

**Keywords** Sport creator, Southeast Asia, Market Opportunity, Sports Broadcasting, Sports Media, Content Creation

## 1. Introduction

Watching sports games from home used to be a generally straightforward process: turning on the TV, switching to the sports channel, and enjoying the live game and commentary. However, over the years, the way sports are consumed from home and anywhere beyond the stadium has changed and deviated from solely traditional television. Nowadays, sports fans are beginning to watch sports on several social media platforms such as YouTube, Instagram, TikTok, Twitch, etc. This switch can be attributed to several factors, one of the most important being changes in the media consumption habits of younger generations, especially Gen Z. With the rise of social media platforms such as TikTok and Instagram, short-form, easily accessible, and on-demand content has become a major part of how Gen Z consumes entertainment (Bleier et al., 2024). This shift has also contributed to the growth of a new area of business: the creator economy (Bleier et al., 2024; Eissa & Eissa, 2022). With social media, influencers and content creators have become increasingly influential voices, as platforms allow creators to target specific communities and audiences who can relate to their interests and perspectives. The reason this economy has become so vast is that there are numerous different categories and sides of social media, and so many opportunities for someone to be seen or find a group of people that relate to them. Sports content creators and influencers are also taking over the sports industry, through analysing players and games, providing coaching, etc. Athletes are also able to communicate with fans first-hand through their own social media accounts. Additionally, content accessed through social media is free, creating an alternative to traditional sports broadcasting and allowing influencers to monetize through advertisements, brand deals, sponsorships, and

more (Bleier et al., 2024). Overall, these changes demonstrate that sports broadcasting is switching from a traditional television-dominated model toward a more digital, interactive, and creator-driven media environment.

Creator-led sports broadcasting refers to sports coverage, commentary, or analysis produced by individual creators who add their own personality, expertise, or perspective(Chiu et al., 2025; Tu, 2026). Some creators analyse specific players' behaviours and technique, while others provide live reactions or commentary to sporting events. These creators differ from traditional broadcasters in the sense that they foster a more personal, informal, and interactive relationship with their audience, while traditional broadcasters tend to be more professional. Uses and Gratifications Theory helps explain why audiences may choose this content, as viewers actively seek media that provides entertainment, information, interaction, or a sense of community(Katz et al., 1973). Repeated exposure can also lead to parasocial relationships, where viewers develop a sense of familiarity or connection with a creator they do not personally know. Credibility is therefore important, particularly when creators present themselves as experts or have professional experience within sports(Chiu et al., 2025; Tu, 2026). Overall, creator-led broadcasting is not replacing traditional sports media, but it is becoming an increasingly important part of it.

Southeast Asia provides a relevant setting for examining creator-led sports broadcasting due to its large population and rapidly developing digital landscape. A significant portion of this region is young, which is the exact audience most likely to be using social media platforms and engaging with creators. A 2024 Pew Research Center study found that adults under 30 were far more likely than older adults to use platforms such as Instagram, Snapchat, and TikTok (Pew Research Center, 2024). Additionally, social media platforms have vast numbers of users and creators from this region. In fact, DataReportal (2025) estimated about 298 million active TikTok users in Southeast Asia in January 2025, representing almost 18.7% of TikTok's global users. On top of the demand for social media, there is a specific demand for sports content. In fact, YouGov's 2023 Global Sports Media Landscape report found that 89% of Indonesian consumers were "sports engagers," meaning they engaged with sports they regularly followed across traditional and digital media platforms (YouGov, 2023). Overall, the combination of sports interest and widespread social media use makes Southeast Asia a relevant region for understanding how creator-led broadcasting may be transforming the sports media industry.

Existing literature (Bleier et al.; Tu, 2026; Chiu et al., 2025) shows that social media and influencers are reshaping the way media, and particularly sports media, is produced and consumed. Previous studies have found that younger audiences increasingly value interactive, on-demand, and short-form content, while influencer characteristics such as authenticity, expertise, and credibility can strengthen emotional attachment and engagement with their content(Tu, 2026; Chiu et al., 2025). Research has also shown that social media can increase audience participation and support the development of creator economies surrounding major sports events. Similar research has already been done in other parts of the world. For example, studies in China have examined how sports livestreamer characteristics influence viewers'

emotional attachment and engagement, while research surrounding the Qatar 2022 World Cup has examined the growing role of influencers alongside journalists and traditional media (Chiu et al., 2025; Pena-Fernandez et al.,2026). This shows that creator-led and influencer-based sports media is already being studied worldwide, but there is still limited research and data focusing specifically on Southeast Asia. Additionally, much of the existing literature focuses on individual aspects of this transformation, such as Gen Z consumption habits, influencer engagement, or the wider creator economy(Eissa & Eissa, 2022; Bleier et al., 2024), rather than examining the commercial potential of creator-led sports broadcasting. In particular, there is limited research examining this opportunity within Southeast Asia. Therefore, the research question for this study is: What is the market opportunity for creator-led sports broadcasting in Southeast Asia? The study aims to assess audience demand, identify the factors that influence engagement with sports creators, determine if there is an existing relationship between social media users in Southeast Asia and sports content creators, and evaluate the potential for creator-led broadcasting to become a viable part of the region's sports media industry.

This paper is structured into several sections. First, the literature review examines existing research regarding digital sports media, creator-led content, and audience engagement, followed by the methodology, which explains the survey conducted among people across Southeast Asia and the approach used to analyse the collected data. There is also a second set of primary research conducted through interviews with specialists in the sports business industry. The results are then presented and discussed in relation to the research question and existing literature, before the paper concludes by summarising the key findings, implications, and opportunities for future research.

### 1.1 Literature Review

There is growing literature (Eissa & Eissa, 2022; Bleier et al.; Tu, 2026; Chiu et al.,2025) available on the creator economy and the opportunity for sports creators in different parts of the world. The researchers have used mixed methods for collecting data and analysing the behavioral patterns of media consumption as well as factors that cause viewers to gravitate towards certain content creators.

Here is the selected literature summary that the author has taken references from:
Smith Mehta (2021), in his doctoral thesis, studied how the internet and streaming platforms changed the media industry in India. The methodology of this study was interviews with creators, executives, and agents, as well as analysis of the media industry as a whole. The main findings were that social media changed content distribution in India, making it easier overall in this industry. The study explains that the advancement of technology and social media has made it easier for any creator to distribute content and has made the different streaming platforms highly intertwined because creators rely on many platforms. While creators have transformed this industry and shown that the barriers to entry are lowered, the paper discusses several issues that exist for creators as well, such as gender inequalities.

RuoNan Tu (2026), in his research, set out to understand if influencer characteristics and characteristics drive/affect engagement in sports livestreaming. The methodology of this study was a statistical approach through analysing data from sports livestream viewers on major platforms. The main findings were that expertise, authenticity, and attractiveness of influencers have a significant positive impact on a viewer's sports livestream engagement. The study also found that emotional attachment is a key factor for loyal viewership and constant audience engagement, which happens as a result of feeling positively toward the influencer and their demeanor. Overall, the study suggests that influencer behavior has a strong influence on audience engagement of sports livestream content and fosters a relationship with viewers, allowing sports media consumption to increase.

Arnaud Zeng (2025) researched how sports leagues have adapted to the changes in media consumption because of Gen Z in order to stay relevant and keep audiences engaged in sports media. The main findings showed that Gen Z often deviates from long matches and gravitates towards short-form content that they can get on demand because of behavioral changes such as shorter attention spans and instant gratification. One other key finding was that Gen Z gravitates towards creator-led content with interactive experiences as opposed to solely watching a game. In response to these changes in demand, sports organizations are focusing on partnering with creators, digital strategies, and exploring new formats of content.

Victoria Mora de la Torre and Laura Martin Lopez (2025) in their journal article examined how streaming platforms and content creators are transforming audiovisual entertainment and challenging traditional television. The methodology of this study used a case study, content analysis, and social media data. Some of the main findings of the study were that nowadays consumers value content that is interactive, features their favorite influencers, and is in real time. An example of this is Instagram or social media live streams, which viewers tend to gravitate towards because of the live and personal experience it gives. Overall, the research suggests that digital platforms are reshaping media consumption to fit the desires and tendencies of the new generation and their tendencies, especially.

Noura Eissa and Eiman Abdel Maksoud Eissa (2022) in their journal article explored the relationship between social media, sports mega events, and the creator economy. The methodology of the study was a literature review of previous studies, the application of Whiting and Williams' Uses and Gratifications Model, and an instrumental case study approach. The main findings of the study were that social media plays a big role in supporting the creator economy surrounding major sports events. It shows that social media helps engage audiences from all over the world, strengthens loyalty, and encourages participation and interaction. That is why it helps event creators build stronger relationships with fans and help meet their wants.

In addition to the literature above, **Table 1** presents the summary of selected papers that were reviewed for finalising the questions of the survey questionnaire.

**Table 1:** Summary of Selected Literature and Theories Used

| Title | What is the main idea? | Variables studied | Theory used |
|---|---|---|---|
| **Live streaming and Livelihoods: Decoding the creator Economy's influence on consumer attitude and digital behavior Author links open overlay panel** | There are a series of factors that a livestreamer has that make viewers want to watch them and trust them enough to buy products or engage with brands that are being advertised. | Credibility, entertainment, informativeness, human interaction, interaction control, positive value, self-brand congruity, social value, consumer attitude, digital behavior, purchase intention, and brand engagement. | Uses and Gratifications Theory (UGT). |
| **Why We Follow: An Examination of Parasocial Interaction and Fan Motivations for Following Athlete Archetypes on Twitter** | Fans follow athletes on social media for different reasons depending on how interactive the athlete is with the fanbase. Greater interactions often create stronger feelings of personal connection (parasocial interaction), which influences follower motivations and engagement. | Athlete interaction style (social and parasocial), interactivity, information seeking, entertainment, personality, fanbase, community, follower motivations, sense of personal closeness to the athlete | Parasocial Interaction Theory and Uses and Gratifications Theory (UGT). |
| **Understanding the appeal of user‐generated media: a uses and gratification perspective** | People use user-generated media to satisfy different needs. They consume content for information and entertainment, participate for social interaction and community, and produce content for self-expression. These motivations explain why user-generated platforms are so appealing. | Information seeking, entertainment, mood management, social interaction, community, self-expression, ease of use, user control, media participation | Uses and Gratifications Theory (UGT). |
| **How Social Media Influencers Foster Relationships with Followers: The Roles of Source Credibility and Fairness in Parasocial Relationship and Product Interest** | An influencer's attractiveness, similarity to followers, and fairness when interacting with followers can strengthen parasocial relationships. These stronger relationships can then increase followers' interest in products promoted by the influencer. | Credibility, similarity, attractiveness, communication fairness, parasocial relationship, and product interest. | Source Credibility Theory, Communication Justice Theory, and Parasocial Relationship Theory. |

## 2. Methodology

This study is mixed-methods and cross-sectional. Both qualitative and quantitative data were collected through a Google Form sent out to 40-50 respondents, as well as interviews conducted with 3 sports management experts. The sample selection method followed for this study is targeted sampling: people who are located in Southeast Asia, because the research aims to analyse the market conditions in the Southeast Asia region only.

### 2.1 Participants and Sampling

The participants for this survey were all located in Southeast Asia, as the study specifically aimed to examine the potential for creator-led sports broadcasting in the region. The survey received responses from Singapore, Malaysia, Thailand, the Philippines, and Indonesia. A total of 44 respondents completed the questionnaire. Convenience Sampling was used, as participants were selected based on their accessibility and willingness to participate in the survey while still meeting the geographical restrictions. The questionnaire was electronic, a Google Form, and was distributed to personal contacts and consumers of sports via WhatsApp and

Gmail. These distribution channels helped the survey easily reach respondents in Southeast Asia in the desired time frame.

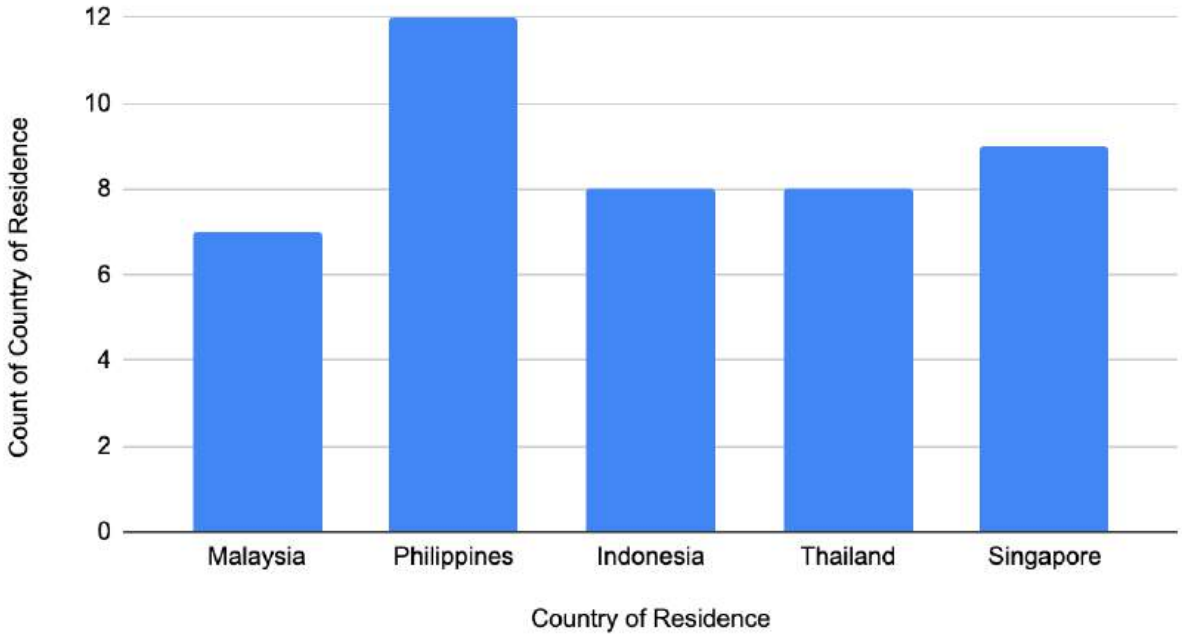



**Figure 1:** Distribution of Respondents' Country of Residence

### 2.2 Data Collection Methods

The data collection process included two main sections: Survey and interviews. First, a survey was conducted electronically through Google Forms and sent out to respondents across Southeast Asia. The target number of responses was 45, of which 44 participants completed the survey. Afterward, interviews were scheduled and conducted over Google Meet with 3 industry professionals who have experience in Sports Media.

### 2.2 Survey

To construct the survey, existing literature regarding creator-led content, sports broadcasting, audience engagement, social media, Gen Z media consumption, parasocial relationships, and the Uses and Gratifications theory was reviewed. The concepts presented in this literature were then used to create the survey questions. The main constructs of the survey are entertainment, credibility, information, market viability, and interaction & connection. The questionnaire primarily used a five-point Likert scale, ranging from Strongly Disagree to Strongly Agree, to measure respondents' level of agreement with the statements related to each construct. Using a consistent scale allowed the responses across different constructs to be compared and analyzed quantitatively. Entertainment was included to measure the existing interest in sports content creators and demand in the market for the content. This construct is based on the Uses & Gratifications theory, which explains that people choose social media to satisfy needs such as participation, social interaction, and relationships (Eissa & Eissa, 2022).

The credibility construct was included to determine whether viewers see content creators as trustworthy, knowledgeable, and authentic (Tu, 2026). The Information construct was included to determine how users consume sports social media content, whether that be through learning about sports, getting commentary and analysis, coaching, etc., and also connects back to the Uses & Gratifications theory. The interaction & connection construct was used to measure how users connect to content creators, interact with the creator and other viewers through chat and comment features, and develop emotional attachments. This concept can be supported by the parasocial relationships theory (Tu, 2026). Finally, the market viability construct is used to determine whether there is demand for creator-led sports content and broadcasting and a market opportunity in Southeast Asia.

A pilot survey was conducted before the main data collection to test the effectiveness and overall structure of the questionnaire. The pilot survey received 6 responses from participants who provided feedback about the wording, organization, and ease of understanding of the questions. Based on this feedback, several changes were made to improve the survey, which is why the pilot was an important step.

### 2.3 Interview

The second section of the data collection method this study relies on is the semi-structured interviews. The questions were framed by the author partially based on the reviewed literature, accompanied by intuitive knowledge of the author. The questions were framed so that the responses are in sync with the constructs of the survey.

## 3. Results and Discussion

The results of this paper are based on the two sources explained above — the survey of 44 participants and 3 interviews conducted with industry professionals. The two sources combined provide insights and a clear understanding of the market opportunity for content creators.

### 3.1 Survey results

The survey respondents represent heterogeneous consumers in Southeast Asia. The sample included respondents of different genders, ages, and employment backgrounds, with the largest number of respondents being general employees. The survey also examined respondents' existing sports viewing habits, including how often they watch sports and what platforms they use. According to the data collected, YouTube was the most popular platform for watching sports, accounting for 41.5% of responses to this question (Figure 2). The remaining responses were distributed across a wide range of platforms including TikTok, Facebook, X, Cable television, Satellite television, OTT services, and other streaming sites. This variety suggests that consumers of sports in Southeast Asia use a wide range of sources as opposed to a single channel or service.

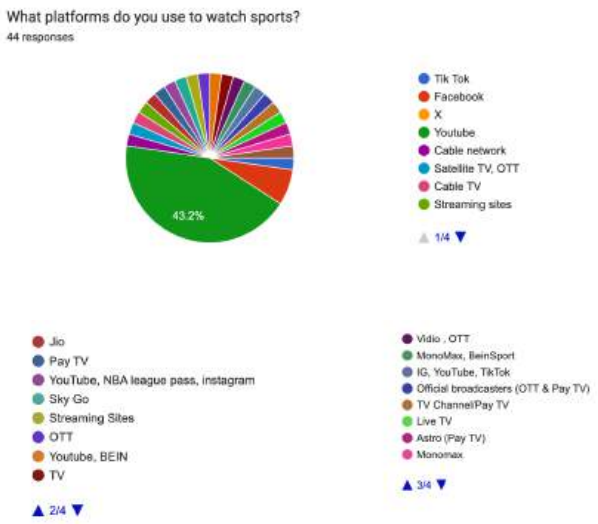



**Figure 2:** Distribution of platforms respondents use to watch sports

The overall response in the Entertainment construct was positive. Most responses point in the same direction, with 59.1% of respondents agreeing/strongly agreed that sports creator content is entertaining, while only 6.8% disagreed. This suggests that content creators clearly have entertainment value for sports audiences. Next, the overall response in the Information construct was moderately positive. 59.1% of respondents said that creators give a different perspective from traditional sports commentary, and 47.7% said they actually learn about sports from influencers. These are related but not equally strong. So, the main conclusion is that content creators may be more useful for extra perspectives or additional commentary than as the main source of sports information.

The Credibility construct, however, is more mixed. 54.5% of respondents believe that sports creators have expertise, but only 34.1% said they trust the information creators share. While half of the respondents having a neutral stance on trusting creators is not negative, 56.8% still prefer pundits/former athletes instead of strangers on the internet. The main conclusion here is that people may think that sports content creators are knowledgeable, but that does not automatically mean they fully trust them.

The Interaction & Connection construct also showed mixed results. 59.1% of respondents agreed that they can relate to sports content creators, while only 43.2% agreed that live chat and comment sections help them interact with creators and other fans. Even fewer, 29.5%, said that they feel heard and understood by the creators they follow. This shows that while people may find sports creators relatable or entertaining, that does not necessarily mean they feel a strong personal connection or sense of community with them.

Finally, the Market Viability construct showed some positive results but also some uncertainty. 52.3% of respondents said they prefer watching sports on social media rather than traditional TV, and 54.5% said that creator-led sports broadcasting could be a more convenient way to consume sports (Figure 3). However, only 40.9% said they would recommend creator-led sports broadcasting to friends, and only 20.5% said they would be willing to purchase products or services promoted by sports creators. This suggests that there is an interest in consuming sports through content creators and social media, but monetization may be an issue.

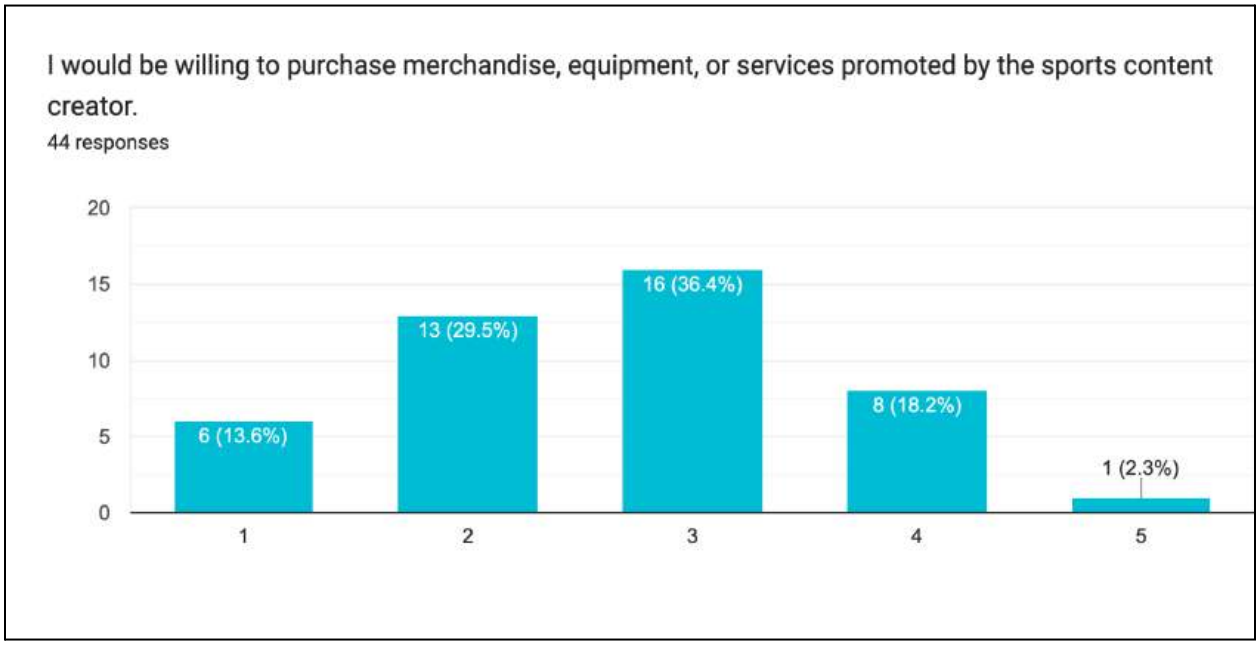



**Figure 3:** Distribution of the responses related to construct 5 (Market Viability)

### 3.2 Interview responses

To understand the mechanisms underlying market functioning, the survey responses were complemented by semi-structured interviews with professionals already engaged in the market. Interviews were held in English and primarily explored sports rights and consumer behavior in Southeast Asia. The interviews lasted roughly 45 minutes and were all recorded, fully transcribed, and analyzed.

The main ideas from the interviews were that there is currently a gap in how sports rights are being distributed and monetized, and that the way sports are currently being consumed in Southeast Asia is changing. A sports rights specialist explained that rights holders traditionally sell their content to television channels or streaming services that pay a licensing fee for the rights. However, platforms such as YouTube and TikTok, which are becoming increasingly popular with Gen Z, generally do not pay for sports rights because they operate as platforms rather than broadcasters. This creates a problem because these platforms are gaining more of the audience and making money from sports content, while the traditional broadcasters who actually pay for the rights are losing viewers and revenue. As one industry professional explained,

“money comes from the eyeballs,” emphasizing that the commercial value of the sports industry depends heavily on the size and engagement of the audience as well as where that audience is consuming content. However, the interviews also showed that large viewership alone is not enough, as creators still need a clear way to generate revenue through sponsorships, advertising, or brand partnerships. Piracy and declining subscriptions in Southeast Asia have made this problem even stronger. Therefore, the main issue is that the platforms with the growing audiences are not necessarily the ones paying for the rights, creating a gap between where sports are being watched and where the money is going. This finding is consistent with previous research that highlights the growing importance of digital and social media platforms in how younger audiences consume and engage with content (Bleier et al., 2024).

Another interviewee gave an example of a creator in Indonesia, Rand, who streamed the Indonesia vs. Singapore football match from the FIFA Azad Cup on YouTube in his own style and reportedly attracted around 9 million viewers. What makes this example especially relevant is that the creator legally sublicensed the rights from the traditional broadcaster, meaning that the two forms of coverage were able to exist alongside each other. The traditional free-to-air broadcast still achieved a very strong market share, while the creator was able to reach an additional audience through YouTube. This suggests that creator-led sports broadcasting does not necessarily have to replace traditional coverage, but could instead increase the overall audience for a sporting event. The interviewee described this as a situation where “one plus one equals three,” as both forms of distribution could potentially create more value together. However, he also explained that the remaining question is how an audience of this size can actually be monetized, showing that strong viewership alone does not automatically make the model commercially successful.

For content creators, the findings suggest that there is a promising opportunity in creator-led sports broadcasting in Southeast Asia, but creators would need to approach it professionally. From the interviews, one of the main points was that creators need to show that they have a large enough audience and strong engagement before rights holders would seriously consider working with them. One way to do this is by persistently posting content on a schedule or at a certain frequency. They would also need to be reliable and consistent, as sports rights involve contacts, sponsors, and specific rules that need to be followed. Another important implication is that creators should not rely only on revenue from platforms such as YouTube, as this may not be enough to support the cost of sports rights. Instead, building direct relationships with brands and sponsors could be a more effective way to generate revenue. Working with an agency or creator network could also help creators manage contracts, sponsorships, and relationships with rights holders more professionally.

### 3.3 Limitations

Despite efforts to combine the consumer as well as the producer side of the market, this mixed-method approach presents certain limitations. First, the sample size was relatively small, meaning that the results may not represent the wider Southeast Asian population. The survey

also used convenience and self-selected sampling, which may have caused some bias in the responses. In addition, the survey was distributed through a single online channel, limiting the range of people who may have seen or responded to it. The study was also cross-sectional, meaning that the data was collected at one point in time and therefore cannot be used to establish a cause-and-effect relationship. Finally, the research was geographically restricted to Southeast Asia, meaning that the results may not apply to other parts of the world.

**4. Discussion**

Overall, the findings suggest that there is a potential market opportunity for creator-led sports broadcasting in Southeast Asia, but there are still some important challenges. The survey results showed that respondents are interested in creator-led sports content and see value in the entertainment, convenience, and different perspectives that creators can provide. This also aligns with Eissa and Eissa (2022), who found that social media plays an important role in building creator economies around major sporting events, particularly through the different information and entertainment needs of audiences. This supports previous research showing that younger audiences are increasingly consuming content through digital and social media platforms rather than traditional media (Bleier et al., 2024). The interviews also supported this idea, especially through the view that audiences are moving towards platforms such as YouTube and TikTok while traditional broadcasters are losing viewers.

However, the findings also show that interest in creator-led sports does not necessarily mean that creators can fully replace traditional broadcasters. The survey showed weaker results in areas such as trust and willingness to spend money, while many respondents still valued professional pundits and former athletes. This connects with the interviews, where industry professionals showed that creators would need to prove that they are reliable, professional, and able to generate engagement before rights holders would seriously work with them. Monetization was also identified as a major challenge, with the interviews suggesting that direct brand partnerships and sponsorships may be more important than relying only on platform advertising, as a large number of views does not always correlate to a high amount of revenue generated. Overall, the strongest opportunity may therefore be for creators to work alongside traditional sports media rather than completely replace it, as there is clear demand from the consumer end but scale and monetization still need to be developed. Regardless, the findings show that the sports media industry is changing and creator-led content is playing a larger role in how sports are consumed as time evolves.

**Works Cited**


Day, R. (n.d.). WHAT THE DATA SAY: 62% enjoy watching sports on social media more than on TV. Stagwell. Retrieved September 2, 2026, from https://www.stagwellglobal.com/what-the-data-say-62-enjoy-watching-sports-on-social-media-more-than-on-tv/

Eissa, N., & Eissa, E. A. M. (2022). Testing Whiting and Williams's social media Uses and Gratifications Model in building Creator Economies through Mega Sporting Events. Res Militares, 12(3), 2047-2060. https://www.researchgate.net/profile/Noura-Eissa-3/publication/365769479_Testing_Whiting_and_Williams%27s_social_media_Uses_and_Gratifications_Model_in_building_Creator_Economies_through_Mega_Sporting_Events/links/63826e32554def6193741c7b/Testing-Whiting-and-Williamss-social-media-Uses-and-Gratifications-Model-in-building-Creator-Economies-through-Mega-Sporting-Events.pdf?__cf_chl_tk=qOo0OwxhE6Pw5uJc19JWsLPkqwjWru2pQEToQMk8RqE-1781176169-1.0.1.1-HsHV2CzvHwJYwa.H9_zWZaetnGrwd4NvDs7wLCz0kvE

How content creators became power players in sports. (n.d.). Ingager Sports. Retrieved September 2, 2026, from https://www.ingagersports.com/how-content-creators-became-sports-power-players

Kost, D. (2025, October 27). User-Generated Content in Sports: Why Fans Trust It More Than Pro Posts. FanZGo. Retrieved September 2, 2026, from https://fanzgo.com/2025/10/27/user-generated-content-in-sports-why-fans-trust-it-more-than-pro-posts/

Mora de la Torre, V., & López Martín, L. (2025). The Transformation of Audiovisual Entertainment From Traditional Television to Live Streaming. VISUAL REVIEW. International Visual Culture Review / Revista Internacional de Cultura Visual, 17(4), 39-54. https://doi.org/https://doi.org/10.62161/revvisual.v17.5826

Smith, M. (n.d.). The 'new screen ecology' in India: A study in digital transformation of media [Doctoral thesis, Queensland University of Technology]. QUT ePrints. https://eprints.qut.edu.au/208077/

Sun, J., Sarfraz, M., Ivascu, L., Han, H., & Ozturk, I. (2024). Live streaming and livelihoods: Decoding the creator Economy's influence on consumer attitude and digital behavior. ScienceDirect, 78. https://doi.org/10.1016/j.jretconser.2024.103753

Tu, R. (2026). Unlocking Audience Engagement in Sports Livestreaming: How Influencer Attributes Cultivate Emotional Bonds and Drive Participatory Behavior. Journal of Physical Education, Recreation, and Dance, 1(2), 1-20. https://doi.org/10.63313/perd.9006

Walsh, P. (n.d.). Why We Follow: An Examination of Parasocial Interaction and Fan Motivations for Following Athlete Archetypes on Twitter. Human Kinetics Journal, 5(4), 481-502. https://doi.org/10.1123/ijsc.5.4.481

Yuan, S., & Lou, C. (2020). How Social Media Influencers Foster Relationships with Followers: The Roles of Source Credibility and Fairness in Parasocial Relationship and Product Interest. Journal of Interactive Advertising, 20(2), 133-147. https://doi.org/10.1080/15252019.2020.1769514

Zeng, A. (2025). Winning Over Gen Z: The Evolving Strategies of Sports Leagues and Media in Response to Changing Youth Habits [Doctoral thesis, Massachussetts Institute of Technology]. MIT Libraries. https://dspace.mit.edu/entities/publication/48484339-e660-4111-af01-cf2b5291bb8e

## Appendix

**Table A1: Summary of the constructs**

| Construct | Definition | Supporting Lit. | Number of questions |
|---|---|---|---|
| Entertainment | The extent to which viewers find creator-led sports broadcasts enjoyable, engaging, and entertaining | **Live streaming and livelihoods: Decoding the creator Economy's influence on consumer attitude and digital behavior**<br>**Author links open overlay panel**<br>**WHAT THE DATA SAY: 62% enjoy watching sports on social media more than on TV** | 3 |
| Information | The extent to which creator-led sports content helps viewers learn about and understand sports further | **How Social Media Influencers Foster Relationships with Followers: The Roles of Source Credibility and Fairness in Parasocial Relationship and Product Interest** | 2 |
| Credibility | The extent to which viewers can trust sports content creators and the information and opinions they share | **How Social Media Influencers Foster Relationships with Followers: The Roles of Source Credibility and Fairness in Parasocial Relationship and Product Interest**<br>**User-Generated Content in Sports: Why Fans Trust It More Than Pro Posts** | 2 |
| Interaction & connection | The extent to which viewers feel connected to the creators they follow and can interact with them, and other viewers as well | **Why We Follow: An Examination of Parasocial Interaction and Fan Motivations for Following Athlete Archetypes on Twitter** | 3 |
| Market Viability | The likelihood that viewers will watch and support creator-led sports broadcasts | Winning Over Gen Z: The Evolving Strategies of Sports Leagues<br>and Media in Response to Changing Youth Habits<br>How content creators became power players in sports | 4 |

**To What Extent is It Feasible for Singapore to Adopt Green Policies used by Norway and China to Increase Consumption of EVs By Sin Eike**

**Abstract**

The purpose of this paper is to analyse the policies that Norway, China, and Singapore have currently adopted to increase the consumption of electric vehicles (EVs). The author focuses their research using primary data collected in Singapore and secondary sources from the Singapore, Norway, and China government websites. Using a regression analysis of new EV fleet percentages and tables comparing the policies of the three countries, the author shows that a series of nudges like bus lane access and financial incentives like reducing parking fees appear more effective for Singapore given the geographical restrictions. Previous research has focused on conducting separate research between the three countries. Attempts at comparing all three policies to determine the next steps Singapore can take were not very successful, given the differences in government targets and geographical conditions. This paper explains the relative economic theories behind why increasing the consumption of EVs provides larger social benefits than traditional gas-powered vehicles. Additionally, the paper analyses existing policies between the three countries before suggesting superior policies tailored to be implemented in Singapore. This study finds that financial incentives that lower the cost of EVs in comparison to traditional gas vehicles, on top of nudges that ease commuting in EVs and building more EV charging ports, result in a positive impact on seeing immediate results in EV consumption, whilst being sustainable and geographically feasible for Singapore.

**Keywords** Carbon Emissions; Electric Vehicles; Climate policy; Nudges; Norway Policies; China Policies; Singapore Policies

**Introduction**

Global temperatures increased rapidly over the past decade, with 2024 being the warmest year, as the Northern Hemisphere saw a 1.67°C increase over historical averages (Climate Institute). To combat climate change, countries around the world implemented several policies with varying levels of effectiveness. Yet, one global method with absolute certainty for reducing $CO_2$ emissions is incentivizing a shift from traditional gas-powered internal combustion engine vehicles (ICEVs) to electric vehicles (EVs). Since 1990, Norway has gradually introduced incentives for zero-emission vehicles, and now leads the world in EV adoption, with 88.9% of all cars sold being electric in 2024 (Adomaitis). China, meanwhile, has the world's largest EV market and was the first country to surpass 10 million annual NEV (New Energy Vehicle) production units (Gong). However, Singapore, a first-world country known globally for sustainability, has been lacking. This paper argues that offering EV consumption incentives yields the greatest economic and social gains. Firstly, this paper discusses the incentives and fiscal benefits that the Norwegian and Chinese governments provided for EVs. Then, the author analyses the elements of a successful policy using empirical evidence and recommends potential

policies Singapore can implement, given its circumstances. Finally, an argument on why a combination of EV cost reductions and nudges is most effective in increasing EV consumption.

The author chose this topic to apply their own economic and research skills, as this market affects the entire Singapore population. This research centers on market failure arising from the production and consumption of ICEVs. Government policies incentivising a shift to EVs reduce market failure by moving the market equilibrium closer to the socially optimal equilibrium.

**Methodology**

This paper adopted a mixed-methods research approach: supply-and-demand analysis, externalities analysis, and primary/secondary data research.

Primary data involves firsthand data collected specifically for the research, meaning up-to-date, easily controlled conditions. Primary data are more trustworthy than secondary data because they are collected personally. For this paper, the author traveled to various parts of Singapore —from the city center to residential areas —to determine the number of EV charging stations, a nudge to increase EV adoption. Observation was used to determine the exact number of EV charging ports, which helped identify ways to increase charging convenience to support greater EV consumption.

Secondary data uses large-scale datasets collected by others, offering greater complexity, convenience, and time efficiency. As traveling to Norway and China to conduct primary research is impossible, this paper will use secondary data collected from Norway and China. Data from Norway primarily consists of metrics from the non-profit Norwegian EV Association (Norsk), which collaborated with the government to ensure accurate data collection. Articles by Lasse Fridstrøm were used to provide a balanced, holistic perspective on the Norwegian EV market by offering opinions on the effectiveness of policies. Data from China consists of empirical evidence by the Australian National University, with differing opinions and metrics from the South China Morning Post.

Visiting every residential area and mall in Singapore to count EV charging stations is not feasible. Paired with the inability to determine national-level emissions through primary research alone, secondary data will also be used to triangulate a balanced argument in Singapore. Secondary data also cross-checks the validity of primary data and broadens the scope without increasing error margins. To minimise bias in the data, a wide range of sources from research papers to local articles were used, and data points were validated against metrics from government websites, such as Singapore's Land Transport Authority.

The author then combined these elements to create their own graphs and tables to determine the policies Singapore can implement to encourage EV adoption.

**Hypothesis**

In 2022, Singapore began actively encouraging the consumption of EVs over ICEVs. The government has consistently enacted policies to make EVs the superior choice over ICEVs, as

they are more sustainable. However, the majority of the policies are temporary and cost-related, aimed at reducing the EV prices. The author believes that implementing permanent financial-based policies and incentive-based nudges from Norway and China—such as installing more EV charging ports in private residences—would accelerate the transition to cleaner, more energy-efficient vehicles. These policies would make EVs Pareto-dominant over ICEVs, shifting the market equilibrium closer to the socially optimal level and reducing the negative externalities associated with ICEVs.

**Theoretical Framework**

Supply and demand curves reveal the quantities of goods/services that producers and consumers are willing and able to purchase at any price. The demand curve is downward-sloping because higher prices decrease quantity demanded, as consumers have limited disposable income and want to maximize utility. The supply curve is upward-sloping because higher prices increase the incentive to produce, as producers are profit-motivated.

Policies that influence demand— such as reducing ownership costs through license plate fees—increase demand for EVs. Nudges —policy or measures that influence behaviour without restricting choice, such as free municipal parking —also increase EV demand. This is shown in Figure 1, which compares EV prices to EV quantity. Such policies shift demand to the right, changing the equilibrium from e to $e_1$ (Q to $Q_1$ and P to $P_1$).

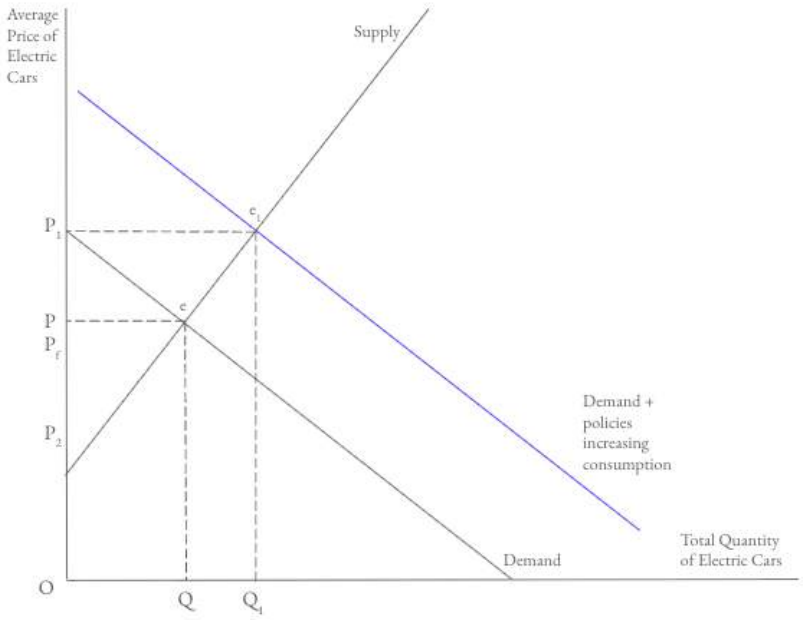



Figure 1:Theoretical results from policies increasing the demand for EVs

Policies around EV production, such as taxes on ICEVs and subsidies on EVs, aim to influence supply. Taxes would impose an additional financial burden on ICEV producers, thereby increasing production costs. Therefore, the greater opportunity costs incentivise a

transition to EVs that yield higher profits. Subsidies are government grants that provide financial assistance to lower the costs of production. Specifically, China's per-unit subsidy of up to 12,600 yuan per new EV battery sold (discontinued in 2022 for purchase tax exemptions) (Yu), as shown in Figure 2, shifts the supply of EVs to the right from e to $e_2$ (Q to $Q_2$ and P to $P_2$) between 2012 and 2015. Other policies increasing EV supply would also see rightward shifts in supply, albeit the exact degree varies based on market reactions.

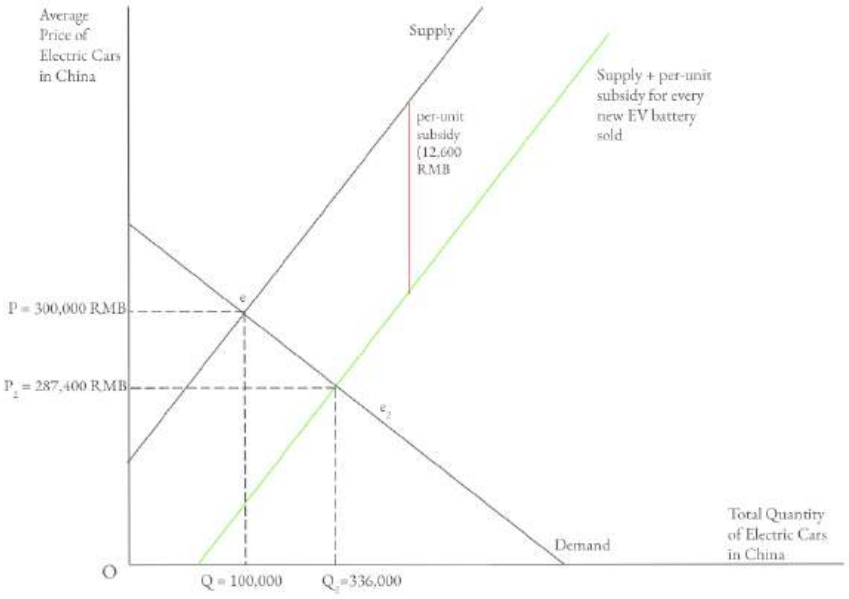



Figure 2: Impact of China's per unit subsidy between 2012 and 2015 as an example illustrating theoretical results from policies increasing the supply of EVs (Interesse),(Yu)

Externalities reveal the differing equilibria between the social equilibrium (marginal social benefit, MSB) and the market equilibrium (marginal private benefit, MPB). In mid-2024, internal combustion engine vehicles (ICEVs) accounted for over 97% of the cars in Singapore (Kuro). ICEVs have negative externalities as consumption has adverse effects on third parties outside the transaction. Therefore, there is overconsumption and market failure where MPB > MSB, leading to welfare loss from allocative inefficiency. This is evident in Figure 3, where the quantity of ICEVs is plotted against their average price in Singapore. There is a welfare loss of aee*, and an overconsumption of QQ*, as the quantity of ICEVs consumed in the free market is above the socially optimal Q*.

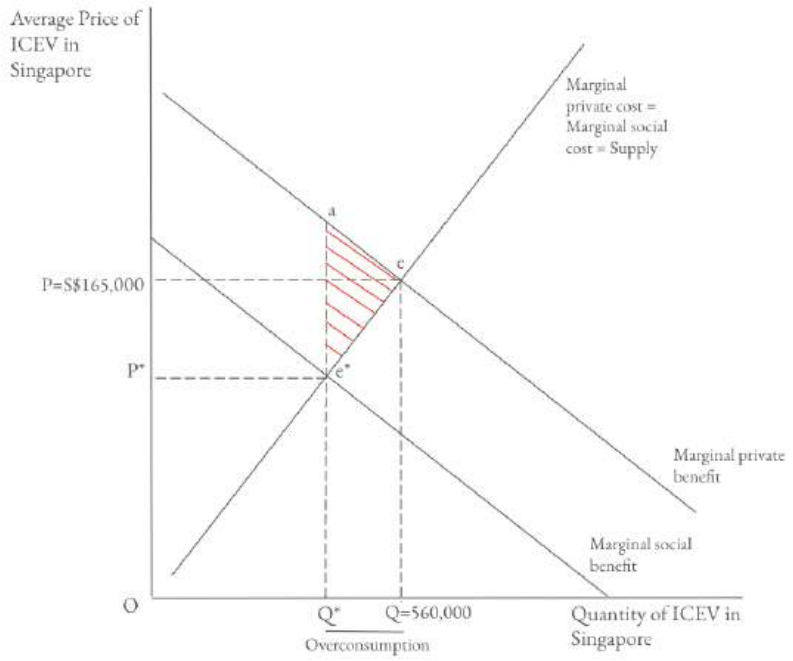



Figure 3: ICEV market externalities in Singapore (Yee),(Joson)

Illustrated in Figure 4, incentivizing consumers to switch from ICEV to EV, be it policies influencing demand or supply, would shift the private demand, MPB, to the left ($MPB_2$). Although EVs are not merit goods, they possess lower welfare losses and can reduce externalities caused by ICEVs. The new equilibrium $e_2$ quantity would reduce to $Q_2$, lowering overconsumption to $Q^*Q_2$, and reducing welfare loss by $acbe_2$ to $be_2e^*$.

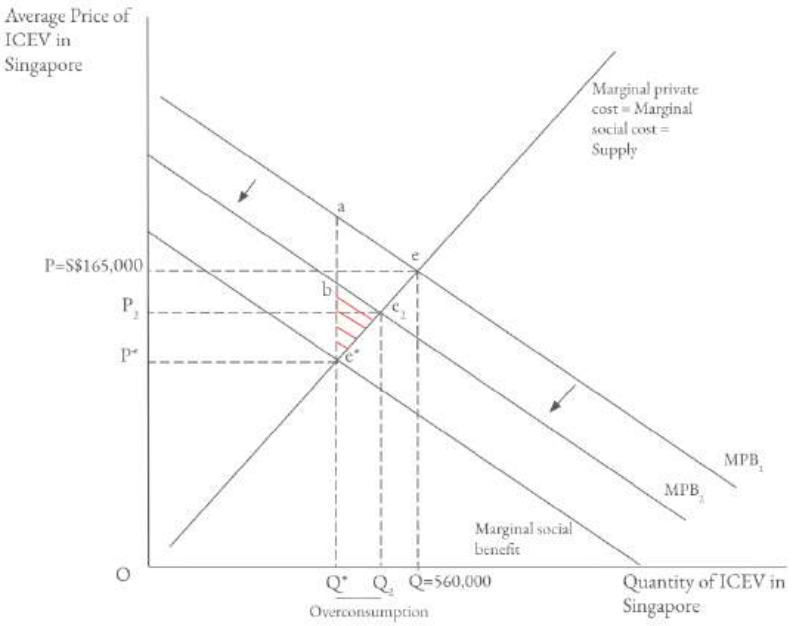



Figure 4: ICEV market externalities in Singapore given a shift to EVs from policies that either increase supply or demand (Yee), (Joson)

Only 4.0% of currently driven cars in Singapore are EVs (Kuro), and replacing all ICEVs with EVs would reduce Singapore's annual emissions by 1.5 to 2 million tonnes of $CO_2$ (Kung). It should be noted that the reduction in emissions stems solely from EVs' higher energy efficiency, as Singapore's electricity is primarily generated from natural gas (ITA). Even so, EVs convert around 60% of grid energy into motion while ICEVs only convert around 20% from gasoline. (EMA). Additionally, EVs have no tailpipe emissions, making them a superior alternative despite Singapore's natural gas energy grid; natural gas also emits 27% less CO2 than gasoline (Swarthmore). Although switching to more renewable energy sources greatly improves EVs' sustainability in Singapore, such investments take far too long and are too expensive.

Instead of creating new policies, Singapore can adopt and adapt already successful policies from countries like Norway and China.

**Economic Analysis**

**A Comparison of EV-Related Policies Between Norway, China, and Singapore**

Norway, China, and Singapore have varying policies that influence the demand and supply in their respective EV markets. The following table outlines their policies.

| | Norway | China | Singapore |
|---|---|---|---|
| Policies aimed at increasing consumption/demand | Rebates to consumers: $27,100 for large EVs and $14,600 for small EVs (Statista) | Vehicle trade-in subsidies: up to RMB 20,000 for scrapping old EVs for new ones (Huang) | EV Early Adoption Incentive (EEAI): rebate of 45% off the Additional Registration Fee (ARF) (IEA) |
| | Tax exemptions from Value Added Tax (VAT) of up to 500,000 Norwegian kroner (Burkandt) | Tax exemptions from purchase tax of 10% for new EVs; no charge for license plate (RMB 50,000+ cheaper) (Zhang, Burke) | Heavy vehicle zero emissions scheme (HVZES): S$40,000 rebate for heavy vehicles paid through 3 years (IEA) |
| | Priority parking, heavily discounted toll rates, and access to bus lanes | Exemptions from parking fees, tolls, and access to bus lanes (Zhang, Burke) | More charging points |
| Policies aimed at increasing production/ supply | National ban on sales of ICEV (Norsk) | Subsidies: RMB 1.65 billion distributed between 2016 to 2020 (Huifeng) | Charging points: new developments need EV charging for at least 20% of parking lots (ITA) |
| | Tighter CO2 emissions standards for manufacturers: average CO2 emission for new car fleets dropped from 60g/km to 45g/km (Fridstrom) | Subsidies: Per unit subsidy of up to 12,600 yuan per EV battery sold from 2012 to 2022 (replaced by dual credit policy and tax exemptions) (Yu) | |

| | | | |
|---|---|---|---|
| | | Dual credit policy: 38% of production must be new EVs, and credits are distributed to be traded. (Cui and Chu) | |

Table 1: Policies aimed at increasing consumption and production for Norway, China, and Singapore

**Norway's Policies**

Norway adopted a mix of subsidies and EV benefits to encourage its consumption. The transition began in the early 1990s and targeted a complete EV transition by 2025 (Norsk). Norway heavily emphasised policies that increased demand, making the utility of EV consumption higher than that of ICEVs.

These policies decrease the fixed costs of purchasing EVs and reduce the variable costs through heavily discounted tolls. Because the actual performance of EVs and ICEVs is similar, these policies led EVs to be the rational superior choice. Additionally, by penalising production of ICEVs, manufacturers are encouraged to use preexisting capital for EVs, shifting supply to the right. The results show a successful transition from ICEV to EV driven by a series of strong government interventions over the last two decades. It is worth noting that Norway's distinct green culture played a role in EV adoption, which is difficult for other countries to replicate.

| Results of Norway's policies | |
|---|---|
| Demand | Monetary incentives like rebates and tax exemptions allow Norwegians to save roughly 30% on vehicle purchase prices (Ayvens) |
| | Heavily discounted toll rates; roughly 25,000 NOK saved annually in running costs (Goehring) |
| Supply | As of 2025, EVs represent 88.9% of new cars sold (Finnerty) |
| | Norway's EV market reached 98% of new car sales in certain months in 2024 (Finnerty) |
| | Global EV production also increased by 25% between 2023 and 2025 (IEA). |

Table 2: Results of Norway's policies in increasing demand and supply of EVs

**China's Policies**

China launched Ten Cities, Thousand Vehicles in 2009 to promote the EV sector (Marquis), but it was further streamlined when President Xi took office. Since then, China has become the global leader in EVs, with policies aimed at increasing both consumption and production.

China's policies are more supply-oriented, aimed at increasing EV production. The most important policy to increase EV supply is the dual credit policy, which encourages innovation and the high-quality/quantity of EVs. Deficits from fuel consumption can only be offset by new

electric vehicle (NEV) credits. To earn credits, firms must fulfil conditions such as a minimum range of 137km for battery-electric vehicles (Autovista). Firms can choose to either bank or trade it among manufacturers, with costs ranging from 2,500 - 3,000 RMB per credit (Autovista). This reduces welfare loss stemming from asymmetric information when governments intervene. Instead, it encourages market forces to determine the equilibrium price of credits based on supply and demand.

| Results of China's policies | |
|---|---|
| Demand | China gives EVs access to certain roads and areas, like the Beijing city center, that ICEVs cannot access on certain days (Xian) |
| | Largest number of charging stations in the world, with each being universally supportive for EVs of different manufacturers (Global Times). |
| Supply | China is the world's largest producer of EVs and the first country to produce 10 million new energy vehicles (NEVs) in 2024 (Autovista). |
| | EV market saw continuous growth, expanding by 38.2% in 2024 alone (Zhang) |
| | EV sales take up 40.9% of total auto sales (a 9% increase from 2023) (Argus) |
| | Dual credit policy is expected to drive NEV market share to 22% by 2025 (Zhang). |

Table 3: Results of China's policies in increasing demand and supply of EVs

**Singapore's Policies**

Whereas Norway and China began shifting from ICEV to EV in the 2000s, Singapore only began pushing policies in 2020. Hence, there has yet to be a significant effect on consumer shifts toward cleaner, energy-efficient cars. The government has attempted to encourage EV adoption, but there are no permanent incentives (HSEV). The Early Adoption Incentive (EEAI), for example, is planned from 1 January 2024 to 31 December 2025 with no announced extension. Granted, this does increase EV consumption, but a lack of direction results in less stability for consumers and producers: Consumers are unsure whether Singapore will continue to promote EVs, and producers reduce capital investment due to higher risk.

Infrastructure-wise, the government targets 60,000 public EV charging stations by 2030 (DHL) and up to S$4,000 per charging station for the first 2,000 charging stations in private residences (Ramamurthy). But many residential areas still lack accessible charging stations, especially public housing (HDB) flats, where over 80% of Singaporeans reside. 2,000 is not enough to meet the growing demand in more populated areas (EVFY). Additionally, in certain older buildings, power supplies are in dire need of an upgrade to support the new-generation EV chargers.

**Data Analysis**

An analysis comparing Norway, China, and Singapore was conducted regarding current EV progress. Various statistical data collected from 2019 to 2024, along with a simple linear regression, were used to assess trends. Linear regression explains the extent to which policies in Norway and China are successful in increasing demand for EVs, and visually illustrates how much Singapore still needs to improve. Descriptive statistics were used to summarize and organize data by reporting annual totals, calculating percentage changes, and providing a description of the changes. Inferential statistics are applied with regression analysis to examine the relationships. Google Sheets and Python were used to calculate linear regression values for R and $r^2$.

**Total EV to ICEV historical proportions**

| Country | Metric | 2019 | 2020 | 2021 | 2022 | 2023 | 2024 |
|---|---|---|---|---|---|---|---|
| Norway | New EV sales share | 55.90% | 74.70% | 86.10% | 88.58% | 82.40% | 88.90% |
| | EV fleet share | 10.00% | 15.00% | 20.00% | 25.00% | 28.10% | 31.00% |
| China | New EV sales share | 4.90% | 5.40% | 15.70% | 26.00% | 33.00% | 47.90% |
| | EV fleet share | 2.50% | 3.00% | 4.50% | 6.00% | 8.00% | 8.90% |
| Singapore | New EV sales share | 0.10% | 0.20% | 0.34% | 11.70% | 18.10% | 32.60% |
| | EV fleet share | 0.03% | 0.07% | 0.18% | 1.00% | 1.83% | 4.00% |

Table 4: Norway, China, and Singapore's new EV sales share and EV fleet share in comparison to ICEV (Kuro) (Straits Times) (Tribdino) (Lim) (Norsk) (Finnerty) (Statista) (IEA) (Andrews)

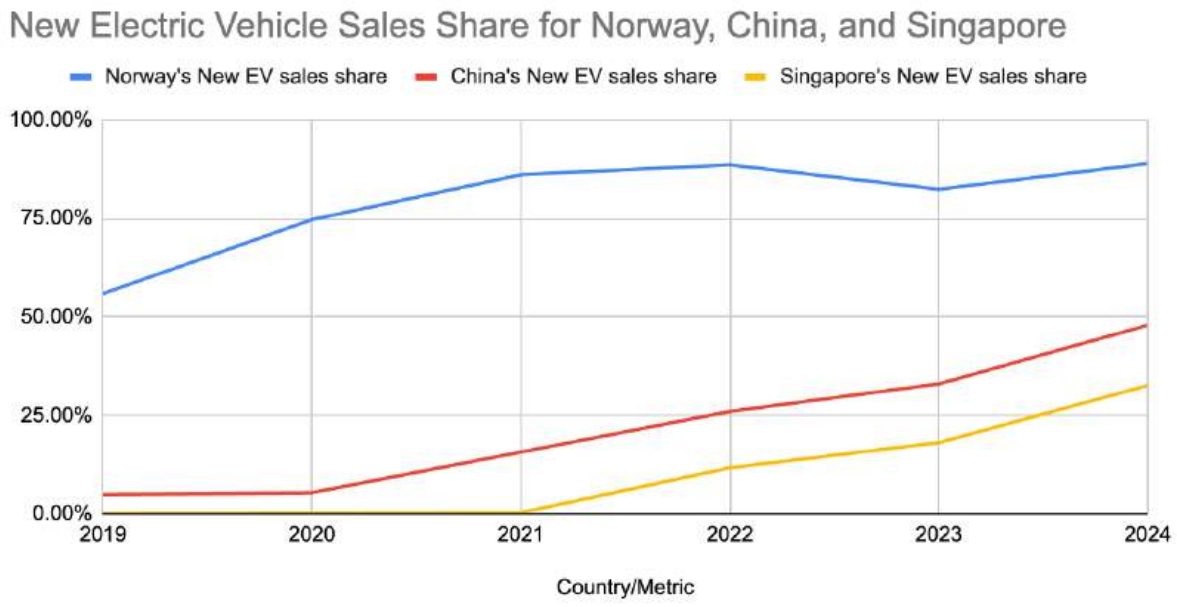



Figure 5: New EV sales share comparison for Norway, China, and Singapore between 2019 and 2024

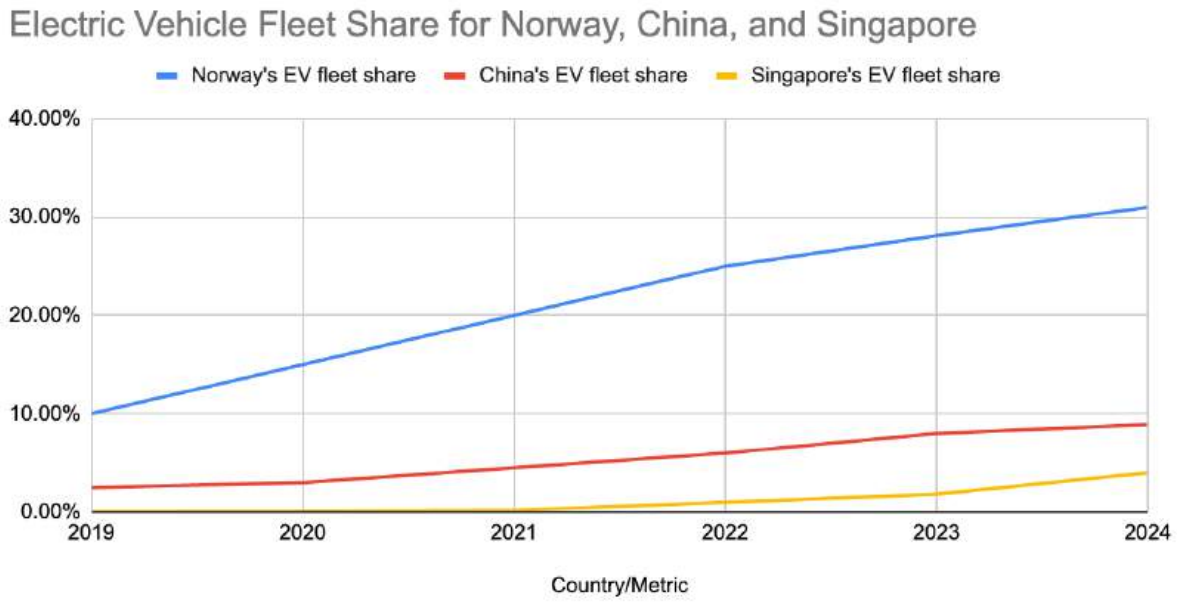



Figure 6: EV fleet share comparison for Norway, China, and Singapore between 2019 and 2024

| Country | Metric | r | $r^2$ |
|---|---|---|---|
| Norway | New EV sales share | 0.80 | 0.65 |
| | EV fleet share | 0.99 | 0.99 |
| China | New EV sales share | 0.98 | 0.96 |
| | EV fleet share | 0.99 | 0.98 |
| Singapore | New EV sales share | 0.93 | 0.86 |
| | EV fleet share | 0.90 | 0.80 |

Table 5: The R and $r^2$ values for New EV sales and EV fleet share in Norway, China, and Singapore using data from 2019 to 2024

Table 4 presents a descriptive analysis of secondary data on EV market demand, expressed as a percentage of sales and fleet share. Figures 5 and 6 are line graphs that visually represent these trends. China had the largest year-on-year increase in EV demand (annual sales share) over the last 5 years, followed by Norway and Singapore. Norway, however, is the only country with an annual sales share higher than ICEV's, which nears 90%. However, both China and Singapore saw large growth in recent years from government intervention. The R and $r^2$ values in Table 5 indicate a strong positive trend across all these countries. The lack of anomalies in Norway and China demonstrates the sustained effectiveness of their policies. As Singapore began its EV transition much later, adopting and adapting proven policies to its own circumstances would help it catch up with the world's EV leaders. Absolute per capita values of

EV adoption would be lower in Singapore due to the smaller population. But high population concentrations suggest that relative percentages could be improved more quickly.

Countries like Norway and China did not publish EV fleet share as a whole number, so a general calculation was conducted with the following formula.

$$EV\ fleet\ share\ (\%)\ =\ \frac{Total\ EV\ stock}{Total\ car\ fleet} \times 100$$

Although having grown by 3.97% over the past 5 years, Singapore has the lowest EV fleet share compared to Norway and China. This occurs because cars are durable goods that can be used for many years, and many Singaporeans do not require replacement. As the new EV sales share increases, the EV fleet share would gradually increase.

**Policy Recommendations Aimed at Increasing EV Demand**

Singapore's land area prevents manufacturing of EVs. Nevertheless, Singapore can adopt and adapt successful policies from Norway and China that have increased EV demand according to its unique capacity. Given the diverse economic and cultural circumstances, recommendations for adapting these policies to Singapore's conditions will be provided.

**Financial benefits**

Norway's financial-driven policies aimed at increasing demand mainly include VAT exemption and purchase tax exemptions. It made EVs a cheaper choice than ICEVs and led to a 33% year-on-year increase in EV demand between 2019 and 2024 (Norsk). In comparison, Singapore has far more taxes than Norway, including the Certificate of Entitlement (COE) required to own a car, with costs ranging from S$98,000 to S$116,000 (Kok). Additionally, there is an Additional Registration Fee and the new road tax that charges vehicles based on the number of miles driven per year.

Meanwhile, China implemented a policy that shifted demand by offering license plates at no charge, reducing the cost of owning an EV by RMB 50,000+ compared to ICEVs (Zhang). If Singapore can lower the COE required for EVs and further reduce the road tax rate, the fixed and variable costs of EVs would be lower than those of ICEVs. This would greatly increase demand for EVs, as they are now the dominant choice.

**Nudges**

Norway and China's incentive-based nudges, such as cheaper parking and bus lane access, could be easily implemented in Singapore. It is especially effective in China, where demand for EVs increased by 43% over 5 years (IEA). Although bus lane access could increase traffic during peak hours, allowing EVs to use bus lanes outside peak hours without fines would improve their convenience for frequent drivers. Parking fees in the city center can sometimes reach S$10 per hour (LTA). As HDBs have free parking on Sundays, extending the benefit to Saturdays for EVs, and further reducing the general EV parking fee by 25% islandwide would be attractive for frequent travelers. At the same time, it does not heavily compromise government

revenue from parking. Free municipal access, in general, reduces the variable costs of owning an EV through small perks and makes it a superior investment for consumers who drive frequently.

**EV Charging Ports**

The biggest weakness of EVs lies in convenience: longer charging times and fewer charging stations compared to ICEVs, making it difficult for consumers without charging access nearby. One solution is to develop batteries that can withstand higher charging voltages for faster charging. However, as Singapore lacks sufficient land, this is not feasible. Another solution that Norway and China have been adamant about involves increasing charging convenience by installing more charging ports in recent years. Because Singapore has a higher population density than Norway and China, installing charging ports should be easier.

| Primary data for charging stations in Singapore's public and private areas through observation | | | |
|---|---|---|---|
| Area | Land use type | Number of charging stations | Usage of charging stations |
| West coast | Residential, a large span of public housing (HDBs) and condominiums (private housing) | 2 EV chargers for the biggest HDB (a marketplace, three food centers, two shopping malls, and over 15 residential buildings) | Always full during 5 separate visits at different times |
| | | 0 EV chargers in 5 separate condominiums, where 3 were built over 10 years ago, and 2 were less than 5 years ago | NA |
| Jurong East | Residential, a large span of public housing (HDBs) | 4 EV chargers for the biggest HDB (a marketplace, one food center, 20 residential buildings) | Mostly full during 5 separate visits at different times |
| Orchard | Commercial/ retail with condominiums | 4 EV chargers in 2 of the biggest shopping malls | Mostly full during 5 separate visits at different times |
| | | 0 EV chargers in 3 separate condominiums that were all built over 10 years ago | NA |

Table 6: Primary data collected through observation on the number of charging ports and their usage in different areas and land types in Singapore

Singapore's EV charging ports expanded from 2.74 stations/$km^2$ to 21.0 stations/$km^2$ between 2019 and 2024 (Mordor). However, primary evidence based on observation from Table 6 indicates that more charging stations are necessary to incentivize ICEV to EV transitions. It should be noted that government-managed public areas have more EV charging stations, with

each station occupying the best parking spaces to incentivise EV consumption for convenience. However, there is a lack of infrastructure due to Singapore's high population density. For private areas outside direct government jurisdiction, the EV Common Charger Grant had co-funded over 1,200 chargers in private residences, such as condominium complexes by 2024 (Loi). Yet, all the private housing visited by the author contains zero EV charging stations.

Singapore has a law requiring 1 EV charger for every 25 parking lots in public or common parking areas, such as HDBs (Choo). However, observation data from the West Coast and Jurong East areas indicate the need for stronger regulations, as they do not comply with the legislation. Furthermore, this law does not affect private residential areas where high-income earners who purchase the most cars live (LTA). Extending it to condominiums would increase EV convenience, leading to greater use and satisfaction, as around 70% of EV owners prefer home charging (Norsk).

Figure 7 depicts private EV charging stations in Singapore. There is a concentration near city centers and far less in residential areas in the north and the west. The primary data the author collected are similar and highlight the need for more spread-out stations.

Figure 7: EV charging stations in Singapore (SPGroup)

The current Singapore EV fleet share is only 4%, 8 times lower than Norway's and 2 times lower than China's (IEA). Furthermore, even though Singapore is more concentrated, meaning it is easier to build chargers around the island, its charging ports per 1,000 people is 2.7 (Yufeng), far behind Norway's 4.9 (ICCT). Empirical evidence suggests that charging availability reduces anxiety by 40-60%, thereby increasing EV satisfaction and the likelihood of transitions from ICEV to EVs (Kazempour). Thus, increasing the number of charging ports per 1,000 people would accelerate EV adoption by increasing convenience and car range, driving higher EV consumption.

In general, primary and secondary data support the hypothesis that government interventions such as COE reductions, free parking, bus lane access, and the installation of more

EV charging ports increase demand for EVs in Singapore. This would then reduce welfare losses from pollution generated by the negative externalities of ICEVs.

**Discussion**

Arising from the investigation are several unresolved questions. Firstly, only policies from Norway and China were adopted and adapted for Singapore; other countries, such as Canada and the US, with differing policies, have proved effective. Investigating more countries' policies would help reach a definitive conclusion on policies to increase Singapore's EV adoption. Given the different geographic and economic circumstances in Norway and China, there is a consistent risk in transferring a successful policy, even with attempts at adaptation. Nevertheless, the policy-recommendation aspect of this research was intriguing, as it involves transferring existing policies and legislation to suit Singapore's distinct circumstances.

Conversely, no primary data were collected in Norway and China due to observation's limitations on movement, leading to a reliance on secondary data that may be prone to error. The number of data points from primary data is lacking due to observation's time-intensive nature, forcing extrapolation when taking into account the entirety of Singapore. This would decrease the accuracy of primary data, even with secondary data to back up observations. However, the sample size could be considered representative, given its relative size compared to Singapore's small land area. Strong benchmarks also increase reliability by reducing personal bias from observation. Nonetheless, expanding the scope of primary data points to 20 condominiums and 5 HDBs in separate areas would improve accuracy. It is worth noting that the majority of secondary data's weaknesses—relevance and bias—were reduced, as most sources were published within the last 3 years and over 40 sources with differing perspectives were used to minimise bias. The central theme of this research is reducing the negative externalities of ICEVs through EVs. Therefore, the focus primarily lies on increasing EV adoption to steer consumers away from less energy- and climate-friendly ICEVs. However, there were no policy recommendations for directly internalising the cost of ICEVs through production taxes. The EV fleet share statistic is also a rough estimate based on the formulas, therefore ignoring uncertainties and assuming ceteris paribus.

Another overlooked issue is the colder climates in Norway and China. Lithium-ion batteries in EVs typically experience chemical wear at temperatures above 35 degrees Celsius (Qu). Hence, it reduces the longevity of EV batteries and negatively influences their demand in Singapore.

Making EVs far cheaper than ICEVs might encourage more citizens who are taking public transportation to purchase a car, causing traffic congestion and loss of productivity in the workforce. Even small nudges, like allowing EVs to access bus lanes during certain non-peak traffic times, could be challenging to regulate and increase the risk of crashes.

**Conclusion**

The author concludes that adapting Norway and China's policies to increase EV demand is a viable strategy and will be largely successful in reducing welfare losses and pollution from ICEVs.

Existing policies in Singapore, especially EV rebates, are similar to those in Norway and China, as all three countries aim to increase EV adoption. Additionally, they all share a similar political will, encouraging a shift in demand from ICEVs to EVs to reduce negative externalities. Consumer perceptions of EVs are also similar, as people have come to realise their environmental benefits and are welcoming these changes. These similarities would ease the adoption of green policies from Norway and China to increase Singapore's EV consumption.

There is a correlation between building more EV charging stations and increased EV adoption due to lower utility costs, fewer worries, and greater convenience. However, certain factors, such as climate conditions, were not considered. While this does not render building more EV stations ineffective, it decreases the correlation between the two. This is evident in Singapore as hotter climates generally reduce battery life of EVs, causing greater maintenance costs - leading to the hypothesis being only partially validated.

The opportunity cost of government spending should also be noted, as these policies could cause budget deficits. Hence, the policy recommended, based on successful implementations in Norway and China and adapted for Singapore's circumstances, has low to no costs. The author recommends financial incentives and nudges: reduced cost of ownership (lower COE cost), less parking fees, bus lane access during non-peak hours, and more EV charging stations for private/public areas. Reducing COE costs lowers EVs' fixed costs, making them a cheaper alternative to ICEVs. Parking fee incentives decrease variable costs of EVs, increasing their attractiveness for frequent drivers. Bus lane access and more EV charging stations make EVs more convenient and incentivise a transition from ICEVs. True, some of these reduce government revenue, but they are effective at reducing pollution. These recommended forms of government intervention would therefore reduce the negative externalities and welfare losses generated by the more polluting ICEV by encouraging a transition towards greater consumption of the more environmentally friendly EV.

**Works Cited**


Australian Fleet Management Association. "Singapore: ICE Vehicle Registrations Ending by 2030?" AFMA, 2024, afma.org.au/singapore-ice-vehicle-registrations-ending-by-2030/.

Ayvens. "Car Cost Index 2025: Total Cost of Ownership." Ayvens, 2025, www.ayvens.com/en-cp/blog/total-cost-of-ownership/car-cost-index-2025/.

Bai, Y. "Life after Subsidies for China's EVs." Dialogue Earth, 2023, dialogue.earth/en/business/life-after-subsidies-for-chinas-evs/.

"Bill Proposes All New Buildings Must Install 1 EV Charger Every 25 Parking Lots." Today, 2023, www.todayonline.com/singapore/bill-proposes-all-new-buildings-must-install-1-ev-charger-every-25-parking-lots-2040276.

ChargeWhere. "The Rise of Electric Vehicles in Singapore: A Look at Ownership and Popular Brands." ChargeWhere, 2024, chargewhere.sg/the-rise-of-electric-vehicles-in-singapore-a-look-at-ownership-and-popular-brands/.

"China Builds World's Biggest EV Charging Network, in Stark Contrast to US Failure." Global Times, 2024, www.globaltimes.cn/page/202406/1314382.shtml.

"China Expands EV Charging Infrastructure in 2024." Argus Media, 2024, www.argusmedia.com/en/news-and-insights/latest-market-news/2650730-china-expands-ev-charging-infrastructure-in-2024.

"China News Service." ECNS.cn, 19 Nov. 2024, www.ecns.cn/cns-wire/2024-11-19/detail-ihekectp0360726.shtml.

"China Reveals Past EV Subsidies: Tesla, State Giants, Where Did the Money Go?" South China Morning Post, 2023, www.scmp.com/economy/china-economy/article/3322666/china-reveals-past-ev-subsidies-tesla-state-giants-where-did-money-go.

DHL. "Singapore's Transition to Electric Vehicles." DHL, 2024, www.dhl.com/discover/en-sg/logistics-advice/logistics-insights/singapores-transition-to-electric-vehicles.

DrWealth. "How Much Does an Electric Car Cost in Singapore?" DrWealth, 2025, drwealth.com/how-much-does-an-electric-car-cost-in-singapore/.

Elbil Norge. "Norwegian EV Market." Elbil Norge, 2024, elbil.no/english/norwegian-ev-market/.

"Electric Vehicle Population in Singapore More than Doubled in 2021." The Straits Times, 2021, www.straitstimes.com/singapore/transport/electric-vehicle-population-in-singapore-more-than-doubled-in-2021.

"Electrified Cars Make Up Nearly 9% of China's Total Car Fleet." CarNewsChina, 2025, carnewschina.com/2025/01/17/electrified-cars-make-up-nearly-9-of-chinas-total-car-fleet/.

"EV Chargers Now Available in About Half of All HDB Carparks." The New Paper, 2023, www.tnp.sg/news/singapore/ev-chargers-now-available-about-half-all-hdb-carparks.

EVfy. "Challenges Facing EV Charging Infrastructure in Singapore and Solutions." EVfy, 2025, evfy.sg/blog_posts/challenges-facing-ev-charging-infrastructure-in-singapore-and-solutions.

Gorozen. "The Norwegian Illusion." Gorozen, 2024, blog.gorozen.com/blog/the-norwegian-illusion.

Gridserve. "Why Norway Leads the World in EV Adoption." Gridserve, 2025, www.gridserve.com/why-norway-leads-the-world-in-ev-adoption/.

Harvard Business School. Electric Vehicles. 2013, www.hbs.edu/ris/Publication%20Files/Electric%20Vehicles_89176bc1-1aee-4c6e-829f-bd426beaf5d3.pdf.

Hirsch, L. "Norway: Nearly All New Cars Sold in 2024 Were Fully Electric." Reuters, 2025, www.reuters.com/business/autos-transportation/norway-nearly-all-new-cars-sold-2024-were-fully-electric-2025-01-02/.

HSEV. "Government Subsidies in Singapore for Electric Vehicles." HSEV, 2024, hsev.co/government-subsidies-in-singapore-for-electric-vehicles/.

International Council on Clean Transportation. "What's Next for China's PHEV Market?" ICCT, 2025, theicct.org/whats-next-for-chinas-phev-market-sept25/.

International Energy Agency. "Trends in Electric Car Markets." Global EV Outlook 2025, 2025, www.iea.org/reports/global-ev-outlook-2025/trends-in-electric-car-markets-2.

---. "Trends in the Electric Car Industry." Global EV Outlook 2025, 2025, www.iea.org/reports/global-ev-outlook-2025/trends-in-the-electric-car-industry-3.

KPMG Singapore. "5 FAQs on EVs in Singapore: KPMG's Viewpoint." KPMG Singapore, 2022, kpmg.com/sg/en/home/media/press-releases/2022/09/5-faqs-on-evs-in-singapore-kpmg-s-viewpoint.html.

Land Transport Authority Singapore. "Guidelines for the Minimum EV Charging Provisions in Developments (as of 9 May 2025)." LTA, 2025, www.lta.gov.sg/content/dam/ltagov/industry_innovations/Technologies/Electric_Vehicles/PDF/Guidelines%20for%20the%20Minimum%20EV%20Charging%20Provisions%20in%20Developments%20(as%20of%209%20May%202025).pdf.

---. "Our EV Roadmap." LTA, 2025, www.lta.gov.sg/content/ltagov/en/industry_innovations/technologies/electric_vehicles/our_ev_roadmap.html.

---. "Parking Rates." OneMotoring, 2024, onemotoring.lta.gov.sg/content/onemotoring/home/owning/ongoing-car-costs/parking/parking_rates.8.html.

---. "Transitioning to EVs." LTA, 2025, www.lta.gov.sg/content/ltagov/en/industry_innovations/technologies/electric_vehicles/transitioning_to_evs.html.

Mordor Intelligence. "Singapore Electric Vehicle Market." Mordor Intelligence, 2023, www.mordorintelligence.com/industry-reports/singapore-electric-vehicle-market.

"More than 15,300 EV Charging Points Installed Islandwide; Target of 60,000 by 2030." The Straits Times, 2025, www.straitstimes.com/singapore/transport/more-than-15300-ev-charging-points-installed-islandwide-target-of-60000-by-2030.

National Alliance on Mental Illness. "Mental Health Conditions." NAMI, 2025, www.nami.org/Learn-More/Mental-Health-Conditions.

"Singapore Certificate to Own Car Now Costs $106,000." Reuters, 2023, www.reuters.com/business/autos-transportation/singapore-certificate-own-car-now-costs-106000-2023-10-04/.

"Singapore's Electric Vehicle Dreams Now Accelerating Faster." CleanTechnica, 2025, cleantechnica.com/2025/06/30/singapores-electric-vehicle-dreams-now-accelerating-faster/.

"Singapore's EV Adoption Rate Nearly Doubles in 2024 with One out of 3 Cars Sold from Jan-May Being Electric." PaulTan's Automotive News, 2024, paultan.org/2024/07/05/singapores-ev-adoption-rate-nearly-doubles-in-2024-with-one-out-of-3-cars-sold-from-jan-may-being-electric/.

SP Group. "Charger Locations." SP Group, 2025, www.spgroup.com.sg/sustainable-energy-solutions/electric-vehicle-solutions/charger-locations.

Statista. "Fleet Share of Plug-in Hybrid and Battery Electric Cars in the Nordics." Statista, 2025, www.statista.com/statistics/1538278/fleet-share-plug-in-hybrid-battery-electric-cars-nordics/.

---. "Norway: Average EV Financial Benefits by Vehicle Size." Statista, 2023, www.statista.com/statistics/1332190/norway-ev-financial-benefits-by-vehicle-size/.

Transportøkonomisk institutt. Sustainability, vol. 13, no. 1346, 2021, www.toi.no/getfile.php/1355696-1616160861/mmarkiv/Forside%202021/sustainability-13-01346.pdf.

WC Shipping. "Norway's EV Market 2025: Policies, Trends, and Global Trade Opportunities." WC Shipping, 2025, www.wcshipping.com/blog/norways-ev-market-2025-policies-trends-and-global-trade-opportunities.

Zhang, J., et al. Australian National University, 2024, crawford.anu.edu.au/sites/default/files/2025-02/acde_td_zhang_burke_wang_2024_01.pdf.

**What Drives Entrepreneurial Intention in Indonesia? The Role of Confidence Among Young Adults By Ahmad Nagara Latief**


## Abstract

Entrepreneurship is often framed as a driver of economic development within developing economies, and still, little evidence on the quantitative side exists on what exactly builds and shapes entrepreneurial intention amongst young adults in Indonesia. This study examines entrepreneurial intention among Indonesians aged 21 to 30 and identifies the factors most strongly associated with it. The study utilizes a cross-sectional survey, distributing it online through Google Forms using both snowball and convenience sampling, which would end up yielding 49 valid responses. Participants rated their interest in entrepreneurship, their perceived likelihood of starting a business, and their confidence in starting and managing one on five-point scales, and selected from a list of motivating and demotivating factors.

The study found Interest($M = 3.69$) as well as likelihood ($M = 3.61$) being moderately high, while other factors like confidence were significantly lower($M = 3.22$). The most frequently endorsed motivators were financial independence (82%) and providing opportunities to others (59%), while the most common barriers related to perceived capability: lack of experience (65%), fear of failure (57%), and lack of capital (49%). A simple linear regression found that confidence was the strongest predictor of likelihood, explaining 43% of its variance ($R^2 = 0.43$, $p < .001$), more than double the variance explained by interest. Respondents from families that own a business also reported significantly higher confidence than those without ($p = .034$).

These findings suggest that entrepreneurial intention in this group is limited less by a lack of interest than by a gap in self-efficacy, and that interventions building confidence may be more effective than motivational or financial incentives alone. Given the small, non-probability sample, results are interpreted as exploratory.



**Keywords** entrepreneurial intention; self-efficacy; Theory of Planned Behavior; Indonesia; young adults; family business background; perceived behavioral control


## 1. Introduction

Micro, small, and medium enterprises (MSMEs) form the backbone of the Indonesian economy, accounting for approximately 98.68% of total businesses and 97% of the workforce (DevTech Systems), yet the country continues to lag its regional neighbors in entrepreneurial activity. Only 3.47% of Indonesia's population are entrepreneurs, compared with 8.6% in Singapore and over 4.5% in Malaysia and Thailand, and the government has set a target of reaching an 8% national entrepreneurship rate (Tempo). Closing this gap is a priority, policy-wise, and one of the most significant reasons in this regard would be the lack of youth employment. Unemployment rates within Indonesians aged 15-24 as of August 2025 had risen to 16.9%, which is far above the national average of 4.85% (Badan Pusat Statistik).

Entrepreneurship is often seen as both an opportunity and a hedge against limited formal-sector employment among younger adults entering the workforce. Understanding what drives, or discourages, entrepreneurial intention in this group is therefore both academically relevant and practically urgent.

According to Ajzen's Theory of Planned Behavior, entrepreneurial intention is regarded as the strongest predictor of whether an individual will actually start a business. This specific framework remarks that intention is shaped by three main factors, namely attitude toward behavior, subjective norms, and perceived behavioral control, the last of which is closely related to self-efficacy, or confidence in one's own ability. Eventual research would ultimately find this component among the strongest predictors of entrepreneurial intention, specifically within countries that are developing countries in Southeast Asia (Virasa et al.). Scholars have also consistently recognized a gap between intention and action. This means that individuals constantly have entrepreneurial intentions yet never act on them, mostly due to obstacles such as lack of experience, limited mentorship, and especially limited capital.

Studies and evidence coming from Indonesia are constantly reflecting this pattern as well. Applying Ajzen's framework towards Indonesian students have only found that entrepreneurship education raises intention indirectly, as they build psychological states, and especially the self-efficacy/confidence that would actually drive it (Mahendra et al.; Khoiriyah et al.), while family background has been identified as one real-world source of that confidence (Georgescu and Herman). These findings, examined in the literature review that follows, point to confidence in one's own capability, rather than knowledge or interest alone, as central to whether interest becomes intention.

Despite all the research done, two gaps still remain. First, motivating and demotivating factors are usually studied separately rather than as a whole, which inevitably makes it difficult to see the overall balance of forces that shape intention. Second, studies done in the past focus mostly on current students, which would ignore young adults aged 21-30, leaving them under-examined in this regard, even though this is the stage at which individuals most actively weigh entrepreneurship against traditional employment. By studying young adults aged 21-30 within Indonesia, this study addresses both of those gaps and tests directly whether confidence, as a measure of perceived behavioral control, is the strongest predictor of their intention to start a business.

### 1.1 Literature Review

The study of entrepreneurial intention rests on the premise that intention is the most immediate antecedent of entrepreneurial behavior, and the dominant framework for explaining how it forms is Ajzen's Theory of Planned Behavior. Ajzen proposes that intention is the joint product of three determinants: attitude toward the behavior, subjective norms reflecting perceived social pressure, and perceived behavioral control, meaning a person's belief in their own capacity to carry out the behavior. Crucially, Ajzen treats perceived behavioral control as functionally equivalent to self-efficacy, which makes the framework directly applicable to

entrepreneurship: an individual's confidence in their ability to start and manage a business is, in Ajzen's terms, a measure of perceived behavioral control. This equivalence is the theoretical anchor of the present study.

Despite the framework specifying three determinants, they don't contribute equally, and the control component usually carries the greatest predictive weight. Zhang, Wang, and Owen tested the full model on 275 university students and found that subjective norms and perceived behavioral control were positively correlated with entrepreneurial intention, while attitude, the evaluative and interest-like component, had little to no significant effect. This would suggest that a favorable disposition toward entrepreneurship is not, on its own, sufficient to generate intention; what distinguishes those who intend to act is their belief in their own capability. From an early stage, then, the literature signals that confidence, rather than mere attraction to the idea of entrepreneurship, may be the operative variable.

Assuming perceived behavioral control is so important, an inevitable question would be where it came from, and much of the literature approached this through entrepreneurship education. In this regard, the evidence from Indonesia is relevant and significant. A study conducted in Malang, Indonesia, by Mahendra, Djatmika, and Hermawan involved 224 management students, finding that entrepreneurship education has little to no direct effect on intention, as its influence operated almost entirely through the psychological status of motivation and attitude, both accounting for a large share of the variance. Remeikiene, Startiene, and Dumciuviene reached a compatible conclusion in a Lithuanian sample, arguing that intention is driven chiefly by traits such as self-efficacy and behavioral control, and that these traits can be cultivated through education. Together, these studies reframe education not as a direct cause of intention, but as a means of building the psychological attributes that actually produce it.

This raises the question of which attribute matters most, and the most closely matched evidence points to self-efficacy. In a 2025 study of 226 vocational students in Malang, Khoiriyah, Pratikto, and Dhewi found that entrepreneurship education again had no significant direct effect on intention ($p = 0.095$); instead, entrepreneurial self-efficacy significantly predicted intention ($\beta = 0.349$, $p < .001$) and fully mediated the education–intention relationship ($R^2 = 0.475$). The same hierarchy appears beyond Indonesia. Examining 327 Chinese college students, Liu, Lin, Zhao, and Zhao found that although education did have a significant direct effect, self-efficacy was by far the stronger predictor. Across these studies, the pattern is consistent: self-efficacy is not merely one determinant among several, but the dominant one, precisely the claim the present study sets out to test.

If self-efficacy is the constraining factor, then it is necessary to comprehend what creates it, and one factor that comes up again and again is the family background of an individual. Studying 617 Romanian students, Georgescu and Herman discovered that students who came from an entrepreneurial family background had significantly higher intentions, which were due to the informal education received by the children of entrepreneurs in the family setting. The researchers also concluded that the importance of formal education was higher for non-entrepreneurial students, which indicated that family background partially replaces

classroom education. Even though intention was measured instead of confidence, this is undoubtedly a self-efficacy pathway, and it confirms the hypothesis of the current research that family business background enhances confidence.

In order to connect these results to the case of Indonesia, it is necessary to look at them on a regional level. In the study of six ASEAN economies, Virasa, Sukavejworakit, and Promsiri revealed that perceived behavioral control was always one of the highest indicators when predicting entrepreneurial intentions in Southeast Asian countries, which included Indonesia.

Taken together, the literature builds a coherent case: intention is governed by perceived behavioral control (Ajzen); the control component consistently outperforms attitude (Zhang et al.); education shapes intention only indirectly, by building psychological traits (Mahendra et al.; Remeikiene et al.); self-efficacy is the trait that does the work (Khoiriyah et al.; Liu et al.); family exposure is one real-world source of it (Georgescu and Herman); and the pattern holds across the region (Virasa et al.). Yet a gap remains. Almost all of this evidence is drawn from student populations, leaving the early-career phase largely unexamined, even though young adults aged 21 to 30 face the entrepreneurial decision under markedly different conditions. Moreover, motivating and demotivating factors are typically studied in isolation from the psychological predictors of intention. The present study addresses both gaps by examining entrepreneurial intention among Indonesian young adults aged 21 to 30 and testing directly whether confidence is the strongest predictor of their intention to start a business.

## 2. Methodology

### 2.1 Objective and Aim

The study's main objective is to measure entrepreneurial intention within individuals between the ages of 21 and 30 in Indonesia, as well as to locate the factors that motivate or demotivate them from pursuing entrepreneurship. Through comparing factors such as motivation, demotivation, and readiness indicators across demographic groups, the study identifies the factors most strongly associated with shaping entrepreneurial intention and which groups face the greatest barriers.

- Acquire survey data from Indonesian young adults aged 21–30 via Google Forms.
- Quantitatively measure the overall level of interest in entrepreneurship and the likelihood of starting a business using 1–5 Likert scale questions.
- Identify and rank the most influential motivating factors (e.g., financial independence, being your own boss, recognition, opportunity creation, government support).
- Identify and rank the most influential demotivating factors (e.g., fear of failure, lack of experience, lack of capital, lack of mentorship, family obligations, lack of institutional support).
- Examine how readiness and exposure (entrepreneurship course taken; confidence to start/manage a business) relate to entrepreneurial intention.

- Assess and compare the differences in intention and factor scores across demographics groups

**2.2 Research design**

In order to locate the factors that affect motivation and demotivation with regard to entrepreneurial intention, this study utilizes a cross-sectional survey design. This specific design captures present interest levels, perceived motivators, perceived barriers, as well as readiness at a single point in time, which would enable comparison across demographic groups and statistical testing of relationships between factors and entrepreneurial intention.

This study uses a Google Form in order to obtain proper responses, distributing it online to participants. All respondents received the same questions with the same wording and response scales, creating a consistent dataset that can be analyzed through descriptive statistics and group comparison.

**2.3 Participants**

Participants were approached to complete the survey.

**2.3.1 Target population**

The population targeted in this study comprises young adults aged 21-30 in Indonesia.

**2.3.2 Sampling Method**

Participants were recruited using non-probability sampling, primarily convenience sampling and snowball sampling. The Google Forms survey link through WhatsApp groups, Instagram, LinkedIn, university/alumni communities, and personal networks. The respondents were also encouraged to share it with their peers in order to get more responses.

**2.3.3 Inclusion Criteria**

All respondents must

- Be aged 21-30
- Indonesian nationality

**2.4 Data Collection**

**2.4.1 Survey Instrument**

Data were collected via Google Forms using a structured questionnaire. The survey's purpose is to measure entrepreneurial interest and intention, including perceived motivators, perceived barriers, and indicators of readiness. The questionnaire contained the following sections:

**a. Demographics**

Age
- 21-24
- 25-27
- 28-30

Economic background (family income per month)
- ≤20,000,000 IDR
- 20,000,000 - 39,999,999 IDR
- 40,000,000 – 59,999,999IDR
- 60,000,000 – 79,999,999 IDR
- 80,000,000 - 100,000,000 IDR
- ≥100,000,000 IDR

Gender
- Male
- Female

Level of Education
- Highschool
- Bachelors
- Masters
- PHD

The family owns a business
- Yes
- No

**b. Entrepreneurial intention**

Entrepreneurial interest
- "On a scale of 1–5, how interested are you in entrepreneurship?" (1 = Not interested at all … 5 = Extremely interested)

Entrepreneurial likelihood
- "Overall, how likely are you to start a business of your own?" (1–5)

**c. Motivating Factors**
- Desire for financial independence
- Desire to be your own boss
- Recognition (status)
- Desire to provide opportunities to others
- Institutional or government support

**d. Demotivating Factors**
- Fear of failure
- Lack of experience

- Lack of capital (money)
- Lack of ambition
- Lack of institutional support
- Lack of mentorship
- Family obligations

**e. Readiness and Exposure**

- "I have taken a course on entrepreneurship" (Yes/No)
- "I feel confident to start and manage my own business" (1–5)

### 2.4.2 Survey Distribution Process

The Google Form link was distributed online. The opening section of the form included a brief description of the study's purpose. Responses were collected over a defined period until a sufficient number of responses were obtained for basic descriptive analysis and subgroup comparisons.

### 2.4.3 Data cleaning

After collection, the dataset was screened for:

- Incomplete submissions (missing key outcome variables)
- Responses outside the target age range (21–30)
- Duplicates or clearly inconsistent responses (if identifiable)
- The final analysis included only cleaned, valid responses.

## 2.5 Data Analysis

### 2.5.1 Descriptive Statistics

In order to analyze the statistical data, all responses were collected into a Google Sheets spreadsheet, where the multi-select motivator and demotivator answers were split into separate columns so each option could be counted individually using COUNTIF formulas. Means and standard deviations were then calculated for the three Likert items (interest, likelihood, and confidence), and the number of respondents selecting each motivating and demotivating factor was summed and converted into a percentage of the total sample. Bar charts were created in Google Sheets to visualize the ranked frequencies of these factors.

### 2.5.2 Subgroup comparisons

In order to compare entrepreneurial intention across demographic groups, pivot tables were built in Google Sheets with each demographic variable (age, gender, income, education, family business background, and course exposure) as the row field and the three Likert outcomes as the value fields, set to average. To test whether observed differences between groups were statistically meaningful rather than due to sampling variation, independent-samples t-tests were

performed using the T.TEST function, specifying two tails and unequal variances. Pearson correlations between interest, likelihood, and confidence were computed using the CORREL function. Finally, a simple linear regression was conducted using the LINEST function, with confidence as the predictor and likelihood of starting a business as the outcome, to test whether confidence significantly predicted entrepreneurial intention. Given the small sample size (N = 49) and the non-probability sampling design, all inferential results are interpreted as exploratory rather than confirmatory.

### 2.6 Ethics

Participation in the survey was entirely voluntary and participants could take part or quit at any point in the process without facing any consequences. Participants received an overview of the survey's objectives and how their responses would be handled when starting to complete the Google Form, and proceeding with completing the survey meant that they had consented to taking part in the survey. Personal identifiers (names, emails, phone numbers, or addresses) were not collected by the survey and the demographic questions were arranged in wide categories (for example age groups and income ranges) in order to prevent individual identification. All responses were kept in confidence and analyzed only as a whole (percentages, means, comparisons between groups) so none of the participants' data was presented in isolation. Due to the fact that some of the questions are related to subjective experience (for instance fear of failure or family responsibility) the survey was constructed in a way to avoid biased language and participants did not have to provide responses for any question they did not feel comfortable with.

### 2.7 Limitations

**Sample size**

With N = 49, statistical power for subgroup comparisons is limited, and effects in small subgroups (e.g., respondents in the 60–79M IDR bracket, n = 2; course-takers, n = 9) should be interpreted with particular caution.

**Self-report bias.**

All outcomes are self-reported and may be subject to social desirability bias, particularly on items related to ambition or fear.

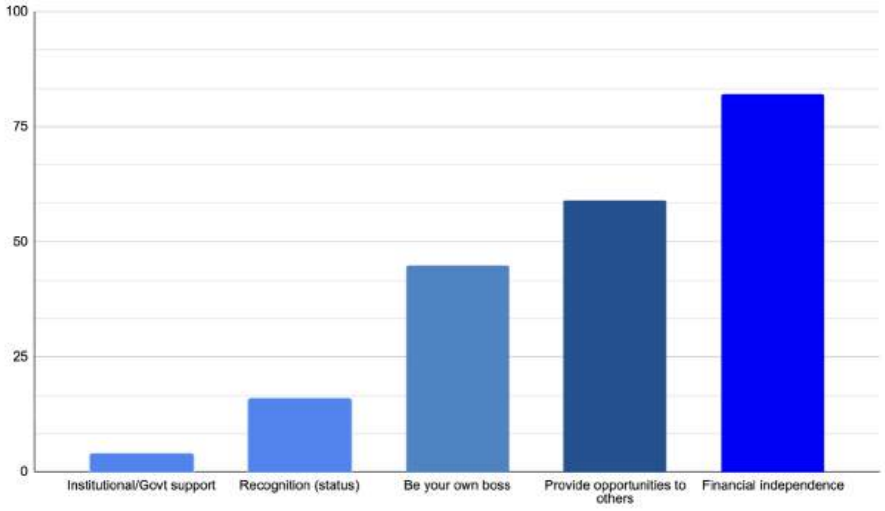
**Cross-sectional design.**

Because data were collected at a single point in time, the study cannot establish causal direction between predictors (e.g., family business background, course exposure) and entrepreneurial intention.

## 3. Results and Discussion

This section will report the findings of the survey analysis, examining entrepreneurial intention among Indonesian young adults aged 21-30 and the factors most strongly associated with it, followed by a discussion of the implications in light of existing literature. The results will

be shown through descriptive statistics, independent-samples t -tests, correlation analysis, and regression analysis. A total of 49 valid responses were analyzed (N = 49). Entrepreneurial intention is captured by two measures: interest in entrepreneurship and self-reported likelihood of starting a business, as well as a measure of confidence (self-efficacy) to start and manage a business, each rated on a Likert scale. This section will open with the overall patterns in interest, confidence, and the factors respondents associated with entrepreneurship, before eventually shifting towards the demographic and psychological factors most strongly linked to entrepreneurial intention.




Most frequently endorsed (a) motivating and (b) demotivating factors (N = 49). Percentages indicate the proportion of respondents selecting each factor; because the items were multi-select, percentages do not sum to 100%.

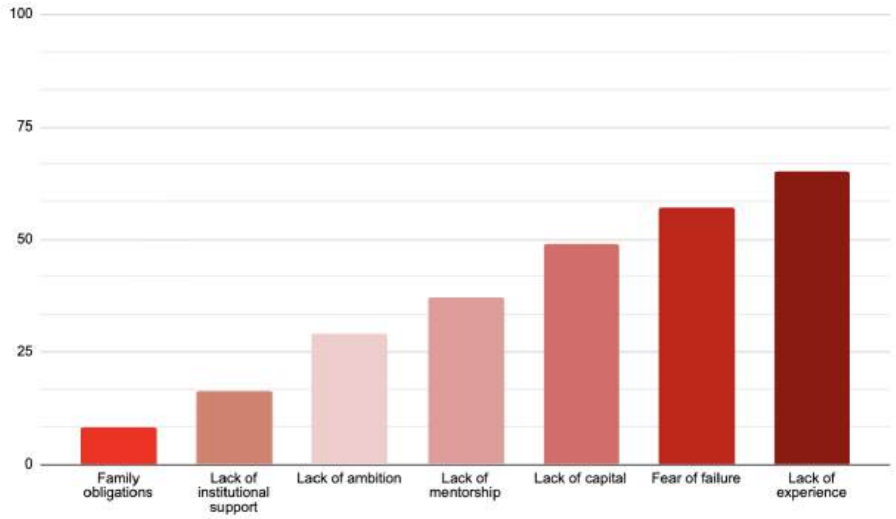


Figure 1a shows that the most frequently endorsed motivator was the desire for financial independence (82%), followed by providing opportunities to others (59%) and being one's own boss (45%), while external motivators such as recognition (16%) and government support (4%) were rarely selected. As shown in Figure 1b, the most common barriers were lack of experience (65%), fear of failure (57%), and lack of capital (49%), whereas structural barriers such as lack of institutional support (16%) and family obligations (8%) were uncommon. This indicates that entrepreneurial motivation in this sample is driven mainly by intrinsic and economic factors, while the principal barriers relate to perceived capability rather than external or structural constraints. These findings align with prior work on entrepreneurial motivation in Indonesia, which similarly identifies autonomy and financial independence as dominant drivers and perceived behavioral control as central to intention (Khoiriyah et al.; Virasa et al.)

**Table 1**

| Variable | Category | N | Mean | SD | t | p-value |
|---|---|---|---|---|---|---|
| **Confidence** | **The family owns a business** | 22 | 3.59 | 1.05 | 2.18 | 0.034 |
| | **No family business** | 27 | 2.93 | 1.07 | | |

**p**<0.05

Table 1 shows that there was a significant difference in confidence to start and manage a business between respondents whose family owns a business ($M = 3.59$, $SD = 1.05$) and those whose family does not ($M = 2.93$, $SD = 1.07$), given $t(47) = 2.18$ and $p < 0.05$. Respondents with a family business background reported notably higher confidence, a difference of 0.66 points on the five-point scale.

This difference becomes clearer when interpreted through the study's theoretical framework. Ajzen's Theory of Planned Behavior treats perceived behavioral control as one of the three main predictors of intention. This refers to a person's belief in their ability to perform a behavior and is often used interchangeably with self-efficacy. The confidence question in this study measures exactly this: how capable a respondent feels of actually starting and running a business. These findings would ultimately suggest that having a family business is linked to stronger perceived behavioral control and that such exposure is one of its real-world sources. Growing up around a family business likely builds self-efficacy through natural, informal observation and seeing what running a business actually involves and having ready access to guidance, which respondents without that background lack.

Georgescu and Herman discovered that students with an entrepreneurial family background expressed much stronger intention to become entrepreneurs because of the influence of family informal education. In turn, the current research extends this finding to the Indonesian youth aged between 21 and 30 and demonstrates the mediating role of self-efficacy in the relationship. Thus, family entrepreneurship experience does not directly influence entrepreneurial intention but does so indirectly through confidence. It is crucial for the overall

discussion as confidence plays the key role instead of the background. The following regression analysis proves this statement directly. In addition, the correlation analysis showed that all the three intention variables were related positively and significantly. The highest correlation coefficient was calculated for the correlation between confidence and the likelihood of becoming an entrepreneur ($r = .66$, $p < .001$) as well as between confidence and the interest ($r = .54$, $p < .001$). However, the weakest correlation coefficient was obtained for the correlation between interest and likelihood ($r = .45$, $p = .001$). Moreover, confidence was correlated stronger with the likelihood than with the interest, which is a preliminary indication that self-efficacy influences the intention more significantly than the interest.

**Table 2:** Regression analysis of the relationship between confidence (self-efficacy) and likelihood of starting a business (N = 49)

| **Model** | **Coefficient (B)** | **SE B** | **t** | **p** |
|---|---|---|---|---|
| **Constant** | 1.34 | 0.40 | 3.32 | 0.002 * |
| **Confidence** | 0.71 | 0.12 | 5.96 | <0.001 * |
| **$R^2$** | 0.43 | | | |
| **F** | 35.48 | | | |

Table 2 presents a simple linear regression using confidence (self-efficacy) as the predictor and the likelihood of starting a business as the outcome. The relationship was strong and highly significant, with $B = 0.71$ ($t = 5.96$, $p < 0.001$), indicating that each one-point increase in confidence was associated with a 0.71-point increase in a respondent's likelihood of starting a business. The constant of 1.34 represents the predicted likelihood for a respondent reporting the lowest possible confidence, and the model as a whole was significant ($F = 35.48$, $p < 0.001$). Most importantly, the $R^2$ of 0.43 indicates that confidence alone explained approximately 43% of the variation in respondents' likelihood of starting a business. For a single predictor of a complex behavior such as entrepreneurial intention, this is a substantial proportion, since individual factors in social science research typically account for far less of the variation in outcomes.

This result is consistent with both the theoretical framework of the study and previous literature on this topic. In particular, according to Ajzen's Theory of Planned Behavior, perceived behavioral control, which is similar to self-efficacy, is one of the biggest predictors of intention, and this result is confirmed repeatedly in entrepreneurial studies on the economies of developing Southeast Asia (Virasa et al.). The confidence indicator used in this research reflects precisely that construct, while the result of regression analysis is exactly what one would expect from the theory. Moreover, there is an important implication of this finding for practice. If confidence is the biggest predictor of entrepreneurial intention, then increasing entrepreneurship among young people in Indonesia will be most effective if their self-efficacy is developed via mentoring and

practical experience of conducting a business rather than stimulating their interest, which appears to be already quite high among the members of this group, according to the descriptive statistics.

## 4. Conclusion

In conclusion, the results from this study show that the entrepreneurial intention of young Indonesians aged between 21 and 30 years is influenced not by the absence of interest but by the absence of confidence. The interest was relatively high and resulted mainly from intrinsic motivators such as financial independence and autonomy, while the most common obstacles, including lack of experience, fear of failure, and lack of capital, were all associated with capability rather than extrinsic factors. Confidence, or self-efficacy, proved to be the most significant factor influencing entrepreneurial intentions, accounting for 43% of variations in their likelihood of becoming entrepreneurs, which is twice as much as the interest accounts for, and having a family business proved to be a predictor of confidence. Given the fact that the interest is already there, these results mean that any initiatives to promote entrepreneurship among this particular group should focus on confidence-building through mentorship and practical experience rather than on financial motivations or information campaigns. These findings can be of use to educators and policy-makers who want to turn this demographic's interest into ventures.

These findings should, however, be interpreted in light of the study's limitations. The sample was small (N = 49) and recruited through non-probability convenience and snowball sampling, limiting how far the results can be generalized. As a cross-sectional study, it can identify associations but not establish causal direction. The analysis is therefore best understood as exploratory, identifying patterns that warrant confirmation in larger and more representative samples. Even so, the consistency of the findings across descriptive, correlational, and regression analyses, together with their alignment with established theory, lends them credibility as an initial account of the factors shaping entrepreneurial intention among Indonesian young adults.

**Works Cited**


Ajzen, Icek. "The Theory of Planned Behavior." Organizational Behavior and Human Decision Processes, vol. 50, no. 2, 1991, pp. 179–211.

Badan Pusat Statistik. "Unemployment Rate by Age Group: National Labour Force Survey (Sakernas)." Badan Pusat Statistik, 2025, www.bps.go.id/en/statistics-table/2/MTE4MCMy/unemployment-rate-by-age-group.html

DevTech Systems. "Formalizing and Strengthening Micro, Small, and Medium Enterprises in Indonesia." DevTech Systems, 25 Mar. 2022, devtechsys.com/insights/2022/03/25/formalizing-and-strengthening-micro-small-and-medium-enterprises-in-indonesia/.

Georgescu, Maria-Ana, and Emilia Herman. "The Impact of the Family Background on Students' Entrepreneurial Intentions: An Empirical Analysis." Sustainability, vol. 12, no. 11, 2020, article 4775.

Khoiriyah, Siti Nur, et al. "The Effect of Entrepreneurship Education and Digital Literacy on Entrepreneurial Intention Mediated by Entrepreneurial Self-Efficacy in Vocational High School Students in Malang City." International Journal of Business and Quality Research, vol. 3, no. 4, 2025, pp. 121–144.

Liu, Xianyue, et al. "Research on the Effects of Entrepreneurial Education and Entrepreneurial Self-Efficacy on College Students' Entrepreneurial Intention." Frontiers in Psychology, vol. 10, 2019, article 869.

Mahendra, Angga Martha, et al. "The Effect of Entrepreneurship Education on Entrepreneurial Intention Mediated by Motivation and Attitude among Management Students, State University of Malang, Indonesia." International Education Studies, vol. 10, no. 9, 2017, pp. 61–69.

Remeikiene, Rita, et al. "Explaining Entrepreneurial Intention of University Students: The Role of Entrepreneurial Education." Active Citizenship by Knowledge Management and Innovation: Proceedings of the Management, Knowledge and Learning International Conference 2013, ToKnowPress, 2013, pp. 299–307.

Tempo. "Indonesia Lags behind Malaysia, Thailand in Entrepreneurship, Says the SME Minister." Tempo, 6 Oct. 2023, en.tempo.co/read/1780656/indonesia-lags-behind-malaysia-thailand-in-entrepreneurship-says-sme-minister.

Virasa, Thanaphol, et al. "Predicting Entrepreneurial Intention and Economic Development: A Cross-National Study of Its Policy Implications for Six ASEAN Economies." Heliyon, vol. 8, no. 5, 2022, article e09435.

Zhang, Pingying, et al. "A Study of Entrepreneurial Intention of University Students." Entrepreneurship Research Journal, vol. 5, no. 1, 2015, pp. 61–82.

**Closing the Loop: The Role of Retailers in California's SB 554 By Siree Chawla**

**Acknowledgements**

Thank you for the guidance of Gregory Robson mentor from The University of Notre Dame in the development of this research paper.

**Abstract**

California's Senate Bill 54 demands "producers to reduce single use plastic packaging and foodware by 25% by 2032" (CalRecycle, 2026), representing one of the most inventive Extended Producer Responsibility laws in the United States, aiming to reduce large amounts of plastic waste while promoting recyclable and compostable packaging. Despite these efforts, recyclable materials such as thermoformed polyethylene terephthalate (PET) continue to experience low recovery rates because existing recycling systems often fail to encourage consistent consumer participation. This paper explores how retailers can play a more active role in improving thermoformed PET recycling through convenient collection infrastructure, financial incentives, and consumer education. Using qualitative research a review of literature on consumer behaviour, circular economy theory, Extended Producer Responsibility policies, and international recycling systems. This paper will examine consumer behaviour at retail locations and incorporate industry perspectives to identify a solution for increasing recycling participation. Based on data collected, this paper proposes the California Retail Recycling Innovation Framework, which follows a retailer-led pilot program This paper discusses multiple retailer-led solutions to SB54 and expands the most effective solutions using evidence collected from pilot programs, academic research and theoretical frameworks.

**Lumiere Research Scholar Final Paper Template**

Despite public efforts to increase individuals' attention towards the negative impacts plastic creates on public health and the environment, the recycling rates in the United States remain extremely low. According to Fukui (2016), the United States discarded approximately 70 million tons of plastic waste in 2019, more than any other country. Of this waste, only about 4.5% was recycled, while 73% ended up in landfills (Fukui, 2016). Additionally, plastics that enter aquatic environments, rather than recycling systems, degrade into microplastic that accumulate in marine ecosystems and have been detected in the human body (Witczak et al., 2024). These statistics illustrate the need to implement policies like California's Senate Bill 54 across more states.

Beyond its environmental and health concerns, plastic pollution represents a significant challenge economically by increasing waste management costs and reducing the availability of recycled materials needed to support and create a circular economy-economic system which keeps products in use as long as possible to eliminate waste-in California.

To address the issues of recycling, Californians in particular have called for more formally placed rules regarding plastic waste and improving the recycling stream. Continued

years of public advocacy and awareness to have companies recycle specifically among community stakeholders, governments and industry representatives have led to California's Plastic Pollution Prevention and Packaging Producer Responsibility Act (SB54). Rather than waiting upon local governments and community stakeholders and taxpayers to take action, SB54 implements the Extended Producer Responsibility model which requires producers to finance recycling programs, redesign methods of packaging, and reduce plastic waste in a products' life cycle. SB54 has the potential to influence future recycling programs and allows California to mirror other countries' progress on recycling projects to better the future of the country. SB54 was created in 2022, with the goal of reducing plastic waste and advancing a circular economy. Its long term success is based upon consumer participation. A recyclable package cannot be refurbished if not correctly sorted or recycled in its correct stream of distribution. As Leonard Read explains in *I, Pencil*, the production of a simple pencil we use everyday is actually made up of countless individual parts and thousands of specialized techniques by various individuals (1958). Likewise, a successful recycling stream requires cooperation among multiple individuals like producers, consumers, recycling facilities, etc. If one area fails to fulfill their responsibilities in the recycling stream the end result will be far from successful. This idea is further referenced in Garret Hardin's concept of the Tragedy of the Commons (1968), which demonstrates how shared resources deteriorate when individuals fail to recognize the importance of collaborative work. Thus the goals of SB54 cannot be fulfilled without active participation from Californians.

Extensive research (Read 1958; Hardin 1968) has sought to understand the Extended Producer Responsibility (EPR), circular economy principles, consumer recycling behaviour and solutions to help SB54 be more effective. Organizations, such as the Organization for Economic Co-Operation and Development (OECD) and the Ellen MacArthur Foundation, have developed frameworks to address the producer's responsibility in recycling. Additionally, behavioral researchers like Ajzen (1991) have focused on consumer behavior and consumer incentives to recycle. Collectively, these studies have established a working ground and map for a successful cycle of recycling by highlighting the importance of collaboration.

Limited research proposes feasible solutions for allowing consumers to participate and follow California's SB54. As the gap remains unfulfilled, retailers are left uneducated on how they can actively encourage consumers to recycle. To address these underlying issues this paper proposes the California Retail Recycling Innovation Framework, a retailer centered solution targeted towards consumer participation in post consumer thermoformed PET recycling under SB54.

The proposed framework supports the body of the research on EPR by emphasizing an area that has received limited attention: the role of retailers in influencing consumer recycling behavior. As California implements SB54, future research should evaluate retailer-led programs to determine which strategies will directly influence consumer participation most effectively in order to improve rates of thermoformed PET recycling to create a successful environmental future.

This paper introduces the Circular Consumer Toolkit, a standardized collection of

educational resources and retail led strategies designed to simplify recycling decisions and increase consumer participation in a cheaper and practical approach. Rather than focusing on producer or policy compliance, this paper addresses practical strategies for positioning retailers as the necessary link between legislation and consumer action. This paper reviews literature on Extended Producer Responsibility, stakeholder theory, consumer behaviour, and the circular economy. It also examines international recycling streams that can be used to inform California's approach. The paper proposes the California Retail Recycling Innovation Framework and the Circular Consumer Toolkit as complementary strategies to strengthen the effectiveness of SB54. Finally, this paper explores the broader policy implications of these recommendations and how we can integrate solutions to creating an equitable and successful circular economy while mainstreaming recycling in California.

## Methodology

This study demonstrated a qualitative research approach to determine possible solutions in which retailers could increase consumer participation in post consumer thermoformed PET recycling under California's SB 54. Rather than collective quantitative data, the paper highlighted past research, public policy and industry practices from a local to global scale in order to identify an improved proposal to California's SB54.

The main research method consists of a literature review of peer reviewed journal articles, California SB54 legislation, CalRecycle publications, Extended Producer Responsibility, stakeholder theory, consumer behavior and the circular economy. These following sources were used to identify the common limitations to recycling of thermoformed PET and decipher the targets needed to address these limits and propose an effective solution to recycling programs in California.

This literature was used to find gaps and weaknesses in past research, particularly addressing the limited programs available to retailer-led consumers engagement strategies. The paper elaborates on an innovative solution called the Circular Consumer Toolkit, a conceptual framework which allows for multiple ideas like consumer incentives, education and convenient recycling programs to create a strong closed loop scaling system in California.

## Literature Review

### *Extended Producer Responsibility*

The Organization for Economic Co-operation and Development (OECD) defines Extended Producer Responsibility as an environmental policy in which producers are responsible for the disposal and collection of products after consumer use (OECD, 2024). Over the past three decades the Extended Producer Responsibility has altered the course of packaging waste by shifting financial and operational responsibility from governments and taxpayers to producers. According to OECD, implementing laws like SB54 and placing financial responsibility on producers,motivates companies to redesign and improve

packaging in order to reduce unnecessary waste (OECD, 2024). According to OECD, implementing laws like SB54 has forced producers to consider the lifespan of packaging and forced companies to find alternative eco-friendly solutions to obey the laws of recycling. This approach is the foundation of California's SB54.

California took inspiration from OECD and the EPR in creating SB54, which requires producers- companies responsible for covered packaging or plastic food service ware sold or distributed in California (CalRecycle, 2026)- to reduce single use plastic and recycle their packaging. This legislation represents a growing advocacy for waste management before it reaches stores. SB54 encourages companies to design packaging which can be recycled. The law represents one of the most progressive policies within a closed loop recycling system without minimal strict enforcements like money fines regarding Extended Producer Responsibility in the United States.

While SB54 is one framework in response to EPR, scholars continue to develop alternative solutions to Extended Producer Responsibility. Nicole Beetcher's research focuses on strengthening the circular economy and steers from traditional EPR policies. Her work emphasizes that successful environmental legislation extends beyond statutory requirements; it heavily depends on participation across supply chains to be effective.

Very little research has found a cost effective approach to environmental policies and consumer behaviour under SB54. This gap is important to establish a relationship between consumers during purchasing and after purchasing to help promote a circular recycling system. Addressing the foundation for retailer-centred implementation supports the fundamentals of this paper.

***Consumer Behaviour and Behavioural Economics***

California's SB54 effectiveness is primarily dependent on consumer participation. Whether a company decides to redesign their packaging and produce recyclable thermoformed PET, the product's value will be limited if consumers do not correctly recycle and return these products. According to CalRecycle (2026), the biggest cause of plastic in landfill is due to a lack of participation in recycling plastic which is a direct result of consumers being heavily uneducated in this area of discussion. If recycling infrastructures can improve and be easier for users, more individuals may make the choice to recycle. Not only are consumers needed to initiate a recycling chain, policy makers must be involved in understanding consumer incentives and influences to improve recycling systems. As Barr (2007) explains, "recycling decisions are shaped by both environmental attitudes and practical barriers." Consumers are often not reluctant to recycle, but they are unaware of the benefits or the methods of recycling. However as previously stated, the responsibility of the repercussions of plastic production and disposal should lie with the companies who produce plastic products.

One of the most common theories of recycling is Ajzen's 1991 Theory of Planned Behaviour, which suggests that people's actions are influenced by three distinct points: stereotypes, their perspective towards a behaviour and their perceived knowledge. Connecting

to PET recycling, consumers are more likely to recycle willingly if they are either shown the benefits on the environment, mirroring a social trend to recycle and or feel recycling is easy to do. If they lack access to a recycling bin or have little to no knowledge of what can be recycled, their behavioural control decreases, lowering the chances of them participating. Increasing recycling rates requires efforts from retailers to clearly identify how, what, and where to recycle to grow consumer confidence on this subject.

Behavioural economics provides another perspective through the Nudge Theory (Thaler and Sunstein, 2008), which proposes that small changes to how choices are portrayed can heavily influence the ideas and behaviour towards a subject within a society or individual. Rather than relying solely on regulations or financial incentives, nudges encourage consumers to make the more desirable decision by making it the simpler choice. Examples such as basic recycling labels, common collection bins and retailer reward programs are just the few which can heavily steer a consumer's participation in recycling. Together, the Nudge Theory and Ajzen's Theory of Planned Behaviour provide the theoretical framework for this paper's proposal of the Circular Consumer Toolkit. The toolkit focuses on the idea that retailers should implement strategies based upon three main factors: education, behavioural design and convenient recycling opportunities to increase consumer participation and successfully support SB54 from a producer and consumer standpoint.

***Stakeholder Theory and Role in Retailers in Consumer Participation***

Although producers can redesign packaging, retailers have a more direct influence on consumer's purchasing decisions. Edward Freidman's stakeholder theory argues that businesses create greater long term value when they consider the interest of all stakeholders rather than prioritizing individual stakeholders (Freeman 1984). Applying this framework to SB54 suggests that retailers can benefit from investing in consumer recycling initiatives, as these programs increase participation and strengthen the relationship between consumers and producers. In order to increase recycling rates programs, digital rewards, and in-store recycling bins must become more widely implemented. These initiatives can support customer education and familiarity with recycling processes and increase compliance with SB54. As a result, these solutions can facilitate the long term movement of a circular economy.

Michael Jensen broadens the discussion by arguing that stakeholder interest should support long term maximization rather than short term performance and benefits (Jensen, 2001). According to Jensen, successful companies are able to balance multiple stakeholders while working with clear objectives and known values (Jensen, 2001). Retailers affected by SB54 should approach sustainability initiatives as strategic business ventures rather than solely viewing them as environmental obligations. Encouraging consumers to return thermoformed PET packaging can reduce future compliance costs, and as a result position retailers to be leaders among sustainable practices. This approach allows for environmentally friendly values and long term profitability to collaborate.

Lynn Stout rejects Jensen's argument that companies should maximize profit. (Stout,

2012). Stout's approach varies from other philosophers, she challenges the traditional beliefs that maximizing short term shareholder value should be the sole purpose of a business. Stout believes that corporations possess the ability to pursue broader social objectives when those actions are able to benefit both society and business. Companies can consider how their decisions affect consumers, employees and communities. By considering these broader interests, businesses can contribute to societal goals while still creating long-term value. In this case, companies can invest in practices that reduce environmental harm and strengthen consumer trust and future success. Retailer recycling can represent investments which support the future of sustainability.

Additionally, philosophical concepts like Leonard Read's *I, Pencil* illustrates how complex systems depend upon countless corporations and individuals who may never directly interact but collaborate to create a singular pencil (Read, 1958). Similarly, a closed loop packaging system cannot function through the legislation alone. Producers, retailers, consumers, recyclers, local governments and policy makers may never meet directly, but must work collaboratively to comply with California's SB54. Retailers who interact with consumers possess knowledge regarding purchasing habits and customer preferences which is useful to policymakers who may overlook that information. The information then allows retailers to spread awareness design incentive programs and abide by SB54. Without one or the other the law is vulnerable to failure.

Overall, these theoretical perspectives suggest that retailers instead of aiming to comply with SB54, attempt to shape consumer behaviour. Adopting ideas from Read, Freeman, Jensen and Stout, while looking at stakeholder theory, positions retailers as the primary coordinator within California's transition towards a circular economy. Their ability to educate consumers and provide convenient recycling opportunities can help align business strategies with environmental objectives making retailers the most influential actors in increasing participation in thermoformed PET recycling. These ideas help strengthen the following sections of the paper by agreeing with the specific barriers that prevent consumer participation and also proposes retailer centred strategies to overcome them.

***Retailer Led Recycling and System Gaps***

Retail stores like grocery stores such as Walmart, Apple, 7-Eleven are one of the few locations that consumers visit on a daily basis making it a perfect opportunity to enforce and introduce recycling standards and education initiatives. Although this may seem like a perfect option to implement recycling programs and integrate recycling, many retailer-led programs do not have such programs yet, making these programs necessary to bridge the gap in California's transition towards a circular economy.

Several organizations have conducted research that shows that retailer participation can increase consumer immersion in recycling. For example, The Recycling Partnership has worked with retailers in several local governments to improve recycling infrastructure while educating and encouraging consumers to actively practice recycling habits. Likewise, Closed Loop

Partners has integrated collection technologies, metrical recovery facilities and new innovations to improve recovery of plastic packaging. Despite these efforts, the need to improve direct engagement with consumers at the point of purchase must be emphasized further.

Additionally, retail take back programs have proven effective for products such as plastic bags, electronic waste, etc. Many grocery stores provide collection bins which allow consumers to return materials during a shopping trip. These programs continue to reduce the stigma associated with recycling inconvenience by making recycling easily accessible to consumers that can simply drop off their items during a shopping trip.

Despite studies showing success, groomed PET scaling, cleaned and sorted thermoformed PET prepared for recycling into new packaging, remains excluded from retailer collection systems. Items like berry containers, salad clamshell, and prepared food containers are commonly used throughout California but are rarely made available for collection programs and recycling programs. As a result consumers are left questioning what is recyclable and what is not. This disconnect limits the effectiveness of SB 54 because consumers are the primary success of a closed loop system and are directly responsible for returning packaging to the recycling stream and thus making the SB54 bill effective.

Retailers serve as the direct point of communication between producers and consumers. Retailers influence purchasing decisions and facilitate collection opportunities. Improving and expanding the role of retailers in California's recycling system can create a more practical strategy for improving recovery rates while maintaining the ideas of SB54.

Addressing a solution to this gap provides the basis for the recommendations proposed in the following sections which introduce a detailed approach targeted to increasing consumer engagement and creating a closed loop packaging system in California.

***The Circular Consumer Toolkit: A Retailer Centred Framework***

The literature reviewed throughout this paper focuses on California's SB54 but also reviews the foundation for Extended Producer Responsibility (EPR), in order to achieve a successful circular packaging system. This paper will discuss a different, more hands on approach to the role retailers can play in influencing recycling behavior, compared to previous research which primarily focuses on producers and policy makers. Retailers directly interact with consumers at the point of purchase making it a perfect opportunity to bridge the gap and develop a strategy combining policy objectives and everyday consumer behavior. This paper takes influences from the ideas of stakeholder theory and consumer behavior to propose the Circular Consumer Toolkit which highlights a retail-centred approach designed to increase participation in thermogenesis PET recycling and effectively involve California's SB54 in an accessible closed loop packaging system.

The Circular Consumer Toolkit begins with consumer education. Many consumers remain uneducated about what specific thermoformed PET packaging can be recycled. This lack of information can often result in contamination or recycling neglect. A few inexpensive ways retailers can improve this problem is by creating standardized recycling labels, QR codes

linking to recycling instructions, in store educational displays and digital resources explaining how to properly recycle. Simple yet effective ways of integrating recycling can be beneficial to the long term success of SB54 by creating small scale solutions to begin and later expanding to larger more costly projects. By improving consumer knowledge consumer confidence is in turn improved.

The second part examines convenient collection infrastructure. Recycling behavior is strongly influenced by convenience making accessibility an important factor in increasing participation. Retailers should install dedicated thermoformed PET recycling bins or on large scale production vending machines near store entrances to allow consumers to return food packaging waste. The idea for "reverse vending machines" is based on policies in several European countries which have proven extremely successful to the recycling chain. For example, TOMRA, a Norwegian manufacturing company for reverse vending machines, has consistently recovered around 92% of plastic packaging like Thermoformed PET (TOMRA, 2024). Integrating policies like this into California complies with the SB54 bill while having the ability to integrate the 75% of typical plastic in landfills into a productive and successful recycling stream. Integrating collection into shopping habits reduces the amount of inconvenience and effort it takes for consumers to recycle.

A third element of the framework introduces consumer incentives. Behavioral research proves that individuals are more likely to participate when recycling provides an immediate and tangible benefit. Retailers could allow for an increase in loyalty programs, digital rewards, coupons, or discounts for returning thermoformed PET back to grocery stores. These incentives allow for a free unrestrictive approach to recycling and allows consumers to decide for themselves whether they want to participate or not without the pressure of mandatory systems. Such ideas allow strong customer support while allowing SB54 to be beneficial in environmental objectives.

Lastly, successful implementation of the Circular Consumer Toolkit requires collaboration between all parties: producers, retailers, policymakers and consumers. Beginning with producers they can finance collection systems through SB54's Extended Producer Responsibility requirements. Retailers can then begin educational programs and establish accessible recycling standards. Finally, consumers can actively partake in returning their plastic to grocery stores and local collection bins. This shared responsibility among groups reflects the key idea of stakeholder theory and allows for a smooth transnational approach towards circular economy.

The Circular Consumer Toolkit does not replace or discourage SB54. It strengthens the legislation by addressing the fundamental issue of consumer participation. By combining education, convenient collection methods, educational resources and multi stakeholder collaboration the proposal provides a strong and clear focus to fully support recycling of thermoformed PET and answers doubts of California's SB54 convenience in the long term.

**Discussion**

### *Evaluating Circular Consumer Toolkit*

The findings of this paper suggest that the greatest limitation to successfully implementing California's SB54 is lack of consumer participation. Despite the legislation requiring an Extended Producer Responsibility framework, which requires producers to finance recycling and reduce packaging waste, greater focus on consumer participation is needed to raise recycling rates. A successful closed loop packing system is reliant on producer compliance but also on consumers consistently returning their recyclable plastic. A primary topic repeated in the literature in order to facilitate consumer participation was convenient and accessible throughout recycling programs and increase of consumer knowledge through education. These ideas suggest consumer engagement is as or even more important than improving recycling infrastructures because without participation, the advances in infrastructure provide no benefit.

The Circular Consumer Toolkit was developed to address these limitations by placing retailers at the centre of consumer engagement. Unlike previous ideas where producers or local governments are the main source towards creating a closed loop recycling structure, this framework recognizes that retailers interact directly with consumers and therefore possess the ability to influence recycling decisions at point of purchase. Through the use of physical and digital educational resources, convenient recycling systems, standardized recycling labels and incentive based programs, retailers have the ability to increase recycling rates starting from increase in consumer participation by clearing confusion and explaining availability of recycling to make it more accessible.

These findings contribute to the idea that increasing recycling participation requires a handful of interventions. Consumers are more likely to participate when the objective is clear and easy to access. Examples like retail collection bins or reverse vending machines provide opportunities to return thermoformed PET packing during routine shopping trips making it part of the everyday shopping trip to consumers. Similarly, educational information specifically shown at the point of purchase can help clarity but also bring concern to consumers in wanting to create a more sustainable future for themselves. Not only that, it helps to clarify what thermoformed PET can be recycled, helping recovery rates increase and contamination rates decrease. This emphasizes convenience as a key driver of recycling participation while extending the literature by proposing a coordinated retailer led approach.

### *Challenges and Future Considerations*

Although the paper proposes solutions to support the great success and long term development of California's recycling system, integrating these ideas would present several practical concerns. One being the cost of such projects. While many of the proposed strategies display a less expensive approach such as recycling labels, digital recycling resources and QR codes, on the other hand, installing reverse vending machines would require substantial financial support from retailers and producers. The cost involved with installing, mastering, purchasing

and operating the reverse vending machines across multiple branches of grocery stores present itself to be a costly approach. While large retail stores may be able to comply, small retail stores will struggle to adopt this approach due to financial logistics. Additionally, if government involvement is not present, the proposed projects may fail to succeed.

Furthermore, even if accessible recycling systems are present, consumer participation is not guaranteed. Recycling habits are influenced by convenience and benefits. Some consumers may continue to follow habits and dispose of recyclable plastic into trash bins rather than recycling bins therefore the suggestion that education alone can solve this issue is unreasonable. Long term success requires continuous public awareness campaigns alongside improvements in collection infrastructure.

In addition to filling a gap in literature on the role of retailers in influencing consumer recycling, these findings also suggest that future research should continue to expand and move beyond conceptual frameworks by evaluating retailer-led pilot programs within California. Measurable and meaningful data could be tested by such strategies dedicated to thermoformed PET, collection stations, digital incentive programs and education campaigns in different retail locations to compare the results in recovery rate, recycling and contamination rates. Collecting data will be necessary to compare success levels among different solutions in order to continue to expand on the most successful implantation long term.

While challenges remain, it is evident that these limitations do not diminish or degrade the importance of implementing and fostering the success of California's SB54. Instead, they highlight the need for effective trial and error, collaboration and innovation. Additionally, these drawbacks can help to create an extremely successful streamline in recycling through progressive approaches. Recycling policies can eventually transition California towards a more sustainable and circular closed loop packaging system.

***Policy Implication and Future Research***

The findings of this paper suggest the long term success of California's SB54 depends entirely on the continued research and improvements to grow the recycling system. As the legislation is implemented, policymakers should closely view consumer participation and identify which of the following strategies propose the greatest undress in thermoformed PET recovery. This evidence can be knowledge for future interventions in improving the recycling system and ensuring it continues to be accessible to everyday consumers and responsive to consumer behavior.

Future research should be conducted to test these solutions in real world settings to ensure the relationship between proposals and outcomes align. This research should explore whether the proposed strategies can be realistically implemented by retailers and consumers to produce long term improvements in recycling participation and thermoformed PET recovery. Pilot programs can evaluate consumer education campaigns, and collection systems can provide valuable data on participants rates, behavioral motivations and material recovery. Furthermore, retailers should pivot across different communities to compare and share results

to determine the most effective strategies for specific populations. Not only that, retailers should look to lend aid or ideas from other successful countries on a global scale that are already successfully producing high recycling rates.

**Conclusion**

California's SB54 represents an important shift towards a circular economy by placing greater responsibility on producers to recycle and recover packaging materials. However, this paper argues that without the participation of consumers this policy alone cannot achieve long term sustainable success. Research on Extended Producer Responsibility, stakeholder theory, and consumer behavior, has found that limitations such as stereotypes on recycling, accessibility and incorrect recycling methods can reduce the recovery of thermoformed PET packaging. These findings suggest that consumer participation should be made the primary focus of successfully implementing SB54 laws to encourage recycling rather than a secondary consideration.

As California continues to grow and strengthen its circular economy goals, this research suggests that consumer participation should remain the central idea around SB54 in order to succeed. By exploring ideas from evidence-based policy and technological advancements and innovations further recycling initiatives can better serve towards consumer behavior and lead to a sustainable circular economy packaging model. Every package recycled, every consumer engaged, and every retailer committed to improving and implementing SB54 is another step closer towards closing the loop. In the end, everyone must work together towards a like minded goal to achieve success and the greatest measure of SB 54's success will be its ability to transform responsibility into collective action.

**Works Cited**

CalRecycle. (2024). Sb 54: Plastic pollution prevention and packaging producer responsibility act. In the CalRecycle Home Page. https://calrecycle.ca.gov/packaging/packaging-epr/ Extended producer responsibility. (2024). In OECD. https://www.oecd.org/en/publications/2024/04/extended-producer-responsibility_4274765d.html

Discover the technology inside reverse vending machines. (2024). In Tomra.com. https://www.tomra.com/reverse-vending/media-center/feature-articles/inside-reverse-ve n ding-machines

Friedman, M. (1970, September 13). The social responsibility of business is to increase its profits. The New York Times. https://www.nytimes.com/1970/09/13/archives/a-friedman-doctrine-the-social-responsibi l ity-of-business-is-to.html: 4.

Fukui, K. (2026, April 5). Global plastic recycling stalls at under 10%. Nikkei Asia. https://asia.nikkei.com/spotlight/datawatch/global-plastic-recyclinhttps://www.oecd.org/e n/publications/2024/04/extended-producer-responsibility_4274765d.html G-stalls-at-under-10 : 3.

Jensen, M. C. (2001). Value maximization, stakeholder theory, and the corporate objective function. SSRN Electronic Journal, 22(1): 6-10.

PET thermoform packaging design resource document. (n.d.). Retrieved August 14, 2026, from https://plasticsrecycling.org/wp-content/uploads/2024/11/RES-PET-01-Thermoforms.pd f Reed, L. (n.d.). “I, pencil: My family tree” as told to Leonard E. Read, dec. 1958 | online library of liberty. In oll.liberty fund.org. Retrieved August 5, 2026, from https://oll.libertyfund.org/titles/read-i-pencil-my-family-tree-as-told-to-leonard-e-read-d e c-1958: 4-11.

Robson, G. (2025). AI and political economy: Citizens beware – the journal of politics blog. Buffalo.Edu. https://ubwp.buffalo.edu/jopblog/2025/08/12/ai-and-political-economy-citizens-beware/ Stout, L. (2016). The dumbest business idea ever. The myth of maximizing shareholder value. In Economics. https://evonomics.com/maximizing-shareholder-value-dumbest-idea/: 10. Witczak, A., Przedpełska, L., Kamila Pokorska-Niewiada, & Cybulski, J. (2024). Microplastics as a threat to aquatic ecosystems and human health. Toxics, 12(8), 571–571. https://doi.org/10.3390/toxics12080571: 3.

**Exploration of Central Bank Independence and its Ties to COVID-19 Related Fiscal Policy**
**By Emma Goga**

**Abstract**

The emergence of post COVID-19 economies has been a determining factor in the development of economics worldwide. The global recession caused by COVID-19 led to fiscal aid becoming a necessity among citizens of every nation, whether they may be nations with independent or orthodox central banking practices. In this paper, the aim is to explore how central bank independence plays a role in fiscal recovery in times of economic downturn, as well as how effective these responses are. Following that, variables across different fiscal response policies are analyzed in order to draw a conclusion on which is the most successful. These datasets will then contribute to the final diagnosis of how central bank independence plays a role in the success of fiscal aid packages in post COVID-19 economies. Through cross references in datasets on fiscal stimulus, GDP forecasts, and pandemic related statistics including hospital beds per capita, the datasets reveal the overarching correlations between high central bank independence levels, high forecasted GDP growth and recovery rates, as well as reduced figures of hospital beds per capita. All in all, with comprehensive data, future pandemic related recession can be mitigated and receive responses that entail maximum efficacy.

**Introduction**

The emerging COVID-19 outbreak in Wuhan, China as of December 2019 negatively impacted economies worldwide. Being a global pandemic, supply and manufacturing chains, workforce quotas, and numerous economic facets were negatively affected. In order to combat these adverse effects, individual governments took to installing fiscal aid packages to benefit the economy as well as citizens. However, many factors play into fiscal aid, namely central bank independence. How could the independence of a central bank impact a government's response to an economic downturn? The common understanding that market and mixed economies are cyclical leads to the conclusion that larger aid packages will therefore lead to greater recovery in times of economic peril, including that of the COVID-19 crisis.

This paper aims to use a wide variety of secondary quantitative data and methods to analyze data regressions on effects of independence in order to communicate the clear positive correlation between central bank independence and fiscal stimulus package size. In short, central bank independence is associated with larger national fiscal aid packages that have increased efficacy in mitigation of economic downturn. First, data is examined through a variety of regressions including independence index, the International Monetary Fund's (IMF) fiscal aid size analysis, and a review of various other empirical studies conducted on the topic. In expeditious terms, this paper differs from others on the subject of fiscal response and central bank independence largely due to the fact that while the relationship between the two is analyzed, overall effectiveness of these conditions on stimulus packages and their effect in turn is observed on a top level in order to draw a conclusion on central bank independence as a factor

that holds a positive relation with the latter. These variables, respectively, correlate with each other to form the underlying conclusion of high central bank independence index figures equating to higher coefficients in expansionary policy, as well as relationships with other variables, including median population age, fiscal policy, and hospital beds per capita during the COVID-19 pandemic.

**Methods and Methodology**

Secondary data analysis is a large portion of this evidence, as it provides information and data regressions available in hindsight beyond firsthand collection.

The first section will detail the data gathered and analyzed, in relation to the topic and thesis claim of the paper through a fixed-structure regression represented as the following equation[1] (Elgin, C, Yalaman, A, Yasar S, Basbug G, 2021).

$$Stimulus_{it} = \alpha_0 + \alpha_1 CBI_i + \alpha_k \sum_{k=2}^{n} X_{k_{it}} + \alpha_j \sum_{j=n+1}^{m} Z_{j_i} + \gamma_t + \epsilon_{it}$$

Within the equation, Stimulusit represents the stimulus, with country i and year t. CBIi represents the core variable, the index of central bank independence within country i as of most recent times. The central bank independence and stimulus size index are regressed together six times in order to draw concise relationships between the variables. These procedures are then compared with our second regressions. The latter regressions involve six policy variables under three categories, monetary policy, exchange rate policy, and fiscal policy. The three markers involve multiple factors coded as percentages of total GDP, including total adopted fiscal measures, interest rate cuts, objective package size (as coded by GDP), among other monetary measures coded as a dummy variable[2] (Elgin, C, Basbug, G, Yalaman, A, 2020).[3] This data is regressed into an overall stimulus index table for expeditious comparison of the previous variables and the second set of variables. Finally, the observed data from both regressions is compared with figures from the international COVID-19 economic stimulus index, or the CESI.[4] Among the CESI, the variable of hospital beds per capita is reviewed in the research to develop a final correlation between stimulus size and case progression.

Subsequently, the methodologies of analysis are discussed to further confirm regressions, data, and connections. The first and second respective regressions are cross referenced first disregarding variables. Afterward, variables are considered in order to delve into descriptive characteristics, while the former grants precedence to inferential characteristics. Furthermore, the different characteristics and regressions are analyzed to draw a concise conclusion from the multitude of data. This is done to compare general conclusions to understand the more specialized cessation.

Finally, overall connections will be drawn for a clear conclusion on the data, topic, and thesis. See table A for the primary regression on central bank independence and stimulus size and all six data sets. See table B for the secondary regression in which policy variables and stimulus size are measured in seven categories. Table C introduces the final variable for the

viewing of all data. Standardized data tables are used to signify the correlations between each data set. Suppose that high central bank independence index scores are cohesive with larger stimulus size. The level comparison of data applied allows a succinct environment for the support of this hypothesis. Beyond that, uniform formatting across tables creates an environment where the next conclusion is supported, as the second regression uses lapidary data to prove the correlation between larger stimulus packages.

**Results**

After running analyses on the first dataset, table A reveals the results of six fixed panel regressions. In all regressions, the CBI index's coefficient is steadily negative[5] (Elgin, C, Yalaman, A, Yasar S, Basbug G, 2021). This indicates that countries linked to higher central bank independence rates are also found to have lower fiscal policy cuts, as well as lower reserve requirements across branch banks. On the contrary, the coefficient for GDP per capita is noted to be consistently positive across every regression, (Elgin, C, Yalaman, A, Yasar S, Basbug G, 2021). This coefficient environment illustrates that the mentioned larger policy rate cuts and reserve requirements are observed in generally richer nations. Additionally, this regression illustrates a negative coefficient for inflation, (Elgin, C, Yalaman, A, Yasar S, Basbug G, 2021). This is key in understanding that countries with higher inflation rates pre COVID-19 implemented smaller expansionary policies.

**Table A**

Panel regressions of rate cuts.

| | Policy rate | Policy rate | Policy rate | Res. Req. | Res. Req. | Res. req. |
|---|---|---|---|---|---|---|
| CBI | −19.79*** | −22.51*** | −32.77*** | −15.43*** | −19.17*** | −25.91*** |
| | (−5.19) | (−5.45) | (−7.36) | (−3.76) | (−4.70) | (−5.72) |
| Infection rate | | 0.72*** | 0.29** | | −0.46*** | −0.71*** |
| | | (3.79) | (1.99) | | (−3.35) | (−4.62) |
| Case-Fatality | | 319.71 | 758.31** | | 728.02* | 1047.22** |
| | | (0.87) | (1.99) | | (1.75) | (2.46) |
| GDP per capita | | 0.29*** | 0.28*** | | 0.35*** | 0.43*** |
| | | (5.03) | (4.40) | | (5.45) | (5.40) |
| Growth Forecast | | −0.22 | −0.72*** | | −0.13 | −0.17 |
| | | (−1.19) | (−3.18) | | (−0.60) | (−0.78) |
| Inflation | | | −0.36*** | | | −0.70*** |
| | | | (−3.95) | | | (−6.38) |
| Stringency | | | 0.08*** | | | −0.04 |
| | | | (2.70) | | | (−0.86) |
| Developed | | | −6.48 | | | −25.33*** |
| | | | (−1.32) | | | (−4.83) |
| Lower Bound | | | −5.84* | | | 7.59** |
| | | | (−1.71) | | | (2.55) |
| Interest Rate | | | −0.63*** | | | 0.35** |
| | | | (−4.19) | | | (2.11) |
| Observations | 2252 | 2052 | 1902 | 2091 | 1948 | 1802 |
| R-squared | 0.06 | 0.12 | 0.25 | 0.09 | 0.12 | 0.17 |
| Exchange Rate Regime | NO | NO | YES | NO | NO | YES |

***p<0.01.
**p<0.05.
*p<0.1.
All regressions include regional, and version dummies. Robust t statistics are in parentheses.

Thereafter, view the regressions in table B. Within this dataset, COVID-19 economic stimulus index or CESI scores were analyzed among population age averages. Furthermore, the dataset reveals that in situations where median age has a positive coefficient, larger economic aid

package size follows. Furthermore, throughout the data regressions in both tables B and C, there is a significant positive relationship between median population age and fiscal stimulus package size.

**Table B**

| Dep. Var. | CESI | CESI | CESI | CESI | CESI | CESI | CESI |
|---|---|---|---|---|---|---|---|
| Median Age | 0.07* | 0.10* | 0.09* | 0.09* | 0.06* | 0.05* | 0.05* |
| | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) |
| Hospital Beds (per-capita) | | -0.15* | -0.13* | -0.12** | -0.11* | -0.11* | -0.11* |
| | | (0.04) | (0.04) | (0.04) | (0.06) | (0.04) | (0.04) |
| Infection Rate (%) | | | 546.25* | 224.83 | -69.56 | -149.34 | -151.30 |
| | | | (211.49) | (237.44) | (196.50) | (220.96) | (225.97) |
| Stringency Index | | | | 0.004 | | | |
| | | | | (0.006) | | | |
| GDP per-capita (000 USD) | | | | | 0.03* | 0.03* | 0.03* |
| | | | | | (0.001) | (0.001) | (0.001) |
| Total Cases | | | | | | 0.007** | 0.008** |
| | | | | | | (0.003) | (0.004) |
| Health Expenditures (% GDP) | | | | | | | -0.03 |
| | | | | | | | (0.04) |
| *R*-squared | 0.27 | 0.31 | 0.34 | 0.30 | 0.43 | 0.43 | 0.43 |
| Observations | 146 | 146 | 143 | 69 | 140 | 140 | 139 |
| F-Test | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |

Within table C, where hospital beds per capita is newly included into the dataset, conclusions slightly differ. It's observed that the value for hospital beds per capita has a significant negative coefficient and correlation with economic stimulus size, with its own negative coefficient, respectively (Elgin, C, Basbug, G, Yalaman, A, 2020).[6] Moreover, within nations where the amount of COVID-19 induced hospital beds is lower, proportionately, more

| | Mean | Median | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| COVID-19 Economic Stimulus Index(CESI) | 0.00 | -0.31 | 1.28 | -4.25 | 4.85 |
| Fiscal Policy Stimulus (%) | 2.09 | 0.48 | 3.60 | -7.20 | 17.80 |
| Interest Rate Cut (%) | 11.63 | 0.00 | 21.47 | -29.73 | 100.00 |
| Macro-Financial Package (% of GDP) | 1.87 | 0.00 | 4.02 | 0.00 | 26.00 |
| Other Monetary Measures(0-1 dummy) | 0.85 | 1.00 | 0.36 | 0.00 | 1.00 |
| BoP Measures (% GDP) | 0.10 | 0.00 | 0.58 | 0.00 | 6.00 |
| Other BoP Measure (0-1 dummy) | 0.19 | 0.00 | 0.40 | 0.00 | 1.00 |

severe fiscal stimulus policies are implemented.

**Table C**

**Analysis**

Beginning with the datasets from table A, the relationship between CBI, reserve requirements, and fiscal policy cuts is revealed. The three have a positive relationship, as they are all tied to each other. In other terms, as the independence of central banks leans in one direction, for example, with a negative coefficient, reserve requirements and fiscal policy cuts follow with in turn, low values and negative coefficients. Furthermore, as seen in table A, countries with higher CBI figures, respectively, experience smaller forecasted economic growth rates. In turn, a negative relationship between central bank independence scores and forecasted economic growth is established within the dataset. Cumulatively, high central bank independence scores are correlated with higher forecasted growth as well as high fiscal policy cuts and reserve requirements. High fiscal policy cuts, respectively, are a key principle of expansionary economic policy, common in economies recovering and growing in post recessionary societies. The presence of such variables and their relationships with central bank independence further supports the principle of central bank independence's inherent correlation with economic growth and recovery in expansionary times.

Among other variables, fiscal policy, however, differs in correlation from its reductional counterpart. As seen in table B, fiscal stimulus itself is found to have a positive relationship with a different factor, in this case, median population age. Expeditiously, fiscal stimulus package size is directly tied to median population age. In that respect, if median population age were to lean on a lower value, the population's fiscal aid package sizes would follow suit, therefore leaning negatively. However, when including regressions from table C, the established relationship with median population age changes in the context of its other key factor, the variable of hospital beds per capita and its surrounding relationships. As this is pandemic based, fiscal data, it is also possible to draw pandemic-based data and conclusions. We then introduced the variable of hospital beds per capita. As hospital beds per capita is introduced, regressions in the dataset reveal that hospital beds per capita, have a significant negative relationship with fiscal stimulus package size. Moreover, higher median population age is correlated with higher fiscal stimulus package size. In turn, high fiscal stimulus package sizes lead to decreased hospital beds per capita. Therefore, successful fiscal aid and pandemic mitigating behaviors directly depend on large fiscal packages.

With tables A, B, and C's data regressions cross referenced, variables correlating with each other form a concise idea. As high central bank independence leads to high fiscal policy variability, a positive relationship is established with median population age. Among median population age, higher fiscal stimulus is recorded, therefore establishing the negative relationship with hospital beds per capita.

**Concluding Remarks**

As post COVID-19 economies continue to recover from times of downturn, it is imperative to analyze past conflicts and the conditions surrounding themselves and their amendments. In this case, with the aim of analyzing the environment in which worldwide

economies attempted to recover from pandemic-induced economic downturn, valuable connections and relationships have been established between variables. With this information in mind, it is imperative to recognize the limitations of our data, as it is secondary. Despite this, strong figures and the nature of the data itself proves reliability across regressions. Moreover, the established relationships prove the conclusion of high central bank independence's direct relation to reduced hospital beds per capita, through extensive correlation and datasets.

Works Cited

Glinavos, Ioannis. A Friend in Need? The Effect of Institutional Independence on Central Bank Responses to the COVID-19 Crises, (November 22, 2020). University of Westminster School of Law Research

Papers.cfm

Key Words: Central Banking, COVID Pandemic, Policy response, COVID-19, Institutional Independence, Economic Emergency

The article mainly explains how central bank independence plays a role in independence of central banks and their policy efficacy in relation to COVID-19 based economic crises. Currently, I intend to use it by examining and referencing their analysis on independence and policy effects. Objectively, it helps me gain an understanding about how fiscal and political policies are tied to bank independence, as well as how they can change economic circumstances across nations during crises.

Elgin, Ceyhun, Abdullah, Yalaman, Yasar, Sezer, Basbug, Gokce. Economic Policy Responses to the COVID-19 Pandemic: The Role of Central Bank Independence, (April 27, 2021). ElSevier Research Letters

View

Key Words: Monetary Policy, COVID-19, Fiscal Policy, Central Bank Independence

Downloadable as PDF file

This article focuses on how central bank independence influences the fiscal responses that different nations adopt in response to COVID-19 fueled crises. I will use this article as a core source for my paper as it almost exactly discusses my most pivotal points. This paper uses direct data regressions and studies to explicate information that supports the thesis. Using this paper and its data as a reference, I will have a multitude of credible and relevant studies to aid my explanation within my paper.

Stottele, Emelle. A Comparative Analysis of the 14 Central Banks' Monetary Responses Before and After the COVID-19 Pandemic in Relation to Central Bank Independence, (2023). CMC Senior Theses

viewcontent.cgi

Key Words: Monetary Policy, COVID-19, Fiscal Policy

Downloadable as PDF file

This article provides a strong reference for a thesis, with clear and concise data and studies to back the explanation. This paper will be the source of a large portion of my data, as well as a concise reference for logical idea and thesis development. Using its original data and scholarly writing, this paper will be an asset when writing and sourcing data for my respective paper.

English, Bill, Forbes, Kristin, Ubide, Angel. Monetary Policy and Central Banking in the COVID Era, CEPR Press

60024-monetary_policy_and_central_banking_in_the_covid_era.pdf

Key Words: Monetary policy, Banking, Central Banks, Fiscal Policy, COVID-19, COVID-19 economics

Downloadable as PDF file

Seeing as this article is extremely content heavy, this paper will give me multiple important examples for theses, idea progression, and idea explanation. I intend to reference this both in my writing and in the process because there is so much reasoning and material to use and build off of. Due to its high content and credible sources, data, and analysis, this paper will be yet another asset to my review paper.

Bernholz, P. Independent Central Banks as a Component of the Separation of Powers, ( September 8, 2013). Const Polit Econ
S10602-013-9142-y
Key Words: Inflation, Separation of Powers, Monetary Institutions
Downloadable as a PDF file

This article details important information about central banks, their independence, and separation of powers prior to COVID-19. With these pre-pandemic data regressions and reasonings, I will be able to accurately read how the pandemic has affected central bank independence as well as separation of powers. As it is credible and relevant, this source will prove to be valuable in both its data as well as comparisons throughout my paper.

Elgin, Ceyhun, Basbug, Gokce, Yalaman, Abdullah. Economic Policy Responses to a Pandemic: Developing the COVID-19 Economic Stimulus Index, ( April 10, 2020). CEPR Press
View
Key Words: COVID-19, Stimulus, Fiscal Aid
Downloadable as a PDF file

This article aims to measure the effectiveness of differently sized fiscal aid packages in response to COVID-19. This will develop my information before the conclusion by providing a clear reference of how the size of aid packages either benefit or hinder their success. It is credible and recent with authority on the subject, which proves to be valuable in this analysis

[1](Elgin, C, Yalaman, A, Yasar S, Basbug G, 10, 2021). All mathematical work accredited to contributors, see bibliography. Equation is referenced again in later data regressions with all work accredited to the same authors, see bibliography.

[2] The respective dummy variable takes the value of 1 if similar measures are present, while taking the value of 0 if they are not.

[3] (Elgin, C, Basbug, G, Yalaman, A, 8, 2020). Regressions and data drawn from cited contributors. All studies are reviewed and accredited to aforementioned authors, see bibliography.

[4] All data is taken from the stimulus index as it is an independent organization with no literature to cite.

[5] (Elgin, C, Yalaman, A, Yasar S, Basbug G, 12, 2021). The data and rates take numerical values along with the variables, meaning that coefficients signify not only the presence but the prevalence of the specified factors.

[6] In this context, negative coefficients are again used to highlight the negative presence of these variables in order for a concise portrayal of the negative relationship to be communicated.

**The Impact of WTI Crude Oil Price Changes on Producer Prices Across U.S. Industries, 2006–2026 By Hridhaan Kataria**


## Abstract

This study examines the relationship between West Texas Intermediate (WTI) crude oil prices and production costs, measured through Producer Price Indices (PPI), in seven U.S. industries: agriculture, textiles, fertilizer, transportation, transportation and warehousing, plastics, and industrial chemicals. WTI crude oil prices and PPI data were retrieved from the Federal Reserve Economic Data (FRED) database, covering December 2006 to May 2026. Pearson correlations were calculated across three analytical windows: the full sample period; a rising oil price window from November 2021 to June 2022; and a falling oil price window from July 2022 to June 2023. Percentage-change ratios were also computed to assess the magnitude of price adjustment relative to oil price movements. Fertilizer was the only industry that showed a clear long-term association with oil prices, exhibiting a moderate positive correlation ($r = 0.49$). When oil prices rose sharply from November 2021 to June 2022, all seven industries showed strong positive correlations, with an average r of 0.89. During the oil price decline from July 2022 to June 2023, six industries still maintained positive correlations. Transportation, however, shifted to $r = -0.91$: its PPI increased by 2.5% even as oil prices fell by 31.0%. Agriculture and fertilizer prices adjusted substantially whether oil prices rose or fell, while several manufacturing sectors adjusted much less when oil prices dropped. These patterns suggest that oil prices are more closely associated with producer prices during rapid market shifts than over the long run, and that this association is not uniform across industries. Because the sub-period findings rely on small samples ($n = 8$ and $n = 12$) and the full-sample correlations are drawn from trending price levels, the results are best read as descriptive associations rather than causal estimates; a robustness check using month-on-month changes is reported in Section 4.7.


## 1. Introduction

Oil prices shape how much things cost to make, since oil fuels machines, moves goods, and supplies the raw materials for plastics and many other factory inputs. Rising crude prices ripple through supply networks, raising the cost of shipping, power, packaging, and base materials. Producers feel these swings sharply, and so do buyers and the officials who monitor economic stability. Industries that rely heavily on energy or petrochemical inputs face the greatest risk when oil prices spike or collapse. In its April 2026 *World Economic Outlook*, the International Monetary Fund warns that the energy price shock linked to conflict in the Middle East is raising commodity prices and adding to inflation, and Federal Reserve research shows that higher oil prices feed into production and transport costs before appearing in the prices of food and other goods (Alp et al.). The World Bank's April 2026 *Commodity Markets Outlook* similarly warns that the surge in energy and commodity prices will push up inflation and weaken growth among both commodity importers and exporters.

This research paper examines how shifts in world oil prices are associated with production costs in agriculture, textiles, fertilizer, transportation, transportation and warehousing, plastics, and industrial chemicals. It proposes that higher crude oil prices tend to accompany higher production costs, especially where petroleum serves as an energy source or a raw material. Still, the strength of this association should vary across industries depending on how much energy they consume, how much they rely on extended supplier networks, and whether companies can pass added costs on to buyers (Jo et al.; Alp et al.).

Although farming's dependence on oil is partly indirect, its effects show up clearly in daily operations: running machinery, pumping water, and moving goods all depend on petroleum. USDA data show that fuel and fertilizer together account for about a tenth of U.S. farm production spending, roughly $49 billion of $478 billion in 2024 (*Farm Production Expenditures*), and an earlier USDA study found that energy-based inputs made up about a third of corn production costs in 2006–2011 (Beckman et al.). When natural gas prices shift, so do the prices of nitrogen fertilizers. Disruptions to shipping routes or limits on imports add further pressure beyond broad energy trends. Transportation and warehousing are exposed in a different way: diesel powers most freight movement, so costs climb quickly when diesel prices jump. Because synthetic materials such as polyester are made from oil, movements in crude markets are also linked to textile production; polyester alone makes up about 60 percent of global fiber output ("Textile-Clothing"). Finally, petrochemical producers turn feedstocks such as naphtha and ethane into the building blocks of plastics and basic industrial chemicals, tying these sectors closely to fossil fuel markets ("How Much Oil"; *Future of Petrochemicals*).

The purpose of this paper is to examine how shifts in oil prices are associated with production costs in specific industrial sectors and to highlight those most at risk. It focuses on production-level costs rather than retail prices, while noting where such costs are passed downstream. After an overview of how crude oil trends relate to operating costs, the analysis moves to a sector-by-sector evaluation in which each industry is assessed for its sensitivity to petroleum market swings. A discussion of these patterns and their limitations follows, and the paper concludes by identifying the industries most exposed to instability in world oil prices.

**2. Literature Review**

Existing research establishes that oil prices have both direct and indirect relationships with production costs. Directly, oil affects the price of fuel, transportation, shipping, and warehousing. Indirectly, oil and gas influence the cost of materials, including fertilizer, plastics, chemicals, and synthetic fibers. Jo et al. explain that oil price shocks can act as supply shocks because energy is an input to production, although they find that this cost channel dominates only in the most oil-intensive U.S. industries, while demand effects dominate in most others. Oil prices are therefore linked to the production costs of many goods and services before being reflected in consumer prices (Alp et al.). Recent reports by the International Monetary Fund and the World Bank also connect energy shocks with inflationary pressure, especially when fuel, fertilizer, and transportation costs rise (*World Economic Outlook*; *Commodity Markets Outlook*).

Agriculture is among the sectors most exposed to crude oil prices. According to a U.S. Department of Agriculture Economic Research Service report, the sector uses energy directly as fuel and electricity and indirectly as an input in the production of fertilizers and pesticides, and its production costs rise when the prices of energy-intensive inputs increase (Beckman et al.). Reports from the USDA National Agricultural Statistics Service show that fuel and fertilizer remain significant shares of farm spending (*Farm Production Expenditures*). However, most existing studies focus on agricultural production and its response to energy prices rather than on the relationship between oil prices and production costs in other industries.

Studies of supply chains also show the relationship between fuel prices and production costs. Milewska and Milewski examine how rising fuel and energy prices affect transport and logistics costs in global supply chains, and whether higher shipping costs make it worthwhile to move production closer to customers. They find that, despite a sharp rise in freight rates, relocating production from the Far East to Europe was still not profitable. This literature, however, tends to analyze transportation costs separately from other production expenditures, even though both are part of total production costs, which leaves a gap in understanding how fuel prices relate to production costs as a whole.

Petrochemical and textile studies also demonstrate the connection between oil prices and production costs. The U.S. Energy Information Administration explains that naphtha and other petroleum products, along with hydrocarbon gas liquids from natural gas processing, are used as feedstocks for plastics ("How Much Oil"). The International Energy Agency emphasizes the significance of petrochemicals as an energy-intensive industry (*Future of Petrochemicals*). Textile Exchange reports that polyester is the most widely produced fiber in the world, accounting for 59 percent of global fiber output, and that most of it is fossil-based (*Materials Market Report*). Similarly, Coface links oil price movements to fluctuations in the prices of synthetic fibers ("Textile-Clothing"). Taken together, these studies tend to discuss the effect of oil prices on individual industries in isolation. The current study addresses this gap by comparing the relationship between oil prices and production costs across the agricultural, textile, transportation, warehousing, plastics, and chemicals sectors to identify which areas of production are most closely associated with oil price changes.

## 3. Methodology

### 3.1 Aims

This study aims to identify how changes in oil prices are associated with production costs across various industries. I focus on sectors that rely on oil either directly, like transportation and warehousing, or indirectly, such as agriculture, plastics, chemicals, and textiles. By examining these differences, I hope to identify which industries are most exposed to oil price swings and explain how these fluctuations can ripple through the broader economy. To achieve this, the research addresses several objectives:

- Gather secondary data on oil prices and production costs from government reports, academic studies, and industry sources.
- Compare how oil prices relate to different industries, including agriculture, transportation, warehousing, plastics, chemicals, and textiles.
- Determine whether oil is linked to each industry directly (e.g., through fuel and energy costs) or indirectly (e.g., through raw materials and supply chain costs).
- Analyze which industries are most susceptible to rising production costs when oil prices increase.

### 3.2 Research Design

This study uses a secondary research design to compare how oil prices are associated with production costs across different sectors. This means I rely on existing data and information that other researchers, government agencies, and international organizations have already collected. I chose this design because oil prices and production costs are such large-scale economic issues that it would be impractical to gather all the necessary data ourselves through surveys or experiments.

I compare the correlation between crude oil prices and the PPI of seven industries over the period 2006–2026, using data from the Federal Reserve Bank of St. Louis's *FRED Economic Data* database (FRED). Six of these industries (agriculture, transportation, transportation and warehousing, plastics, industrial chemicals, and textiles) are described in Table 1; fertilizer is treated as a closely related seventh series and analyzed alongside them throughout the Results section (Section 4), because fertilizer is where energy prices are believed to affect farming costs most directly. Each industry was selected because of its connection to oil prices, either through consuming oil directly or relying on oil-based materials. For instance, transportation depends heavily on fuel, while plastics and chemicals use oil and gas as raw materials. Agriculture was included because of the link between energy prices and both fuel and fertilizer costs, and textiles were included because many synthetic fibers are fossil-fuel based.

The research is mainly quantitative, using correlation analysis to examine the relationship between WTI crude oil prices and the PPI of these industries from 2006 to 2026. However, I also use qualitative information from existing literature to help explain the patterns found in the quantitative data. This blended approach allows us to identify the statistical relationships between oil prices and production costs while also explaining the underlying reasons for these patterns. Because correlation analysis alone cannot establish causation, all quantitative findings in Section 4 are framed as associations rather than causal effects, and I discuss this limitation further in Section 5.

### 3.3 Sectors Studied

I selected the following sectors for comparison in this research.

#### *3.3.1 Industries*

Table 1: The industries studied and their connection to oil prices

| Sector | Reason for inclusion | Link to oil prices |
|---|---|---|
| Agriculture | Farms use fuel, machinery, transport, and fertilizer. | Oil and energy prices can increase fuel and fertilizer costs. |
| Transportation | This sector depends directly on fuel for moving goods and people. | Rising oil prices can increase transport and delivery costs. |
| Transportation and warehousing | Storage and distribution involve energy use and the movement of goods through supply chains. | Higher fuel and energy costs can raise distribution and storage expenses. |
| Plastics | Many plastics are made using oil and natural gas feedstocks. | Oil prices can affect the cost of raw materials used in plastic production. |
| Industrial chemicals | Petrochemicals are used in many chemical products. | Oil and gas prices can affect the input costs of chemical production. |
| Textiles | Many synthetic fibers, such as polyester, are fossil-fuel based. | Oil prices can influence the cost of producing synthetic fibers. |

Table 1 outlines the sectors included in this study. I chose these industries because of their diverse connections to oil prices. Some sectors, like transportation, are directly affected by fuel costs. Others, such as plastics, chemicals, and textiles, are affected through raw materials derived from oil and gas. Agriculture is influenced by both direct fuel use and fertilizer costs, making it a crucial sector for comparison.

### *3.3.2 Secondary Sources*

Table 2: The main secondary sources used in the research and their purpose

| Source | Type of source | Purpose in the research |
|---|---|---|
| Federal Reserve Bank of St. Louis (FRED) | Economic database | Source of the monthly WTI crude oil price and PPI series analyzed in Section 4. |
| Jo, Karnizova, and Reza | Federal Reserve working paper | Used to explain how oil price shocks affect industries through supply (cost) and demand channels. |
| Alp, Klepacz, and Saxena | Federal Reserve economic note | Used to support the link between oil prices, production costs, and later consumer prices. |
| USDA Economic Research Service (Beckman, Borchers, and Jones) | Government report | Used to explain how energy, fuel, and fertilizer affect agricultural costs. |
| USDA National Agricultural Statistics Service | Government data report | Used to support recent farm expenditure data. |
| U.S. Energy Information Administration | Government information source | Used to explain how oil and gas are connected to plastic production. |
| International Energy Agency | International report | Used to support information about petrochemicals and energy use. |

| Source | Type of source | Purpose in the research |
|---|---|---|
| International Monetary Fund | International economic report | Used to explain how oil shocks affect inflation and economic stability. |
| World Bank | International economic report | Used to support information about energy prices and commodity markets. |
| Milewska and Milewski | Academic journal article | Used to explain how fuel costs affect supply chains and logistics. |
| Textile Exchange | Industry report | Used to explain the importance of synthetic fibers such as polyester. |
| Coface | Industry risk report | Used to support information about risk in the textile and clothing sector. |

Table 2 presents the sources used in this study. The research draws on a mix of academic studies, government reports, international reports, and industry sources. This diverse approach makes the research more reliable by not relying on a single type of source.

## 3.4 Data Collection

### *3.4.1 Secondary Research*

I collected data for this study solely through secondary research, meaning I did not conduct any surveys, interviews, or experiments ourselves. Instead, I gathered information from existing reports, articles, and studies that discuss oil prices, energy costs, supply chains, and production costs.

For the quantitative part of the research, I collected data from the FRED database. I downloaded monthly data from 2006 to 2026 for the West Texas Intermediate (WTI) crude oil price and the Producer Price Index (PPI) for the industries studied: agriculture, transportation, transportation and warehousing, plastics, industrial chemicals, textiles, and fertilizer. The data were organized in a spreadsheet by month so that the WTI oil price could be compared with the PPI for each industry, and then used to calculate the correlation between WTI crude oil prices and the PPI of each industry.

Our key sources included government agencies such as the USDA and the U.S. Energy Information Administration, international organizations such as the IMF, the World Bank, and the International Energy Agency, and academic studies such as Jo et al. and Milewska and Milewski. I selected these sources because they are reliable in providing information on how oil prices relate to various aspects of production.

The data I collected focused on three main areas:

a.Oil price effects

1.How changes in oil prices are associated with production costs.

2.How oil price shocks can act as supply shocks.

3.How higher oil prices can then be linked to consumer prices.

b.Industry links to oil

1.Agriculture's dependence on fuel and fertilizer.

2.Transportation's reliance on fuel.

3.How warehousing is affected by energy and supply chain costs.

4.How plastics and chemicals depend on oil and gas feedstocks.

5.How textiles are affected through synthetic fibers such as polyester.

c.Sector comparison

1.Identifying sectors directly affected by oil prices.

2.Identifying sectors indirectly affected by oil prices.

3.Identifying industries likely to see the largest increases in production costs.

### *3.4.2 Source Selection*

To ensure the research was accurate and relevant, I chose sources based on their reliability, relevance, and direct connection to our research question. Government and international organization sources provided official economic and industry data. Academic sources explained the theory of oil price shocks and production costs. Industry sources offered specific details about sectors such as textiles and petrochemicals. I only included sources that were directly related to oil prices, energy costs, production costs, or industry-specific impacts, keeping the focus on the main research question.

## 3.5 Data Analysis

### *3.5.1 Qualitative: Comparative Analysis*

I used a comparative analysis to examine the information I collected. This involved comparing each sector based on how oil prices are associated with its production costs, and whether oil is linked to it directly through fuel and energy costs or indirectly through raw materials and supply chain costs. For instance, I analyzed transportation as a sector with direct oil dependence because fuel is a primary cost. Plastics and chemicals were analyzed differently, as oil and natural gas serve as raw materials in their production. Agriculture was compared using both categories, as it directly uses fuel and also relies on fertilizer, which is tied to energy prices.

### *3.5.2 Quantitative: Descriptive Data*

I also incorporated quantitative data, such as percentages, cost categories, and expenditure data from sources like the USDA, wherever it was available. This numerical information helped support the written analysis with concrete evidence.

The quantitative data were analyzed using correlation analysis to measure the strength and direction of the relationship between crude oil prices and the Producer Price Index (PPI) of each industry. This statistical method helped identify patterns and compare how strongly different industries were associated with changes in oil prices. However, the study did not use a full

predictive statistical model, as the main goal was to describe and compare these associations rather than to make exact predictions about future production costs, and the correlational design means the results cannot establish that oil prices caused the observed changes in producer prices.

### 3.6 Ethics

Since this paper is based on secondary research, no human participants were involved, which means I did not collect any personal data, names, emails, or private information. As a result, the ethical risks were significantly lower than in studies involving surveys or interviews. Nevertheless, ethical considerations were still important. All information used in the paper was drawn from published sources, and I referenced these sources using MLA (9th edition) style. I summarized findings from other researchers in our own words rather than copying them directly. Data integrity was another ethical consideration. To maintain accuracy and transparency, all data were used as published by the original sources, without manipulation or selective reporting. I also used a diverse range of sources, including government reports, academic studies, and international organizations, so that the findings rest on reliable evidence and do not lean too heavily on a single source or viewpoint.

## 4. Results

This section reports the relationship between crude oil prices and the Producer Price Index (PPI) for each industry. The analysis uses 234 months of data, from December 2006 to May 2026, with the West Texas Intermediate (WTI) crude oil price as the independent variable and each industry's PPI as the dependent variable. The findings are presented in three parts:

- the long-term relationship over the whole period,
- a period when oil prices rose sharply (November 2021 to June 2022), and
- a period when oil prices fell (July 2022 to June 2023).

Results are presented using descriptive statistics, Pearson correlation coefficients, and graphs. This section reports the findings only; their interpretation follows in the Discussion (Section 5). Seven industry series are analyzed: the six described in Section 3.3 plus fertilizer, which was collected separately because it is the channel through which energy prices affect farming costs most directly.

A note on interpretation: the two sub-period windows are short ($n = 8$ for the rising window and $n = 12$ for the falling window), so their Pearson correlations, while statistically significant, rest on very few observations and should be read as suggestive rather than robust. In addition, the full-sample correlations in Section 4.2 are calculated on price levels, and oil prices and PPI series can share a common upward trend over 20 years regardless of any short-run relationship. Section 4.7 therefore reruns the full-sample analysis on month-on-month changes to check how much of the level-based correlation reflects shared trends rather than genuine co-movement.

### 4.1 Descriptive Statistics

Table 3 shows the descriptive statistics for WTI crude oil prices and the PPI of each industry for the entire study period.

Table 3: Descriptive statistics for WTI crude oil prices and industry Producer Price Indices, December 2006 – May 2026 (n = 234 monthly observations)

| Series | Mean | SD | Min | Max | First value (Dec 2006) | Last value (May 2026) | Total change |
|---|---|---|---|---|---|---|---|
| WTI crude oil (US$/barrel) | 72.83 | 21.56 | 16.98 | 133.96 | 62.03 | 104.07 | +67.8% |
| Agriculture (PPI) | 127.8 | 15.5 | 103.6 | 191.7 | 103.6 | 130.5 | +26.0% |
| Textiles (PPI) | 152.1 | 20.4 | 125.0 | 195.8 | 125.0 | 195.8 | +56.6% |
| Fertilizer (PPI) | 211.2 | 52.3 | 122.0 | 360.9 | 122.0 | 320.6 | +162.8% |
| Transportation (PPI) | 122.5 | 12.1 | 103.8 | 148.2 | 104.8 | 148.2 | +41.4% |
| Transportation and warehousing (PPI) | 135.1 | 22.5 | 100.0 | 194.6 | 100.0 | 194.6 | +94.6% |
| Plastics (PPI) | 174.2 | 30.2 | 132.8 | 236.1 | 134.3 | 236.1 | +75.8% |
| Industrial chemicals (PPI) | 283.7 | 48.1 | 197.0 | 384.2 | 197.0 | 384.2 | +95.0% |

*Source: Author's calculations using monthly data from the Federal Reserve Economic Data (FRED) database.*

Over the study period, the WTI crude oil price averaged US$72.83 per barrel, with a standard deviation of US$21.56. It reached its highest point, US$133.96, in July 2008 and its lowest, US$16.98, in April 2020, a swing of more than US$115 per barrel. Figure 1 shows how the oil price moved during the whole period.

The PPI series behaved quite differently. Fertilizer recorded the largest overall increase, up 162.8%, and also the most variation, with a standard deviation of 52.3; it started at 122.0 in December 2006 and reached a high of 360.9 in October 2008. Industrial chemicals and transportation and warehousing nearly doubled, rising 95.0% and 94.6% respectively. Agriculture had the smallest increase, at 26.0%, and transportation the second smallest, at 41.4%, although the transportation index rose steadily rather than spiking. Most of the series (textiles, transportation, transportation and warehousing, plastics, and industrial chemicals) reached their highest point in May 2026, the last month of the sample.

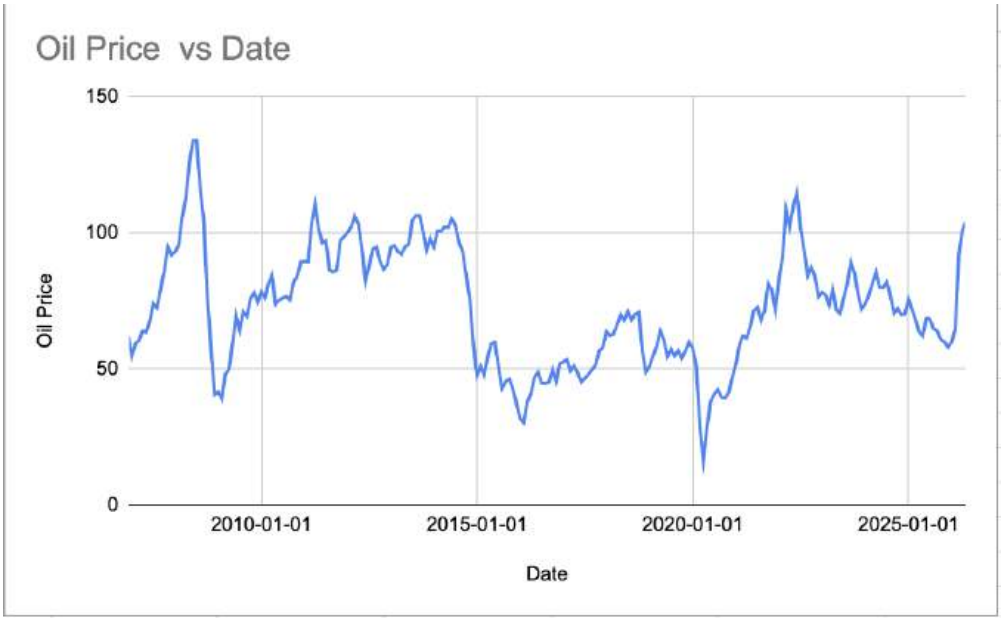



Figure 1: WTI crude oil price, December 2006 – May 2026

### 4.2 Long-Run Correlations, December 2006 – May 2026

Table 4 reports the Pearson correlation coefficients between the WTI crude oil price and each industry PPI across the full sample.

Table 4: Correlation between WTI crude oil prices and industry PPI, full sample, December 2006 – May 2026 (n = 234)

| Industry | r | p value | Strength of relationship |
|---|---|---|---|
| Fertilizer | 0.494 | < 0.001* | Moderate positive |
| Agriculture | 0.082 | 0.214 | Little to none |
| Plastics | 0.050 | 0.449 | Little to none |
| Textiles | 0.042 | 0.528 | Little to none |
| Transportation and warehousing | 0.038 | 0.558 | Little to none |
| Industrial chemicals | −0.006 | 0.922 | Little to none |
| Transportation | −0.121 | 0.065 | Little to none |

**p < 0.05. Industries are ordered from the strongest to the weakest correlation.*

Over the full period, only one industry showed a correlation of meaningful size. Fertilizer and oil prices had a moderate positive correlation ($r(232) = 0.49$, $p < 0.001$), the only relationship in Table 4 that was statistically significant at the 5% level. The other six industries all fell between −0.3 and +0.3, indicating little to no linear relationship with oil prices in the long run. Industrial chemicals had a correlation of almost zero ($r = −0.006$, $p = 0.922$), and transportation had a weak negative coefficient ($r = −0.121$, $p = 0.065$).

Figures 2 and 3 plot the two most distinct cases against the oil price for the whole period: fertilizer, which had the strongest long-run correlation, and transportation, which had the lowest. Charts for the other five industries appear in Appendix A.

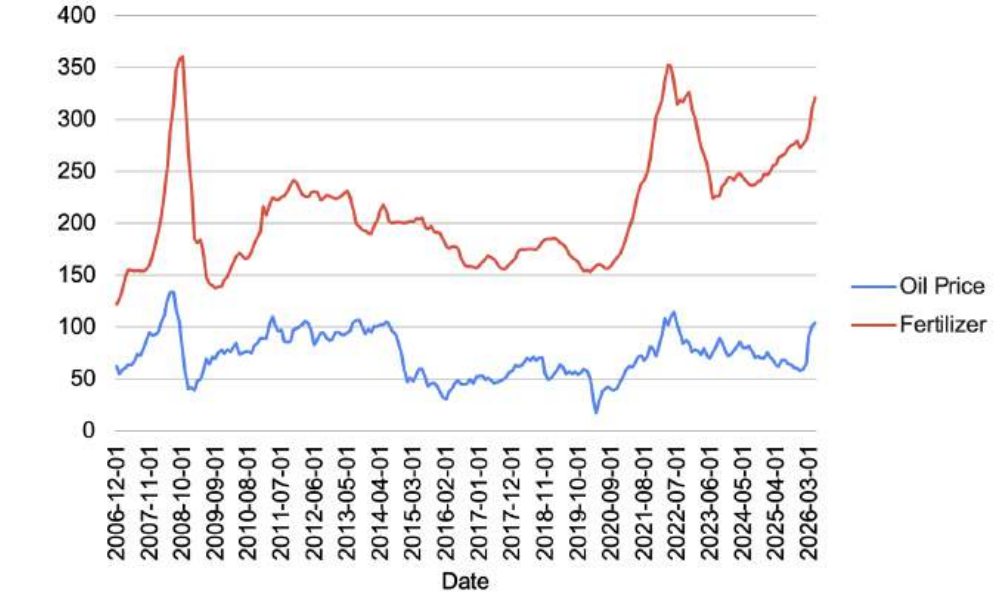



Figure 2: WTI crude oil price and fertilizer PPI, December 2006 – May 2026

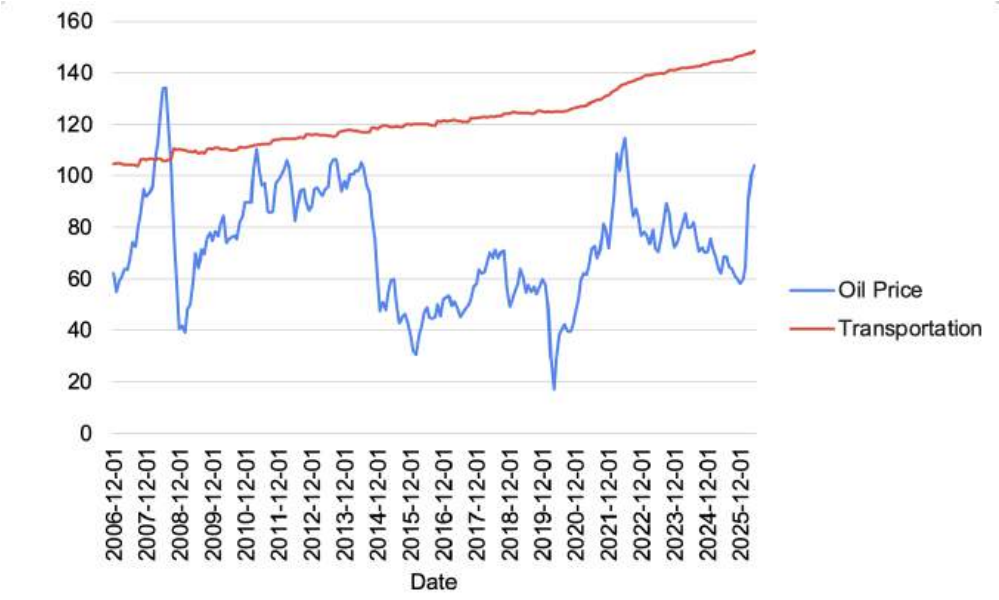



Figure 3: WTI crude oil price and transportation PPI, December 2006 – May 2026

Figures 2 and 3 help explain these coefficients. Oil prices were volatile, with a peak in 2008, a drop from 2014 to 2016, a crash in 2020, and a recovery from 2021 to 2022. The fertilizer index followed these cycles closely: its two biggest peaks, in 2008 and 2022, coincided with oil's two biggest peaks in the sample. The transportation index behaved differently, rising

almost continuously over the whole period with relatively little short-term variation, so a cyclical oil price series produced a near-zero long-run coefficient against it. The same upward trend without matching cycles is visible in the industrial chemicals, plastics, textiles, and transportation and warehousing charts in Appendix A. Because both fertilizer's cycles and transportation's trend unfold over many years, these level-based coefficients partly reflect long-run trend alignment rather than month-to-month co-movement; Section 4.7 examines this using month-on-month changes instead of levels.

### 4.3 The Rising Oil Price Window, November 2021 – June 2022

From November 2021 to June 2022, the WTI crude oil price rose from US$79.10 to US$114.68 per barrel, an increase of 45.0%. Over these eight months the price averaged US$95.07 and ranged from US$71.80 to US$114.68. Table 5 shows the correlations for this period.

Table 5: Correlation between WTI crude oil prices and industry PPI during the rising oil price window, November 2021 – June 2022 (n = 8)

| Industry | r | p value | Change in PPI over window | Strength of relationship |
|---|---|---|---|---|
| Transportation | 0.940 | < 0.001* | +3.7% | Strong positive |
| Transportation and warehousing | 0.937 | < 0.001* | +13.1% | Strong positive |
| Textiles | 0.926 | < 0.001* | +7.7% | Strong positive |
| Industrial chemicals | 0.926 | 0.001* | +8.7% | Strong positive |
| Plastics | 0.916 | 0.001* | +8.0% | Strong positive |
| Fertilizer | 0.855 | 0.007* | +19.5% | Strong positive |
| Agriculture | 0.748 | 0.033* | +34.1% | Strong positive |

**p < 0.05. The WTI crude oil price rose 45.0% over the same window. With only eight monthly observations, these coefficients should be treated as indicative rather than precise estimates.*

Every industry showed a strong positive correlation with oil prices during this window, and all seven coefficients were statistically significant at the 5% level. The coefficients were tightly clustered, ranging from r = 0.75 to r = 0.94, with an average of 0.89. Transportation had the highest coefficient ($r(6) = 0.94$, $p < 0.001$), followed by transportation and warehousing (r = 0.94), textiles (r = 0.93), industrial chemicals (r = 0.93), and plastics (r = 0.92). Agriculture had the lowest coefficient ($r(6) = 0.75$, $p = 0.033$), which still indicates a strong positive relationship.

However, the size of the price changes did not mirror the strength of the correlations. Oil rose 45.0% during this window, while agriculture PPI rose 34.1% and fertilizer PPI rose 19.5%. In contrast, transportation PPI rose only 3.7%, textiles 7.7%, plastics 8.0%, and industrial chemicals 8.7%. The industries whose prices followed oil's direction most closely were therefore

not necessarily those with the largest price movements. Figure 4 illustrates this for transportation: its index rose steadily but only slightly while the oil price climbed.

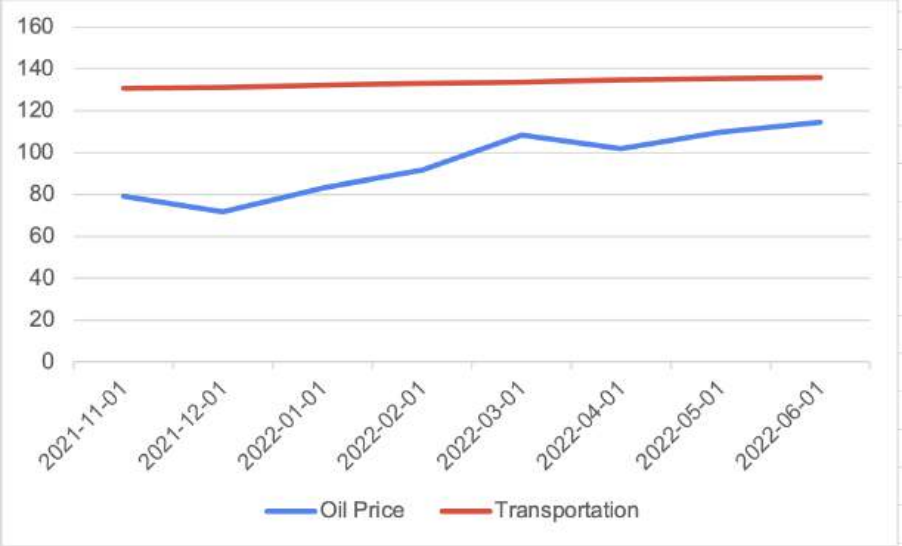



Figure 4: WTI crude oil price and transportation PPI during the rising oil price window, November 2021 – June 2022

### 4.4 The Falling Oil Price Window, July 2022 – June 2023

The next window covers the twelve months from July 2022 to June 2023, during which the WTI crude oil price fell from US$101.92 to US$70.31 per barrel, a decrease of 31.0%. The average price during this window was US$81.47. Table 6 shows the correlations for this period.

Table 6: Correlation between WTI crude oil prices and industry PPI during the falling oil price window, July 2022 – June 2023 (n = 12)

| Industry | r | p value | Change in PPI over window | Strength of relationship |
|---|---|---|---|---|
| Agriculture | 0.876 | < 0.001* | −21.8% | Strong positive |
| Industrial chemicals | 0.828 | < 0.001* | −2.6% | Strong positive |
| Plastics | 0.782 | 0.003* | −1.8% | Strong positive |
| Transportation and warehousing | 0.767 | 0.004* | −3.1% | Strong positive |
| Fertilizer | 0.724 | 0.008* | −22.8% | Strong positive |
| Textiles | 0.678 | 0.016* | −1.2% | Moderate positive |
| Transportation | −0.909 | < 0.001* | +2.5% | Strong negative |

**p < 0.05. The WTI crude oil price fell 31.0% over the same window. With only twelve monthly observations, these coefficients should be treated as indicative rather than precise estimates.*

Six of the seven industries still showed a positive correlation with oil prices in this window, with coefficients ranging from 0.68 for textiles to 0.88 for agriculture, and all six were

statistically significant. Because oil prices were falling, a positive coefficient here means that these industries' prices were also falling.

Transportation was the only exception. Its correlation reversed sign ($r(10) = -0.91$, $p < 0.001$), meaning that its PPI kept rising as oil prices fell.

The size of the price changes again varied widely between industries. While oil prices fell 31.0%, agriculture PPI dropped 21.8% and fertilizer PPI fell 22.8%. The other four industries with positive correlations changed very little: textiles fell 1.2%, plastics 1.8%, industrial chemicals 2.6%, and transportation and warehousing 3.1%. Transportation PPI, by contrast, rose 2.5%. Figure 5 shows this reversal.

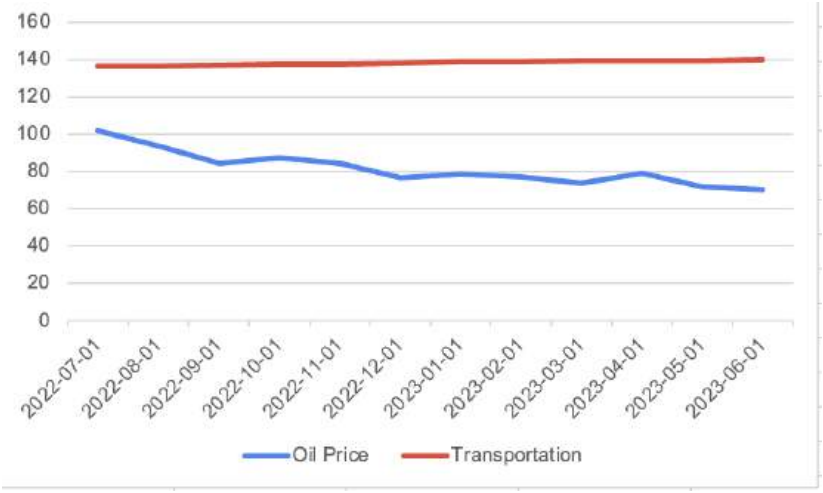



Figure 5: WTI crude oil price and transportation PPI during the falling oil price window, July 2022 – June 2023

### 4.5 Comparison Across the Three Windows

Table 7 brings together all three sets of results so that each industry can be compared across the full sample and the two shorter windows.

Table 7: Summary of correlations between WTI crude oil prices and industry PPI across the three analysis windows

| **Industry** | **Long run (n = 234)** | **Rising oil (n = 8)** | **Falling oil (n = 12)** | **Change from long run to rising window** |
|---|---|---|---|---|
| Fertilizer | 0.494 | 0.855 | 0.724 | +0.361 |
| Agriculture | 0.082 | 0.748 | 0.876 | +0.666 |
| Plastics | 0.050 | 0.916 | 0.782 | +0.866 |
| Textiles | 0.042 | 0.926 | 0.678 | +0.884 |
| Transportation and warehousing | 0.038 | 0.937 | 0.767 | +0.899 |
| Industrial chemicals | −0.006 | 0.926 | 0.828 | +0.932 |
| Transportation | −0.121 | 0.940 | −0.909 | +1.061 |
| **Mean across industries** | **0.083** | **0.893** | **0.535** | **+0.810** |

*Industries are ordered by the size of their long-run correlation.*

Every industry showed a much stronger correlation with oil prices in the rising window than over the full sample. The average correlation rose from 0.08 in the long run to 0.89 in the rising window, an increase of 0.81. Fertilizer had the smallest increase (+0.36), but it already had a moderate correlation in the long run. Transportation had the largest increase (+1.06), moving from a weak negative coefficient to the strongest positive coefficient in the rising window.

The average correlation fell to 0.54 in the falling window, largely because of transportation's negative coefficient; excluding transportation, the average for the other six industries was 0.78.

Figure 6 illustrates this comparison.

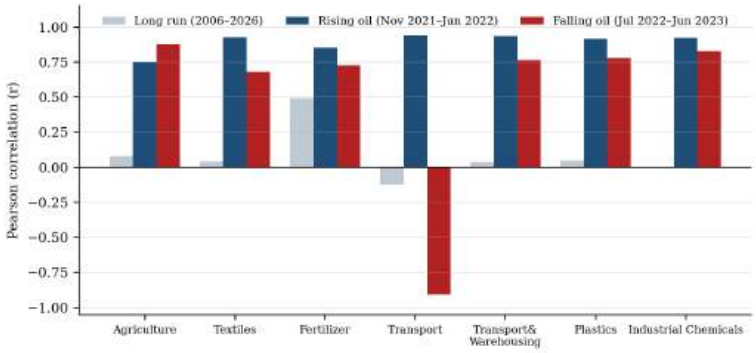



Figure 6: Correlation between WTI crude oil prices and industry PPI across the three analysis windows

### 4.6 Magnitude of Price Movement

Correlation shows whether two series move in the same direction, not by how much. Table 8 therefore compares each industry's PPI change with the oil price change in each window. A ratio of 1.00 would mean that the industry price moved by exactly the same proportion as oil.

Table 8: Percentage change in industry PPI relative to the percentage change in WTI crude oil prices within each sub-period

| **Industry** | **Rising window: change in PPI** | **Ratio to oil (+45.0%)** | **Falling window: change in PPI** | **Ratio to oil (−31.0%)** |
|---|---|---|---|---|
| Agriculture | +34.1% | 0.76 | −21.8% | 0.70 |
| Fertilizer | +19.5% | 0.43 | −22.8% | 0.73 |
| Transportation and warehousing | +13.1% | 0.29 | −3.1% | 0.10 |
| Industrial chemicals | +8.7% | 0.19 | −2.6% | 0.08 |
| Plastics | +8.0% | 0.18 | −1.8% | 0.06 |
| Textiles | +7.7% | 0.17 | −1.2% | 0.04 |
| Transportation | +3.7% | 0.08 | +2.5% | −0.08 |

*Ratios were calculated as the percentage change in industry PPI divided by the percentage change in the WTI crude oil price over the same window. A negative ratio means the industry PPI moved in the opposite direction to oil prices.*

Table 8 shows two clear patterns. First, agriculture and fertilizer adjusted substantially in both directions. Their ratios were 0.76 and 0.43 when oil rose and 0.70 and 0.73 when oil fell, so neither showed the downward stickiness seen in the other industries; fertilizer in fact fell proportionally more than it rose. Second, textiles, plastics, industrial chemicals, and transportation and warehousing had much smaller ratios when oil prices fell than when they rose. For example, textiles rose by 0.17 of the oil price increase but fell by only 0.04 of the decrease. Plastics rose by 0.18 and fell by 0.06, and industrial chemicals rose by 0.19 and fell by 0.08. Transportation was the only industry with a negative ratio in the falling window, as its PPI rose while oil prices fell.

### 4.7 Robustness Check: Month-on-Month Changes

The correlations above use price levels, but both oil prices and PPI series tend to trend over time. Table 9 therefore reruns the full-sample analysis using month-on-month percentage changes instead of levels, which helps separate the influence of shared trends from genuine short-run co-movement.

Table 9: Correlation between month-on-month percentage changes in WTI crude oil prices and industry PPI, full sample (n = 233)

| Industry | r (monthly change) | p value | r (levels, from Table 4) |
|---|---|---|---|
| Transportation and warehousing | 0.343 | < 0.001* | 0.038 |
| Industrial chemicals | 0.259 | < 0.001* | −0.006 |
| Fertilizer | 0.147 | 0.025* | 0.494 |
| Textiles | 0.145 | 0.027* | 0.042 |
| Plastics | 0.106 | 0.107 | 0.050 |
| Transportation | −0.036 | 0.584 | −0.121 |
| Agriculture | −0.062 | 0.343 | 0.082 |

*$p < 0.05$.

Using monthly changes, four of the seven industries showed a statistically significant positive correlation with oil prices: transportation and warehousing ($r = 0.34$, $p < 0.001$), industrial chemicals ($r = 0.26$, $p < 0.001$), fertilizer ($r = 0.15$, $p = 0.025$), and textiles ($r = 0.15$, $p = 0.027$). The coefficients were small, and their ranking differed from the level-based results. Transportation and warehousing and industrial chemicals had almost zero correlation in levels but the two largest coefficients in monthly changes. Fertilizer's coefficient, on the other hand, fell from 0.49 to 0.15, which indicates that a meaningful part of fertilizer's long-run level correlation reflects the two series trending together over 20 years rather than fertilizer prices reacting to oil month by month.

### 4.8 Summary of Findings

The main findings of the analysis are as follows:

- Over the full sample (December 2006 – May 2026), fertilizer was the only industry with a moderate, statistically significant association with oil prices ($r = 0.49$, $p < 0.001$). The other six industries had correlations between −0.12 and 0.08, none of which were significant at the 5% level.
- During the rising window (November 2021 – June 2022), all seven industries had strong, statistically significant positive correlations with oil prices, ranging from 0.75 to 0.94 with an average of 0.89.
- During the falling window (July 2022 – June 2023), six industries kept significant positive correlations of between 0.68 and 0.88. Transportation, however, flipped to a strong negative correlation of −0.91: its PPI rose 2.5% even as oil prices fell 31.0%.
- Every industry's correlation was higher in the rising window than over the full sample, and every industry except transportation also had a higher correlation in the falling window. The average correlation rose from 0.08 in the long run to 0.89 in the rising window, although both sub-period estimates rest on very few monthly observations ($n = 8$ and $n = 12$).
- Agriculture and fertilizer prices moved the most, whether oil prices were rising or falling. Textiles, plastics, industrial chemicals, and transportation and warehousing saw their PPIs move less when oil prices fell than when they rose.
- In month-on-month changes, transportation and warehousing ($r = 0.34$) and industrial chemicals ($r = 0.26$) showed the largest significant associations, and fertilizer's correlation shrank considerably (from 0.49 to 0.15), pointing to a shared long-run trend rather than a purely month-to-month relationship.

## 5. Discussion

The results suggest that, for most of these industries, oil prices are only loosely associated with producer prices over the long run but move closely with them during sharp market shifts. Over nearly two decades, only fertilizer showed a meaningful level-based correlation, yet all seven industries moved strongly with oil during the 2021–2022 price surge. This pattern is consistent with Jo et al. 's finding that the cost (supply) channel of oil price shocks dominates only in the most oil-intensive industries, and with evidence that oil price increases spread through production and transport costs into the prices of other goods (Alp et al.). The rising window also overlaps with a period of broad-based inflation in the U.S. economy, so the strong sub-period correlations may partly reflect general price pressures rather than oil alone. Fertilizer's position as the only industry with a long-run association fits its dependence on energy as both fuel and feedstock (Beckman et al.), and the World Bank expects fertilizer prices to rise sharply alongside energy prices in 2026 (*Commodity Markets Outlook*). However, the robustness check in Section 4.7 shows that much of this association reflects the two series sharing long cycles rather than fertilizer responding to oil month by month. Agriculture and fertilizer also

adjusted substantially in both directions, which suggests that their prices respond to changing market conditions whether oil is rising or falling.

In contrast, textiles, plastics, industrial chemicals, and transportation and warehousing raised their prices more when oil rose than they lowered them when oil fell. This pattern is consistent with asymmetric price adjustment, whereby producer prices respond more readily to oil price increases than to decreases. The month-on-month results add nuance: transportation and warehousing and industrial chemicals showed the strongest short-run associations with oil, which fits their direct exposure to fuel costs and to oil- and gas-based feedstocks, respectively ("How Much Oil"; *Future of Petrochemicals*).

Transportation is the clearest outlier. Its PPI rose steadily across the whole sample, including during the 2022–2023 decline in oil prices, and it was the only industry whose correlation reversed sign in the falling window. Although fuel is a major input for transportation, this pattern suggests that other cost drivers, such as labor, equipment, and contract pricing, shaped the index more strongly over this period. This study cannot test these explanations directly, but they are a natural focus for future work.

Several limitations should be kept in mind when interpreting these results. First, this is a correlational study: it identifies associations between oil prices and producer prices, not causal effects, since no lagged regression or multivariate model controlling for other cost drivers (such as labor, exchange rates, and demand shocks) was used. Second, the rising- and falling-window correlations are based on only eight and twelve monthly observations respectively; although statistically significant, coefficients from such small samples can be sensitive to one or two unusual months and should be treated as indicative rather than precise. Third, the full-sample, level-based correlations in Section 4.2 are affected by the fact that both oil prices and several PPI series trend upward over the 20-year sample. The month-on-month robustness check in Section 4.7 shows that level-based correlations are not a reliable guide to short-run co-movement: they overstate it for fertilizer and understate it for transportation and warehousing and industrial chemicals. The sub-period findings should therefore be treated as descriptive patterns worth testing further with a longer or lagged dataset.

## 6. Conclusion

This study examined how West Texas Intermediate (WTI) crude oil prices and the Producer Price Index (PPI) were related in seven U.S. industries from December 2006 to May 2026, with the goal of identifying which industries moved most closely with oil prices and whether these relationships changed when oil prices rose or fell. Over the full period, only fertilizer showed a moderate long-run correlation with oil prices ($r = 0.49$), while the other six industries showed little to no long-run relationship. When oil prices rose rapidly from November 2021 to June 2022, all seven industries showed strong positive correlations (average $r = 0.89$), consistent with producer prices moving more closely with oil during rapid increases. When oil prices fell from July 2022 to June 2023, six industries remained positively correlated, but transportation swung to $r = -0.91$, with its PPI rising 2.5% as oil fell 31.0%. Agriculture and

fertilizer adjusted substantially in both directions, whereas textiles, plastics, industrial chemicals, and transportation and warehousing lowered their prices far less when oil fell than they raised them when it rose.

With the caveats discussed in Section 5 in mind, companies in agriculture and fertilizer should keep a close eye on oil prices when managing their costs, and supply chain managers should remember that different industries react to oil price changes in different ways. Policymakers concerned about inflation may want to pay particular attention when oil prices move sharply, since producer prices in most of these industries became far more closely associated with oil prices during those rapid changes than they were over the long run. Future studies could examine countries that import large amounts of oil, such as South Korea or Turkey, to see whether greater reliance on imported energy strengthens the link between oil prices and industry prices. Researchers could also study earlier oil shocks, such as those of the 1970s and 1980s, using more detailed data on industry input costs or statistical models that account for delayed effects and other confounding factors. This would show whether the patterns found here hold in different economic periods and would move the analysis from correlation toward a more causally informative design.

**Appendix A: Long-Run Charts for the Remaining Industries**

The charts below show the WTI crude oil price alongside the Producer Price Index for the five industries not shown in Section 4.2, covering the full study period from December 2006 to May 2026.

***Figure A1: WTI crude oil price and agriculture PPI***

***Figure A2: WTI crude oil price and textiles PPI***

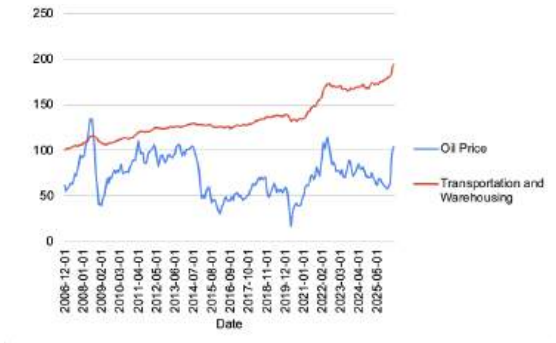



***Figure A3: WTI crude oil price and transportation and warehousing PPI***

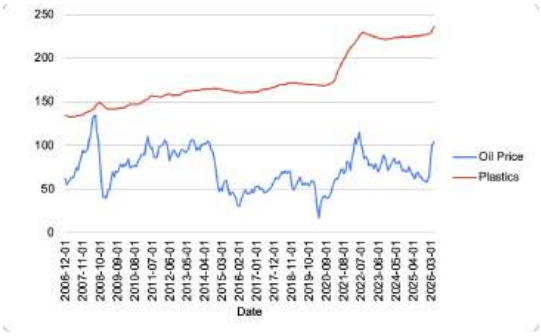



***Figure A4: WTI crude oil price and plastics PPI***

***Figure A5: WTI crude oil price and industrial chemicals PPI***

*All charts cover December 2006 – May 2026.*

**Works Cited**


Alp, Harun, et al. “Second-Round Effects of Oil Prices on Inflation in the Advanced Foreign Economies.” FEDS Notes, Board of Governors of the Federal Reserve System, 15 Dec. 2023, https://doi.org/10.17016/2380-7172.3401.

Beckman, Jayson, et al. Agriculture’s Supply and Demand for Energy and Energy Products. United States Department of Agriculture, Economic Research Service, May 2013. Economic Information Bulletin 112.

Commodity Markets Outlook, April 2026. World Bank, 28 Apr. 2026.

Farm Production Expenditures 2024 Summary. United States Department of Agriculture, National Agricultural Statistics Service, July 2025.

FRED Economic Data. Federal Reserve Bank of St. Louis, fred.stlouisfed.org.

The Future of Petrochemicals: Towards More Sustainable Plastics and Fertilisers. International Energy Agency, Oct. 2018.

“How Much Oil Is Used to Make Plastic?” U.S. Energy Information Administration, 10 July 2024, www.eia.gov/tools/faqs/faq.php?id=34.

Jo, Soojin, et al. Industry Effects of Oil Price Shocks: Re-examination. Federal Reserve Bank of Dallas, 5 Dec. 2017. Working Paper 1710.

Materials Market Report 2025. Textile Exchange, Sept. 2025.

Milewska, Beata, and Dariusz Milewski. “Implications of Increasing Fuel Costs for Supply Chain Strategy.” Energies, vol. 15, no. 19, 2022, article 6934, https://doi.org/10.3390/en15196934.

“Textile-Clothing.” Coface, 2025, www.coface.com/news-economy-and-insights/business-risk-dashboard/sector-risk-files/textile-clothing.

World Economic Outlook, April 2026: Global Economy in the Shadow of War. International Monetary Fund, 14 Apr. 2026.

# Prior AI Access and Perceived Loss: An Exploratory Study of Loss Aversion and Psychological Ownership By Ahaana Bisarya


## Abstract

Artificial intelligence (AI) tools are increasingly becoming part of people's daily activities, which may influence how people value access to these tools. This study examines whether prior access to an AI tool is associated with stronger feelings of loss when that access is threatened. A quantitative between-groups survey was conducted with 104 participants. The participants were divided into 81 frequent AI users, who were presented with a loss-framed scenario involving the possible loss of AI access, and 23 infrequent or non-users, who were presented with a gain-framed scenario involving access to a new AI tool. Participants responded to Likert-scale statements measuring perceived loss aversion, along with questions about emotional responses, willingness to pay, and intended behavior. The results showed that frequent AI users reported substantially higher average loss-aversion scores than infrequent or non-users, with mean scores of 3.28 and 1.91, respectively. These findings suggest that prior experience with AI may be associated with stronger perceptions of loss when access is threatened. The findings are consistent with concepts from loss aversion, the endowment effect, and psychological ownership, although the study does not establish a causal relationship between prior AI use and loss aversion.



## Keywords

Artificial Intelligence, Loss Aversion, Endowment Effect, Psychological Ownership, Prior Access, Technology Reliance, Prospect Theory, Behavioral Economics.


## Introduction

The application of artificial intelligence has moved quickly from an innovation to an integral part of decision-making, learning, and work. Individuals use applications of artificial intelligence in writing, researching, programming, problem-solving, and many other operations. As people continue to adopt these technologies in their daily routines, access to artificial intelligence may no longer be considered just a technological service but a valuable resource that people psychologically internalize.

Traditional economic theories assume that the evaluations of gains and losses made by individuals are objective. The behavioral economics theory shows that the reaction of people to losses is more emotional compared to gains. In their theory called Prospect Theory, Kahneman and Tversky suggested that the psychological pain associated with the loss of something is usually stronger than the pleasure from acquiring that thing (Kahneman & Tversky). The concept of the endowment effect closely connected to it says that people generally give higher value to the objects or resources that they possess (Thaler).

Loss aversion has been traditionally researched in financial decisions and consumer goods as well as physical property, but the development of digital technologies raises a question

whether access to something can also become something that people psychologically internalize? To be more precise, does having sustained access to something change people's reaction to the threat of losing it?

This issue appears to be especially relevant in the case of artificial intelligence. Unlike other goods, the access to artificial intelligence is mostly provided on the basis of subscription, platform, or temporary permission rather than possession. As a result, people may face an unusual experience of losing access to the cognitive resource. Therefore, loss of the access to artificial intelligence may lead to emotional and behavioral reactions that may be different from the response of people who did not have the experience of accessing this kind of technology.

Past research on artificial intelligence has considered the issues of adoption, trust, acceptance, productivity, ethics, and human dependence on AI systems. Still, little consideration has been dedicated to the issue of behavior when the access to AI is being taken away from people who have become used to such access. This is the issue especially relevant in light of growing prevalence of subscription services for AI tools.

In the current study, the researcher aims to investigate whether previous experience with access to an AI system is associated with greater perceived loss when access to that AI system is threatened, compared with limited or no prior access. The loss is defined in terms of losing the access to the AI tool that has been already accessed previously by the participant.

To answer the question set above, the researcher aims at comparing two groups of participants. In the first group, prior access to the AI tool is experienced before the participants are exposed to the threat of the loss of access to it. In the second group, no prior access to the same AI tool takes place. Thus, the researcher tries to find out whether the difference in responses to the threat of losing the access is associated with differences in perceived value of AI due to prior experience

This study is important for the literature on behavioral economics because it adds knowledge to the theory of loss aversion and the endowment effect, considering their applicability to digital intelligence. Moreover, this study extends knowledge of the interaction of human beings with artificial intelligence by treating AI not only as a technology of adoption but as a valuable resource whose loss creates certain psychological impact. The importance of understanding the issue is increasing as AI becomes a part of our education, work, and decision making.

Thus, the primary question for investigation in the study is the following one:

Is prior access to an AI tool associated with greater perceived loss when access to that tool is threatened, compared with limited or no prior access?

**Literature Review**

The theory of loss aversion, one of the main concepts of behavioral economics, states that people feel losses more than an equivalent gain. Prospect Theory assumes that individuals evaluate outcomes relative to a reference point, and not on their own absolute value, and therefore, the same outcome can be perceived differently depending on whether it was

experienced previously or expected (Kahneman & Tversky). Later research showed that the concept of loss aversion is applicable to riskless choices as well; people are especially sensitive to outcomes represented as losses relative to an established reference point (Tversky & Kahneman).

Similar to the concept of loss aversion, the endowment effect suggests that people will value an object more after having it. This idea was first introduced by Thaler (1980), who assumed that the process of owning an object changes its valuation; later experiments proved that individuals are asking for more compensation to give away an object than they would ask for buying it (Kahneman, Knetsch, & Thaler). The endowment effect is tightly related to the loss aversion effect because giving away an owned object is perceived as a loss, and not as the absence of a gain (Kahneman, Knetsch, & Thaler).

Nevertheless, possession does not have to be necessarily physical or legal. There exists a notion of psychological ownership that presupposes development of a sense of "mineness" of the resource due to its familiarity, control and interactions with it by the individual without possessing the resource (Pierce, Kostova, & Dirks; Pierce, Kostova, & Dirks). The notion is especially topical in the context of virtual worlds when an individual tends to interact with digital resources and consume them based on their access rather than actual possession. The analysis of valuation and attachment to non-physical resources and digital goods reveals the possibility of extending these notions beyond tangible assets despite differences in mechanisms of valuation (Morewedge & Giblin; Ericson & Fuster).

The body of research on human-AI interaction is very extensive regarding the study of trust, adoption, and reliance on automated technology. An individual is more likely to rely on AI when he/she perceives it as useful, reliable and helpful in task performance (Lee & See; Parasuraman & Riley). Recently it has been shown that people can develop substantial reliance on algorithmic suggestions and automated systems, which might sometimes lead to overreliance on automated suggestions (Logg, Minson, & Moore).

The current study brings together the two separate literatures above by investigating the potential role of having had access to an AI tool before in serving as a psychological reference point that affects behavior toward its withdrawal. Although past studies have each focused separately on loss aversion, endowment effect, psychological ownership, and dependence on AI, there is little work devoted to whether the possession of access to AI is itself an endowment that can serve to make individuals loss averse. In this study, we investigate whether individuals with pre-existing access to an AI tool respond more emotionally and behaviorally to a threat to their access than individuals with no prior access to the tool.

## Method

### *Research Design*

This study used a quantitative, between-groups survey design to examine whether prior experience with AI was associated with differences in perceived loss aversion. In the experiment,

a total number of 104 participants was used. The participants were grouped according to the frequency of their use of AI.

### *Classification of Participants and Groups*

From the 104 participants, 81 were considered to be frequent AI users since they had indicated that they frequently used or sometimes used AI. 23 out of the 104 participants were considered infrequent or non-AI users since they rarely or never used AI.

### *Survey Scenarios*

The two groups received different scenarios. Frequent AI users were presented with a loss-framed scenario in which they were asked to imagine losing access to an AI tool that they had previously been able to use. Infrequent or non-users were presented with a gain-framed scenario involving access to a new AI tool.

### *Measures*

The participants had to answer a series of Likert scale questions that helped to assess their feelings regarding losses from the absence of access to the AI tool.

### *Data Analysis*

Descriptive statistics and comparison of the average answers of both groups were used for the data analysis. The analysis was carried out, in particular, with regard to the difference in the answers related to the feeling of losses, and the demographic and behavioral responses helped to interpret the results.

## Results

For this study, the participants' demographic and behavioral data, their feelings of loss from the absence of access to AI and willingness to pay, were analyzed graphically. Also, there is the comparison of participants with the experience of working with AI and those who have no such experience.

The survey consisted of 104 participants from different age groups. Figure 1 presents the age distribution of the respondents. The majority of participants belonged to the 45–54 years age group (41 participants; 39.4%), followed by the 35–44 years group (27 participants; 26.0%) and the 25–34 years group (21 participants; 20.2%). A smaller proportion of respondents belonged to the 55–64 years group (12 participants; 11.5%) and the 18–24 years group (3 participants; 2.9%). Overall, the sample was predominantly composed of working-age adults.

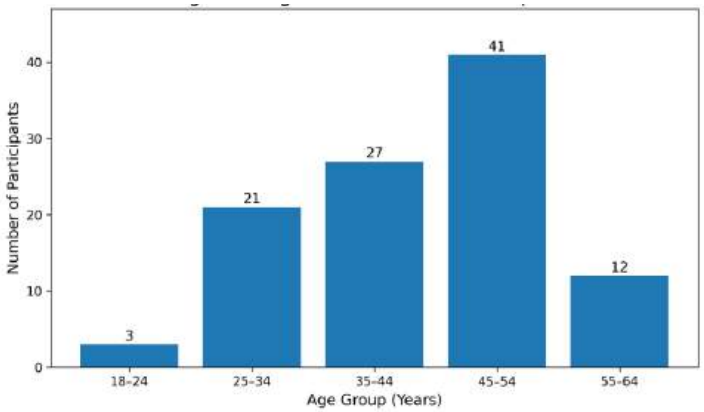



Fig 1. Age Distribution of Participants

Figure 2 presents the gender distribution of the participants. The sample consisted of 50 females (48.1%) and 45 males (43.3%), while 6 participants (5.8%) identified as non-binary and 3 participants (2.9%) preferred not to disclose their gender. The distribution provides representation across different gender categories.

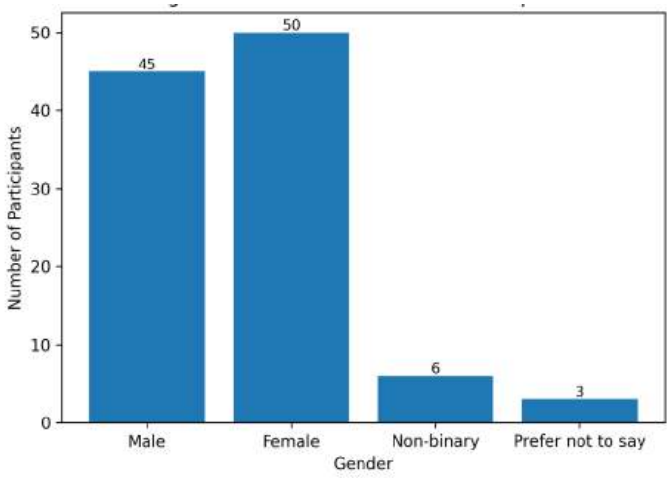



Fig 2. Gender Distribution of Participants

The respondents also differed in their frequency of AI usage. As shown in Figure 3, 55 participants (52.9%) reported using AI often, while 26 participants (25.0%) reported using AI sometimes. In comparison, 16 participants (15.4%) used AI rarely and 7 participants (6.7%) reported never using AI. Thus, the majority of respondents had prior experience with AI tools, providing a basis for comparing participants with different levels of prior AI exposure.

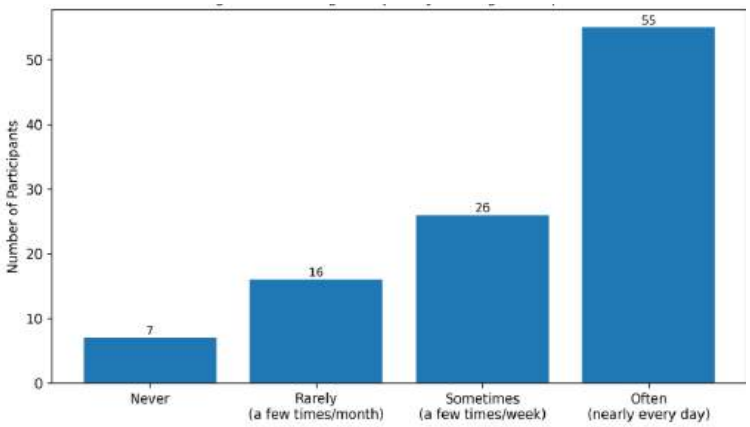



Fig 3. AI Usage Frequency Among Participants

Figure 4 presents household digital subscription ownership among the participants. A large majority, 87 participants (83.7%), reported that their household subscribed to at least one digital service, whereas 17 participants (16.3%) reported having no digital subscriptions. This indicates that most respondents had some familiarity with subscription-based digital services, which provides relevant context for their responses to questions concerning paid AI access.

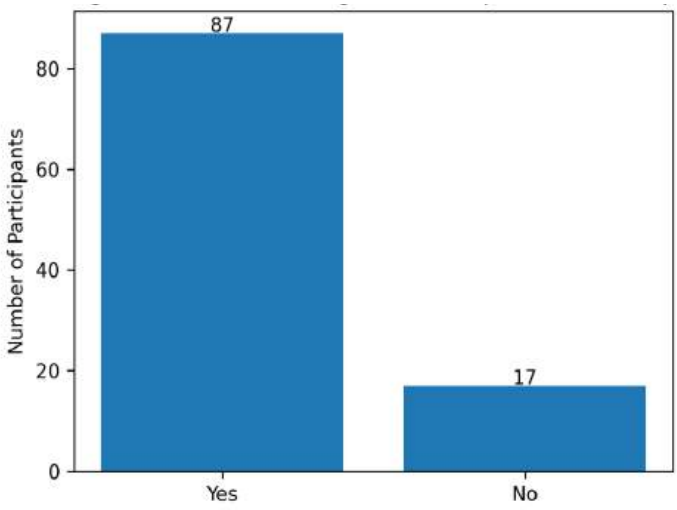



Fig 4. Household Digital Subscription Ownership

In addition, participants were also requested to indicate the maximum amount they would be willing to pay per month for AI. From Figure 5, it can be observed that the most preferred category is $21–30 and the number of participants in this category is 27 (26.0%). Second in line, the category is $11–20 which has 22 participants (21.2%). Third is $0–10, which has 21 participants (20.2%). Very few participants indicated their willingness to pay more than $100 per month. These figures suggest that people are not very willing to pay for AI services, but only at a lower and moderate price level.

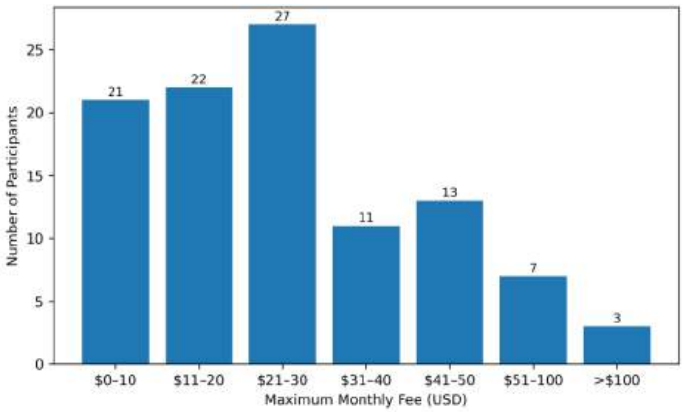



Fig 5. Distribution of Maximum Monthly Fee Participants Are Willing to Pay

Figure 6 shows the average agreement with the three statements that have been developed to measure loss aversion towards AI access. It has been observed that the statement which has the highest mean value is "Loss of access to AI feels like a loss" (Mean = 3.15), then comes "Loss of AI affects daily life activities" (Mean = 2.98) and finally comes "Loss of AI feels more than gain of equal amount" (Mean = 2.80).

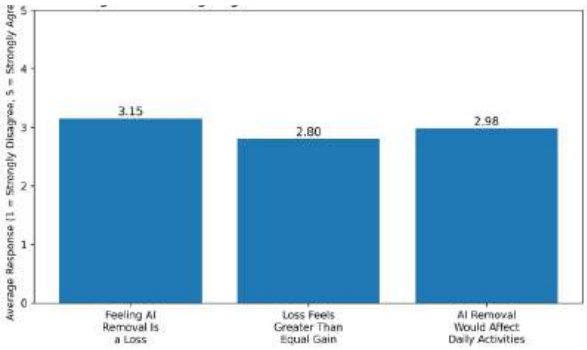



Fig 6. Average Agreement with Loss Aversion Statements

Additionally, the participants were asked to describe how they felt about paying for artificial intelligence services. As shown in Figure 7 below, the most common answer was acceptance with 52 participants choosing this option. Frustration came second with 31 participants, anxiety third with 25 participants and indifference fourth with 24 participants. Relief was chosen by 10 participants and other emotions by 4 participants.

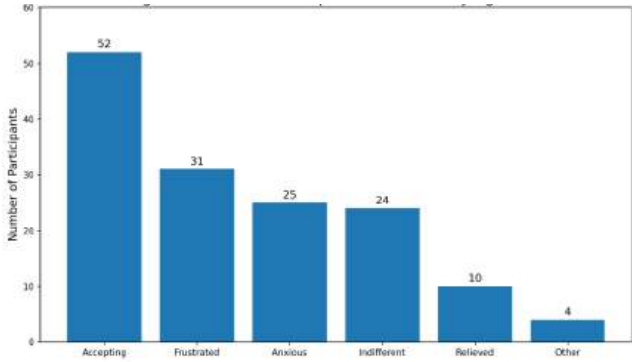



Fig 7. Emotional Responses Towards Paying for AI

Figure 8 below shows participants' likelihood to pay for artificial intelligence services. In this figure, the highest number of respondents chose the neutral category, which comprised 31 participants (30.1%). The second highest number was those likely to pay and this group had 25 participants (24.3%). Those unlikely to pay comprised 19 participants (18.4%). There were 14 participants who felt very likely to pay and also another 14 felt very unlikely to pay.

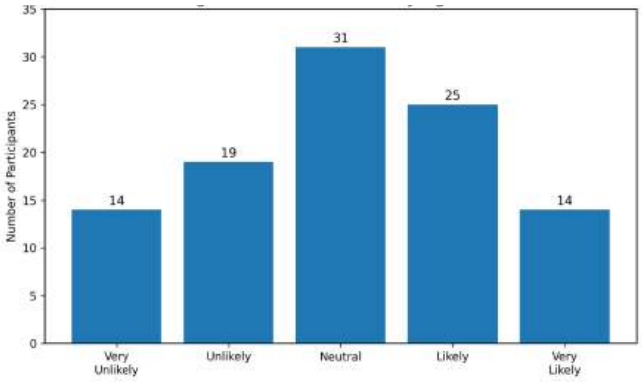



Fig 8. Likelihood of Paying for AI

Furthermore, participants were asked how they will react if they choose not to pay for AI-based services. As per Figure 9, it can be observed that 48 participants (46.2%) will adapt to the scenario and go ahead, while 40 participants (38.5%) will look for some free solution to their problem. Out of the total participants, only 10 participants (9.6%) will try accessing AI for free, whereas 6 participants (5.8%) will simply discontinue the use of AI solutions.

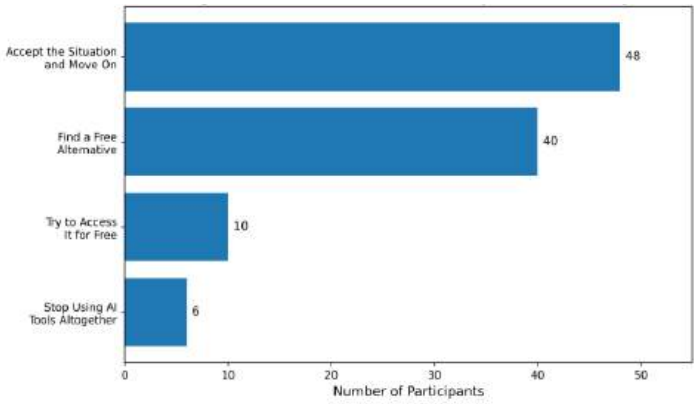



Fig 9. Intended Action if Participants Do Not Pay

The central analysis of the study compares participants with previous AI experience and those with limited or no prior AI experience. Figure 10 compares the average responses of the two groups across the three loss-aversion statements. Previous AI users reported higher agreement across all three measures. For the statement that losing access to AI feels like a loss, the mean score was 3.47 for previous users compared with 2.04 for first-time users. For the statement that the loss feels worse than an equivalent gain, the corresponding scores were 3.14 and 1.61. Similarly, the mean score for the statement that losing AI would affect work or daily life was 3.24 for previous users compared with 2.09 for first-time users. Across the three loss-aversion statements, the average score was 3.28 for previous AI users and 1.91 for first-time or infrequent users.

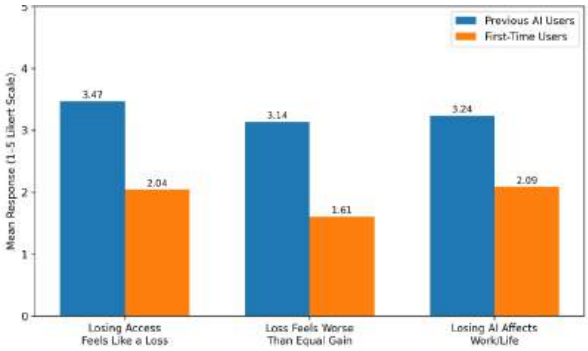



Fig 10. Comparison of Previous AI Users vs First-Time Users

**Discussion**

The key findings of this study were that more frequent users of AI had a higher mean score in terms of loss-aversion (3.28) compared to infrequent and non-users (mean score 1.91). This implies that those subjects having more experience in using AI were affected more by a possible loss of access to AI.

The observed differences are consistent with theories of loss aversion, the endowment effect, and psychological ownership. However, because prior AI experience and scenario framing

were not independently controlled, the findings should be interpreted as an association rather than evidence that prior AI access causes greater loss aversion.

The results may also be related to the phenomenon of psychological ownership. Pierce et al. state that humans can experience ownership towards objects that they cannot really own(Pierce et al.; Pierce et al.).

Nevertheless, there are a number of limitations that must be taken into consideration. Both groups used in the current study were provided with different scenarios: frequent AI users – loss frame, and infrequent or non-AI users – gain frame. In other words, AI use prior to the experiment and scenario framing were not controlled in this case. The difference in loss-aversion scores can be caused by prior AI use, by framing of the scenarios, or even by the combination of these two factors. In this regard, it may be said that there is an existence of correlation but not causation.

**Conclusion**

The current research investigates whether prior exposure to an artificial intelligence tool is connected with the increased loss-aversion when access is threatened in comparison with the situation when such a tool is experienced for the first time.

The results obtained demonstrate a descriptive difference between participants with prior experience and participants with little or no prior experience in the use of artificial intelligence tools. Participants who used AI tools before reported higher average scores on all the three indicators of loss aversion.

This finding is consistent with the theories of loss aversion, endowment effect, and psychological ownership. The prior use of AI may form a reference point, so that access to it subsequently becomes familiar and valuable. Thus, the threat of losing access to it may become something different compared to obtaining access to a new AI system.

On the other hand, it is clear from the findings that people do not have to pay for retaining access to AI even if it is valuable for them. Many respondents stated that they would rather lose the access or find alternatives that can be obtained without costs.

Hence, it can be concluded that although psychological losses and willingness to pay cannot be considered the same phenomena, they are related. Thus, this finding provides descriptive evidence for the proposed research hypothesis that the prior access to AI is related to perceptions of loss. But further statistical analyses are required to confirm this connection.

The paper adds to the discussion related to the topic of behavioral economics, adding an analysis of loss aversion and endowment effect in relation to digital resources, that is, AI. The role of AI becomes more prominent in everyday life activities; thus, the evaluation of how the loss of access to such resources will influence the consumer's decision making becomes highly important.

**Limitations**

Several limitations must be taken into account when analyzing the results of the study.

First, the sample used for the experiment included 104 people. Thus, although the obtained results are valuable to consider, they cannot be generalized to the whole population because of the relatively small sample size.

Second, the study involves the use of the non-probability convenience sampling method. The participants were selected based on online and personal/professional contacts. Therefore, they do not necessarily reflect the features and attitudes of the general population.

Third, the study mainly involves self-reported responses. The participants were expected to report their perceptions, emotions, willingness to pay, and intentions of behavior. Intentions may not necessarily be reflected in actual behavior.

Fourth, this research relies on an experimental manipulation based on a hypothetical situation in which participants could potentially lose access to AI or obtain new access to AI. While the use of such a technique permits one to compare the psychological framing of the two situations, a hypothetical situation does not guarantee that the response will match a real loss of access.

Fifth, participants were grouped depending on how often they used AI in the past. While frequency of use is a good variable to compare groups, the frequency of use of AI may not account for the degree of dependence, types of AI being used, or even the duration of such use. It is possible that the perception of loss depends on those factors independent from the frequency of use.

Finally, the current analysis gives only a descriptive comparison of the two groups. Even though the differences between the users who used AI before and users who never used AI before are quite large, descriptive differences alone do not give us the causality. Statistical significance and regression analyses are needed.

In spite of all its flaws, however, the current study is a framework for analyzing the extent to which access to digital technology like AI is capable of taking on ownership-like qualities and loss.

In the future, one might conduct research in the form of a 2×2 design, where pre-experience with AI will be distinguished from scenario framing. Under this design, both frequent and infrequent users of AI will be confronted with the scenarios framed both by loss and gain. This will help in establishing whether the stronger responses to loss aversion are due to pre-experience with AI, loss framing, or both.

**Works Cited**


Ericson, Keith M. Marzilli, and Andreas Fuster. “The Endowment Effect.” Annual Review of Economics, vol. 6, 2014, pp. 555–579. https://doi.org/10.1146/annurev-economics-080213-041320.

Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler. “Experimental Tests of the Endowment Effect and the Coase Theorem.” Journal of Political Economy, vol. 98, no. 6, 1990, pp. 1325–1348. https://doi.org/10.1086/261737.

Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler. “Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias.” Journal of Economic Perspectives, vol. 5, no. 1, 1991, pp. 193–206. https://doi.org/10.1257/jep.5.1.193.

Kahneman, Daniel, and Amos Tversky. “Prospect Theory: An Analysis of Decision under Risk.” Econometrica, vol. 47, no. 2, 1979, pp. 263–291. https://doi.org/10.2307/1914185.

Lee, John D., and Katrina A. See. “Trust in Automation: Designing for Appropriate Reliance.” Human Factors, vol. 46, no. 1, 2004, pp. 50–80. https://doi.org/10.1518/hfes.46.1.50_30392.

Logg, Jennifer M., Julia A. Minson, and Don A. Moore. “Algorithm Appreciation: People Prefer Algorithmic to Human Judgment.” Organizational Behavior and Human Decision Processes, vol. 151, 2019, pp. 90–103. https://doi.org/10.1016/j.obhdp.2018.12.005.

Morewedge, Carey K., and Colleen E. Giblin. “Explanations of the Endowment Effect: An Integrative Review.” Trends in Cognitive Sciences, vol. 19, no. 6, 2015, pp. 339–348. https://doi.org/10.1016/j.tics.2015.04.004.

Parasuraman, Raja, and Victor Riley. “Humans and Automation: Use, Misuse, Disuse, Abuse.” Human Factors, vol. 39, no. 2, 1997, pp. 230–253. https://doi.org/10.1518/001872097778543886.

Pierce, Jon L., Tatiana Kostova, and Kurt T. Dirks. “Toward a Theory of Psychological Ownership in Organizations.” Academy of Management Review, vol. 26, no. 2, 2001, pp. 298–310. https://doi.org/10.5465/amr.2001.4378028.

Pierce, Jon L., Tatiana Kostova, and Kurt T. Dirks. “The State of Psychological Ownership: Integrating and Extending a Century of Research.” Review of General Psychology, vol. 7, no. 1, 2003, pp. 84–107. https://doi.org/10.1037/1089-2680.7.1.84.

Thaler, Richard. “Toward a Positive Theory of Consumer Choice.” Journal of Economic Behavior & Organization, vol. 1, no. 1, 1980, pp. 39–60. https://doi.org/10.1016/0167-2681(80)90051-7.

Tversky, Amos, and Daniel Kahneman. “Loss Aversion in Riskless Choice: A Reference-Dependent Model.” The Quarterly Journal of Economics, vol. 106, no. 4, 1991, pp. 1039–1061. https://doi.org/10.2307/2937956.

# Attack or Explain? Debate Rhetoric and Low-Dollar Giving in the 2020 Democratic Presidential Primary By Danial Yessentay


## Abstract

During the 2020 Democratic presidential primary, an unusually large field of candidates competed not only for votes but for small online contributions, and the Democratic National Committee (DNC) tied debate eligibility directly to the number of unique donors a campaign could report. This paper asks whether opponent-directed negative rhetoric and constructive policy messaging in primary debates were associated differently with candidates' short-run, low-dollar fundraising response. Building on political-economy accounts of campaign giving as expressive participation and on the genuinely mixed findings of the negative-campaigning literature, the paper develops a comparative framework distinguishing “attack” and “explain” rhetorical strategies. Using publicly reported, campaign-disclosed 24-to-72-hour fundraising figures alongside a close reading of the debate moments that produced them, it examines five widely covered cases from the 2019–2020 race. High-salience attacks generated the fastest and most visible short-run spikes in the sample, but constructive policy messaging also moved substantial money, and the financial beneficiary of an attack was not always the candidate who delivered it.


## Introduction

The 2020 Democratic presidential primary was, by design, a contest decided as much by online small donors as by primary voters. When the Democratic National Committee (DNC) set the qualification rules for the first debates in February 2019, it gave candidates two paths onto the stage: a minimum share of support in early polling, or contributions from at least 65,000 unique donors (PBS NewsHour, 2019). By the sixth debate that December, the donor threshold had risen to 200,000, and by January 2020 it stood at 225,000 (NBC News, 2019). A campaign's national visibility was, quite literally, gated by its ability to convert speech into small, fast contributions – nearly all of them processed through ActBlue, which by 2020 handled the overwhelming majority of Democratic online giving (Bouton, Cagé, Dewitte, & Pons, 2022).

This arrangement makes the 2020 primary an unusually clean setting for a basic political-economy question: what kind of debate speech actually moves small donors, and how quickly? Candidates split their speaking time between two broad strategies – attacking a rival's record or judgment, and advancing their own policy proposals, as in Senator Elizabeth Warren's “I have a plan for that” brand. Both are investments of scarce speaking time, and campaigns presumably expect some return in votes, attention, or dollars.

This paper asks: during the 2020 Democratic primary debates, how did opponent-directed negative rhetoric and constructive policy messaging differ in their short-run association with low-dollar ActBlue giving – defined, following recent campaign-finance research, as individual contributions at or below $200 (Bouton et al., 2022)? A fully rigorous answer would require matching sentence-level coding of all eleven debates to reconciled, transaction-level records for

every candidate and date – a data-engineering effort beyond a single paper. This paper instead uses a transparent comparative case-study design: it identifies the primary's most heavily covered rhetorical moments for which campaigns disclosed short-run (24-to-72-hour) fundraising figures, classifies the precipitating rhetoric, and compares the pattern of response. The result is informative about the political economy of debate speech in the small-dollar era, and about where a larger study should look next.

The paper's guiding hypothesis is stated as an association, not a causal claim:

H1. Candidates' rates of opponent-directed negative rhetoric and constructive policy messaging are associated with short-run changes in the number and value of low-dollar contributions during the 2020 Democratic presidential primary.

## Literature Review

### Why People Give

Political giving has long puzzled economists: an individual donation has a vanishingly small chance of changing an outcome, and no campaign can promise a twenty-dollar donor a meaningful return on that dollar. Ansolabehere, de Figueiredo, and Snyder (2003) argued that campaign giving is better understood as consumption or expressive participation than as investment – a framework that fits small donations far better than the large checks written by organized interests. Hill and Huber (2017), using merged survey and administrative data, found donors are ideologically more extreme than non-donors and that their giving tracks perceived electoral stakes rather than a calculated assessment of influence. Both accounts suggest giving is closer to a rapid, emotionally inflected act than a deliberate investment – exactly the behavior a single debate exchange could plausibly move within hours.

### Small, Digitally Mediated Giving, and Does Campaign Text Move It?

That a debate could move money quickly is a recent possibility. Bouton, Cagé, Dewitte, and Pons (2022) built a dataset of all 340 million itemized federal contributions from 2005 to 2020 and showed small donations – facilitated first by ActBlue – grew rapidly and became demographically more representative of the country than large donations, though not a mirror of the electorate. Their central fact, that a large share of federal money now moves through a single, near-instantaneous digital conduit, is the infrastructural precondition for this paper's question. A smaller literature asks whether communication content, not just its existence, explains giving: Mummalaneni, Wang, and Isaac (2024) analyzed 764 fundraising emails from nineteen candidates in this same 2020 primary and found textual features such as empathy and vulnerability associated with meaningfully different donation outcomes, estimating a roughly nine-percent average gain from optimized language; Gaynor and Gimpel (2021) reached a similar conclusion for a single campaign's email program. Neither study examines debates, but both establish, in this very election, that campaign text is not fundraising-neutral – motivating whether debate speech, a far less controllable medium, behaves the same way.

**Does Going Negative Pay Off?**

The negative-campaigning literature is, if anything, a caution against assuming attacks pay financially. Lau, Sigelman, and Rovner's (2007) meta-analysis of 111 studies found no reliable electoral advantage from negativity and only modest, inconsistent turnout effects, alongside a consistent erosion of trust and efficacy. Niebler and Urban (2017) followed primary donors into the general election and found intraparty negative advertising reduced the probability a primary donor gave again later – a financial cost this paper's short window cannot observe. Jacob, Lelkes, and Westwood (2026) examined over 2.2 million statements by members of the 118th Congress and found personal attacks reliably bought media attention but no reward in fundraising, vote margin, or legislative success – attention and money can diverge even where negativity draws attention.

**Conditional Mechanisms and the Debates-as-Treatment Question**

Other work complicates a simple story in either direction. Valentino, Brader, Groenendyk, Gregorowicz, and Hutchings (2011) showed experimentally that anger has a distinctive mobilizing effect, a plausible channel for attacks to raise money even where negativity in general does not. Krupnikov (2011) found negativity's demobilizing effect on turnout is conditional on timing and target, not universal. Mutz (1995), studying the 1988 Democratic primary, found horse-race coverage could pull contributions in either direction depending on whether a candidate was framed as underdog or frontrunner. Feigenbaum and Shelton (2013) showed fundraising and perceived viability reinforce each other bidirectionally, so a single debate's apparent effect is entangled with a candidate's pre-existing trajectory. It is also worth taking seriously that debates may not matter much at all: Le Pennec and Pons (2023), using data from sixty-two elections and fifty-six debates across ten countries, found no detectable average effect of debates on vote-choice formation. Giving, however, is more elastic than vote choice – lower-cost, reversible, closer to an impulsive act – so a null result for votes does not settle the question for money, though it is a useful caution.

Taken together, this literature establishes that small-dollar giving is expressive, that campaign text demonstrably moves donors in this election, and that negative rhetoric's financial payoff is genuinely contested. No study we located directly compares attack and policy-advocacy debate rhetoric against short-run low-dollar giving within a single primary race – the gap this paper addresses, exploratively.

**Theoretical Framework and Hypotheses**

Consider a candidate's debate performance as serving two audiences at once: primary voters, whose decision is filtered through polling and coverage over weeks or months; and online donors, who can act within minutes at a cost – a five- or twenty-dollar gift – low enough to resemble a social-media share more than a financial decision. The DNC's rising donor-count thresholds made this second audience unavoidable: after the November 2019 debate, a campaign needed 165,000 unique donors to stay on stage in December, and 200,000 after that (Democratic

National Committee, 2019, as reported in contemporaneous coverage). Every debate doubled as a fundraising deadline.

Two rhetorical strategies could plausibly satisfy that incentive, through different mechanisms. An attack draws a sharp, easily narrated contrast, is more likely to be excerpted into a shareable clip, and can trigger the fast emotional response – anger, rivalry – that the literature on emotional mobilization treats as a genuine spur to action (Valentino et al., 2011). But an attack on a fellow Democrat also risks the backlash documented by Lau, Sigelman, and Rovner (2007) and, per Niebler and Urban (2017), risks depressing the same donor's willingness to give later – a cost this paper's short window cannot observe. Constructive messaging offers a different route: congruence with a giver's views, a signal of competence, and, for a distinctive proposal, its own shareable moment that needs no opponent. Its risk is the opposite: policy detail can be dense enough for coverage to skip past it in favor of a cleaner conflict narrative.

The paper's guiding hypothesis is deliberately agnostic about direction:

H1. Candidates' rates of opponent-directed negative rhetoric and constructive policy messaging are associated with short-run changes in the number and value of low-dollar contributions during the 2020 Democratic presidential primary.

Stated this way, H1 is a claim that an association exists, not a claim about which rhetorical mode has the larger or more favorable association. To make H1 discussable against evidence, this section unpacks it into three more specific expectations, offered as genuinely open questions rather than settled predictions:

H1a. Because an attack that becomes a recognizable “moment” is structurally more likely to be amplified into a shareable clip than a policy explanation of comparable length, opponent-directed rhetoric will show a larger and faster short-run low-dollar response than constructive policy messaging, holding the level of media attention a moment receives roughly constant.

H1b. Constructive policy messaging will nonetheless be positively associated with giving, and its financial return will show up less as single-event spikes than as gradual, cumulative growth in a candidate's donor base over the course of the campaign.

H1c. The financial response to any given rhetorical moment will not necessarily accrue to the candidate who produced it: attacks in particular can generate spillover giving to a third candidate whose supporters are quickest to convert the moment into a donation appeal of their own.

These sub-expectations concern short-run association, not causal identification, and are consistent with several mechanisms this paper's design cannot separate – attention, anger, policy congruence, and the surrounding media narrative – which, per Feigenbaum and Shelton's (2013) account of the feedback between viability and fundraising, are unlikely to operate independently of one another.

**Methodology**

A fully rigorous test of H1 would need the kind of data infrastructure Bouton, Cagé, Dewitte, and Pons (2022) built for their study of small donors: itemized ActBlue and FEC records reconstructed at the candidate-day level, matched to independently double-coded content analysis of all eleven debates, analyzed with candidate and event fixed effects. Reconciling that dataset – parsing bulk filings, resolving renamed committee identifiers, running a two-coder reliability process on a corpus of several hundred thousand words – is a large-scale project that cannot be claimed complete without the underlying raw downloads to support it. This paper does not attempt that reconstruction, and treats any research claiming exact, reconciled totals without showing verified data with real skepticism – a standard it holds itself to as well. It instead uses a comparative case-study design built from four kinds of public material: debate transcripts, a rhetoric-coding procedure, campaign-disclosed contribution data, and a fixed short-run window.

The 2020 primary included eleven sanctioned debates between June 2019 and March 2020, transcribed in full and archived by nonpartisan repositories such as the American Presidency Project. Given the scope of a single-author study on a short timeline, this paper does not attempt a full sentence-level pass over all eleven transcripts; it instead works from the specific exchanges independent news coverage flagged as each debate's most consequential moments, returning to the transcript or broadcast record to confirm wording and context.

Each exchange was classified as predominantly an attack, predominantly constructive policy messaging, or mixed. Following established content-analysis practice, the coding rests on communicative function – an unfavorable evaluation of an identifiable rival versus affirmative advocacy of the speaker's own position – rather than tone or truth value (Krippendorff, 2004). This coding was performed by a single reviewer and not independently cross-checked.

The outcome is low-dollar ActBlue giving, defined as individual contributions at or below $200 (Bouton et al., 2022). The contribution data used here are the specific figures campaigns disclosed to the press – a donor count, an average gift size, a total dollar figure – treated as campaign-reported claims corroborated across outlets where possible, not as independently audited totals. The short-run window is the 24-to-72-hour period following a debate, matching the window campaigns themselves used; it reflects what is checkable in the public record, not a theoretical claim about the correct horizon, and cannot detect longer-run effects such as the general-election backlash Niebler and Urban (2017) documented.

The independent variables, following H1, are a candidate's rate of attack rhetoric and rate of constructive messaging in a given exchange; the dependent variables are the number and dollar value of qualifying contributions afterward. Given the small number of cases and the absence of reconciled transaction data, this paper does not fit a regression or compute an inferential statistic – doing so on five non-randomly selected, self-reported cases would create a false impression of precision. The approach is a structured descriptive comparison: classify rhetoric, record the disclosed figures, and compare the pattern against H1 and H1a–H1c.

**Results and Discussion**

Table 1 summarizes the five cases. Two are single, dateable exchanges with an attack-dominant character (Harris–Biden; “wine cave”); one is a single, constructive debate moment (Yang's pilot announcement); two are cumulative, non-debate-specific trajectories (Warren; Sanders) serving as a baseline for steady, brand-driven giving. Reported figures range from just over one million dollars (Yang, within days) to just over three million (Harris, that weekend), while the cumulative cases report tens of millions over a full quarter – a reminder that the two kinds of case are not on the same time scale.

Table 1. Summary of the five cases discussed, with reported short-run fundraising figures and their primary sources.

| **Candidate** | **Debate / Date** | **Rhetoric Type** | **Reported Short-Run Response** | **Primary Source(s)** |
|---|---|---|---|---|
| Harris | June 27, 2019 | Attack (Biden, busing) | ~$2M/24 hrs, 63,277 donors, 58% first-time; +$1.2M that weekend | CNN/KPAX; PBS NewsHour (2019) |
| Buttigieg / Sanders | Dec. 19, 2019 | Attack (Warren on "wine cave") | Viral moment; redirect captured by Sanders surrogates, not Warren | BuzzFeed News; NBC News (2019) |
| Yang | Sept. 12, 2019 | Constructive (Dividend pilot announcement) | $1M+ within days; Q3 $10M (+257%); 99% gifts <$200, avg. $30.18 | CBS News; Yahoo Finance (2019) |
| Warren | Cumulative, Q1–Q3 (no single debate) | Constructive ("plan for that" brand) | $6M → $19.1M → $24.6M; 509,000 donors in Q3 | PBS NewsHour (2019) |
| Sanders | Cumulative, 2019 (baseline) | Constructive / movement-branded | ~$6M in first 4 hrs post-launch; steady top-tier fundraiser | IVN (2019) |

Each case below reports both how the precipitating exchange was coded (rhetoric pattern) and what the campaign subsequently disclosed about giving (fundraising pattern), since separating the two would obscure the timing relationship between them.

At the first Democratic debate in Miami on June 27, 2019, Senator Kamala Harris directed a sustained critique at frontrunner Joe Biden over his past collaboration with segregationist senators and his record opposing school busing. Widely covered as the debate's defining moment, it was followed by more than two million dollars raised within twenty-four hours from over 63,000 donors, 58 percent first-time givers, averaging roughly thirty dollars a gift (CNN/KPAX, 2019); a further $1.2 million arrived that weekend, near $3.2 million total (PBS NewsHour, 2019) – the sample's sharpest example of an attack producing an immediate response, consistent with H1a. But Harris's polling gains proved temporary and her later fundraising did not sustain Warren's or Sanders's growth (CNBC, 2019); if Niebler and Urban's (2017) longer-run finding applies here, this spike may have coexisted with costs a short window cannot see.

At the December 19, 2019 debate, Senator Warren attacked Mayor Pete Buttigieg over a fundraiser held that month in a Napa Valley wine cave, arguing that "billionaires in wine caves should not pick the next president of the United States" (BuzzFeed News, 2019) – the debate's most replayed moment. The clearest financial beneficiary was neither Warren nor Buttigieg, whose campaign kept relying on small gifts averaging about $32 (Boston Globe, 2019); instead, Bernie Sanders surrogates wore shirts referencing the story at the debate itself, redirecting attention to a Sanders donation page built around the moment (NBC News, 2019) – a clean illustration of H1c.

Not every spike followed an attack. At the September 12, 2019 debate, Andrew Yang announced his campaign would give ten American families $1,000 a month for a year, dramatizing his basic-income proposal rather than criticizing a rival (Yahoo News, 2019). Within days his campaign reported over a million dollars in new donations (CBS News, 2019), part of a third-quarter haul of $10 million – a 257 percent jump, 99 percent of gifts under $200, averaging $30.18 (National Review, 2019; Yahoo Finance, 2019). This supports H1b but complicates H1a: the operative variable looks less like "policy versus attack" than which candidate manufactured a shareable moment.

The clearest contrast comes from the candidates whose fundraising grew steadily rather than jumping after one debate. Warren's operation, built on refusing high-dollar fundraisers and her "I have a plan for that" brand, grew from $6 million (Q1) to $19.1 million (Q2) to $24.6 million (Q3), the last from 509,000 donors making 943,000 donations (PBS NewsHour, 2019). Sanders raised nearly $6 million in the first four hours after launch alone (IVN, 2019) and stayed a top fundraiser every quarter with no single debate credited – consistent with H1b's expectation of gradual payoff.

Read together, these cases offer qualified support for H1a: the largest, fastest spike followed an attack, fitting the attention- and anger-based channels the literature suggests (Valentino et al., 2011). The support is qualified three ways. The "wine cave" case supports H1c: reward can accrue to a third party, a more interesting fact than a simple negativity premium. Yang's case supports H1b and complicates H1a: a well-designed constructive moment can spike just as fast, so the operative variable may be "shareable moment" more than "attack" as such. Warren's and Sanders's trajectories, also consistent with H1b, show much of 2020's giving grew around a sustained policy brand rather than any single spike – a pattern this spike-selecting method is poorly positioned to weigh against the attack-driven cases, since campaigns announce dramatic spikes far more readily than steady growth. Overall, the five-case pattern is consistent with H1's claim that rhetoric type is associated with short-run giving, but its direction and beneficiary vary in ways this comparison can describe but not explain with confidence.

This paper is built from self-reported, publicity-selected fundraising figures rather than reconciled transaction data, and from five cases chosen for documentation rather than random or exhaustive sampling of the full debate calendar. It cannot separate rhetoric's effect from the media coverage it generates, a candidate's pre-existing trajectory, or simultaneous solicitations outside the debate. Krippendorff's (2004) standards guided how exchanges were classified, but

no second coder reviewed the classifications, so the labels here reflect this paper's own careful judgment rather than a validated instrument.

**Conclusion**

This paper asked how opponent-directed attacks and constructive policy messaging in the 2020 Democratic primary debates related to candidates' short-run, low-dollar fundraising. Using the best-documented public cases from that race, it finds that the fastest and largest short-run spike followed an attack, that a well-designed constructive moment could produce a comparably fast spike, that an attack's financial reward is not guaranteed to reach the candidate who delivered it, and that a substantial share of 2020's small-dollar giving grew steadily around policy-based brands rather than around any single debate.

The broader point is not that negativity simply pays and policy does not. The 2020 primary's institutional structure – an unusually large field, a debate stage gated by rising donor-count thresholds, and a platform that processed a five-dollar gift in seconds – rewarded whichever rhetorical choice produced the most shareable, decision-forcing moment for a national audience. Attacks suit that structurally, consistent with why this sample's clearest spike followed one; but Yang's announcement shows distinctive policy content can do the same, and the "wine cave" case shows the reward can be captured by a bystander rather than its author. A campaign-finance system built substantially on virality is not neatly biased toward attacks over substance – it is biased toward whichever content, of either kind, travels fastest.

This paper's design cannot turn that qualified reading into a causal claim. The natural next step is the transaction-level study deliberately not attempted here: independently verified, itemized ActBlue and FEC records reconstructed at the candidate-day level for all eleven debates, matched to sentence-level coding with genuine inter-coder reliability statistics and analyzed with candidate and event fixed effects. Such a study could also test whether the short-run gains documented here carry the longer-run costs Niebler and Urban (2017) found when they followed primary donors into the general election – a question this paper's short window cannot answer, but a consequential one for the small-dollar era's incentives. The rhetoric-coding procedure developed here could also inform a lightweight applied tool that classifies and visualizes debate rhetoric in near real time.

**Works Cited**


Ansolabehere, S., de Figueiredo, J. M., & Snyder, J. M., Jr. (2003). Why is there so little money in U.S. politics? Journal of Economic Perspectives, 17(1), 105–130.

Bouton, L., Cagé, J., Dewitte, E., & Pons, V. (2022, revised 2024). Small campaign donors (NBER Working Paper No. 30050). National Bureau of Economic Research. https://doi.org/10.3386/w30050

Feigenbaum, J. J., & Shelton, C. A. (2013). The vicious cycle: Fundraising and perceived viability in US presidential primaries. Quarterly Journal of Political Science, 8(1), 1–40.

Gaynor, S. W., & Gimpel, J. G. (2021). Small donor contributions in response to email outreach by a political campaign. Journal of Political Marketing, 23(1), 51–73. https://doi.org/10.1080/15377857.2021.1910611

Gilardi, F., Alizadeh, M., & Kubli, M. (2023). ChatGPT outperforms crowd workers for text-annotation tasks. Proceedings of the National Academy of Sciences, 120(30), e2305016120.

Hill, S. J., & Huber, G. A. (2017). Representativeness and motivations of the contemporary donorate: Results from merged survey and administrative records. Political Behavior, 39(1), 3–29.

Jacob, M. S., Lelkes, Y., & Westwood, S. J. (2026). Entrepreneurs of conflict: A descriptive analysis of when and how political elites use divisive rhetoric. PNAS Nexus. As described in University of Pennsylvania Annenberg School for Communication, “Insults get attention, not results, for some U.S. lawmakers.”

Krippendorff, K. (2004). Reliability in content analysis: Some common misconceptions and recommendations. Human Communication Research, 30(3), 411–433.

Krupnikov, Y. (2011). When does negativity demobilize? Tracing the conditional effect of negative campaigning on voter turnout. American Journal of Political Science, 55(4), 797–813.

Lau, R. R., Sigelman, L., & Rovner, I. B. (2007). The effects of negative political campaigns: A meta-analytic reassessment. Journal of Politics, 69(4), 1176–1209.

Le Pennec, C., & Pons, V. (2023). How do campaigns shape vote choice? Multicountry evidence from 62 elections and 56 TV debates. Quarterly Journal of Economics, 138(2), 703–767.

Mummalaneni, S., Wang, R. J.-H., & Isaac, M. S. (2024). Email campaigns that suit the candidate: Leveraging automated text analysis to increase political donations. Journal of Interactive Marketing, 59(4), 385–399.

Mutz, D. C. (1995). Effects of horse-race coverage on campaign coffers: Strategic contributing in presidential primaries. Journal of Politics, 57(4), 1015–1042.

Niebler, S., & Urban, C. (2017). Does negative advertising affect giving behavior? Evidence from campaign contributions. Journal of Public Economics, 146, 15–26.

Valentino, N. A., Brader, T., Groenendyk, E. W., Gregorowicz, K., & Hutchings, V. L. (2011). Election night's alright for fighting: The role of emotions in political participation. Journal of Politics, 73(1), 156–170.

Axios. (2019, June 30). Kamala Harris sees a polling bump after taking on Joe Biden.
Boston Globe. (2019, December 21). Buttigieg backers defend 'wine cave' fundraiser.
BuzzFeed News. (2019, December 19). The fight between Elizabeth Warren and Pete Buttigieg blew up over a “wine cave.”
CBS News. (2019, October 2). Andrew Yang raises $10 million in the third quarter.
CNBC. (2019, July 5). Kamala Harris raises $12 million in second quarter, trailing Biden.
CNN/KPAX. (2019, June 29). Kamala Harris raises $2 million in 24 hours after breakout debate.
Democratic National Committee qualification criteria, as reported by PBS NewsHour (2019, May 9) and NBC News (2019, December 20).
IVN. (2019). The DNC debate rules: What you need to know.
National Review. (2019, October 2). Yang raised $10 million in third quarter, easily outpacing two sitting U.S. senators.
NBC News. (2019, December 19). 'Wine cave': Inside Buttigieg's high-priced underground fundraiser.
PBS NewsHour. (2019, May 9). What do Democratic candidates need to make the first 2020 debates?
PBS NewsHour. (2019, July). Harris raised nearly $12 million in the second quarter.
PBS NewsHour. (2019, October). Warren raises $24.6 million, largely from small donors.
Yahoo Finance. (2019, October 2). Andrew Yang's 257% fundraising surge blows away all other Democratic candidates.
Yahoo News. (2019, September 13). Andrew Yang reveals his debate surprise — $1,000 a month for 10 more families.

**Impact of Analytical AI on Equity Volatility By Aliyaan Noorani**


## Abstract

This study specifically examines how the news of publicly traded companies implementing analytical AI can affect investor sentiment and behaviour, as measured through short term changes in return volatility of stocks and average returns. Using an event study methodology, this study analyzes three publicly traded companies—Walmart, Hilton, and Amazon—each with a specific and dateable disclosure of analytical AI implementation. For each company, daily returns and return volatility were compared across pre and post event windows of 1, 5, and 7 trading days, using independent two sample t-tests and Levene's tests of equality of variances. Ordinary least squares regressions were also estimated to assess whether the event date predicted returns beyond what the volatility would suggest. Contrary to the study's hypothesis that AI implementation disclosures would be associated with increased short term volatility, no company exhibited a statistically significant change in volatility at any window length. However, Hilton demonstrated a statistically significant decline in average returns at both the 5 day and 7 day windows, a result not detected in the regression analysis, suggesting that any sentiment effect present may be short-lived and concentrated immediately around the event. Given the small sample size and the number of statistical comparisons conducted, these findings are best interpreted as exploratory into potential effects of news about AI implementation by certain corporations rather than conclusive evidence of a broader relationship between analytical AI disclosures and investor sentiment.


## Introduction

Artificial Intelligence continues to take various industries in today's economy by storm. Different types of AI have been implemented in order to develop efficiency and solutions for problems. This has also changed the way investors and the overall public view corporations, as many businesses begin to incorporate AI in their practices. This sentiment, especially when talking about publicly traded companies, can vastly impact the market. Sentiment is largely caused by news on these corporations implementations and endeavours. To that end, this paper attempts to answer the following question: How does news about publicly traded companies implementing analytical Artificial Intelligence, as constituted by machine learning systems that analyzes data to support decision making in order to increase revenue and efficiency for said companies, impact investor sentiment and behaviour around the companies' stock price? This study measures investor sentiment indirectly through observable price behaviour, specifically short term stock return volatility and average returns, rather than through direct sentiment indicators such as options based measures or text based sentiment analysis. This approach followed Barunik (2020), who demonstrates that a stock's volatility relative to its historical average can itself serve as a meaningful substitute for underlying shifts in market sentiment. Implementation of analytical Artificial Intelligence by publicly traded companies is hypothesized to overall cause higher amounts of measurable, short term stock return volatility during the

trading days around a specific event, as investors may reassess positioning and risk with changes in technology and decision making. Narrower event windows are also hypothesized to show an overall higher volatility and more pronounced impact after a specific news event has occurred, and will gradually diminish as the market adapts and adjusts to the news of AI implementation over time.

**Literature Review**

In order to assess how the incorporation of analytical AI can actually impact sentiment, we must first incorporate a variety of existing sources that include information on the importance of sentiment as well as how it is tracked, and the ways in which analytical AI can be implemented to generate growth. Sentiment can be defined as investors' collective optimism or pessimism about a stock's future cash flows and risk that is not fully justified by fundamentals. Unlike idiosyncratic noise trading, which reflects random errors across individual investors, sentiment is systematic: it moves broadly across investors at the same time, often driven by many cognitive biases. There can be different outlooks on the way that investor sentiment is approached as well, such as bottom up and top down (Baker 2007). Similarly, a study by Ballinari and Behrent (2021) focuses on a comparison of 9 different sentiment estimation techniques, specifically based on that of social media. It concentrates more specifically on text, and the two ways that it can be analyzed in order to gauge sentiment. One tracks negative and positive wording in order to gain a better idea of sentiment on platforms such as Twitter and Instagram, while the other trains an AI model in order to track and recognize tone. This study does not adopt text based or social media sentiment measures directly, primarily due to the practical difficulty of obtaining consistent and comprehensive social media data across an 8 year sample period for the three different companies. A study by Arindam Bandopadhyaya (2008) demonstrates that the Put-Call ratio is the best method for studying investor sentiment, as it specifically demonstrates whether investors believe a certain stock's price will tend to increase or decrease. Once companies decide to utilize Artificial Intelligence to make decisions and analyze data to a higher extent, the change in PCR can be analyzed in order to see how sentiment stayed consistent or changed. Jozef Barunik (2020) states in his study that looking specifically at how a stock price fluctuates, or its volatility compared to its overall average can provide insight on sentiment. Based on sentiment, the market would project Bullish or Bearish prices compared to previous averages.

According to a study by Yu et al. (2026), different types of AI are overall implemented by companies in the S&P 500 index. For example, about 10% of companies implement AI for actual production uses, and those companies with a deep implementation of Artificial Intelligence have higher profit margins than those who are recently implementing AI in their operations. AI implementation has also led to higher firm valuation, although this does not necessarily directly tie to investor sentiment or stock valuation. It is worth noting that this relationship may reflect reverse causality: profitable, well capitalized firms may simply have greater resources available to invest in AI, rather than AI adoption itself driving the observed margin and valuation gains.

The study does not establish which direction the relationship runs. Businesses utilize analytical Artificial Intelligence in order to make projections and decisions for the future. As demonstrated in a study conducted by Sareen Rachakatla (2023), analytical AI can be used for various decisions, including analyzing target markets, customer behaviour, and even using it to analyze data and make decisions for supply chain optimization and operational uses by corporations.

**Methodology**

In order to test the hypothesis of publicly traded companies implementing analytical AI's association with short term increases in stock return volatility, this study employs an event study methodology applied to three publicly traded companies: Walmart (WMT), Hilton (HLT), and Amazon (AMZN). These businesses were specifically selected due to the fact that they each have specific, dateable news events in which they clearly disclose AI implementation. Furthermore, the companies have variability in reasons and methods of implementation as well as overall impact of news. Walmart's disclosure, made on October 25, 2023, centered on a "next generation supply chain" system, which uses real time data and analytical AI in order to optimize inventory placement across stores and regions, ensuring that regional demand patterns are met more precisely. Amazon's disclosure, on the other hand, occurred on June 11, 2025. Their system focused on an AI powered demand forecasting model which included time bound data as well as allowing warehouse robots to interpret commands. Hilton's disclosure, made on February 11, 2026, differed from the other two both in format and scope. The news of implementation emerged in the yearly Q4 2025 earnings call, where AI was said to be incorporated in order to drive operational and distribution efficiencies. Hilton aims to make supported decisions regarding customer relations as well partnerships using analytical AI models. For each company, event windows were specifically constructed by identifying the trading day nearest to the news event date and extracting a symmetric number of trading days immediately before (pre-event) and after (post-event) that date. In order to actually assess whether the length of the trading window has a substantial impact on the return volatility for each company after a news event, three window lengths were tested for each event: 1 trading day, 5 trading days, and 7 trading days pre and post event. This multi-window approach allows for comparison of whether any effect is concentrated immediately around the event or persists over a longer adjustment period, which is consistent with the hypothesis that a narrower window could show a more pronounced effect. For each company and each window length, average daily returns and return volatility (measured as the standard deviation of daily returns) were calculated separately for the pre-event and post event periods. In order to test for a statistically significant difference in average returns between the pre and post event periods, an independent two sample t test was conducted. To test for a statistically significant difference in volatility between the two periods, Levene's test of equality of variances was conducted. A significance threshold of $\alpha = 0.05$ was used for both tests. In addition to the pre and post window comparisons, ordinary least squares (OLS) regressions were estimated for each company in order to explore the relation between volatility and returns around the event date. Specifically, daily returns were regressed on two lagged values

of the 21 day rolling volatility (1 and 2 trading days lagged) and an indicator variable equal to 1 on the specified event date and 0 otherwise. The latter specification captures a differential effect on the daily return on the specified event date beyond the relation to volatility estimated over the preceding 21 day period. Two design limitations of this regression specification are worth noting up front. First, this regression does not include a market or benchmark return, such as the daily return on the S&P 500 index, as a controllable variable, which is the standard baseline for a regression of this kind. Second, because this regression evaluates a single day event indicator against several years of daily trading data, the model has limited statistical power to detect a short lived effect, even if one may potentially be present. Given these limitations, the full sample regression is expected to show low explanatory power almost by design, since a short lived, event specific effect is unlikely to be detectable when compared against years of ordinary daily trading data. Accordingly, this regression is best understood as a comparison check against the more sensitive pre/post comparison reported later in this paper.

**Results**

Before actually looking at the effect of the news events, baseline volatility was calculated for each corporation across the full sample period (January 2018 through August 2026) to contextualize any event-window findings. Measured as the Standard Deviation of daily returns, Amazon demonstrated the highest overall volatility (0.0218), followed by Hilton (0.0191), and least of all Walmart (0.0142). This order is generally consistent with each company's sector and market capitalization profile. For example, Walmart’s larger and more stable retail operation system generally corresponds to a lower baseline volatility, while Amazon, with its rapid growth orientation and constant development, shows a much higher baseline volatility. This baseline volatility reflects each company’s overall return dispersion across the full sample period but is not adjusted for market wide movements. As shown in Figure 3 below, all three companies experienced sizable, broad based volatility spikes in early 2020 and 2022 that were unrelated to AI announcements. Without subtracting a market benchmark such as the S&P 500s from each company’s daily return, it is difficult to fully separate firm specific reactions to the AI disclosures from these broader market wide conditions.

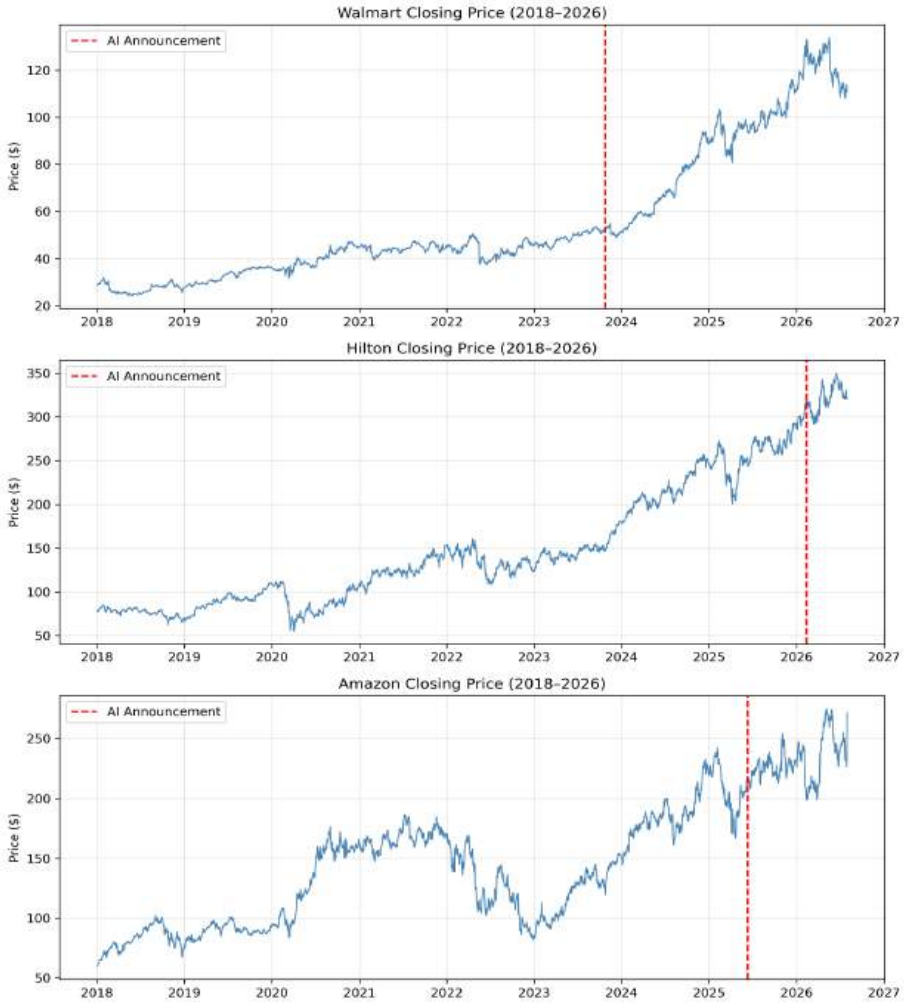



*Figure 1. Closing stock price history with AI implementation announcements marked for each company.*

In order to test whether the identified implementation event dates had a statistically significant differential impact on daily returns, beyond that explained by recent volatility patterns, an OLS regression for each company was estimated using two lagged values (1-day and 2-day) of 21 day rolling volatility and a binary indicator for the event date. This specification is designed to isolate the event date's effect from the broader volatility environment surrounding it. If investor sentiment really did shift meaningfully due to the AI disclosure, the event date coefficient would be expected to capture a return effect that is different from what general recent volatility trends would predict. Across all three companies, the regression models did not show any statistically significant explanatory power overall. This outcome is consistent with the low statistical power inherent in comparing single day indicator against years of daily trading data, as noted in the Methodology section, rather than indicating a genuine absence of any effect. For Walmart, the model produced an $R^2$ of 0.000 ($F = 0.139$, $p = 0.937$), with the event date coefficient (beta = -0.0036, $p = 0.801$) not reaching significance. This may show that Walmart's

October 25, 2023 disclosure of AI powered inventory management system did not have any clear measurable effects on returns once the model adjusted and accounted for recent volatility, although the announcement was concrete and precise itself. Hilton's model also did not show a significant fit ($R^2 = 0.000$, $F = 0.250$, $p = 0.861$), with an event date coefficient of beta = 0.0040 ($p = 0.833$). This result is notable given that Hilton's disclosure, embedded within a Q4 2025 earnings call rather than a standalone explicit AI announcement by the company, was the only one to produce a statistically significant result discussed in the pre/post window comparison below. This suggests that Hilton's regression may not have detected the effect as the model compared the single event day against the company's entire return history. A short shift in returns may easily be minimized when being measured against a large amount of data, even if the effect becomes clear when the analysis is narrowed. Amazon's model showed higher explanatory power ($R^2 = 0.002$, $F = 1.696$, $p = 0.166$), though still not statistically significant. Notably, Amazon's two lagged volatility terms were independently significant at the 5% level (Lag 1: beta = -0.6088, $p = 0.048$; Lag 2: beta = 0.6266, $p = 0.042$), which suggests a short term mean-reverting relationship between recent volatility and returns that exist independently of the AI announcement. This indicates that Amazon's returns are sensitive to its own recent volatility patterns. However, Amazon's event-date coefficient (beta = -0.0212, $p = 0.333$) was not statistically significant, indicating that the AI disclosure itself did not produce a substantial deviation in returns compared to what the stock's ordinary volatility dynamics would already predict. Taken together, these regression results do not provide evidence that the AI implementation event date, evaluated as a single day indicator against each company's multi-year return history, was associated with a measurable shift in daily returns. This also points to a key methodological point carried into the event window analysis: a single day indicator measured against years of data may be too diluted to actually detect a short lived sentiment reaction, even if one is present in the days immediately around the event. The pre/post window comparison, conducted at 1, 5, and 7 day trading windows for each company, provides more depth and nuance than the general regression analysis. At the narrowest window (1 day), pre and post-event trading volatility, t statistics, and p values could not be determined as a 1 day trading window and single day sample does not provide basis for calculating variance; only the directional change in average return could be assessed at this window length. Table 1 summarizes the results for the 5 and 7 day windows, where full statistical testing was possible.

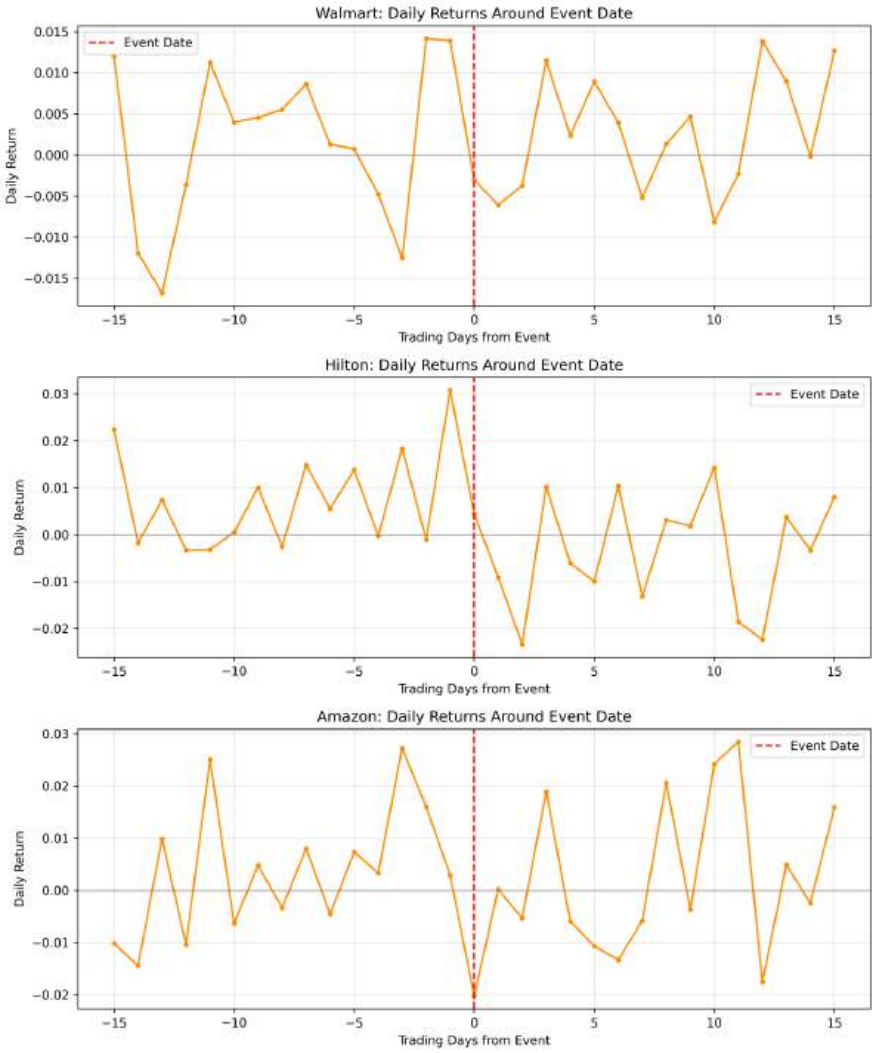



*Figure 2. Daily returns in the window immediately surrounding each event date.*

| Company | Window (days) | Pre Avg Return | Post Avg Return | Return Change | t stat | t p-value | Returns Sig. | Levene p-value | Vol. Sig. |
|---|---|---|---|---|---|---|---|---|---|
| Walmart | 5 | 0.0023 | 0.0026 | 0.0003 | 0.047 | 0.9642 | No | 0.3662 | No |
| Walmart | 7 | 0.0031 | 0.0017 | -0.0014 | -0.305 | 0.7664 | No | 0.4522 | No |
| Hilton | 5 | 0.0123 | -0.0077 | -0.0200 | -2.484 | 0.0382 | Yes | 0.6242 | No |
| Hilton | 7 | 0.0117 | -0.0059 | -0.0176 | -2.783 | 0.0166 | Yes | 0.9696 | No |
| Amazon | 5 | 0.0114 | -0.0006 | -0.0119 | -1.725 | 0.1232 | No | 0.9631 | No |
| Amazon | 7 | 0.0086 | -0.0032 | -0.0118 | -2.103 | 0.0573 | No | 0.8450 | No |

*Table 1. Pre/Post Event Window Comparison (5-day and 7-day windows)*

Walmart's stock did not show any statistically significant change in average returns at either the 5 day window ($t = 0.047$, $p = 0.964$) or 7 day window ($t = -0.305$, $p = 0.766$), and volatility did not significantly change at either window (Levene $p = 0.366$ and $p = 0.452$ respectively). This result is consistent with the findings stated above, where Walmart's event date coefficient was also far from significance ($p = 0.801$). Taken together, both methods suggest that Walmart's October 25, 2023 disclosure of AI implementation was most likely treated by investors as a routine operational efficiency measure rather than a large signal warranting a reassessment of the company's risk. Hilton, on the other hand, was the only company to show a statistically significant difference in average returns between the pre and post event periods, at both the 5 day window ($t = -2.484$, $p = 0.038$) and the 7 day window ($t = -2.783$, $p = 0.017$). In both cases, average daily returns declined following the event: from 0.0123 to -0.0077 at the 5 day window, and from 0.0117 to -0.0059 at the 7 day window. Despite this significant shift in average returns, Hilton's volatility did not significantly change at either window (Levene $p = 0.624$ and $p = 0.970$, respectively), which indicates that the detected effect was concentrated in the direction and magnitude of returns rather than in the degree of return dispersion. Amazon's average returns declined numerically following its event date at both windows (5 day: 0.0114 to -0.0006; 7 day: 0.0086 to -0.0032), and the 7 day window approached but did not cross the significance threshold ($t = -2.103$, $p = 0.057$). Amazon's volatility did not change significantly at either window (Levene $p = 0.963$ and $p = 0.845$). At the 7 day window, the t-test failed to reject the null hypothesis at the conventional 0.05 threshold ($p = 0.057$). The direction of the estimated effect was negative, consistent with the pattern observed at the 5 day window, though this result does not provide evidence of a shift in sentiment. The significant lagged volatility terms found in Amazon's regression reflect a mean reverting relationship that is a well documented and generic feature of daily returns for most liquid large-cap stocks, rather than something unique to Amazon or to AI related news specifically. This regularity should therefore be treated as a separate feature of the data, unrelated to the event date finding discussed above.

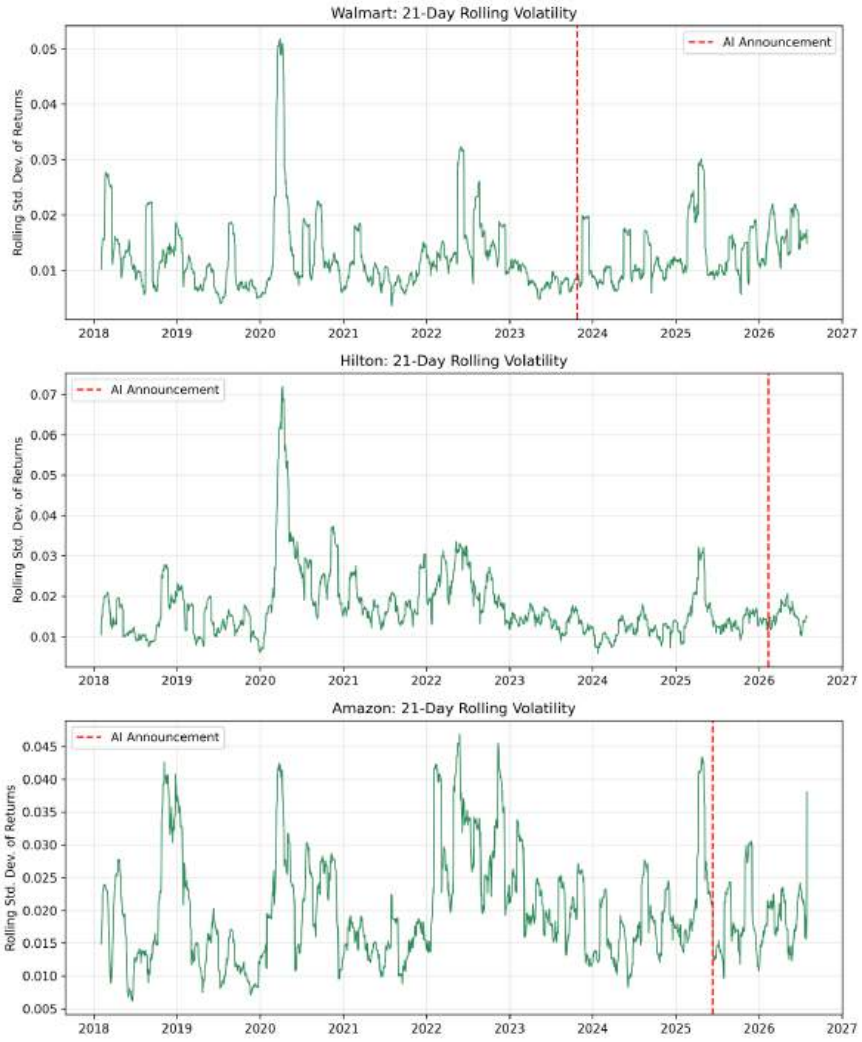



*Figure 3. Rolling Volatility with AI implementation announcement days marked.*

As shown in Figure 3, rolling volatility for all three companies fluctuates considerably across the full sample period, driven largely by broader market events. None of the three companies show a distinct and sizable spike in volatility coinciding with their AI disclosure date, visually enforcing the Levene's test results above.

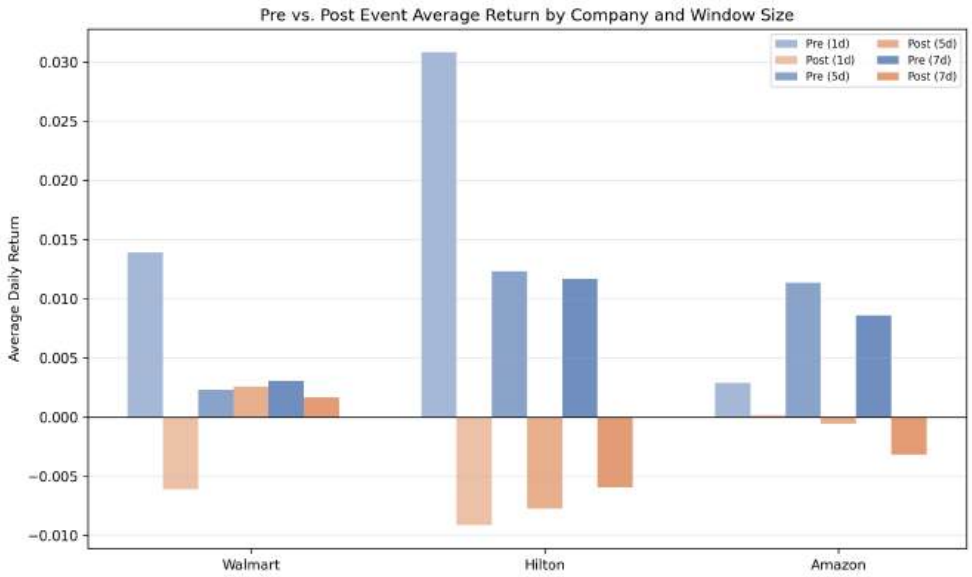



*Figure 4. Pre vs. Post event average return by company and window length.*

Figure 4 visualizes the return comparisons summarized in Table 1. Walmart's and Amazon's pre and post event bars remain close in height along all window lengths, consistent with the absence of a significant shift for each company. Hilton's bars, in contrast, show a visible decline from a positive pre-event average to a negative post-event average at both the 5 and 7 day windows.

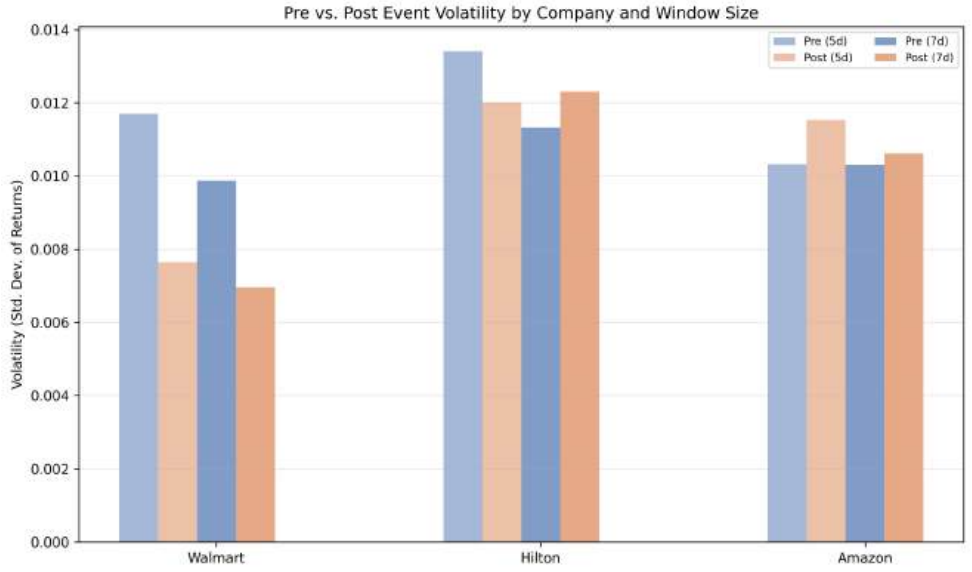



*Figure 5. Pre vs. Post event volatility by company at each window length.*

Figure 5 presents the corresponding comparison for volatility. Unlike Figure 4, the pre and post event bars remain broadly similar in height across all three companies, including Hilton, reinforcing that the detected effect was concentrated in returns rather than return dispersion.

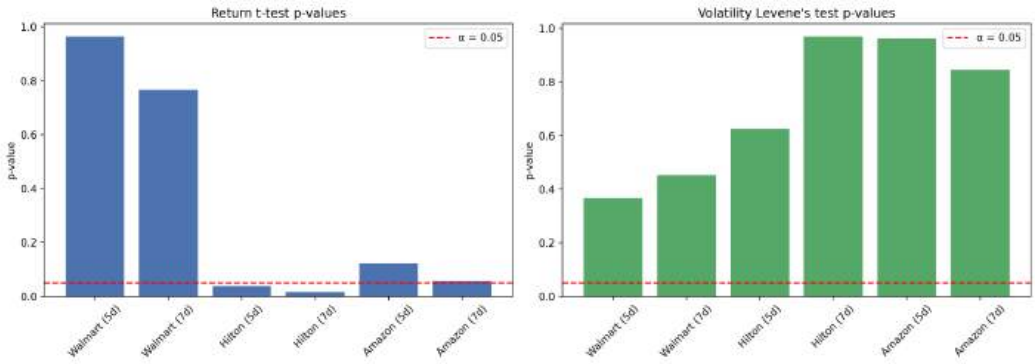



*Figure 6. Summary of statistical significance across all tested companies and windows.*

Figure 6 consolidates the significance of every comparison conducted in this study. Across both panels, only Hilton's two return based comparisons fall below the alpha = 0.05 line, while every other bar remains above it, underscoring that significant evidence was limited to a single company and a single measure out of eighteen total comparisons.

Across the three companies and three window lengths that were tested, Hilton was the only company in which a statistically significant result in average returns was demonstrated, and this effect was present at both the 5 day and 7 day windows but not measurable at the 1 day window. No company exhibited a statistically significant change in return volatility, as measured by Levene's test, at any window length where testing was possible. The regression results, which isolated the event date as a single day indicator across the full sample period, did not find a statistically significant event day effect for any company, suggesting that the effect detected in the window comparison for Hilton is better characterized as a short-term shift concentrated in the days surrounding the event, rather than a persistent or broadly measurable relationship across the full historical sample.

It is worth noting that this pattern runs counter to the study's hypothesis that narrower windows would show a more pronounced effect that diminishes as the window widens. In practice, the opposite occurred: Hilton's 7 day t-statistic was larger in magnitude than its 5 day t-statistic, and Amazon's 7 day result was closer to significance than its 5 day result. This suggests that a full return reaction may take several trading days to materialize, or that 5 versus 7 days is too narrow a comparison to reveal the diminishing pattern the hypothesis anticipated.

**Limitations**

Several limitations of this study should be acknowledged when these results are being interpreted. First, the sample size is small, only three publicly traded corporations. Therefore, these findings are best understood as an exploratory case that could be looked at for what impacts of AI in the market could occur, but should not be generalizable to the market as a whole about how analytical AI disclosures affect investor sentiment for all publicly traded companies. A much larger sample would be necessary in order to determine whether the significant return

shift detected for Hilton reflects a broader phenomenon or is specific to these events. Second, this study conducted eighteen statistical tests in total across the t-test and Levene test analyses, in addition to three separate OLS regressions. When a threshold of alpha = 0.05 is applied across this many tests, some significant results would be expected to occur by chance alone. The single significant finding for Hilton should therefore be interpreted with caution, and would benefit from replication using a larger sample. Third, Hilton's event date is a less clean measure of analytical AI implementation than Walmart's or Amazon's. While the other two are demonstrated in an explicit AI announcement, Hilton's announcement was in a Q4 2025 earnings call. Because the announcement was disclosed simultaneously with other variables, the significant shift around this may reflect the general earnings results. This study specifically measured raw stock returns rather than market adjusted, or abnormal, returns. A more standard event study approach would compare each company's return to a market benchmark, such as the S&P 500, over the same period, to isolate the company specific reaction from broader market movement. This limitation is relevant due to the broad market wide volatility spikes visible in Figure 3 during early 2020 and 2022, which reflect general market conditions rather than firm specific reaction. Without market adjustment, such conditions cannot be cleanly distinguished from a genuine firm specific response to AI disclosures using the approach utilized here.

**Conclusion**

This paper set out to examine how specifically the news of publicly traded companies implementing analytical AI affects investor sentiment and behaviour, as measured through short term stock changes in volatility and returns. Using an event study methodology that assessed the companies of Walmart, Amazon, and Hilton, the results provided mixed but informative evidence. Notably, this study's original hypothesis, that news about AI implementation by publicly traded companies would be correlated to increased volatility, was not supported. No company exhibited a statistically significant change in volatility, measured using Levene's test, at any window length. This suggests, for the events examined here, an AI adoption disclosure may not be likely to introduce genuine uncertainty about how the stock trades. The study's related hypothesis that narrower event windows would show a more pronounced effect that gradually diminishes as the window widens was also not supported: where a significant return effect was present, it was slightly stronger at the 7 day window than at the 5 day window for both Hilton and Amazon, the opposite predicted pattern. This indicates that, within the range of windows tested, any return reaction of AI disclosure may take a few trading days to develop rather than being strongest immediately after the event. Taken as a whole, this study's findings are best understood as suggestive rather than conclusive. The significant result exhibited by Hilton may be due to a small sample size as well as confounding variables exhibited in the Q4 earnings statement. Ultimately, the results show that in the three corporations studied, AI disclosure will not show significant results in volatility in the ways in which the stock trades. In the future, studies with greater sample sizes could look at industry and market wide data in order to generalize further to the market.

**Works Cited**


Baker, Malcolm, and Jeffrey Wurgler. “Investor Sentiment in the Stock Market.” Journal of Economic Perspectives, vol. 21, no. 2, Apr. 2007, pp. 129–51, https://doi.org/10.1257/jep.21.2.129.

Ballinari, Daniele, and Simon Behrendt. “How to Gauge Investor Behavior? A Comparison of Online Investor Sentiment Measures.” Digital Finance, vol. 3, no. 2, June 2021, pp. 169–204, https://doi.org/10.1007/s42521-021-00038-2.

Bandopadhyaya, Arindam. Measures Of Investor Sentiment: A Comparative Analysis Put-Call Ratio Vs. Volatility Index. University of Massachusetts-Boston, Aug. 2008, https://scispace.com/pdf/measures-of-investor-sentiment-a-comparative-analysis-put-2mc222wfxn.pdf.

Barunik, Jozef, et al. “Investment Disputes and Abnormal Volatility of Stocks.” 2006.10505v1, 18 June 2020, arXiv, https://arxiv.org/abs/2006.10505v1.

Garfinkle, Alexandra. “Exclusive: How Walmart Is Using AI to Supercharge Its Holiday Plans.” Yahoo Finance, 31 Oct. 2023, https://finance.yahoo.com/news/exclusive-how-walmart-is-using-ai-to-supercharge-its-holiday-plans-190849969.html.

Garland, Max. “Amazon Touts AI Upgrades for Forecasting, Deliveries and Robotics.” Supply Chain Dive, 20 June 2025, https://www.supplychaindive.com/news/amazon-ai-supply-chain-usage-upgrades/75071/.

“Hilton (HLT) Q4 2025 Earnings Call Transcript.” The Globe and Mail, 2025, https://www.theglobeandmail.com/investing/markets/stocks/HLT-N/pressreleases/160824/hilton-hlt-q4-2025-earnings-call-transcript/.

Rachakatla, Sareen Kumar, et al. AI-Driven Business Analytics: Leveraging Deep Learning and Big Data for Predictive Insights. Journal of Deep Learning in Genomic Data Analysis, Dec. 2023.

SA News Team. “Amazon Unveils 3 AI-Driven Tools to Get Packages to Customers Faster.” Seeking Alpha, 11 June 2025, https://seekingalpha.com/news/4457344-amazon-unveils-3-ai-driven-tools-to-get-packages-to-customers-faster.

Transcribing, Motley Fool. “Hilton (HLT) Q4 2025 Earnings Call Transcript.” The Motley Fool, 11 Feb. 2026, https://www.fool.com/earnings/call-transcripts/2026/02/11/hilton-hlt-q4-2025-earnings-call-transcript/.

Vuocolo, Alex. “Walmart Taps Inventory Management System to Predict Demand Ahead of Holidays.” Retail Brew, 23 Oct. 2023, https://www.retailbrew.com/stories/2023/10/23/walmart-inventory-management-system-demand-holidays.

Yu, Yang, et al. “AI Adoption in S&P 500 Firms.” 2607.08920v1, 9 July 2026, arXiv, https://arxiv.org/abs/2607.08920v1.

**Does the Tax Code Push Workers Away from Independent Contracting?**
**FICA, SECA, §199A, and Transitions Between W-2 and 1099 Work By Nihit Yeruva**


## Abstract

Unincorporated independent contractors pay self-employment (SECA) tax at roughly twice the employee payroll (FICA) rate, a difference often described as a tax penalty on contracting. This paper asks whether that penalty exists in the economically relevant sense and whether it affects workers' choices between W-2 employment and 1099 work. Simulating 2024 federal liabilities shows that the answer to the first question depends on the comparison benchmark: at equal gross pay the contractor pays 4.8–5.6 percentage points more of income before the §199A deduction, but at equal total employer cost, the benchmark implied by standard incidence analysis, the payroll-tax structure is neutral to within 0.1 point and §199A tips the balance toward contracting by 1–4 points. For the second question, I link Current Population Survey ASEC records across adjacent years (2015–2026; 311,610 matched person-year pairs) and estimate transition models. About 1.9% of private W-2 workers become unincorporated contractors each year and 28.8% of unincorporated contractors return to W-2 work. Cross-sectional associations between the simulated tax differential and switching have the wrong sign and reflect income rather than taxes. A difference-in-differences design using the 2018 introduction of §199A, which lowered the contractor premium far more for middle and upper earners than for low earners, finds no increase in entry where the tax cut was larger and rules out responses larger than roughly a quarter of the baseline rate. By contrast, 72% of W-2 workers but only 39% of unincorporated contractors have employer-based health coverage. The evidence suggests that benefits, not taxes, are the binding cost of contracting, and that policy aimed at the independent workforce should focus on portable benefits rather than SECA reform.



**Keywords** self-employment, independent contracting, payroll tax, SECA, §199A, occupational choice, tax neutrality.


## 1. Introduction

A worker offered the same $80,000 as a W-2 employee or as a 1099 independent contractor faces different tax bills. The employee owes 7.65% of wages in Social Security and Medicare payroll taxes under the Federal Insurance Contributions Act (FICA), with the employer remitting a matching 7.65%. The contractor owes both halves, 15.3%, under the Self-Employment Contributions Act (SECA), softened by two adjustments: the tax applies to 92.35% of net profit, and half of it is deductible from income. Since 2018 the contractor may also deduct 20% of qualified business income under Internal Revenue Code §199A. Commentary on the growth of gig and contract work frequently treats SECA as a penalty that discourages self-employment and calls for "neutrality" between employment types.

This paper examines that claim in two steps. The first is conceptual: whether contracting is tax-penalized depends on what is held constant. Comparing a contractor earning $80,000 to an

employee earning $80,000 in wages ignores the employer's 7.65% contribution, which the incidence literature treats as borne by the worker through lower wages. Comparing instead to an employee whose total cost to the employer is $80,000 changes the answer. I formalize both benchmarks and simulate liabilities under 2024 federal rules across the income distribution.
The second step is empirical: whether the differential, under either benchmark, measurably affects transitions between employment types. The literature on self-employment and taxes has focused on entrepreneurial entry and the "success tax" created by progressive rates (Gentry and Hubbard 283), on non-pecuniary returns to self-employment (Hamilton 604), and on the measurement of alternative work arrangements (Katz and Krueger, "Rise and Nature" 382; Katz and Krueger, "Understanding Trends" 132). What is missing is a direct test of whether the FICA/SECA wedge, computed at the individual level, changes the probability that a wage worker becomes a contractor or a contractor returns to wage work. I provide one using matched Current Population Survey (CPS) Annual Social and Economic Supplement (ASEC) panels and a difference-in-differences design built around §199A, whose 2018 introduction reduced the contractor premium by an amount that varies sharply with income.

The findings are as follows. At equal gross pay, contractors pay 4.8–5.6 points more of income before §199A and 0.7–4.6 points more after it. At equal employer cost, the payroll-tax structure is neutral to within 0.1 point, and §199A gives contracting a 1–4 point advantage. In the data, 1.9% of private W-2 workers move to unincorporated contracting each year, a rate that is flat from 2014 to 2024. Cross-sectional regressions of switching on the tax differential yield positive coefficients, the opposite of the predicted sign, because the differential is a nearly deterministic function of income and lower earners switch more for unrelated reasons. The §199A difference-in-differences estimate is small, negative, and statistically indistinguishable from zero; its confidence interval excludes effects larger than about 0.5 percentage points at the maximum treatment intensity, roughly a quarter of the baseline rate. The one large difference between the two groups is employer-based health coverage: 72% of W-2 workers versus 39% of unincorporated contractors.
The paper proceeds as follows. Section 2 describes the tax institutions. Section 3 sets out a discrete-choice framework and the two neutrality benchmarks. Section 4 describes the data, Section 5 the empirical strategy, and Section 6 the results. Section 7 interprets the findings against competing explanations and draws policy implications. Section 8 discusses limitations and Section 9 concludes.

## 2. Institutional Background

**FICA.** Employees pay 6.2% of wages up to the Social Security wage base ($168,600 in 2024) plus 1.45% of all wages for Medicare, and an Additional Medicare Tax of 0.9% on wages above $200,000 (single filers). Employers remit a matching 6.2% and 1.45%, with no match on the 0.9% surtax.

**SECA.** Self-employed individuals pay 12.4% Social Security tax up to the wage base and 2.9% Medicare tax, both applied to 92.35% of net self-employment profit, plus the 0.9% surtax

above the threshold. The 92.35% factor equals $1 - 0.0765$ and exists to approximate the employee's wage net of the employer's share. One half of the Social Security and Medicare portion of SECA is deductible in computing adjusted gross income, mirroring the fact that the employer's share of FICA is not part of the employee's taxable wages.

**§199A.** The Tax Cuts and Jobs Act of 2017 created, for tax years 2018 through 2025, a deduction of 20% of qualified business income from pass-through businesses, including sole proprietorships. The deduction is capped at 20% of taxable income before the deduction. For taxpayers above a taxable-income threshold ($191,950 for single filers in 2024), the deduction is limited by W-2 wages paid and depreciable property; for a sole proprietor with no employees it phases out to zero over the next $50,000 of income. For specified service trades (health, law, accounting, consulting, financial services, performing arts) it phases out entirely above the threshold. W-2 wage income never qualifies.

**Benefits.** Employees commonly receive employer-sponsored health insurance (excluded from both income and payroll tax), retirement contributions, unemployment insurance, and workers' compensation. Contractors must purchase coverage individually (the premium is deductible from income but not from SECA), fund their own retirement, and are statutorily ineligible for unemployment insurance and workers' compensation.

## 3. Theoretical Framework

### 3.1 Worker utility and employment-type choice

Worker $i$ chooses an employment type $k$ from W-2 employment, incorporated self-employment, unincorporated self-employment, and non-participation, maximizing separable utility over consumption $c$ and leisure $l$ given a preference parameter $\theta_i$ that captures taste for autonomy, risk tolerance, and schedule flexibility:

$$U_i = u(c_i, l_i, \theta_i), \qquad c_i = y_i - T_k(y_i) - B_k, \qquad l_i = L - h_i, \qquad (1)$$

where $T_k$ is total tax liability under type $k$ and $B_k$ is the value of benefits forgone relative to the W-2 baseline ($B_W = 0$; $B_C > 0$ for contractors).

### 3.2 Tax liability functions

Let $\tau(\cdot)$ denote the federal income-tax schedule, $D$ the standard deduction, the wage base, and $m$ the Additional Medicare threshold. For a W-2 employee with wage $w$,

$$F_E(w) = 0.062 \min(w,) + 0.0145\, w + 0.009 \max(w - m, 0), \qquad (2)$$

$$F_R(w) = 0.062 \min(w,) + 0.0145\, w, \qquad (3)$$

$$T_W(w) = \tau(w - D) + F_E(w), \qquad (4)$$

where $F_E$ is the employee share and $F_R$ the employer shares. For an unincorporated contractor with net profit $y$,

$$SE(y) = 0.124 \min(0.9235y,) + 0.029(0.9235y) + 0.009 \max(0.9235y - m, 0), \qquad (5)$$

$$T_C(y) = \tau\big(y - D - \tfrac{1}{2} SE_0(y) - Q(y)\big) + SE(y), \qquad (6)$$

where $\frac{1}{2}SE_0$ is the deductible half of the Social Security and Medicare portion and $Q(y)$ is the §199A deduction described in Section 2, modeled for a non-service sole proprietor with no employees.

### 3.3 The tax differential and the incidence benchmark

The quantity of interest is the excess tax burden of contracting, $\Delta\tau(y) = T_C(y) - T_W(\cdot)$. Its sign and size depend on which W-2 earnings level is held fixed.

Under **Benchmark A (equal gross pay)**, the contractor with profit $y$ is compared to an employee with wage $w = y$:

$$\Delta\tau_A(y) = T_C(y) - T_W(y). \qquad (7)$$

This is the comparison a worker makes when offered "the same number" on a W-2 or a 1099, and it corresponds to statutory incidence.

Under **Benchmark B (equal employer cost)**, the employer's share of FICA is treated as borne by the worker through lower wages, as the incidence literature generally finds for payroll taxes (Rothstein 177). The comparable W-2 wage is the $w$ satisfying $w + F_R(w) = y$, and the worker's total burden includes both payroll shares:

$$\Delta\tau_B(y) = T_C(y) - \left[\tau(w - D) + F_E(w) + F_R(w)\right], \qquad w + F_R(w) = y. \qquad (8)$$

Below the wage base, $w = y/1.0765 \approx 0.9289y$, while SECA applies to $0.9235y$. The 92.35% factor and the half-SECA deduction were designed to make Benchmark B approximately neutral; whether they succeed is a quantitative question answered in Section 6.1. Tax policy is employment-type neutral at income $y$ under benchmark $b$ if and only if $\Delta\tau_b(y) = 0$.

### 3.4 Switching threshold

For the W-2 versus unincorporated-contracting margin, with $\theta$ expressed in dollars and hours approximately equal across types, a first-order expansion of the indifference condition gives a threshold rule: the worker switches to contracting if and only if

$$\theta_i \geq \theta^*(y) \equiv \Delta\tau(y) + B_C - \Delta h(y)\,\frac{\partial u/\partial l}{\partial u/\partial c}. \qquad (9)$$

The threshold rises with the tax differential and the benefit gap and falls with any flexibility premium $\Delta h(y)$. The testable prediction is that a larger $\Delta\tau$ lowers the probability of switching into contracting and raises the probability of switching out, holding income and other determinants fixed.

### 3.5 Transition model

Letting indirect utility include an i.i.d. Type-I extreme value shock, the probability that a worker of type $k$ in year $t$ is of type $j$ in year $t+1$ takes the McFadden conditional-logit form (105–42). For the binary entry margin studied here,

$$\Pr(\text{SE}_{i,t+1} = 1 \mid \text{W-2}_{it}) = \Lambda\big(\alpha + \beta\,\Delta\tau_{it} + \gamma' X_{it} + \lambda_t + \mu_s\big), \qquad (10)$$

with $X_{it}$ individual controls, $\lambda_t$ year effects, $\mu_s$ state effects, and $\Lambda$ the logistic function. Equation 9 predicts $\beta < 0$. Because the annual entry rate is small, marginal effects are reported both in percentage points and relative to the baseline rate.

## 4. Data

### 4.1 CPS ASEC matched panel

The Current Population Survey is a monthly survey of about 60,000 households conducted by the Census Bureau and the Bureau of Labor Statistics. Its Annual Social and Economic Supplement, fielded each March, records income by source and class of worker for the prior calendar year. Because of the CPS's 4–8–4 rotation design, roughly half of ASEC respondents in year $t$ are interviewed again in the ASEC of year $t+1$. I use the harmonized IPUMS CPS extract of ASEC samples 2015–2026 for persons aged 18–64 (Flood et al.) and link individuals across adjacent years using the validated person identifier CPSIDV, retaining a link only if reported age advances by zero to two years and sex is unchanged. This yields 311,610 validated person-year pairs from 1,160,423 person-years.

Employment type is taken from CLASSWLY, the class of worker on the longest job held in the prior calendar year, which shares its reference period with the income variables. I collapsed it to five states: private W-2 employee, government employee, incorporated self-employed, unincorporated self-employed, and not working. Because CLASSWLY describes the prior calendar year, a pair of ASEC interviews in years $(t, t+1)$ measures a transition between calendar years $t-1$ and $t$. Prior-year wage and salary income (INCWAGE) and net self-employment income (INCBUS plus INCFARM) are used to compute tax differentials. Employer-based group health coverage in the prior year (GRPCOVLY) is available for the 2019–2026 samples. All statistics are weighted by the ASEC person weight.

### 4.2 Simulated tax differentials

For each private W-2 worker I compute $\Delta\tau_A$ and $\Delta\tau_B$ from Equations 2–8 at the worker's prior-year wage, converted to 2024 dollars using annual CPI-U averages, treating the worker as a single filer taking the standard deduction. The §199A deduction is included for destination years 2018 and later. Both differentials are expressed in percentage points of income. The §199A treatment intensity used in Section 5.2 is the reduction in $\Delta\tau_A$ produced by the deduction at the worker's income. The regression sample consists of private W-2 workers with prior-year wages of at least $5,000 and at least one week worked ($N = 174{,}855$).

## 5. Empirical Strategy

### 5.1 Transition logit

I estimate Equation 10 by weighted maximum likelihood with year, state, major industry (1990 Census groups), and major occupation (2010 Census groups) fixed effects, and controls for age, age squared, and sex, with and without log income. Standard errors are clustered by state. The same specification is estimated on the return margin, with unincorporated contractors as the origin state and return to private W-2 employment as the outcome.

The identification problem with this design should be stated at the outset. Within a year, $\Delta\tau(y)$ is a deterministic function of income with only a handful of kinks (Figure 1). Once income is controlled flexibly, $\beta$ is identified only from the kinks, and any omitted determinant of switching that varies with income will load onto it. The cross-sectional estimates are therefore reported as descriptive, and causal inference rests on the design in the next subsection.

### 5.2 §199A difference-in-differences

The 2018 introduction of §199A reduced $\Delta\tau_A$ by $\tau'(y)\,Q(y)$: zero for W-2 earners, rising with the marginal income-tax rate, and phasing out above the threshold for sole proprietors without employees. This produces treatment intensity that varies with income: in the regression sample it averages 0.01 points in the lowest income decile, 1.5 points in the fifth, 3.6 points in the ninth, and 3.0 points in the top decile where the phase-out begins to bind. If the tax differential affects switching, entry into contracting should have risen after 2018 by more where the intensity was larger. I estimate

$$Y_{ist} = \alpha + \beta\,(\text{Post}_t \times \text{Intensity}_i) + \delta\,\text{Intensity}_i + \gamma' X_{ist} + \lambda_t + \mu_s + \varepsilon_{ist}, \qquad (11)$$

where $Y_{ist}$ indicates a transition from private W-2 employment to unincorporated self-employment, $\text{Post}_t = 1$ when the destination calendar year is 2018 or later, and $X$ includes income-decile, industry, occupation, age, and sex controls. A second specification adds income-decile-by-post interactions so that $\beta$ is identified from the within-decile variation in intensity, and a third restricts the sample to destination years through 2019 to exclude the pandemic. The predicted sign is $\beta > 0$. Parallel pre-trends are assessed with an event-study version that interacts intensity with each destination year, 2017 as the reference. Estimation is by weighted least squares with state-clustered standard errors.

## 6. Results

### 6.1 Is the tax code neutral? Simulation results

Figure 1 and Table 1 report simulated 2024 liabilities. Under Benchmark A the contractor premium is 5.6 points of income at $30,000 and $50,000, falling to 4.8 points near the wage base, before §199A; the premium shrinks as the marginal rate rises because the deductions become more valuable. Including §199A, the premium falls to 4.6 points at $30,000, 1.6 points at $80,000, and 0.7 points at the wage base.

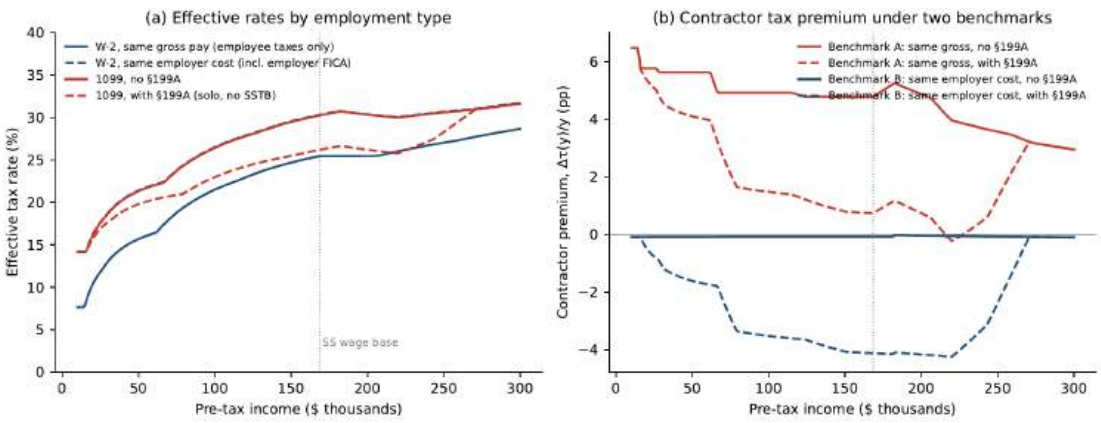



Figure 1. Simulated effective tax rates (a) and contractor premium (b) by employment type, single filer, 2024 federal rules. Benchmark A compares a contractor to an employee with the same gross pay; Benchmark B to an employee with the same total employer cost. The §199A lines assume a non-service sole proprietor with no employees. State taxes, the self-employed health insurance deduction, and Schedule C expenses are excluded.

Table 1. Simulated 2024 federal liabilities and contractor premium, single filer. Premium in percentage points of income.

| Income | W-2, same gross ($) | W-2, same cost ($) | 1099, no §199A ($) | 1099, §199A ($) | Prem. A, no §199A | Prem. A, §199A | Prem. B, no §199A | Prem. B, §199A |
|---|---|---|---|---|---|---|---|---|
| $30,000 | 3,911 | 5,624 | 5,601 | 5,301 | +5.6 | +4.6 | −0.1 | −1.1 |
| $50,000 | 7,841 | 10,696 | 10,657 | 9,892 | +5.6 | +4.1 | −0.1 | −1.6 |
| $80,000 | 15,561 | 19,560 | 19,501 | 16,872 | +4.9 | +1.6 | −0.1 | −3.4 |
| $120,000 | 27,518 | 33,420 | 33,331 | 29,067 | +4.8 | +1.3 | −0.1 | −3.6 |
| $168,600 | 42,900 | 51,090 | 50,966 | 44,146 | +4.8 | +0.7 | −0.1 | −4.1 |
| $250,000 | 67,543 | 76,653 | 76,478 | 70,705 | +3.6 | +1.3 | −0.1 | −2.4 |

Under Benchmark B the picture reverses. Before §199A, the contractor's total burden is within 0.1 point of the employee's at every income level: the 92.35% base and the half-SECA deduction do what they were designed to do, and the "double" payroll tax on contractors is simply the employer's share made visible. Including §199A, contracting is *favored* by 1.1 points at $30,000, 3.4 points at $80,000, and 4.1 points at the wage base, with the advantage disappearing only once the deduction phases out above about $240,000.

The commonly stated proposition that the tax code penalizes self-employment is therefore true only under statutory incidence and only before §199A. Under economic incidence the code was approximately neutral before 2018 and has favored unincorporated contracting since. A full-deductibility reform for SECA, often proposed as a neutrality fix, closes only 15–35% of the Benchmark A gap in these simulations and widens the Benchmark B advantage.

### 6.2 Transition patterns

Table 2 reports the annual transition matrix. Private W-2 employment is highly persistent (84.5% remain), while unincorporated self-employment is not: 51.4% remain, 28.8% return to private W-2 work, and 8.2% incorporate. The W-2-to-unincorporated-SE rate is 1.9% per year. Among workers, 77.3% are private W-2 employees, 14.2% government employees, 3.3% incorporated self-employed, and 5.2% unincorporated self-employed.

Table 2. Annual employment-type transition probabilities (%), CPS ASEC 2015–2026 matched panel, ages 18–64, ASEC-weighted. Rows: type in year $t$; columns: type in year $t+1$.

| | W-2 private | Government | SE inc. | SE uninc. | Not working | N (unwtd.) |
|---|---|---|---|---|---|---|
| W-2 private | 84.5 | 4.9 | 1.7 | 1.9 | 6.9 | 183,274 |
| Government | 25.7 | 67.3 | 0.8 | 0.8 | 5.4 | 39,143 |
| SE incorporated | 39.3 | 3.3 | 40.3 | 12.9 | 4.2 | 9,087 |
| SE unincorporated | 28.8 | 1.9 | 8.2 | 51.4 | 9.8 | 13,874 |
| Not working | 17.5 | 2.4 | 0.5 | 1.9 | 77.8 | 66,232 |

Figure 2 plots the entry rate by calendar year. It is flat at 1.8–2.2% throughout, with no visible break in 2018. This is a raw series, not a causal estimate, but it indicates that any aggregate effect of §199A on entry was small.

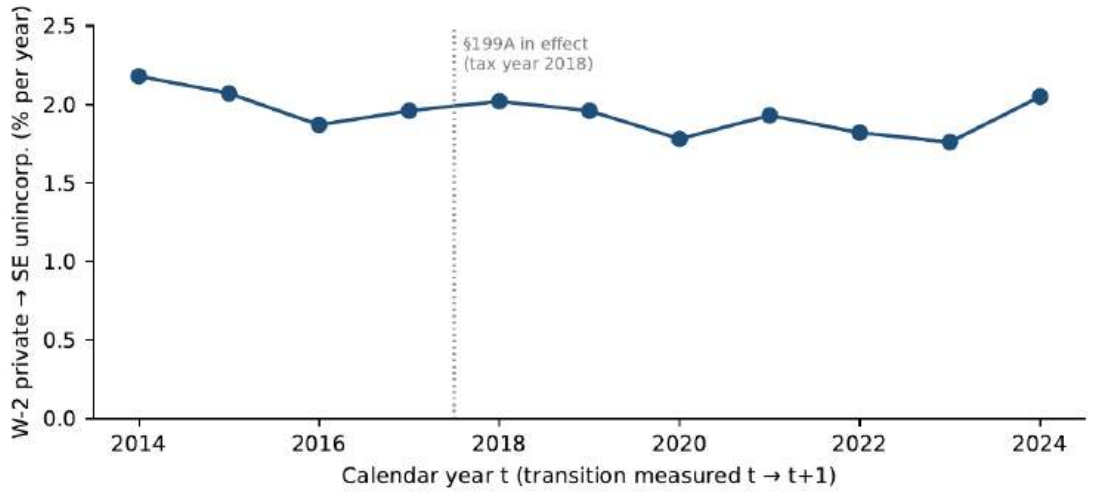



Figure 2. Private W-2 to unincorporated self-employment transition rate by calendar year, CPS ASEC matched panel, ASEC-weighted.

### 6.3 The benefits gap

Among workers aged 18–64 in the 2019–2026 samples, the share with employer-based group health coverage in the prior year (as policyholder or dependent) is 83.7% for government employees, 71.8% for private W-2 employees, 53.8% for the incorporated self-employed, and 39.4% for the unincorporated self-employed. The 32-point gap between private W-2 workers and unincorporated contractors is the empirical counterpart of $B_C$ in Equation 1. Because the measure counts dependent coverage, the gap in access through one's *own* job is larger.

### 6.4 Transition logit

Table 3 reports the coefficient on the simulated contractor premium from Equation 10. Under both benchmarks and with or without a log-income control, the coefficient is positive and statistically significant: a one-point *larger* tax penalty on contracting is associated with a 0.16 -- 0.31 point *higher* probability of entering it, 8–16% of the 1.92% base. This is the opposite of the sign predicted by Equation 9.

Table 3. Weighted logit estimates, probability of transition from private W-2 employment to unincorporated self-employment (entry) and from unincorporated self-employment to private W-2 employment (return). All models include year, state, industry, and occupation fixed effects and controls for age, age squared, and sex. Standard errors clustered by state.

| Measure | Income control | β̂ | SE | AME (pp per pp) | AME, % of base |
|---|---|---|---|---|---|
| *Entry margin* (N = 174,855; base 1.92%) | | | | | |
| Δτ_A (same gross) | none | 0.140 | 0.016 | 0.26 | 13.6 |
| Δτ_A (same gross) | log income | 0.085 | 0.023 | 0.16 | 8.3 |
| Δτ_B (same employer cost) | none | 0.168 | 0.017 | 0.31 | 16.3 |
| Δτ_B (same employer cost) | log income | 0.103 | 0.023 | 0.19 | 10.0 |
| *Return margin* (N = 12,220; base 29.1%) | | | | | |
| Δτ_A (same gross) | log income | 0.006 | 0.041 | 0.13 | 0.4 |

The positive sign should not be read as a behavioral response. Figure 1(b) shows that within a year the premium is highest at low incomes and declines with income, and it is well documented that lower-wage workers move into self-employment more frequently than higher-wage workers. The premium is therefore proxying for low income even after a log-income control, which captures only the linear component. The cross-sectional design cannot separate the tax wedge from the other correlates of income, which is why the causal evidence rests on the §199A design. On the return margin, where the contractor's own choice is more plausibly primary, the coefficient is a precisely estimated zero.

### 6.5 §199A difference-in-differences

Table 4 reports the Post × Intensity coefficient from Equation 11. The predicted sign is positive. The estimates are negative in all three specifications, small, and in the preferred specification with income-decile-by-post interactions indistinguishable from zero ($-0.078$, SE $0.115$).

Table 4. §199A difference-in-differences: effect of a one-point reduction in the contractor premium on the probability of transition from private W-2 employment to unincorporated self-employment, in percentage points. Weighted least squares with income-decile, year, state, industry, and occupation fixed effects; standard errors clustered by state.

| Specification | Coefficient | SE | *N* |
|---|---|---|---|
| Baseline | −0.145 | 0.062 | 174,855 |
| Income-decile × post interactions | −0.078 | 0.115 | 174,855 |
| Pre-pandemic only (destination years ≤ 2019) | −0.117 | 0.085 | 89,113 |

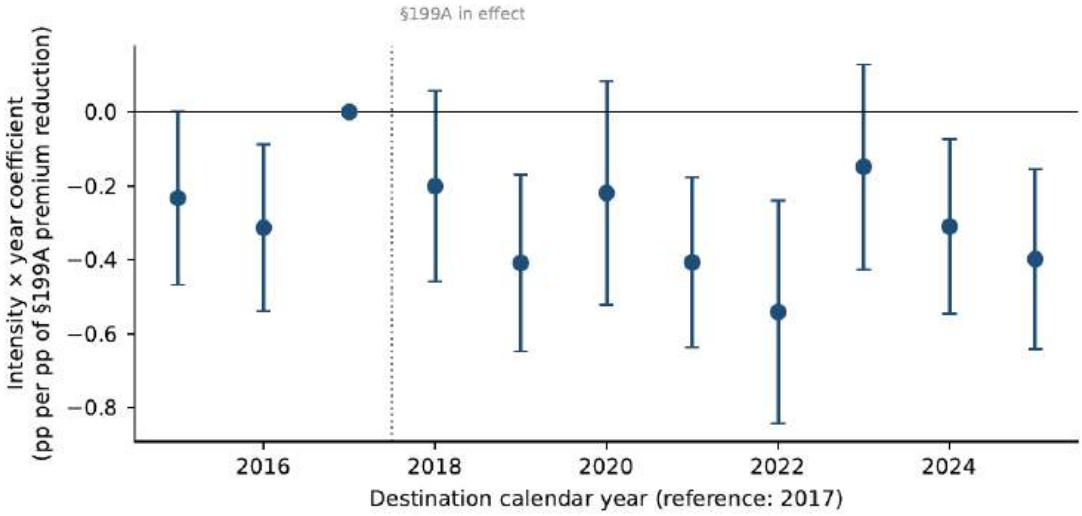



Figure 3. Event study: coefficients on Intensity × destination year (reference year 2017) with 95% confidence intervals, state-clustered.

The event study in Figure 3 shows why neither the sign nor the magnitude of the point estimates should be over-read. The coefficients for 2015 and 2016 are negative and of similar size to those after 2018, so the parallel-trends condition relative to 2017 is not cleanly satisfied; 2017 appears to be an unusually high year for high-intensity workers rather than 2018 an unusually low one. The informative content of the exercise is the upper bound. The 95% confidence interval of the preferred specification extends to about $+0.15$ points per point of intensity. At the maximum intensity of roughly 3.6 points, that implies §199A raised annual entry into contracting by at most about 0.5 percentage points on a 1.9% base. A response of the

size that would make SECA reform a first-order policy lever, say a doubling of entry among the workers who received the largest tax cut, is ruled out.

## 7. Discussion

### 7.1 What the evidence supports

Taken together, the simulation and the panel evidence point the same way. Under the economically relevant benchmark the tax code does not penalize unincorporated contracting and since 2018 has favored it, and a tax change that moved the contractor premium by up to 3.6 points of income produced no detectable movement into contracting. The threshold rule in Equation 9 says that switching depends on $\Delta\tau$, $B_C$, and preferences; the data suggest the first term is small and inert and the second is large.

Three alternative explanations for a null tax response, each with a different policy implication, are consistent with these results.

**Preferences dominate.** If workers choose employment type for autonomy, flexibility, or task variety, a few points of income will not move them (Hamilton documents median self-employment earnings well below comparable wages, implying large non-pecuniary returns, 604–31). The low persistence of unincorporated self-employment, with 28.8% returning to W-2 work each year, suggests considerable experimentation consistent with learning about one's own taste.

**Classification is employer-driven.** If employers decide who is a contractor, the tax position of the worker is irrelevant to the transition. The null result on the entry margin, where employers plausibly initiate the change, and the equally null result on the return margin, where the worker's own optimization is more plausibly primary, do not distinguish this from the preference story. A classification shock such as California's AB5 (2020), which changed who may legally be a contractor without changing the tax differential, would be the natural test.

**Benefits are the binding constraint.** The 32-point gap in employer health coverage dwarfs the tax differential under either benchmark. For a worker earning $60,000, a family health plan purchased individually can cost a quarter of income before subsidies, against a Benchmark A tax premium of about 5 points and a Benchmark B premium near zero. If this is the operative margin, interacting the tax differential with alternative insurance access (a spouse's plan, Medicaid, the ACA marketplace) should show responses only where coverage is not at stake, and the 2014 Medicaid expansion provides a quasi-experiment.

### 7.2 Connection to optimal tax theory

Saez shows that when labor supply responds at the extensive margin, optimal tax rates at a participation threshold depend inversely on the participation elasticity (1039–73). Applied to the employment-type margin, the relevant parameter is the switching elasticity between W-2 and contractor work. The estimates here imply that elasticity is small, which by the same logic means the efficiency cost of whatever differential exists is also small, and the Harberger

deadweight-loss approximation (25–80), which scales with the elasticity and the square of the wedge, is correspondingly negligible. The welfare argument for SECA "neutrality" reform is therefore weak on both counts: the wedge is near zero under economic incidence, and behavior does not respond to it.

### 7.3 Policy implications

The analysis suggests three conclusions for policy. First, proposals to make SECA fully deductible, or otherwise to "equalize" payroll taxation across employment types, are solving a problem that largely does not exist under economic incidence and that §199A has already overcorrected for sole proprietors; full deductibility closes a minority of the statutory gap while widening the economic one. Second, the question of whether §199A should be extended beyond 2025 is in part a question about employment-type neutrality: its expiration would return the code to approximate Benchmark B neutrality, and its extension preserves a 1–4 point advantage for unincorporated contracting that this paper finds workers do not act on. Third, if the goal is to make independent work viable for those who want it, the lever is the benefits gap: portable health, retirement, and unemployment coverage that does not depend on classification. Shifting attention from the tax differential to the benefits differential is the paper's main practical message.

## 8. Limitations

Four limitations should be weighed. First, tax differentials are computed under 2024 federal rules for all years, with incomes converted to 2024 dollars. Pre-2018 law had a lower standard deduction, personal exemptions, and different bracket thresholds, so the cross-sectional *levels* of $\Delta\tau$ before 2018 are approximate. This does not affect the difference-in-differences design, whose identifying variation is the income gradient of the §199A deduction, a feature of the law from 2018 onward that is captured correctly; and because the cross-sectional estimates are already discounted as non-causal, approximation error in their regressor does not change the paper's conclusions. Second, all workers are treated as single filers taking the standard deduction. Joint filers face different thresholds and the §199A phase-out begins at a higher income, which would shift intensity for some high earners but not the pattern that low earners received essentially nothing. Third, the CPS class-of-worker item is self-reported and is known to classify many 1099 arrangements as wage jobs, particularly among platform workers; this misclassification attenuates estimated responses toward zero and is a reason to interpret the upper bound of the confidence interval, rather than the point estimate, as the paper's finding. Fourth, the panel is two years long, so responses that take longer than a year to materialize, which Chetty argues are common in the presence of adjustment frictions (969–1018), are not captured. Administrative tax data linking Form W-2 and Schedule C filings over longer horizons would address the third and fourth limitations directly.

## 9. Conclusion

The claim that the U.S. tax code penalizes independent contracting through the FICA/SECA asymmetry is true only when the employer's share of payroll tax is ignored. Once that share is attributed to the worker, as incidence analysis implies, the payroll-tax structure is neutral to within a tenth of a percentage point of income, and the §199A deduction has since 2018 tilted the balance toward contracting by one to four points for most sole proprietors. Matched CPS panels show that about 1.9% of private wage workers become unincorporated contractors each year, that this rate did not change when §199A arrived, and that the workers who received the largest reduction in the contractor premium did not enter contracting at higher rates than those who received almost none. Cross-sectional correlations between the tax differential and switching are driven by income, not taxes. What does differ sharply between the two forms of work is access to employer-provided benefits, with a 32-point gap in health coverage. The evidence favors portable benefits, not payroll-tax reform, as the policy response to the independent workforce.

**Acknowledgments**

I thank Professor Laurence Ales of Carnegie Mellon University's Tepper School of Business for guidance on topic selection and research design. All analysis and writing are my own, as are any errors. Replication code for the tax simulator and all estimates are available on request.

**Works Cited**


Chetty, Raj. “Bounds on Elasticities with Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply.” Econometrica, vol. 80, no. 3, 2012, pp. 969–1018.

Flood, Sarah, et al. IPUMS CPS: Version 12.0 [dataset]. IPUMS, 2024, https://doi.org/10.18128/D030.V12.0.

Gentry, William M., and R. Glenn Hubbard. “Tax Policy and Entrepreneurial Entry.” American Economic Review, vol. 90, no. 2, 2000, pp. 283–287.

Hamilton, Barton H. “Does Entrepreneurship Pay? An Empirical Analysis of the Returns to Self-Employment.” Journal of Political Economy, vol. 108, no. 3, 2000, pp. 604–631.

Harberger, Arnold C. “Taxation, Resource Allocation, and Welfare.” The Role of Direct and Indirect Taxes in the Federal Revenue System, Princeton UP, 1964, pp. 25–80.

Katz, Lawrence F., and Alan B. Krueger. “The Rise and Nature of Alternative Work Arrangements in the United States, 1995–2015.” ILR Review, vol. 72, no. 2, 2019, pp. 382–416.

Katz, Lawrence F., and Alan B. Krueger. “Understanding Trends in Alternative Work Arrangements in the United States.” RSF: The Russell Sage Foundation Journal of the Social Sciences, vol. 5, no. 5, 2019, pp. 132–146.

McFadden, Daniel. “Conditional Logit Analysis of Qualitative Choice Behavior.” Frontiers in Econometrics, edited by Paul Zarembka, Academic Press, 1974, pp. 105–142.

Rothstein, Jesse. “Is the EITC as Good as an NIT? Targeting, Design, and Distributional Effects.” American Economic Journal: Economic Policy, vol. 2, no. 1, 2010, pp. 177–208.

Saez, Emmanuel. “Optimal Income Transfer Programs: Intensive Versus Extensive Labor Supply Responses.” Quarterly Journal of Economics, vol. 117, no. 3, 2002, pp. 1039–1073.

U.S. Internal Revenue Service. Publication 334: Tax Guide for Small Business, 2024. Instructions for Schedule SE (Form 1040), 2024. Instructions for Form 8995: Qualified Business Income Deduction, 2024.

**The Effects of AI in the Stock Market and How Machine Learning Models are Applied to Predict Future Prices By Daniel Hong**


## Abstract

The increasing use of artificial intelligence in the stock markets has created new opportunities for predicting stock-prices using historical data. This study looks into the effectiveness of artificial intelligence in predicting future stock prices. It also details whether increasing complexity leads to improved accuracy. Numerical stock-market data, including variables such as opening price and closing price, were gathered and used to train models. The models started with a simple methodology all the way to a retaining system made to look at data sets as a whole. Model performance was evaluated using mean squared error (MSE), inverse mean squared error (IMSE), and mean absolute error (MAE). These predictions and models were tested in a simulated trading system to examine their potential in application. The results demonstrate differences in predictive performance among models, as seen when predicting stocks such as KO, and highlight challenges associated with actually predicting the stock market. Overall, this research weighs both the potentials and limitations of artificial intelligence as a tool for stock market predictions and provides a foundation for further research when incorporating additional economic variables.


## Introduction

The stock market is an open market in which anyone can buy and sell ownership of a public company. These portions of ownership within the company are known as shares or stocks. In the context of the global economic system, the stock market is roughly 100%-120% of the entire globe's GDP. This makes the stock market highly profitable, or extraordinarily loss-making. The reason prices move is because buyers and sellers constantly disagree about what companies are worth. institutions, economic conditions, interest rates, news, and expectations about the future all affect pricing.

Depending on how you use the stock market, anything can happen. Having a strategy is important but understanding how the stock market works can make you money in the long run. The stock market is almost if not entirely based on mathematical patterns dealing with statistics and probability. When looking at prices, we can see fluctuations and bull and bear seasons all explained through the concept of a distribution within a distribution. A distribution shows where the general market may go as it typically stays in the middle; however, when one end of the prices whether up or down has a surge, it pushes it constantly to the extremes in which it can go purely because of its backing of probability. When it has "run out of probability" to keep going in a certain direction, it is only a matter of time until it reverses its trends and heads the other way because it cannot either go down anymore or go up. The distribution within a distribution part comes in when there are smaller fluctuations of where it is going. The microsecond changes that happen are due to probability fluctuating on the distribution within a distribution. Although

this is all probability and patterns, there are still outside effects such as the economic conditions and news that people follow to affect the market.

AI in previous years has started to become well known for finding patterns. It has a system of math that even humans cannot understand anymore to make decisions and produce outputs that are capable of answering the input's prompt. If the input prompt were to predict the next day's stock price, it would use pattern recognition to estimate how much the price will change. Given that the stock market is all patterns, AI would be extraordinarily useful in making predictions and trading. A trained bot could easily make close predictions for the next day leading to potential profit. As seen with the research paper written by Fatima Dakalbab, the study reviews 143 research papers on the impacts of AI trading finding that deep leaning is the most commonly used AI approach when testing AI in the stock markets, and that people use technical analysis. Another research paper by Oliveira reviews 2,326 papers over a twenty year time period on AI in the stock market and then studies these papers in categories such as portfolio optimization, stock prediction, financial sentiment analysis, etc. The study then details the findings that AI-based financial research has increased over the years. Ozdemir, another researcher, reviews 54 empirical studies to examine the fundamental indicators and different machine-learning models are used for stock prediction. He specifies the usage of LSTMs when predicting prices, Random Forests with return predictions, and Transformers and LLMs being on the rise of usage.

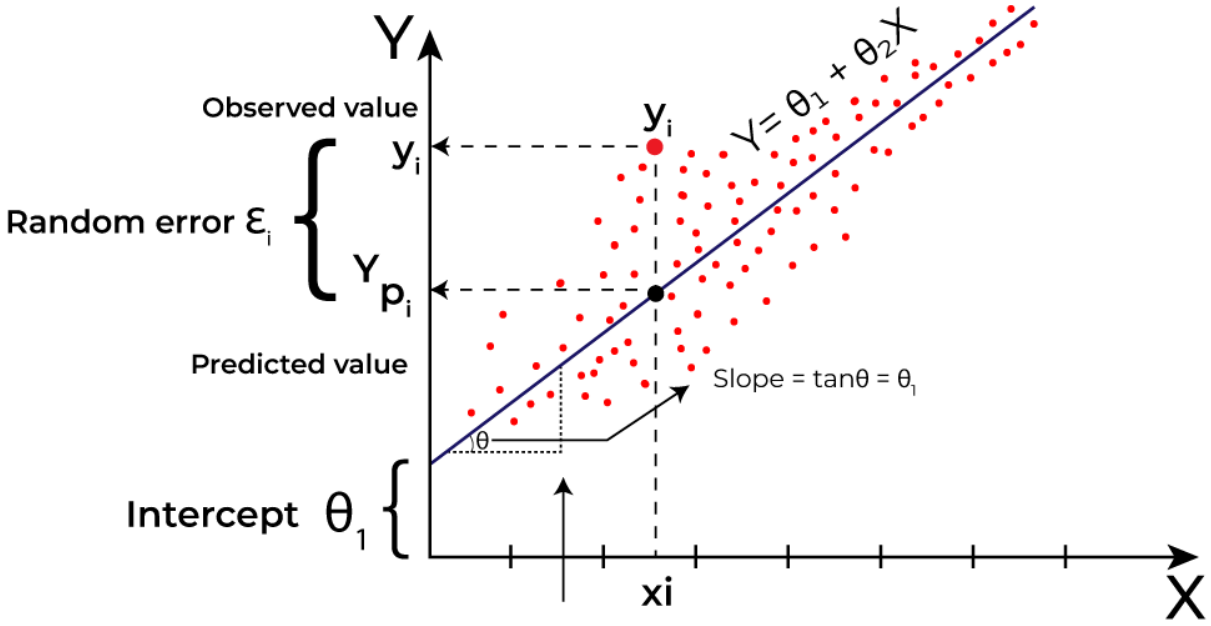
**Materials and Methods**

To train an AI model, you first need data to back it up. Yahoo Finance is a free data center that provides prices and data history over a stock's course of time. To get results that are good for using, we picked a 5 year data set of the Tesla (TSLA), Apple (AAPL), and Google (GOOG), Nasdaq (QQQ) to train my model. This data set consists of thousands of prices and patterns so a model could train and produce accurate results. The data set can be changed at any time and is crucial for being able to find large patterns that consistently line up with each other. This allows for predictions to be accurate no matter what stock you choose to predict with.

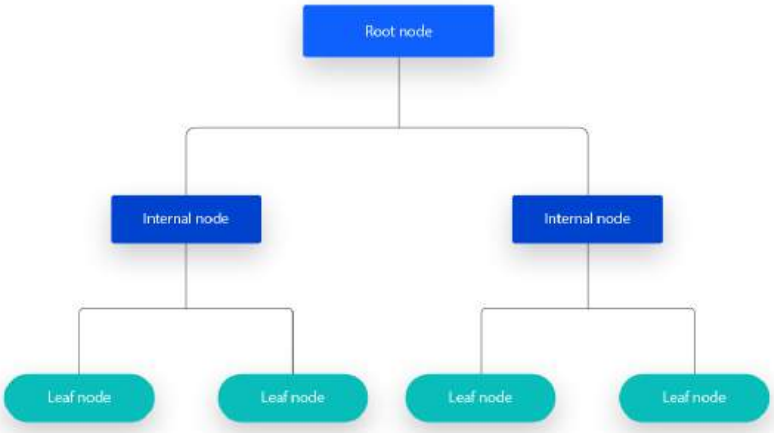
When one is training and testing a model, they have to use the different data sets so the model can make answers on its own. Rather than already knowing the answers from the training set, you must have separate data for testing. When using more complex models that run risks such as overfitting, you may also have to split the data into a validation set as well. If you do not split the data set, the model will already have the actual results of the day that is selected. This means it will not actually be predicting the price on its own. After we split the data into two thirds training and one third testing, we began to train different models based on this data split.

The first model we trained and tested was the most simple model out there, a linear regression. This turned the prices and the days and mapped them onto a coordinate plane. It then found the line of best fit and gave results based on that equation. Taking the line of best fit, it had found short term trends in the stock market. This is useful for finding different lines of equilibrium that the stock market would revolve around. By using this model, the stock market

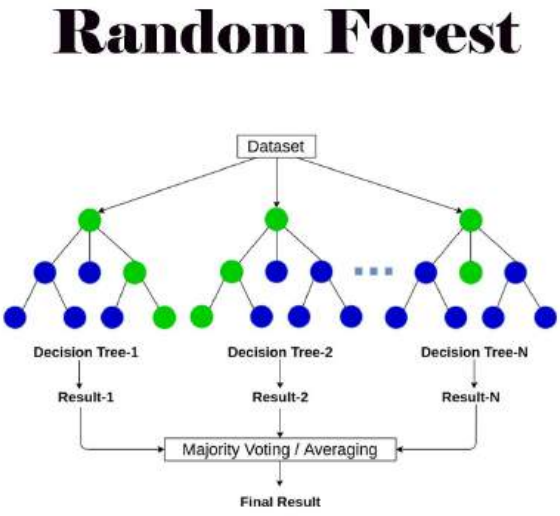
was able to be predicted over specific periods of time. The biggest drawback with this method is that it is unpredictable with larger data sets and longer periods of time. We used this for predicted short term trends; although, it was not the most efficient compared to other models.



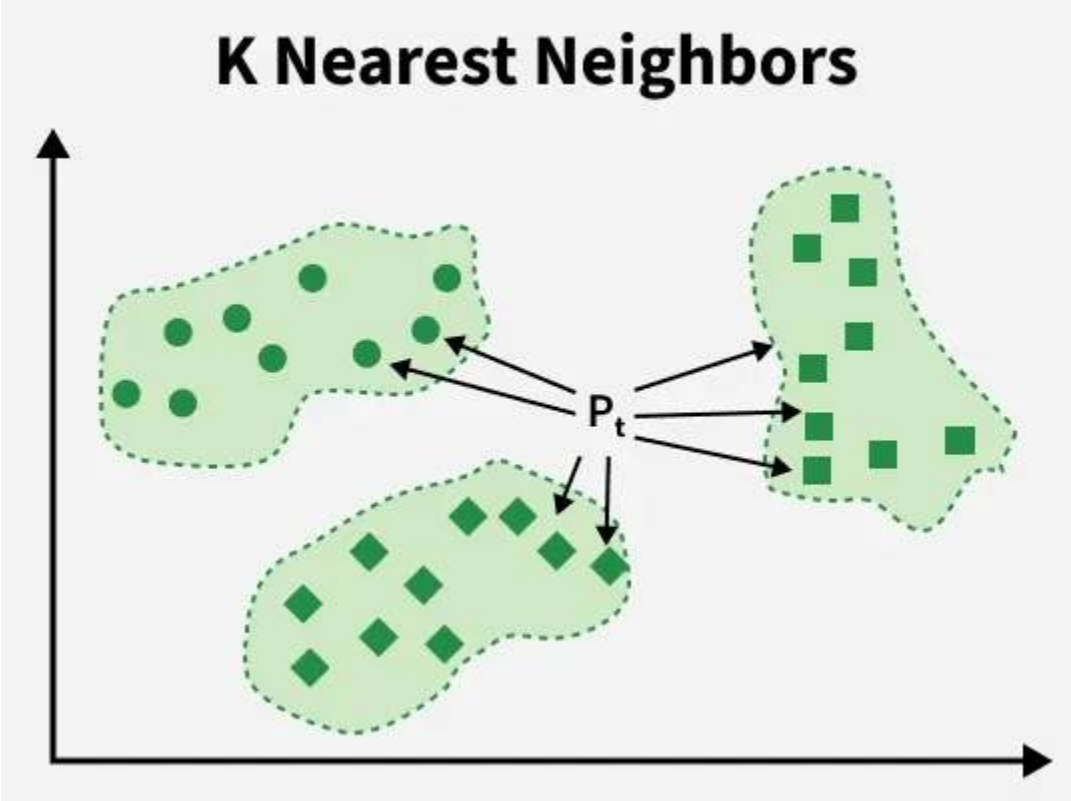

The second model we trained and tested was the decision tree. A decision tree makes multiple nodes and decisions and evaluates each decision. After it goes through the best decisions and makes a prediction based on that. It weighs the options and makes its optimal path with this. On paper, this model should be very accurate in finding the most optimal path to get to the price for the next day. However, the one major flaw that it could not account for were the random external effects that it could not predict and adjust for. Because the stock market also has external economic factors affecting it, the decision tree can not predict these. The model did not have enough information to weigh the economic factors to give paths accurate weights. This led to this model overall being extraordinarily inaccurate and unusable in the final predictor.



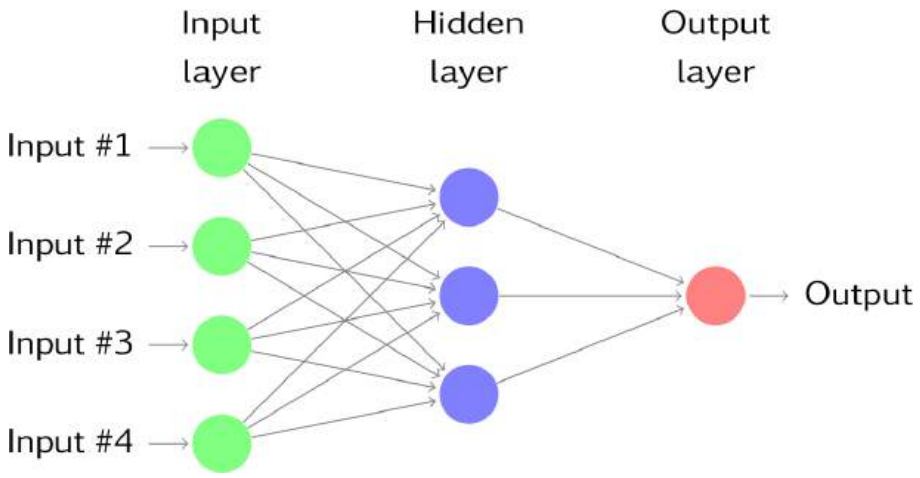

The next model that was coded was the random forest model. This is a collection of decision trees that then leads to a single output. It puts multiple weights on each decision tree with separate weights on their most optimal paths for predictions. This is able to take into account more factors, yet it ran into the same problem as before. Because of its inability to make predictions with external effects, the decision trees in the random forest were not accurate enough to lead to a close final prediction. Random forest also had decision trees that were set in a single way of thinking without considering different types of trends in different long term scenarios. In a bull and bear market, this model would not be useful as it was not given enough data to train on both long term market scenarios. The decision tree in the random forest also always goes with the most optimal method rather than considering other factors that may lead to different paths being more optimal in different situations.



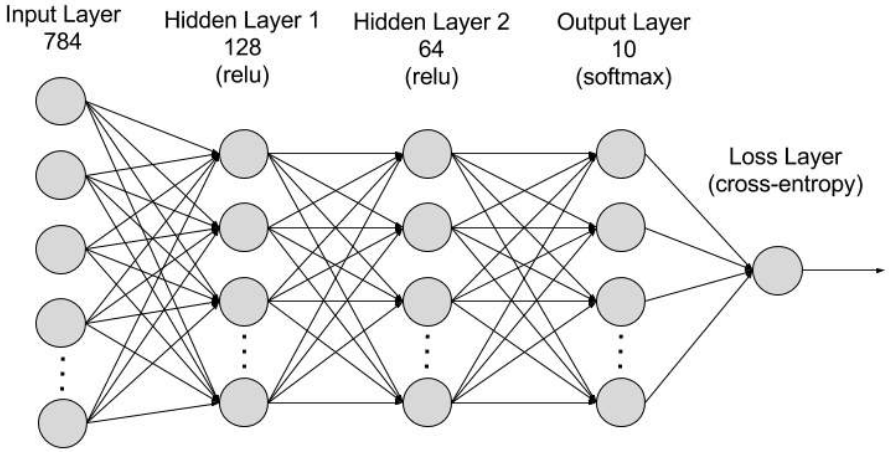
Going back to a graphing system, we trained a K-Nearest Neighbor model. This used numerous dimensions to plot different aspects of a data point. In this case, it would categorize things such as increase or decrease and the price itself to predict where the next point will be plotted. This prediction method was significantly better than the previous models as it accounted for more details and patterns the stock market has. The KNN model was locating specific patterns within how the stock market trends run while balancing the complexity of the different patterns. It was great for categorizing different trends and predicting what the next price will be, especially when it can triangulate the next point with the past three days of data.



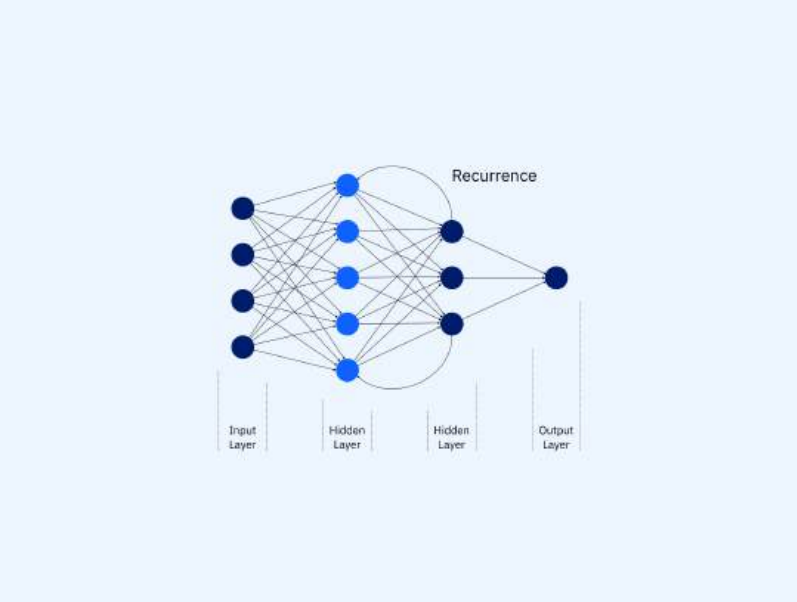

We experimented with a neural network system after this. A neural network uses hidden layers to encode and process the user's input. In the neural network, we implemented a small random state that was able to ensure it did not go to the same problems as the decision tree. This made it so that some parts were randomized so that it was not giving only linear predictions that would not work in the constantly changing stock market. The neural network was only semi-accurate. These hidden layers connected together and accounted for a large amount of data and stock market trends. Each extra layer added more possibilities, yet also ran the risk of overfitting which ultimately happened. This was useful for taking outside factors as its random state made it have predictions that could be similar to the effects of an unpredictable event.



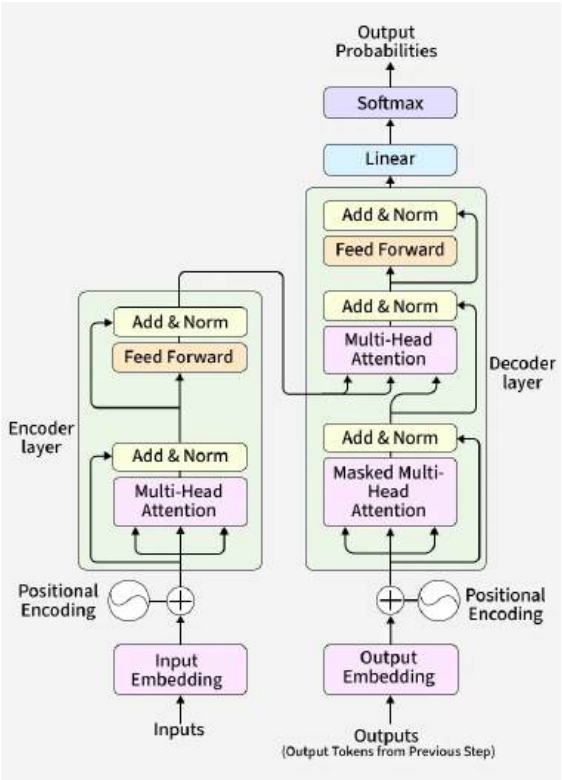

A more complex model that was tested is a Recurrent Neural Network (RNN). This model was able to retain past information and use it in its predictions. This became very useful for making predictions based on the current status of the stock market over a long term outlook. This came with the drawback of overfitting. Other than that, this model had a very good key concept that would become important later to making accurate predictions.




One issue with the RNN model was that it considered all previous data given. An LSTM fixes this by only using data that is relevant. It uses recent data and only lets certain parts of data into its consideration. It still uses data in a sequential order, but it is more short term than the RNN. This allows for relevant data and closer predictions in the model.




Recently, transformers have been changing how AIs are made and run on. We had coded a transformer that was also able to take data and interpret it based on the surrounding context. If a price was down by a significant margin for a long period of time, it would be able to interpret that there is a bear market leading to predictions on what the price would be. This ran into the same problem of overfitting and being unable to accurately determine what the price should be based on context.



To test the model's combined form, we made a simulation bot using different weights on the models to play in the stock market's past data. We coded the simulation bot with a simple trading strategy and used the last four years of available Yahoo Finance data. The simulation bot traded every single day on one stock.

**Overfitting**

To fix the overfitting on both the RNN and the LSTM, we limited the network that we were using to make sure it did not include too much data that was not as relevant. We then also used a random dropout system to force the model to learn more general patterns rather than relying on a couple of strategies. We then also used L2 regularization to penalize excessively large weights and discourage the model from being too complex. One of the more important changes to the original model came with early stopping as it did not need to train for all 100 epochs. A system was added to make sure the validation loss was always improving. If it did stop, the learning rate would get cut in half so it could make smaller adjustments. We also disabled shuffling so that it can only see past data when making decisions and it can therefore accurately make predictions on previous trends. This led to the LSTM and RNN to be the most accurate models of all the models tested.

Now when combining the models together, the linear regressor, decision tree, and random forest had been too inaccurate to use in the final model that is producing the output. The Neural Network and K-Nearest Neighbors however had a very nice synergy together allowing for extraordinarily accurate predictions that allowed for an, on average, 400% return on investment over 1200 simulated days. The RNN and LSTM were tested with trading in the simulations; however, it had run into the problem of predicting such minor changes in which the program would make too large of a move for that small change.

**Results**

This training was done under the basis that the models would eventually become accurate enough to predict the open and closing pricing of the stock for the next day. Numerous models were coded such as a linear regression, neural networks, recurrent neural network, etc to help analyze how models work and what would be the most accurate in daily usage. To test the accuracy of the models in a real stock, we put them into a simulation test with $10,000 fake dollars to play around with. We also provided a 90% weight on the KNN model and a 10% weight on the Neural Networks model for this specific test.

| Testing MSE | Google | Starbucks | MacDonalds | BMW |
|---|---|---|---|---|
| Linear Regression | 22.38 | 2.96 | 9.75 | 0.27 |
| Decision Tree | 46.32 | 6.06 | 19.03 | 0.64 |
| Random Forest | 483.04 | 15.31 | 94.31 | 0.44 |

| | | | | |
|---|---|---|---|---|
| K-Nearest Neighbors | 29.40 | 4.33 | 17.38 | 0.36 |
| Neural Network | 23.99 | 3.85 | 12.81 | 0.45 |
| Recurrent Neural Network | 4376.4398 | 10.0296 | 108.4971 | 0.5211 |
| Recurrent Neural Network (Fixed for Overfitting) | 0.00042547 | 0.00028450 | 0.00017173 | 0.00040749 |
| Long Short-term Memory | 405.5944 | 14.0783 | 28.1908 | 1.351 |
| Long Short-term Memory (Fixed for Overfitting) | 0.00042149 | 0.00028427 | 0.00016131 | 0.00037526 |
| Transformer | 1.1028 | 0.6768 | 1.0564 | 1.0459 |

| Training MSE | Google | Starbucks | MacDonalds | BMW |
|---|---|---|---|---|
| Linear Regression | 14.63 | 2.72 | 9.01 | 0.24 |
| Decision Tree | 0.00 | 0.00 | 0.00 | 0.00 |
| Random Forest | 318.61 | 16.78 | 102.14 | 0.13 |
| K-Nearest Neighbors | 7.07 | 1.33 | 4.73 | 0.35 |
| Neural Network | 20.14 | 3.56 | 11.88 | 0.20 |
| Recurrent Neural Network | 449.9120 | 6.9597 | 29.7684 | 0.7966 |
| Recurrent Neural Network (Fixed for Overfitting) | 0.00040173 | 0.00040657 | 0.00012957 | 0.00037705 |
| Long Short-term Memory | 34.5554 | 10.3944 | 38.9440 | 0.9662 |
| Long Short-term Memory (Fixed for Overfitting) | 0.00041189 | 0.00042006 | 0.00012220 | 0.00036756 |
| Transformer | 1.171 | 0.9946 | 1.0364 | 1.0273 |

| Simulation Test | Google | Starbucks | MacDonalds | BMW |
|---|---|---|---|---|
| Starting/Final Amounts | 10000/ ~36556.74 | 10000/18850.968 | 10000/26170.585 | 10000/12166.66 |

We have highlighted the results were statistical outliers due to the fact that the model had overfit the data.

**Discussion**

The result of these models was clearly that there is a pattern to the stock market. If an AI is able to take the past three days' prices and then predict the fourth day's price, there must be a consistent pattern for it to be making money in the first place. Based on the stocks tested, there is always a correlation on what AI can do with a stock that is predicted to grow in a stable ETF such as the Nasdaq.

When considering all of the impacts on the stock market through AI, it is clear the future changes the entirety of the stock market will have with the growth of AI as it becomes a better predictor. There were some very clear limitations such as the amount of data and the processing ability of the machines used. We also were unable to implement present context, such as political and economic issues, within the price predicting, leading to some incorrect predictions. Our models could only train on one stock at a time leading to it being very slow on predicting multiple stocks. These issues could be solved in future research with better materials. This study had been done with limited resources, so with a better backing and more resources, these models could potentially be combined in a better way leading to more accurate results. In two years, we hope to keep working on these issues and making the model better.

Future research could improve upon this study by making the AI model account for real world events in predictions. Outside factors that revolve around the economy are known to make an impact on stock prices. This model primarily uses historical stock-market data such as past prices to predict future prices. However, future models could incorporate additional information such as financial news sentiment, company earnings reports, interest rates, inflation, and other market indicators and economic changes. These factors may provide the model with information about market movements that would not have been identified from price markers alone.

Future studies and models could also explore the optimization potential. This study's results were based on a manual optimization method. Manual optimization was the easiest way to get weighings, but they may not be the most effective. Since there are infinitely many possibilities of weights, it may be easier if a different method was used. If this were to be worked on in the future, different methods such as a bayesian optimization method or linear weighing may have worked better to combine the models more effectively.

Finally, another direction may be to improve the trading simulation's trading algorithm. The current algorithm runs on a simple buy or sell one or five stocks system. This method is an extraordinarily basic strategy that does not always lead to the best results. When predictions indicated that the market would jump up or down, the bot was not able to buy or sell more stocks. These potential losses are a potential limitation of the simulation bot and something that could be worked on in future studies.

**Works Cited**


Buczynski, Wojtek, et al. “A Review of Machine Learning Experiments in Equity Investment Decision-Making: Why Most Published Research Findings Do Not Live Up to Their Promise in Real Life.” International Journal of Data Science and Analytics, vol. 11, 2021, pp. 221–242. https://doi.org/10.1007/s41060-021-00245-5.

Cao, Lijuan, and Francis E. H. Tay. “Financial Forecasting Using Support Vector Machines.” Neural Computing & Applications, vol. 10, 2001, pp. 184–192.

Dakalbab, Fatima, et al. “Artificial Intelligence Techniques in Financial Trading: A Systematic Literature Review.” Journal of King Saud University - Computer and Information Sciences, vol. 36, no. 3, 2024, article 102015. https://doi.org/10.1016/j.jksuci.2024.102015.

Fischer, Thomas, and Christopher Krauss. “Deep Learning with Long Short-Term Memory Networks for Financial Market Predictions.” European Journal of Operational Research, vol. 270, no. 2, 2018, pp. 654–669. https://doi.org/10.1016/j.ejor.2017.11.054.

Jiang, Zhengyao, Dixing Xu, and Jinjun Liang. “A Deep Reinforcement Learning Framework for the Financial Portfolio Management Problem.” arXiv, 2017. https://doi.org/10.48550/arXiv.1706.10059.

Khemka, Moksh, et al. “Machine Learning in Stock Market Forecasting: A Comprehensive Review.” Discover Computing, vol. 29, 2026, article 599. https://doi.org/10.1007/s10791-026-10491-5.

Kumbure, Mahinda Mailagaha, et al. “Machine Learning Techniques and Data for Stock Market Forecasting: A Literature Review.” Expert Systems with Applications, vol. 197, 2022, article 116659. https://doi.org/10.1016/j.eswa.2022.116659.

Moghar, Adil, and Mhamed Hamiche. “Stock Market Prediction Using LSTM Recurrent Neural Network.” Procedia Computer Science, vol. 170, 2020, pp. 1168–1173. https://doi.org/10.1016/j.procs.2020.03.049.

Oliveira, et al. “Artificial Intelligence Applied to Stock Market Trading: A Review.” IEEE Access, 2021. https://doi.org/10.1109/ACCESS.2021.3058133.

Olorunnimbe, Kenni, and Herna Viktor. “Deep Learning in the Stock Market—a Systematic Survey of Practice, Backtesting, and Applications.” Artificial Intelligence Review, vol. 56, 2023, pp. 2057–2109. https://doi.org/10.1007/s10462-022-10226-0.

Ozdemir, Ozancan, et al. “A Survey of Fundamental Indicators, Machine Learning Techniques and Feature Extraction Techniques for Stock Price Prediction.” Discover Artificial Intelligence, vol. 6, 2026, article 1087. https://doi.org/10.1007/s44163-026-01635-z.

Pippas, Nikolaos, Elliot A. Ludvig, and Cagatay Turkay. “The Evolution of Reinforcement Learning in Quantitative Finance: A Survey.” ACM Computing Surveys, vol. 57, no. 11, 2025, article 295, pp. 1–51. https://doi.org/10.1145/3733714.

Qi, Wenjun. “AI-Empowered Financial Stock: A Survey.” AIP Advances, vol. 16, 2026, article 020702. https://doi.org/10.1063/5.0322804.

Rouf, Nusrat, et al. “Stock Market Prediction Using Machine Learning Techniques: A Decade Survey on Methodologies, Recent Developments, and Future Directions.” Electronics, vol. 10, no. 21, 2021, article 2717. https://doi.org/10.3390/electronics10212717.

Strader, Troy J., et al. “Machine Learning Stock Market Prediction Studies: Review and Research Directions.” Journal of International Technology and Information Management, vol. 28, no. 4, 2020, article 3. https://doi.org/10.58729/1941-6679.1435.

Wang, Feng, et al. “A Survey on Recent Advances in Reinforcement Learning for Intelligent Investment Decision-Making Optimization.” Expert Systems with Applications, vol. 282, 2025, article 127540. https://doi.org/10.1016/j.eswa.2025.127540.

# Individual Differences and Financial Education: Examining the Role of Big Five Personality Traits and Financial Literacy among High School Students in Hong Kong

By Anna Chandler

**Abstract**

As adolescents begin to make their own financial decisions, financial literacy becomes increasingly important, however, previous research done on the subject has focused solely on adults and more sophisticated investors. This study explored the relationship between the Big Five personality traits (Conscientiousness, Openness to experience, Agreeableness, Neuroticism, and Extraversion) and three dimensions of financial literacy, namely financial knowledge, financial attitude, and financial risk tolerance, using a sample of fifty high school students from international schools in Hong Kong. Participants answered questions on various topics on an online questionnaire to assess their personality traits, financial knowledge, financial attitude, financial risk tolerance, and previous financial education. A Pearson correlation analysis and independent-samples t-tests were used to evaluate relationships among the variables. The results demonstrated that the personality traits, Conscientiousness and Openness to experience were both negatively correlated with financial risk tolerance and financial attitude, while extraversion showed a positive correlation with financial risk tolerance and financial attitude. On the other hand, both Neuroticism and Agreeableness did not show any significant correlations with the financial literacy variables. Additionally, financial attitude was positively correlated with financial risk tolerance, but financial knowledge was not correlated with either variable. Additionally, while previous personal finance education correlated with higher financial knowledge compared to without previous personal finance education, there was no significant difference between the two groups regarding financial attitude or financial risk tolerance. Overall, these results indicate that the three dimensions of financial literacy represent different aspects of financial decision-making, and personality is more closely correlated with financial attitudes and risk tolerance than with financial knowledge.

**Keywords** Big five personality traits, financial literacy, financial risk tolerance, financial attitude, financial knowledge, high school students

## 1. Introduction

### 1.1 Why Financial Decision Making Matters

Financial decision-making, which relates to the management of money, is an integral part of people's everyday lives. For studying effective methodologies for financial decision-making, adolescence is especially important, as many young adults begin to earn their own money and transition to financial autonomy (Comon, 2022). During this transformative period, adolescents often find themselves deciding how to allocate their funds for saving, spending, credit, and financial risk (Nguyen, 2013; Comon, 2022).

The financial knowledge and habits developed during adolescence lay the groundwork for future financial behavior and decision-making. As Nguyen (2013) points out, financial literacy is particularly relevant for high school students, who are increasingly faced with situations that require them to make independent financial choices. Common (2024) emphasizes that financial literacy is key to ensuring students' financial well-being down the line. Financial literacy encompasses three main components: financial knowledge, which is what someone understands about financial concepts; financial attitude, which reflects how a person views financial situations; and financial behavior, which is how these elements come together when making financial decisions (Yahaya et al., 2019).

Even though financial literacy is important, it doesn't fully explain the wide range of financial choices people make (Yahaya et al., 2019).People with the same financial knowledge may still choose different financial options due to their varying risk tolerance levels (Raheja & Dhiman, 2017; Kaur & Goel, 2022). This means that understanding one's financial risk tolerance is crucial in the decision-making process, as it influences how much risk a person is willing to face in different financial options (Raheja & Dhiman, 2017).

### 1.2 Personality traits - Influence on Financial Decision Making

Different personality traits influence how we think, react to situations, and make decisions (Kaur & Goel, 2022). Personality can be defined with the help of the Big Five model that includes these five traits:

- Conscientiousness, the degree of an individual's achievement orientation and responsiveness;
- Openness to experience, the extent of a person's innovativeness, creativity, and curiosity;
- Extraversion, the tendency to seek the company of other people;
- Agreeableness, the tendency to be compliant and considerate; and
- Neuroticism, an individual's tendency to experience negative emotions, emotional instability, and stress (Kaur & Goel, 2022).

In 2017, Raheja and Dhiman conducted a study involving 500 investors and identified significant correlations between the Big Five personality traits and risk tolerance. Later, Kaur and Goel’s (2022) research reiterated these findings by identifying relationships between Big Five personality traits and investors' financial risk tolerances. These findings showed that variations in risk tolerance play a significant role in shaping how people make financial decisions, even when they find themselves in comparable financial situations (Raheja & Dhiman, 2017; Kaur & Goel, 2022).

### 1.3 Personality and Financial Risk Tolerance

Financial risk tolerance refers to how much risk a person is comfortable taking when it comes to making investment choices (Raheja & Dhiman, 2017). Specifically, those who have a

higher financial risk tolerance tend to embrace uncertainty and the possibility of losing money. In contrast, individuals with a lower risk tolerance usually lean towards making financial decisions that come with less uncertainty (Raheja & Dhiman, 2017). People exhibit varying levels of readiness towards risk tolerance. This makes financial risk tolerance a crucial factor to consider when examining how investment decisions are made (Raheja & Dhiman, 2017).

Previous studies found evidence of links between personality traits and financial risk tolerance. Raheja & Dhiman (2017) investigated how different personality characteristics of investors relate to their risk tolerance. Similarly, Kaur & Goel (2022) examined the personality traits associated with investors' attitudes towards financial risk. More recently, Mukhdoomi and Shah (2023) found that traits like Extraversion and Openness to experience had a positive correlation with risk tolerance, while neuroticism seemed to have the opposite effect. Additionally, by using regression analysis, they identified that both Agreeableness and Conscientiousness didn't significantly predict attitudes toward risk.

Overall, previous research suggests some consistent findings for certain personality traits and inconsistent findings for others, indicating that relationships between Extraversion, Openness, and Neuroticism are fairly consistent with financial risk tolerance, whereas the results for Conscientiousness and Agreeableness are mixed (Raheja & Dhiman, 2017; Kaur & Goel, 2022; Mukhdoomi & Shah, 2023). Moreover, most of the current research focuses mainly on adults and investors, which leaves us wondering if these same dynamics apply to younger individuals (Raheja & Dhiman, 2017; Kaur & Goel, 2022; Mukhdoomi & Shah, 2023).

### 1.4 Differentiating Financial Knowledge, Financial Attitudes, And Financial Behavior.

It's evident that personal risk tolerance plays a significant role in how people make financial decisions. However, we can't overlook the impact of financial knowledge, attitudes, and education on the financial behaviors of adolescents (Yahaya et al., 2019; Aschenwald et al., 2023).

Yahaya et al. (2019) observed a significant association between financial knowledge and financial attitude, but no direct link between financial knowledge and financial behavior. This suggests that just having financial knowledge doesn’t necessarily lead to better financial habits. However, financial education was observed to have some relevance by Aschenwald et al. (2023). According to Aschenwald et al. (2023), financial education, specifically financial education received from parents has some influence on financial behaviors.

These findings suggest that there are factors beyond just personality that might play a role in the financial choices teenagers make (Yahaya et al., 2019; Aschenwald et al., 2023).

### 1.5 Present Research

Overall, the current research indicates that a person's financial risk tolerance is influenced by various factors, such as their personality traits, financial knowledge, and attitudes towards money. However, most of this research has primarily focused on adults, leaving a gap when it comes to understanding adolescents (Raheja & Dhiman, 2017; Kaur & Goel, 2022; Mukhdoomi

& Shah, 2023). Additionally, the research showed inconsistent findings between risk tolerance and personality traits (Raheja & Dhiman, 2017; Kaur & Goel, 2022; Mukhdoomi & Shah, 2023).

This study aims to fill gaps by examining how the Big Five personality traits correlate to financial knowledge, financial attitude, and financial risk tolerance among high school students. The main question driving this research is: HHow do these personality traits relate to various aspects of financial literacy, including financial knowledge, financial attitude, and financial behavior measured by financial risk tolerance, and how do these elements influence financial decision-making among high school students?

Based on previous research, this study hypothesises that:

Hypothesis 1:
Extraversion and Openness to experience will be positively correlated with financial risk tolerance, while Neuroticism will be negatively correlated with financial risk tolerance.

Hypothesis 2:
This study also hypothesized Conscientiousness and Agreeableness will have a significant correlation with financial risk tolerance, but the direction is not yet known, as previous research shows inconsistent findings.

Hypothesis 3:
The Big Five personality traits (Conscientiousness, Extraversion, Openness to experience, Agreeableness, and Neuroticism) will be significantly correlated with financial knowledge.

Hypothesis 4:
The Big Five personality traits (Conscientiousness, Extraversion, Openness to experience, Agreeableness, and Neuroticism) will be significantly correlated with financial attitude.

Hypothesis 5:
Financial knowledge will be positively correlated with financial attitude, and financial attitude will be positively correlated with financial risk tolerance. Financial knowledge will also be significantly correlated with financial risk tolerance, although no specific direction is predicted.

Hypothesis 6:
Students who have previously received personal finance education will have higher levels of financial knowledge than students who have not received previous financial education. Personal finance education will also be significantly associated with financial attitude and financial risk tolerance.

## 2. Methodology

## 2.1 Aims

This study aims to explore how the Big Five personality traits connect with financial literacy in high school students.. Financial literacy was measured in terms of the following three dimensions: financial knowledge, financial attitude, and financial risk tolerance. This study investigates how each of the Big Five personality traits relates to these three important aspects.

## 2.2 Participants

The participants were approached to respond to the survey. This study included 50 high school students who are currently studying in Hong Kong. Each student was selected through convenience sampling. The participants were personally approached and invited to participate in the study. Participation in the study was completely voluntary, and participants were informed about the purpose of the research before deciding to take part.

The sample of the participants included students from different grade levels ranging from grade 10-12; gender identities, the sample included 36 female participants (72%), 13 male participants (26%), and 1 non-binary participant (2%); age, participants ranged in age from 15 to 19 years, with a mean age of 16.68 years and a median age of 17 years;  and ethnic background, the sample included 32 Asian participants (64%), 11 participants of mixed ethnic backgrounds (22%), and 7 White/Caucasian participants (14%).

## 2.3 Materials and Measures

Data were collected through an online Google Form consisting of 5 different sections: demographics, personality traits, financial attitudes, financial knowledge, and financial risk tolerance. Financial literacy was measured through three dimensions: financial knowledge, financial attitude, and financial risk tolerance.

### a. Demographics

Questions were asked to identify different demographic information, including age, gender, grade level, ethnicity, previous experience studying finance, whether they received an allowance, and how they access their allowance, such as bank transfers, debit cards, credit cards, cash, and mobile payment apps. These variables were collected in order to help describe the sample and provide context of the participants.

### b. Personality Traits

Personality traits were measured using the Big Five Inventory-10 (BFI-10), a brief measure developed by Rammstedt and John (2006). The BFI-10 assesses the five major dimensions of personality: Extraversion, Agreeableness, Neuroticism, Conscientiousness, and Openness. Participants were asked to rate their agreement with ten statements on a five-point Likert scale ranging from 1 - Strongly Disagree to 5 - Strongly Agree. “I see myself as someone outgoing, sociable” is an example of a statement asked. Higher scores for each personality trait indicated a higher presence of the particular trait (Rammstedt & John, 2006).

**c. Financial Attitudes**

Financial attitudes were measured using three statements, a subsection taken from the Financial Management Behaviour Scale developed by Lusardi and Mitchell (2011). Participants responded using a 5-point Likert scale, ranging from 1 - Strongly Disagree to 5 - Strongly Agree. The statements assessed the participants' attitudes towards spending and saving money. An example of a statement asked is “Money is made to spend”. Higher scores indicated greater agreement with spending-oriented statements, while lower scores indicated more savings-oriented attitudes.

**d. Financial Risk Tolerance**

Financial risk tolerance was measured using the 13-item risk tolerance scale developed by Grable and Lytton (1999). The questions assess individuals’ willingness to take financial risks through a series of hypothetical financial decision-making scenarios like investment choices, reactions to gains and losses, and preferences for different levels of financial risk. Participants selected the option that they agreed with most. An example is “In general, how would your best friend describe you as a risk taker?” and “If you unexpectedly received $20,000 to invest, what would you do?”. A higher score indicates a high tolerance for risk, whereas a lower score indicates a lower tolerance for risk (Grable & Lytton, 1999).

**e. Financial Knowledge**

Financial knowledge was assessed using five multiple-choice questions taken from the Global Financial Literacy Excellence Center. The questions measured understanding of key financial concepts including compound interest, inflation, mortgages, bonds, etc. An example of a question asked of the participants was what would happen to purchasing power when inflation exceeds interest earned on savings. Correct responses were totalled to create an overall financial knowledge score.

**2.4 Procedure**

Potential participants were approached and invited to participate in the study using email. Students who expressed interest were provided with a link to the online survey. Upon opening the survey, participants viewed an information page detailing the research's purpose, emphasizing that participation was entirely voluntary, providing an estimate of how long it would take to complete, assuring them that all responses would remain anonymous, and explaining how their data would be utilized.

After providing informed consent, participants completed the survey independently, answering the following question sets:

1. Demographics
2. Personality traits
3. Financial attitudes
4. Financial risk tolerance

5. Financial knowledge

The survey took each participant around 10-15 minutes to complete. After submitting their responses, participants exited the survey.

### 2.5 Ethics

Several ethical considerations were taken into account throughout the study. Firstly, informed consent was given before each participant was sent the survey. Participants were provided with an information page at the beginning of the survey indicating the purpose of the research and what participation involved. Secondly, participation was voluntary, and the participants were told they could withdraw from the survey at any time before submission. No names or other personally identifiable information was collected, ensuring anonymity and confidentiality. Finally, the survey questions were designed to explore personality traits, financial attitudes, financial knowledge, and hypothetical financial decisions to gauge risk tolerance, all of which posed no risk to the participants. The data collected was strictly for educational and research purposes.

## 3. Results

### 3.1 Effects of Personality Traits on Financial Literacy and Financial Risk Tolerance

Pearson correlation analyses were conducted to examine the relationships between personality traits and financial variables. Conscientiousness was significantly negatively correlated with both financial risk tolerance, $r(48) = -.325$, $p = .021$, and financial attitude, $r(48) = -.354$, $p = .012$, but was not significantly correlated with financial knowledge. In contrast, Extraversion was significantly positively correlated with financial risk tolerance, $r(48) = .500$, $p < .001$, and financial attitude, $r(48) = .394$, $p = .005$, while showing no significant relationship with financial knowledge. Openness to experience was significantly negatively correlated with financial risk tolerance, $r(48) = -.404$, $p = .004$, and financial attitude, $r(48) = -.386$, $p = .006$, but was not significantly correlated with financial knowledge. Finally, neither Neuroticism nor Agreeableness showed significant correlations with financial knowledge, financial attitude, or financial risk tolerance.

### 3.2 Relationship between Financial Literacy and Financial Risk Tolerance

A Pearson correlation analysis revealed that financial attitude was positively correlated with financial risk tolerance, $r(48)= .557$, $p<.001$. However, financial attitude was not significantly correlated with financial knowledge. Similarly, financial knowledge was not significantly correlated with financial risk tolerance. Overall, financial risk tolerance was significantly associated only with financial attitude, while financial knowledge showed no significant relationships with the other variables.

### 3.3 The Effect of Personal Finance Education on Financial Literacy

Independent-samples t-tests were conducted to examine differences in financial attitude, financial risk tolerance, and financial knowledge between students who had previously taken a personal finance class and those who had not. There was no significant difference in financial attitude between the students who had taken a personal finance class and those who had not, $t(48) = -1.381$, $p = .174$. Similarly, financial risk tolerance did not differ significantly between students who had taken a personal finance class and those who had not, $t(48) = -1.459$, $p = .151$. However, financial knowledge differed significantly between the groups, $t(44) = -3.180$, $p = .003$, with students who had taken a personal finance class demonstrating higher levels of financial knowledge than those who had not.

**4. Discussion**

The findings gathered from the high school students revealed significant correlations between personality traits and both financial risk tolerance and financial attitude. In particular, traits like conscientiousness, extraversion, and openness to experience stood out in these correlations.

In line with the hypothesis, a negative correlation was found between conscientiousness and both financial risk tolerance and financial attitude; This suggests that individuals who score higher in conscientiousness tend to have lower financial risk tolerance and financial attitude. This finding is consistent with some prior research, which have shown mixed results regarding the relationship between conscientiousness and financial risk tolerance.. While some studies reported a positive correlation, others reported a negative correlation (Pinjisakikool, 2018; Rabbani et al., 2019; Kaur & Goel, 2022; Mukhdoomi & Shah, 2024).

The negative correlation observed can be attributed to traits associated with a conscientious personality, such as being careful, responsible, and organized (Rammstedt & John, 2006). These traits might lead students to shy away from taking financial risks. This is particularly significant for conscientious adolescents, who are still in the process of building their financial knowledge and tend to approach unfamiliar financial choices with caution. This finding also serves as evidence of the direction of the conscientious personality in relation to financial risk tolerance in high school students, especially given the inconsistent results from earlier research.

The positive correlation found between extraversion and financial risk tolerance and financial attitude was in line with the hypothesis and previous research. This means that students who are more extroverted tend to report a greater risk tolerance and a higher financial attitude compared to their introverted peers. This finding is in line with past research that has also highlighted a connection between extraversion and a higher tolerance for financial risk(Pinjisakikool, 2018; Rabbani et al., 2019; Kaur & Goel, 2022). Finding the same results as previous studies suggests that this relationship holds true not just for adults but also for adolescents.It could indicate that the tendency to embrace financial risks is something that exists even before individuals gain substantial financial experience. This might mean that students with

higher levels of extraversion are less intimidated by uncertainty when it comes to making financial decisions.

On the other hand, the findings regarding openness to experience didn't quite align with the positive correlation we expected based on our hypothesis, which suggested a link between openness to experience and financial risk tolerance, as seen in earlier research (Kaur & Goel, 2022); Instead, the results revealed a negative correlation between openness to experience and both financial risk tolerance and financial attitude. This unexpected negative correlation might indicate that the dynamics between openness to experience and financial decision-making vary depending on the demographic (in this case, high school students) since most prior studies have centered on adults. One possible explanation could be that adolescents generally have less financial experience than adults. Therefore, the adolescents who exhibit higher openness to experience might develop greater financial risk tolerance as they mature. Overall, the negative relationship observed in this study adds to the current scientific understanding of the relationship between openness and financial risk tolerance specifically among teenagers.

The absence of any significant correlation for neuroticism and agreeableness is also relevant when interpreting the results of the study. This doesn't imply that these traits are completely unrelated to financial risk tolerance and attitudes in adolescents; instead, it means that there just might not be any significant correlations in the present sample. This differs from some previous research which did identify a negative link between neuroticism and financial risk tolerance (Mukhdoomi & Shah, 2023).

One of the key takeaways from the study is that the connections between personality traits were only evident when it came to financial risk tolerance and financial attitude, but not financial knowledge. This finding partially supports the hypothesis that the Big Five personality traits would be significantly correlated with financial knowledge and financial attitude, as significant relationships were found between conscientiousness, extraversion, and openness to experience and financial attitude; however, no significant relationship was found between personality traits and financial knowledge. This suggests that personality can indeed serve as a predictor for the financial attitudes and risk tolerance of high school students, but it doesn't hold the same weight for financial knowledge. Essentially, students with different personality traits might have similar levels of financial knowledge, yet their financial attitudes and risk tolerance can vary widely.

The correlations found between financial attitude, financial risk tolerance, and financial knowledge were only partially in line with the hypothesis. However, financial knowledge and financial attitude were not significantly correlated and did not support the positive relationship predicted in the hypothesis. Financial attitude was positively correlated with financial risk tolerance, which supports the hypothesis. Financial knowledge was also not significantly correlated with financial risk tolerance; this finding did not support the hypothesis that these two variables would be significantly related.

The positive correlation found between financial attitude and financial risk tolerance indicates that students with higher financial attitude scores (reflecting more agreement to

spending-oriented statements) also display greater levels of financial risk tolerance. This suggests that those who are more inclined to spend their money now are also more open to taking on larger financial risks (Shanmugam et al., 2023).

On the other hand, the lack of a significant relationship between financial knowledge and financial risk tolerance shows that having higher levels of financial knowledge doesn't change a person's risk tolerance. This indicates that financial risk tolerance might be influenced by personal preferences and other factors that go beyond just knowledge (Dhiman & Raheja, 2018).

Additionally, the fact that there is also no significant relationship between financial knowledge and financial attitude suggests that individuals with similar levels of financial understanding can still have varying perspectives on how they choose to manage their spending.

Altogether, these findings demonstrate that one dimension of financial literacy can't predict the other two dimensions. Rather, financial knowledge, financial attitude, and financial risk tolerance are separate elements of financial literacy (Huston, 2010). This finding is also consistent with Yahaya et al. (2019), whose findings also suggested that financial knowledge does not necessarily translate into financial behavior.

Moreover, the lack of significant relationships between financial knowledge and the other two variables is especially valuable because financial knowledge is commonly highlighted as the most important element of financial literacy. This means that possessing information about finance does not mean that the student will demonstrate a certain financial attitude or take risks in their decisions.

The findings regarding the relationship between previous personal finance education and financial knowledge showed that the hypothesis was correct in predicting that students with prior personal finance education have higher levels of financial knowledge. However, the hypothesis saying that personal finance education would be related to financial attitude and financial risk tolerance could not be confirmed by these results. This aligns with the findings of Yahaya et al. (2019), which suggest that just having financial knowledge doesn't automatically lead to better financial behaviors. In summary, while personal finance education does boost students' financial knowledge, that knowledge by itself doesn't seem to change their financial attitudes or financial risk taking.

From these findings, we can conclude that financial education has to go beyond simple financial facts; rather, it should also include opportunities for more practical applications of financial principles, evaluating different financial decisions, and the possible risks of such decisions.

## 5. Limitation

A few limitations should be noted in regard to the interpretation of the results of this research.

The first limitation is that the study used a small sample size found using a convenience sampling method. Instead of randomly selecting participants, individuals were chosen based on who was readily available, primarily focusing on students from private international high schools

in Hong Kong. This raises concerns about the generalizability of the findings; they may not accurately reflect the broader high school population or the entire adolescent demographic when it comes to their backgrounds and financial literacy. For future research, it would be beneficial to include a larger and more diverse sample to see if the relationships identified here hold true for the general adolescent population.

Another limitation to consider is that the financial knowledge questionnaire might not have been completely age-appropriate for the participants. Given that they were high school students, some of the financial questions posed could have been more relevant to adults, which means the students might not have fully grasped them. These concepts include terms like mutual funds, bonds, and mortgages. While the questionnaire did manage to highlight differences in financial knowledge among the students, it might not have fully captured the types of financial understanding that are most pertinent to adolescents. This could have affected how accurate the measurement of the students' financial knowledge was, as well as its correlations with financial attitudes and risk tolerance.

Another limitation to consider is that in this study, financial risk tolerance was used as a behavioral indicator of financial literacy. However, it's important to note that financial risk tolerance might not capture the full spectrum of financial behavior. Financial decision-making involves a variety of behaviors, and financial risk tolerance is just one of many. Future research could delve deeper into other behavioral indicators that relate to the financial habits of adolescents.

**6. Directions for future research**

The findings of this study reveal areas for research moving forward.

Firstly, future research would benefit from developing financial education methods that improve financial literacy beyond just the single dimension of financial knowledge. The findings from this study showed that students who had prior financial education tended to have higher levels of financial knowledge compared to those who hadn't received any personal financial education. However, this didn't translate into better financial attitudes or risk tolerance. This finding is supported by the Yahaya et al. (2019) study, which states that having financial knowledge does not necessarily mean that one would engage in corresponding actions or attitudes. Therefore, it might be beneficial for future studies to also look at how practical applications of financial concepts, such as making financial decisions and evaluating the possible risks associated with those decisions can be integrated into the learning process.

It would also be useful for future research to consider students' financial attitudes and personality traits when developing new styles of financial education. The results from this study revealed that some of the Big Five personality traits, such as Conscientiousness, Extraversion, and Openness to experience, were significantly correlated with financial attitude. This suggests that a high schooler's personality can influence their approach to financial decisions. Therefore, it would be valuable for future studies to investigate how financial education can be tailored to align with the variations in personality traits and financial attitudes.

Thirdly, it would be useful for future research to examine whether the results found in this study can be replicated in a larger sample of adolescents. This study utilized a small convenience sample. An expanded sample involving students from different school environments might give more insight into the relationship among personality, financial literacy, and financial risk tolerance in a general population of adolescents.

Future research may also benefit from using financial knowledge measurements specifically designed to question adolescents. Specifically, a measure of financial knowledge that addresses the financial concepts and issues more familiar to high schoolers, such as savings, spending, and budgeting, as opposed to those that are more adult-oriented. Using questions that are closer to the real-world experience of high schoolers may be a more reliable indicator of financial knowledge, which will help in future research to understand whether or not the relationship between the variables is affected when financial concepts are age-appropriate.

Lastly, future research would benefit from considering alternative ways of measuring financial behavior other than financial risk tolerance. Since this study relied on a small convenience sample, expanding the research to include students from various school settings could provide deeper insights into how personality, financial literacy, and financial risk tolerance interact within a broader adolescent population.

**7. Conclusion**

This study examines the correlations between personality traits and various aspects of financial literacy of high school students as well as the relationships between the students' previous personal finance education and financial knowledge, financial attitude, and financial risk tolerance. Even though earlier personal finance education was linked to a better increased financial knowledge, financial knowledge was not related to financial attitude and risk tolerance. Interestingly,there was a positive correlation discovered between financial attitude and risk tolerance, which suggests that financial attitudes are more closely related to financial risk tolerance than financial knowledge. Personality traits also showed relationships with financial attitude and financial risk tolerance; however, no significant relationships were found between personality traits and financial knowledge. The findings indicate that it's crucial to grasp all three aspects of financial literacy along with individual personality traits. We shouldn't just assume that one aspect mirrors another. This understanding is essential when looking at the factors that shape how adolescents make financial decisions.

**Works Cited**


Comon, J. (2022). Social support and financial literacy for senior high school learners. Psychology and Education: A Multidisciplinary Journal. https://doi.org/10.5281/zenodo.6728323

Dhiman, D. B. (2017). Does Investor Personality Determine their Risk Tolerance?

Global Financial Literacy Excellence Center (GFLEC). (2021, October 13). The Big Three and Big five | Global Financial Literacy Excellence Center (GFLEC). https://gflec.org/education/questions-that-indicate-financial-literacy/

Grable, J. E., H., & Lytton, R., H. (n.d.). Financial risk tolerance revisited: The development of a risk assessment instrument. Financial Services Review, 8, 163–181. https://static1.squarespace.com/static/5a14be9fd0e62827bd667d1f/t/5f1afc8de5b6684a6cc62bb0/1595604109809/2019-12-31+Grable-Lytton-Risk-Assessment.pdf

Huston, S. J. (2010). Measuring financial literacy. Journal of Consumer Affairs, 44(2), 296–316. https://doi.org/10.1111/j.1745-6606.2010.01170.x

Mukhdoomi, A. M., & Shah, F. A. (2023). Risk tolerance in investment decisions: Are personality traits the real triggers? NMIMS Management Review, 31(4), 256–264. https://doi.org/10.1177/09711023241230433

Nguyen, L. T., Gallery, G., & Newton, C. (2016). The influence of financial risk tolerance on Investment Decision-Making in a Financial Advice context. Australasian Accounting Business and Finance Journal, 10(3). https://doi.org/10.14453/aabfj.v10i3.2

Potrich, A. C., Vieira, K. M., & Paraboni, A. L. (2024). Youth Financial Literacy Short Scale: Proposition and validation of a measure. Social Sciences & Humanities Open, 11, 101214. https://doi.org/10.1016/j.ssaho.2024.101214

Pinjisakikool, T. (2018). The influence of personality traits on households' financial risk tolerance and financial behaviour. Journal of Behavioral and Experimental Finance, 18, 42–50.

Rabbani, A. G., Yao, Z., & Wang, C. (2019). Does personality predict financial risk tolerance of pre-retiree baby boomers? Journal of Behavioral and Experimental Finance, 23, 124–132. https://doi.org/10.1016/j.jbef.2019.06.001

Rammstedt, B., & John, O. P. (2006). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. Journal of Research in Personality, 41(1), 203–212. https://doi.org/10.1016/j.jrp.2006.02.001

Razen, M. a. &. A. H. &. M. K. &. M. (2023). Determinants of Financial Literacy and Behavioral Bias among Adolescents. ideas.repec.org. https://ideas.repec.org/p/inn/wpaper/2023-01.html The Society of Economics and Development. (n.d.). https://www.soed.in/article/560

Shanmugam, K., Chidambaram, V., & Parayitam, S. (2023). Relationship between Big-Five personality traits, financial literacy and risk propensity: Evidence from India. IIM Kozhikode Society & Management Review, 12(1), 85–101. https://doi.org/10.1177/22779752221095282

Yahaya, R., Zainol, Z., Abidin, J. H. O. @. Z., & Ismail, R. (2019). The Effect of Financial Knowledge and Financial Attitudes on Financial Behavior among University Students. International Journal of Academic Research in Business and Social Sciences, 9(8). https://doi.org/10.6007/ijarbss/v9-i8/6205

**Expanding Capabilities: The Role of AI in Global South Development By Kai Yui Lee**


## Abstract

The rapid integration of artificial intelligence across different countries and sectors presents both unprecedented opportunities and significant challenges for economic development and human well-being. This paper examines the integral role of AI in three critical sectors, agriculture, medicine, and education, within Global South communities, utilizing Amartya Sen and Martha Nussbaum's Capability Approach as a theoretical framework. Through analysis of diverse case studies including precision farming in Nepal, AI-powered diagnostic tools in India and Uganda, and adaptive learning platforms in Nigeria, Ghana, and India, this research demonstrates that AI applications function not merely as efficiency tools but as capability multipliers that expand individual freedoms and substantive opportunities. The findings reveal that AI-driven precision agriculture enhances resource efficiency and smallholder farmer incomes; AI-enabled medical diagnostics significantly improve early disease detection and healthcare accessibility in resource-constrained settings; and AI-powered personalized learning platforms generate learning gains equivalent to 1.5–2 years of traditional instruction. However, the research also identifies critical limitations, including digital divides, infrastructure deficits, data privacy concerns, and the risk of exacerbating existing inequalities without deliberate policy intervention. The paper concludes that while AI offers transformative potential for addressing persistent developmental challenges, its success ultimately depends on contextual appropriateness, community co-design, and a commitment to expanding human capabilities rather than merely increasing productivity metrics. This framework provides a more holistic evaluation of technological progress, emphasizing that true development is measured by whether individuals gain the substantive freedoms to live lives they have reason to value.


## Introduction

With the rise of AI throughout recent years, sectors and countries from around the world have rapidly started to adapt to this historic technology. The world has seen unprecedented economic efficiency, where many menial tasks and labors are now done in a matter of seconds. Tech giants Apple, Amazon, Microsoft and Meta now use AI for greater efficiency in handling said tasks. For example, according to the Journal of Agriculture and Food Research, in sectors like agriculture, there is enhanced activity and efficiency due to the use of precision farming, addressing key issues like food security. In healthcare, as stated by the European Journal of Medicine Research, AI improves diagnostics, access, and efficiency for managerial work, thus allowing people to receive medical treatment efficiently. In education, AI can personalize each student’s learning experiences, helping each student grow in their own way, as well as reach students who are unable to go to school through the use of a device, giving everyone access to educational services. By addressing these obstacles people face every day, AI can make the economy more efficient and stronger, which also allows for growth in the future. The historic rise and implementation of artificial intelligence in sectors such as agriculture, medical sciences,

and education displays a positive correlation with growth and development, which is vital in addressing problems regarding food security, healthcare access and educational disparities while mitigating higher potential unemployment rates.

This will help us explore various case studies. For example, for agriculture there is a case study from Harvard Advanced Leadership Initiative Social Impact Review, where AI Companies are using Optical Character Recognition (OCR) technology to digitize paper records for farmers in Nepal, increasing efficiency for the farmers. In the realm of education, Ashesi University in Ghana implemented the strategic use of AI in addressing teaching and learning challenges with technology within the Ghanaian education system. One final example can be seen in the medical sector, as Tsinghua University in China created an AI-powered virtual system that can significantly speed up treatment processes and also decrease the wait time for people to see medical professionals. Using the capability approach will serve as a theoretical foundation, as we are highlighting the enhancement of the community's capabilities through the development of AI and various other technological developments. Moreover, some relevant economic theories, for example, productivity growth and innovation diffusion, will be discussed to contextualize how AI can transform some agricultural practices, improve healthcare delivery, and enhance educational outcomes significantly efficiency in a community and country and thus economic growth, but also to increase one's limits of capability. Understanding the relation between AI and agricultural growth is crucial, as it not only talks about the immediate effects and challenges but also promotes long-term sustainability. I will be talking about these 3 resources because they are all heavily impacted by the emergence of AI, by making each sector more efficient with technological developments, which will help us prove that By addressing AI can make the economy more efficient and stronger which also allows for growth in the future through these sectors . Ultimately, this thesis relates to how the use of AI benefits not only one's own capability but their community as a whole. As we delve into the critical role that these sectors have on the growth of the economy, it is necessary to develop a framework that focuses on the well-being of humans and encapsulates their opportunities to seize their goals and live happier. This framework, representative of this essay, will be the Capability Approach.

The Capability Approach, a framework utilised by famous philosophers and scholars such as Amartya Sen, a pioneer in the economic and philosophical world, whose version of the capability approach focuses on individuals' reality and their actual abilities and liberties to achieve well-being. It makes a clear distinction between functions and capabilities, where functions are what people do and states of being that allow people to flourish,while capabilities are what opportunities are available to people and utilising those opportunities to increase their quality of life. While his central idea emphasizes agency and individual liberties, it simultaneously advocates for social justice by expanding people's capabilities. Martha Nussbaum on the other hand highlights the importance of one's wellbeing in regard to one's freedoms and functions, which ultimately allow people to be more content with their lives. She focuses on how every person should be given the right to have a certain threshold in their capabilities, which should all be provided universally. Her aim is more social justice-oriented, allowing individuals

to pursue their own images of what a good life is for them.For example, an artist who has the freedom to explore alternative careers or creative paths is more likely to find satisfaction in their work, leading to greater happiness and productivity.The idea is applicable in a variety of professional fields. People who are capable of genuinely pursuing different career routes, whether in agriculture, medicine, or education, are more likely to find professional satisfaction, increasing their efficiency and morale. Higher productivity and a larger economic contribution follow from this improved well-being. Therefore, expanding capabilities promotes both macroeconomic efficiency and individual flourishing.This framework is essential in understanding how the economy grows because it emphasizes the enhancement of individual people's capabilities. I will be using the capability approach in this paper to discuss how the emergence of AI connects with the capability approach to make the medical, educational, and agricultural sectors more efficient, while making all individual persons happier with their lives. This is a necessary topic to discuss due to the relevance of the topic in current society as it addresses concerns about quality of life standards for many individuals and citizens across the world. This in turn, hinders the economy and efficiency of multiple sectors.

**Agriculture**

Precision farming, which is an approach of utilising data from drones to decide when and where to plant, heavily requires the use of AI and technology. AI-powered tools help farmers make well-informed decisions regarding planting schedules by combining data using satellite imagery. Targeted interventions are guided by predictive models, which maximize crop yields and resource efficiency while reducing waste. Optical Character Recognition (OCR) technology, for instance, allows for digital agricultural records, making it easier to access crucial crop and market data. Additionally, AI is crucial in advancing sustainable resource management in agriculture. Water shortage is a serious problem, with up to 80% of illness is due to inadequate amounts of water in developing nations, thus AI driven irrigation systems are needed in these communities (Global Affairs Canada). AI-driven irrigation systems make sure that water is used effectively by closely monitoring soil moisture levels and weather forecasts (OĞUZTÜRK). To avoid overwatering and save valuable resources, smart irrigation systems can automatically modify water distribution depending on real-time data (US EPA, "Weather-Based Irrigation Controllers"). This helps farms lower operating expenses while simultaneously promoting environmental sustainability. The United Nations Trade and Development maintains that smallholder farmers often faced significant challenges namely, digital divides; specifically, restricted access to cash, infrastructure, and technology exacerbates digital inequalities, which are major obstacles faced by smallholder farmers. Even as precision farming develops, this disparity poses a threat to a significant portion of the world's agricultural population. But specialized AI solutions are starting to appear to close this gap. It supports that without requiring costly hardware or high-bandwidth internet, mobile-first platforms that provide insights via straightforward SMS or lightweight apps can give small-scale farmers vital information on weather, pest outbreaks, and ideal market prices. Additionally, community-based models are

working well (Kamal and Bablu 7). For a whole cooperative, these hubs can integrate satellite and drone data and provide tailored, practical recommendations, like the exact quantity of organic fertilizer needed for a particular plot.crucial is the small-scale, more economical integration of AI with the Internet of Things (IoT). Smallholders' resource limitations can be directly addressed by low-cost soil sensors and AI-guided automated irrigation kits that significantly lower water and input costs (Gaitan et al.). These tools enable farmers to reduce the risks associated with climate unpredictability and enhance food security by democratizing access to predictive analytics. In the end, precision agriculture needs to be inclusive if it is to live up to its worldwide promise. To ensure that the advantages of data-driven farming reach every field and promote equity and resilience in the agricultural landscape, ongoing innovation in accessible, reasonably priced AI tools must be combined with supportive legislation and training initiatives (Ending Learning Poverty 1, 3-4).

Moreover, by evaluating data on crop performance and soil health, AI can support crop rotation and diversification initiatives (Schweng et al.).AI helps farmers improve soil quality and fight pests and illnesses by suggesting particular crop combinations based on environmental variables (Sabeena et al.). This knowledge guarantees healthier crops and strengthens the farming system's resilience (Bowles et al.). AI can also assist with crop rotation and diversification programs by analyzing data on soil health and crop performance. By recommending certain crop combinations depending on environmental factors, AI assists farmers in enhancing soil quality and combating pests and diseases. This information increases the resilience of the farming system and ensures healthier crops. AI can also simulate the long-term effects of various rotating techniques across several seasons, taking into consideration factors like shifting market prices and climate trends (Li and Liu 1, 2). This enables farmers to make proactive, strategic choices that maintain economic stability and preserve the vitality of their land for future generations (Grewer et al.). A constant feedback loop is created by the system's capacity to learn from each harvest and improve its recommendations, turning static farming practices into dynamic, adaptive ones (Awais et al.).

Beyond just increasing output, AI in agriculture has a big impact on nearby populations. Farmers can earn more money and make more significant contributions to their local economies by improving agricultural efficiency (Li et al. “Agricultural Technology Extension”). For farmers and their families, this income boost may result in better living conditions, access to healthcare, and educational possibilities (Rosnah et al.). AI promotes greater inclusivity in the agricultural industry in addition to its economic advantages (OECD). AI solutions can help smallholder farmers, who frequently struggle with issues including limited access to resources and knowledge (OECD). For example, farmers can have access to previously unattainable knowledge through mobile applications that offer real-time agricultural guidance. These resources can assist them in making well-informed choices regarding market access, pest control, and crop management (Li et al. “Agricultural Technology Extension”).

This technological change has significant and varied socioeconomic consequences. Increased productivity and production directly result in better and more consistent incomes for

individual farmers (Putri and Apriliansyah 3-5). This increase in income has the potential to improve living conditions by giving families improved access to nutrition and healthcare. At the local level, rural towns' economic vibrancy increases with farm profitability due to higher investment and spending (Fan). AI-powered smartphone apps can provide farmers with localized, real-time guidance on market prices, insect diagnosis, and crop management. This feature has the potential to level the playing field and is a formidable weapon for inclusivity (Brini).

Thus, AI is the driving factor behind precision farming's transformation from a specialized idea into a universal necessity. It tackles the twin problems of feeding an expanding population and protecting our planet's ecosystems by optimizing inputs, maximizing yields, and conserving essential resources ("Digital Agriculture and AI"). However, bushels per acre or water saved will not be the only metrics used to assess its final performance. When all farmers, regardless of size or location, can benefit from AI-driven agriculture, true success will be achieved. We can direct this technological revolution toward a future that is not only more productive and sustainable but also more just and inclusive by proactively addressing the obstacles of cost, knowledge, and infrastructure. This will reinforce the foundation of our global food system.

**Medicine**

For billions of people in the Global South, having access to high-quality healthcare is still a risky goal rather than a certainty. The problems are systemic, including a severe lack of qualified medical personnel. According to the World Health Organization, there are estimates that there is a global shortage of 10 million health workers, mostly in low-income areas, infrastructural deficiencies that make it a day's travel for rural residents to get to a clinic, and expenses that drive families into poverty. Although governments and non-governmental organizations have made progress, such as Rwanda's innovative drone delivery system for blood and vaccines or India's extensive network of ASHA (Accredited Social Health Activist) community health workers, the scope of the problem still exceeds conventional solutions (Paul et al.). In this context, artificial intelligence becomes a practical tool for overcoming decades of infrastructural neglect rather than a speculative luxury.

For example, Diabetic Retinopathy (DR) is the leading cause of preventable blindness in the world (Agarwal et al.), but in order to prevent this, it is necessary to have a specialist catch it early on, which costs a lot of money, and also, with a severe lack of ophthalmologists (around 12,000) in a country with over 1 billion people, rural patients likely need to go to a large city and wait long lines in order to be diagnosed, which will cost a fortune and put them into poverty, as a result, many lose their vision unnecessarily. Now the solution is having local health workers in smaller hospitals use a low cost portable smartphone based camera (Remidio) to capture retinal images, using Google's ARDA, they can analyze the image locally, where the AI matches or exceeds the diagnostic accuracy of human specialists. By 2025, ARDA has screened over 600,000 patients in India alone with a 97% sensitivity and 96.4% for specificity (Brant et al.).

Another example is the utilisation of AI when identifying Rheumatic Heart Disease (RHD). Currently 23% of all global RHD cases are from Sub-Saharan Africa alone ("Rheumatic Heart Disease"). However, RHD deaths had been entirely preventable by the 1950s, but many still die from this disease due to RHD being invisible until it is too late. This means that when symptoms finally arise, it will require open-heart surgery, which is something that is inaccessible in rural areas due to costs and lack of medical professionals. Thus, AI becomes essential in this area since it is more accessible in rural areas in global south communities where there may be very limited specialists in the field in the whole country. Just by using a phone, people can check, analyze and spot microscopic leaks that RHD causes. When catching RHD early, in nearly 48% of the children (classified as ages 5 to 17) caught RHD early, the heart valve damage regressed back to normal over a period of 2 years (Beaton et al.). In Uganda, there are 7 paediatric cardiologists (Mbabazi et al.) thus in a case study presented in Primary schools in Uganda, using image acquisition AI's, non specialist nurses are able to successfully capture diagnostic quality cardiac images in over 90% of cases (Innovation District).

Finally, more than 20 low-resource settings in Asia and Africa are using Qure.ai's qXR tool in India. With accuracy on par with a radiologist, this AI system analyzes chest X-rays for indications of tuberculosis (TB) and other lung abnormalities in a matter of seconds on a regular laptop (A Case Study on Global Deployment of qXR for TB). There is a severe lack of radiologists, particularly in district hospitals, in a nation like India that accounts for 25% of the world's tuberculosis cases (Global Gains). When Qure.ai's technology is incorporated into portable X-ray machines used in community screening camps, technicians with little training can get readings right away. As a result, early TB case detection in underprivileged communities has dramatically increased, allowing for quicker treatment and the disruption of transmission chains (India: Qure.Ai's qXR Boosts TB Detection & Saves Costs).

Thus, we can utilize the capability approach as a means to discuss how AI contributes to society, in particular medicine, to protect an individual's free liberties (Nussbaum). The ability to live a long, fulfilling life and pursue other important activities like education or employment is largely dependent on one's health. These AI interventions support this capability in two important ways: by improving the institutional and social environment and by strengthening individual agency. They transform healthcare into a more immediate and comprehensible resource from a remote, frequently inaccessible service. Moreover, the capability approach allows us to think and look beyond the immediate health effects to the broader factors that AI can influence. For example, for a TB screening result or a disease prediction to translate into genuine freedoms and liberties, individuals need resources that are complementary: accessible clinics, or affordable medication. Therefore, the true measure of these AI tools is not merely their diagnostic accuracy, but how they can help contribute to one's freedoms and liberties to expand their capabilities and enable truly happier lives.

As a result, even if AI in medicine has great potential to improve health capacity in the Global South (Sinha), its integration must be guided by the concepts of justice, equity, and contextual appropriateness. In order to achieve the ultimate aim of giving every person the true

freedom to be healthy, it is necessary to carefully cultivate systems that enhance local capability for all.

**Education**

Schooling is a vital experience for all kids. However, kids that live within the global south have a lack of available access to schooling compared to those in privileged, more developed countries. Those without the available schooling are unable to get the necessary steps in today's society to be able to get jobs with higher pay and increasing their wages, and are instead forced to work laborious jobs that pay little in comparison. There has been a slow rise in literacy rates from 65% to 69% in sub saharan africa from 2015 to 2024 (International Literacy Day 2025), is a testament to teacher efforts and students' determination to get a strong education. "Adult literacy is the percentage of people ages 15 and above who can both read and write a short simple statement about their everyday life." (World Bank Group). Although literacy rates have gone up in recent years, things aren't as perfect as they seem. In fact, the quality of education is a massive problem, where many students are still unable to speak languages, read or write fluently, leaving them still stuck in the same position as before. However, with the emergence of AI as a tool for success, it allows children from all over the world to be able to access education more freely.

This advantage isn't merely about attendance; it's about the chasm between merely having a schoolhouse and having a transformative educational environment. In many communities across the Global South, schools exist, but they are often overburdened with packed classrooms, with textbooks being shared by groups of 12 to 14 in countries like Cameroon (International Literacy Day 2025).Thus explaining the problem of education presented in many global south communities, despite the increasing numbers. It represents millions of individual stories of parents sending a child to school for the first time, of NGOs building classrooms in remote villages, and of a growing global recognition that education is a fundamental right. Yet, this statistical victory can mask a more insidious crisis: the crisis of quality. The problem of quality manifests as a silent epidemic. Students might attend school for years yet may emerge unable to comprehend a local newspaper, follow basic instructional manuals, or articulate a thought in a complete paragraph. This "learning poverty" traps them just as effectively as no schooling at all. They are equipped with a certificate but not with competency, leaving them still sidelined from the formal economy, leaving them in the cycle of poverty. The dream of education as a great equalizer turns brittle when the foundation is built on sand.

The utilization of AI into global south classrooms becomes more essential with this ongoing crisis regarding the learning of students. AI can become an extremely useful resource for students to be able to learn as a one on one tutor that can easily adapt to the students level without having to pay extensive fees. Instead, students will be able to have access to these AI tools through devices and the internet.

In a study in the Edo State in Nigeria, the government partnered with international education research and piloted an AI tutoring system designed to help secondary students. A

World Bank evaluation found that secondary students who participated in the six week AI tutoring achieved significant growth in mathematics and science equivalent to 1.5 to 2 years of typical learning compared to their fellow classmates. This shows the potential of AI teaching students in these communities, since the AI was also able to pinpoint individual learning gaps that allowed students to understand grade level math (Elias).

In another case study, MIT's Abdul Latif Jameel Poverty Action Lab (J-PAL) evaluated Mindspark, an AI-Driven personalized adaptive software, was used and implemented in government and low income schools in Rajasthan in India. In these schools, students utilised shared tablets for 45 minutes a day, where the platform analyzed student errors and diagnoses misconceptions. As a result of J-PAL's trials, students using mindspark made 2-3 times higher learning growth compared to students who did not use MindSpark in Mathematics and Language (Abdul Latif Jameel Poverty Action Lab).

Finally, Rising Academies created an AI-powered math tutor which specifically aims to run on chat based channels like whatsapp or basic local networks called Rori. Rori was then used in primary and middle schools in Ghana, where students interact with Rori for two 30 minute sessions per week. The chatbox then checks their answers in real time and provides hints tailored to the mistake a student makes. Thus, an evaluation by researchers from Oxford and J-PAL across 11 schools in Ghana showed that students using Rori achieved markedly higher scores in math, with a growth of one additional year of gains compared to students not using Rori (Henkel et al.).

All 3 case study examples show that the utilisation of AI tools in the education sector in global south communities is essential. AI can help expand the capabilities of students by increasing their content knowledge in common subjects (Math, Science and Language), outpacing peers that do not have these resources. With these tools, students will be able to read and write better, do math problems, and be fluent enough to be able to open their doors to more than just jobs that require labor and brute strength, and go into fields that they aspire to go into. This is consistent with Amartya Sen's Capability Approach, which contends that development should be assessed based on whether people truly have the substantive freedoms to be and do what they have reason to value rather than just the availability of resources like schools, teachers, or enrollment rates (Sen). In the context of education, pupils who do not graduate with functional literacy or numeracy are not guaranteed actual competency simply by attending school. The ability to read fluently, solve issues, and think critically enough to pursue worthwhile career and life choices is what increasing capability means. AI tools are essential in assisting students in transforming their education into practical skills through the integration of AI combined with classroom learning. AI fulfills the fundamental promise of Sen's vision for human development by giving pupils practical skills in science, math, and language. This gives them more agency and freedom to break away from the cycle of low-wage labor and pursue occupations they truly want.

However, this AI-augmented future is not automatic, it requires gradual integration and its pitfalls are serious. The problem is that without investment in electricity, internet connectivity,

and affordable devices into these global south schools, AI could exacerbate the very inequalities it seeks to solve, creating a new sub-class of the digitally excluded. Data privacy for vulnerable children is highly important in their learning journeys and must not be commodified. Crucially, AI tools must be designed by both educators and communities in the Global South in order to include possible obstacles such as local languages, values, and curricula.

Ultimately, the goal is to use machines to unleash human potential and expand one's capability to achieve things that allow them to live a satisfactory life, regardless of one's personal background or income. The rising literacy rate can be seen as a misleading statistic, and the importance is the actual learning that the students are making, which is much more efficient with the help of AI. AI, as a tool, can help ensure that they are advancing their conceptual knowledge in standard subjects. It offers a chance to move beyond just counting how many children are in school, to ensuring that every single one of them is truly learning, equipped with the cognitive tools to build a better life for themselves. It also allows students to explore what subjects they are more interested in, giving them a sense of direction into what jobs the students want to pursue in the future. The teacher's role then evolves from a strained distributor of information to a nurtured cultivator of minds, working jointly with technology to ensure that the promise of a strong educational foundation becomes a reality for many, not just the privileged few. Thus in the end, AI becomes a reliable source for students to increase their understanding, paired with teacher assistance.

**Conclusion**

The integration of artificial intelligence across agriculture, medicine, and education in the Global South represents more than technological advancement, it embodies a fundamental shift on how people can expand their own capabilities, and open more doors for themselves. As this essay has demonstrated through diverse case studies through multiple sectors, it is clear that AI applications are not merely efficiency tools but capability multipliers that expand what individuals can actually do and become. This transformation is particularly significant because it addresses the persistent difficulty and obstacles of delivering services to those most in need, on the other hand, it also shows how AI can be used to significantly improve the standard of living for many people. By placing agricultural expertise, strong medical tracking, and integration into the education of children directly through mobile devices, these obstacles that the people of the global south face, while not entirely gone, diminishes to some extent.

In agriculture, AI has the ability to transform how farmers grow food by shifting traditional farming methods to the utilisation of AI and precision farming to ensure that resources are not wasted. Through satellite imagery, weather data, and soil health monitoring, AI helps farmers make smarter decisions about crop rotation which can help increase crop yields while reducing waste. However, these benefits cannot be limited to large-scale farms. Smallholder farmers, who face barriers such as limited access to technology and money must also benefit. With the implementation and integration of mobile-first platforms, community-based data hubs, and low-cost sensors are helping bridge this digital divide by making predictive analytics more

accessible to the general public. Beyond increasing productivity, AI-driven agriculture brings socioeconomic benefits for farmers, for example, increased standard of living, higher profits, and a quality increase in crops. In the end, the true success of AI in agriculture will not be measured solely by crop yields or water saved, but instead, allowing the farmers to be more efficient with their resource use, expanding their capability and standard of living. This connects the capability approach because according to Amartya Sen's Capability Approach, which states that development should be evaluated not by outputs like total crop production or GDP growth, but instead if individuals gain real liberties and have abilities to live their lives with satisfaction and have reason to value.

In the medical department, AI can be life changing technology for millions of people living in the global south and rural areas. Utilising AI to do scans as well as identify complex diseases early on, it will help save millions of lives each year. This will help people of the global south in terms of capability because being healthy allows them to be more efficient and productive, and allows them to live a productive life and achieve the basic thresholds that Martha Nussbaum argues regarding specific liberties that should be met in order for someone to have the capability to do tasks and open more doors for themselves. This also connects to Amartya Sen's version of the capability approach because one of the fundamental parts of Sen's capability approach's liberties is the ability to be healthy. Whether it is individuals using AI diagnostics and doing checkups at their local smaller hospitals, or non-specialist nurses using AI to check for RHD cases, these AI tools have a high percentage of identifying these diseases early on, making them much more treatable.

Education, perhaps most profoundly, illustrates AI's potential to transform statistical progress into genuine empowerment. Rising enrollment rates across the Global South mask the crisis of learning poverty, where students attend school yet emerge without functional literacy or numeracy. AI-powered adaptive learning platforms offer personalized scaffolding that overburdened classrooms cannot provide, diagnosing individual knowledge gaps and delivering instruction in local languages. When a child in rural Ghana can access the same quality of foundational education as their urban counterpart, the capability for meaningful participation in economic and civic life ceases to be predetermined by geography. This directly connects to Sen's capability approach, since it also argues that development should be measured not by resources like attendance in school, but whether they have the freedom to learn and to be able to seek out knowledge, making AI the perfect tool to assist students in the global south with their educational journey.

Although there are significant opportunities and untapped potentials within these sectors in the global south, there are some limitations we must address. Firstly, one could argue that this paper assumes that a significant amount of technological resources are needed (good wifi, regular access to mobile phones or tablets) and that the costs of these technological devices are too expensive for villagers or people in the global south, making this unfeasible. However, this limitation does not undermine the premise of the thesis, because the case studies used in this paper specifically references low resource uses of AI, for example, the Rori platform mentioned

earlier in the paper utilises Whatsapp or basic local networks. Another limitation that one could present is that this paper overestimates the ability of local healthcare workers, teachers, and farmers to utilize these AI tools. In spite of this, the paper acknowledges that this is a valid argument, however, based on the case studies the paper has used, despite the lack of digital literacy in these rural areas, it is proven that they can still implement and utilise these AI tools effectively. For example, in the Uganda nurses example earlier, when using image acquisition AI's, non specialist nurses are able to successfully capture diagnostic quality cardiac images in over 90% of cases, thus proving that digital literacy may be less of a problem than one might think. One last limitation I would like to present is the displacement of human workers. One may argue that if AI can help find illness early on, can tell us exactly when and what to plant at the perfect time, and can teach students at an astonishing rate, what should we do about the teachers, the doctors, the nurses, and the farmers. However, the argument does not state that we should solely rely on the use of AI, instead, it should be the integration of AI with these professions. For example, in the agricultural sector, it is not robots or AI farming, it is the farmers using their professional expertise along with helpful AI data to help them maximize their crop yield while reducing waste.

The capability approach articulated by Amartya Sen and Martha Nussbaum provides the ethical framework for evaluating these developments. True development, as Sen states, is not measured by GDP growth alone but by expansion freedoms and liberties. AI interventions succeed when they enhance what people can actually do, for example in this paper, farm productively, access healthcare, and learn effectively, so that they can help themselves grow and live satisfactory lives based on their own beliefs and values. Nussbaum's emphasis on threshold capabilities reminds us that these technologies must reach those that need it the most, not just benefiting those already advantaged. Yet this technological advancement holds various responsibilities. Without deliberate attention to digital divides, and contextual appropriateness, AI risks creating new hierarchies of exclusion. The tools must be co-designed with local communities, reflect indigenous knowledge systems, and remain subordinate to human dignity. When thoughtfully implemented, however, AI offers the Global South not a shortcut to development but something more valuable: the means to pursue development on its own terms, expanding capabilities and allowing them to achieve goals and aspirations that they have for themselves. The future belongs not to those who adopt technology fastest, but to those who wield it most wisely in service of human freedom, so that schools in these communities don't have to choose which student to give the textbook to, to ensure that there is readily treatment for people before symptoms get serious, before people are unable to make money during droughts or famine. The best path forward uses AI to give people more choices while also making sure no one gets left behind. When technology does both things, it becomes more than just a helpful tool. It becomes a true partner in helping people live better lives.

**Works Cited**


Agarwal, Manisha, et al. 'Diabetic Retinopathy Screening Guidelines for Physicians in India: Position Statement by the Research Society for the Study of Diabetes in India (RSSDI) and the Vitreoretinal Society of India (VRSI)-2023'. International Journal of Diabetes in Developing Countries, vol. 44, no. 1, Jan. 2024, pp. 32–39, https://doi.org/10.1007/s13410-023-01296-z.

'AI in Agriculture: Progress in Implementing the European Union Coordinated Plan on Artificial Intelligence (Volume 2)'. OECD, 2026, https://www.oecd.org/en/publications/progress-in-implementing-the-european-union-coordinated-plan-on-artificial-intelligence-volume-2_3ac96d41-en/full-report/ai-in-agriculture_c9ac6d24.html.

Aijaz, Nazish, et al. 'Artificial Intelligence in Agriculture: Advancing Crop Productivity and Sustainability'. Journal of Agriculture and Food Research, vol. 20, Apr. 2025, p. 101762, https://doi.org/10.1016/j.jafr.2025.101762.

'AIR Research | AIR Creates a Virtual Hospital, Enabling AI Doctors to Self-Evolve-AIR'. Tsinghua.Edu.Cn, 2024, https://air.tsinghua.edu.cn/en/info/1007/1872.htm.

'Artificial Intelligence Adoption in African Higher Education - Ashesi University'. Ashesi University, 7 Dec. 2025, https://ashesi.edu.gh/ashesi-spotlights/ai-adoption-in-african-higher-education/.

Awais, Muhammad, et al. 'Advancing Precision Agriculture Through Digital Twins and Smart Farming Technologies: A Review'. AgriEngineering, vol. 7, no. 5, May 2025, p. 137, https://doi.org/10.3390/agriengineering7050137.

Beaton, Andrea, et al. 'Secondary Antibiotic Prophylaxis for Latent Rheumatic Heart Disease'. New England Journal of Medicine, vol. 386, no. 3, Jan. 2022, pp. 230–40, https://doi.org/10.1056/nejmoa2102074.

Bowles, Timothy M., et al. 'Long-Term Evidence Shows That Crop-Rotation Diversification Increases Agricultural Resilience to Adverse Growing Conditions in North America'. One Earth, 2020, https://agris.fao.org/search/en/providers/122535/records/65df9b190f3e94b9e5da6a1e.

Brant, Arthur, et al. 'Performance of a Deep Learning Diabetic Retinopathy Algorithm in India'. JAMA Network Open, vol. 8, no. 3, Mar. 2025, pp. e250984–84, https://doi.org/10.1001/jamanetworkopen.2025.0984.

'Digital Agriculture and AI| Innovation at FAO | FAO'. OpenInnovationNetwork, 2022, https://www.fao.org/innovation/digital-agriculture-and-ai-innovation/en.

'Disrupting Education? Evidence on Technology-Aided Instruction in India | the Abdul Latif Jameel Poverty Action Lab'. The Abdul Latif Jameel Poverty Action Lab (J-PAL), 2015, https://www.povertyactionlab.org/evaluation/disrupting-education-evidence-technology-aided-instruction-india.

Innovation District. 'Novel AI Platform Detects Rheumatic Heart Disease - Children's National'. Innovation District, 16 Jan. 2024,

https://innovationdistrict.childrensnational.org/novel-ai-platform-detects-rheumatic-heart-disease/.

‘Ending Learning Poverty’. Worldbank.Org, World Bank Group, 12 Oct. 2022, https://www.worldbank.org/en/brief/2022/10/12/ending-learning-poverty?utm_campaign=X+Report&utm_source=hs_email&utm_medium=email.

Fahim, Yosri A., et al. ‘Artificial Intelligence in Healthcare and Medicine: Clinical Applications, Therapeutic Advances, and Future Perspectives’. European Journal of Medical Research, vol. 30, no. 1, Sept. 2025, https://doi.org/10.1186/s40001-025-03196-w.

Fan, Shenggen. ‘How Agriculture Can Improve Health and Nutrition’. World Economic Forum, 9 Apr. 2015, https://www.weforum.org/stories/global-cooperation/how-agriculture-can-improve-health-and-nutrition/.

Gaitan, Nicoleta Cristina, et al. ‘Integrating Artificial Intelligence Into an Automated Irrigation System’. Sensors, vol. 25, no. 4, Feb. 2025, pp. 1199–99, https://doi.org/10.3390/s25041199.

Global Affairs Canada. ‘Water in Developing Countries’. GAC, Government of Canada, 2017, https://www.international.gc.ca/world-monde/issues_development-enjeux_developpement/environmental_protection-protection_environnement/water-eau.aspx?lang=eng.

Grewer, Uwe, et al. ‘Crop Models as Decision Support Tools in Digital Agriculture: The APSIM-Based RotationARM Tool in Australia’. Computers and Electronics in Agriculture, vol. 247, June 2026, p. 111725, https://doi.org/10.1016/j.compag.2026.111725.

Henkel, Owen, et al. ‘Effective and Scalable Math Support: Evidence on the Impact of an AI-Tutor on Math Achievement in Ghana’. arXiv.Org, 5 May 2024, https://arxiv.org/abs/2402.09809.

‘India: Qure.Ai’s qXR Boosts TB Detection & Saves Costs’. Qure.Ai, 2025, https://www.qure.ai/news-press-coverages/Qure.ai-increases-TB-detection-while-saving-costs-shows-evaluation-in-India.

International Literacy Day 2025 Promoting Literacy in the Digital Era FACTSHEET Globally, Steady Progress Has Been Made Since World Leaders Adopted the Sustainable Development Goal 4 (SDG 4) on Education and Lifelong Learning.

Kamal, Mostafa, and Tarek Bablu. ‘Mobile Applications Empowering Smallholder Farmers: A Review of the Impact on Agricultural Development’. International Journal of Social Analytics.

Li, Fan, et al. ‘Agricultural Technology Extension and Farmers” Income: Evidence From China’. Economies, vol. 13, no. 11, Nov. 2025, p. 331, https://doi.org/10.3390/economies13110331.

Li, Ruozhu, and Qingzhi Liu. Reinforcement Learning for Timing-Optimized Greenhouse Climate Control: Integrating Growth Prediction, Market Dynamics, and Economic Constraints.

Mbabazi, Nestor, et al. ‘Congenital Heart Disease Cardiac Catheterization at Uganda Heart Institute, a 12-Year Retrospective Study of Immediate Outcomes’. BMC Cardiovascular Disorders, vol. 24, no. 1, Aug. 2024, https://doi.org/10.1186/s12872-024-04085-6.

Natchev, Vesselin. ‘Harvard ALI Social Impact Review’. Harvard ALI Social Impact Review, 12 Nov. 2024, https://www.sir.advancedleadership.harvard.edu/articles/harnessing-ai-empower-smallholder-farmers-bridging-digital-divide-sustainable-growth.

OECD. ‘About the OECD’. OECD, 2023, https://www.oecd.org/en.html.

OĞUZTÜRK, Gülcay ERCAN. ‘AI-Driven Irrigation Systems for Sustainable Water Management: A Systematic Review and Meta-Analytical Insights’. Smart Agricultural Technology, vol. 11, Aug. 2025, p. 100982, https://doi.org/10.1016/j.atech.2025.100982.

Paul, Marie, et al. Effect of Unmanned Aerial Vehicle (Drone) Delivery on Blood Product Delivery Time and Wastage in Rwanda: A Retrospective, Cross-Sectional Study and Time Series Analysis.

Putri, Aerin Regina, and Satya Apriliansyah. ‘Farming Efficiency and Productivity in Increasing Farmers' Income’. Journal of Agro Complex Development Society, Apr. 2026, pp. 18–23, https://doi.org/10.62012/agrocomplex.vi.33.

‘Rheumatic Heart Disease and Rheumatic Fever Programmes in the WHO African Region: Review of Reviews’. WHO | Regional Office for Africa, 2025, https://www.afro.who.int/publications/rheumatic-heart-disease-and-rheumatic-fever-programmes-who-african-region-review.

Sabeena, Anseena Anees, et al. ‘AI-Driven Predictive Analytics for Crop Rotation and Soil Health Management’. Advances in Economics, Business and Management Research, 2026, pp. 254–65, https://doi.org/10.2991/978-94-6239-624-1_18.

Schweng, Stefan, et al. ‘What Can Artificial Intelligence Do for Soil Health in Agriculture?’. Computer Science Review, vol. 59, Feb. 2026, p. 100832, https://doi.org/10.1016/j.cosrev.2025.100832.

unctad. ‘Why Family Farmers Need Greater Access to Technology | UNCTAD’. Unctad.Org, 13 Apr. 2023, https://unctad.org/news/why-family-farmers-need-greater-access-technology.

University Communications. ‘Microsoft Competition at Ashesi Showcases Student-Led AI Solutions Across Sectors’. Ashesi University, 27 Apr. 2026, https://ashesi.edu.gh/microsoft-competition-at-ashesi-showcases-student-led-ai-solutions-across-sectors/.

US EPA, OW. ‘Soil Moisture-Based Irrigation Controllers’. Www.Epa.Gov, 25 Jan. 2017, https://www.epa.gov/watersense/soil-moisture-based-irrigation-controllers.

———. ‘Weather-Based Irrigation Controllers’. Www.Epa.Gov, 17 Mar. 2021, https://www.epa.gov/watersense/weather-based-irrigation-controllers.

US EPA,OW. ‘Outdoors | US EPA’. US EPA, 22 Aug. 2016, https://www.epa.gov/watersense/outdoors.

WHO. “Health Workforce.” Www.Who.Int, https://www.who.int/teams/health-workforce. Accessed 6 Oct. 2026.

“1.1 TB Incidence.” Who.Int, 2024, https://www.who.int/teams/global-programme-on-tuberculosis-and-lung-health/tb-reports/global-tuberculosis-report-2025/tb-disease-burden/1-1-tb-incidence. Accessed 6 Oct. 2026.

“A Case Study on Global Deployment of qXR for TB.” Qure.Ai, 2017, https://www.qure.ai/blog/scaling-up-tb-screening-with-ai-deploying-automated-x-ray-screening-in-remote-regions. Accessed 6 Oct. 2026.

Sinha, Chaitali. “Global South-Led Responsible AI Solutions to Strengthen Health Systems: An Emergent Research Landscape.” Oxford Open Digital Health, vol. 3, Oxford University Press (OUP), 2025, https://doi.org/10.1093/oodh/oqaf016.

Elias, Martin. “From Chalkboards to Chatbots : Evaluating the Impact of Generative AI on Learning Outcomes in Nigeria.” World Bank, 2025, https://documents.worldbank.org/en/publication/documents-reports/documentdetail/099548105192529324. Accessed 7 Oct. 2026.

Brini, Marco. The Digital Transformation of Food Systems. v3, Technical report, 7 Sept. 2023.

Sen, Amartya. Development as Freedom.

Nussbaum , Martha. Creating Capabilities: The Human Development Approach. 2011. Harvard University Press, 2011.

Rosnah, U. S., F. C. Liufeto, and S. P. Manongga. “Family Farming’s Contribution in Increasing Household Income and Stunting”. Interdisciplinary Social Studies, vol. 2, no. 4, Jan. 2023, pp. 1922-9, doi:10.55324/iss.v2i4.418.

‘Global Gains in Tuberculosis Response Endangered by Funding Challenges’. Who.Int, World Health Organization: WHO, 12 Nov. 2025, www.who.int/news/item/12-11-2025-global-gains-in-tuberculosis-response-endangered-by-funding-challenges. Accessed 9 Oct. 2026.

**Has the Bank of England Regained Credibility? By Vinoothna Bondalapati**


## Abstract

The 2021- 23 inflation surge had damaged the Bank of England's credibility. This paper asks how much has been regained, testing the hypothesis ($H_1$) that the recovery is only partial. Using the bank's household Inflation Attitudes Survey data, it tracks three measures before, during and after the price shock: the gap between households' five-year inflation expectations and the 2% target, how sensitive those expectations are to current inflation, and net satisfaction with the Bank. The results are mixed. Long-run expectations stayed anchored while inflation was high, yet they have since moved significantly further above target, and net satisfaction has won back only about a fifth of its losses. Sensitivity to current inflation is the only measure that has returned to its pre-surge level. Credibility has only partly recovered, but not by enough to be taken for granted, and that makes it harder to justify looking through the 2026 energy shock.


## 1. Introduction

In September 2026, the Bank of England's (BoE) Monetary Policy Committee (MPC) voted 6-3 to hold the bank rate at 3.75 with 3 of its members preferring immediate rise. The split arose from an energy shock where oil and gas prices rose sharply and CPI inflation at 3.1% in August 2026 is projected to exceed 4% by early 2027. The committee saw little evidence of second-round effects in prices and wage setting materialising, however they do recognize the growth in risk as energy prices stay high over time. The majority argued that a possible second-round could be contained through tighter financial conditions with higher market rates' impact on borrowing and economic slack as a result of weak labour demand, so the bank can look through for clear evidence before acting. The minority of the members argued that the inflation would peak around the same time as the 2027 wage round, so an earlier rise in interest rates will be needed to keep expectations anchored. The dilemma in action is present because for a central bank to pursue short-run stabilizing policy, long run inflation expectations have to be anchored to the target (Carlos Carvalho et al 2023). This is more about the credibility and how firmly expectations are anchored than about the energy shock in itself and the safety in making the decision to look through inflation depends on whether the credibility lost during the 2021-2023 inflation surge has been rebuilt. This paper explores whether the Bank of England regained the credibility lost in 2021-2023. It tests 3 possible outcomes. The main hypothesis being credibility is partly restored but remains fragile as the long-term household expectations move towards the target but remain above it. The null hypothesis developed is that credibility has been fully recovered where economic agents perceive inflation as transitory and sensitivity of expectations are back to prior 2021. The alternative hypothesis is that the United Kingdom (UK) has recovered more slowly when compared to the eurozone that experienced the same previous inflation surge. These hypotheses predict different patterns from collected data, pointing to different answers for the 2026 dilemma.

## 2. Credibility and Inflation expectations

Credibility in central banks refers to economic agents' belief that a policymaker will adhere to set inflation targets over time, overcoming any fundamental problem with time inconsistency where discretionary monetary policy faces incentive to induce inflation for temporary boost in output and reduce unemployment. Extending this meaning of central bank credibility, private agents observe decisions over time where a commitment to price stability over a set target builds institutional credibility.

In the context of the UK, the BoE uses an inflation targeting framework that utilizes public announcement of a set numerical target for the rate of inflation it plans to maintain over the medium-run. Here, the primary role of credibility is to anchor long-term inflation expectations, prevent short-term price shocks from triggering second round effects of wage-price spirals and inflation persistence. If the central bank has a reputation for effective communication and accurate forecasting that builds public confidence in the bank's ability to stay committed to inflation targets in the medium run, any risks associated with the self-fulfilling, expectations driven inflation spirals will be decreased. This impacts wages and price bargaining considering nominal prices and wages are considered sticky where inflation erodes the purchasing power of wages, consequently, shaping the labour negotiation and pricing behaviour. There is also risk associated with unanchored expectations, often driven by lack of confidence and credibility in monetary authority, where they can convert transitory price shocks into prolonged inflation surges.

Macroeconomic models tend to rely on financial market yield curve data, based on the behavior of investors' change in demand to bonds as a compensation for decreased purchasing power caused by inflation, to proxy market inflation expectations. Olivier Coibion and Yuriy Gorodnichenko (2015) prove how household expectations are a better reflection of real-world firm pricing behaviour and explain nuances in price dynamics that markets often fail to account for. This considering historical data, regressed philips curve or economic downturns like Great Recession, supporting firm pricing being better aligned with household expectations rather than price experts and accounting that most of the price setters in the economy being small and medium sized enterprises that don't operate with a professional forecasting team.

### 2.1The reason behind chosen factors to provide direct evidence towards credibility.

The gap of medium-term expectations from the target is used because expectations remain well anchored at the official target regardless of temporary disruptions if the central bank still has credibility in the eyes of households. Rudolfs Bems et al. includes deviation of long-term mean inflation forecasts of agents from targets as a component in constructing international credibility index.

Sensitivity to CPI in regards with the effect on current inflation on movement of long-term expectations. This choice for proxy is supported by theory on how CPI inflation spikes or forecasts errors do not lead to increase in long-run inflation expectations if credibility is

sustained. Higher sensitivity to CPI inflation or forecast errors signals towards unanchored beliefs and poor historical policy performance.

Net Satisfaction with the BoE referring to public trust in the central bank. Households expectations dictate wages and prices and considering household expectations tend to be more reliable that market expectations, the credibility of institutions relies on public perception and understanding in the central bank's decision. This choice is evidenced by Andrew Haldane and Michael McMahon (2018) where they evaluate BoE survey data to prove positive correlation between public satisfaction and credibility.

## 3. Data and Methodology

This paper primarily relies on secondary data gathered from official sources including Bank of England inflation attitudes survey, Office of national statistics CPI, and European central bank consumer expectations survey. Additionally, it relies on academic literature including but not limited to research papers, review articles and journals for more comprehensive understanding and utilization of economic theory behind credibility, framework, mechanisms and methodology development.

The factors are constructed in data ranges of 2015 to 2021, 2021 to 2024, and 2024 to 2026. These specific boundaries exist because to determine the extent to which credibility has been regained, credibility has to be compared with credibility under normal conditions and credibility when damaged due to inflation surge. 2015 Q1 to 2021 Q2 is considered as the pre-shock baseline which is used as a benchmark for determining recovery. Inflation was last at 2% target in July of 2021 and did not return to there until may of 2024. The gap can be a direct measure of lost credibility as anchors can be lost when expectations differ from that of target. 2021 Q3 to 2024 Q2 represents the damage to the credibility from which this paper attempts to measure if credibility has recovered. This range covers policy tightening reaching interest rates up to 5.25% and CPI peak at 11.1% that is 9.1% higher than target. The third period is from 2024 Q3 to 2026 Q1 with the first cut in bank rate since the inflation surge and period closes before 2026 energy shock hits, data is focused on measuring recovery.

### 3.1 Formulas used for calculations

$$Gap_t = E_t[\boldsymbol{\pi},\ 5\ years\ ahead] - 2.0$$

The factor compares the constricted average gap for each period in inflation expectations from target. .

$$Sensitivity = E_t[\boldsymbol{\pi},\ 5\ years\ ahead] = \alpha_p + \beta_p \times CPI_t + \varepsilon_t$$

$CPI_t$ is the latest annual rate published before the survey was conducted. A $\beta$ close to zero indicates long-term inflation expectations ignoring current inflation, meaning anchored, credible expectations. A positive and rising $\beta$ means that the public takes into consideration the current inflation when forming long-term inflation expectations.

Net satisfaction is collected from an inflation attitudes survey published by the Bank. Using this data, recovery ratio using mean net satisfaction, and trend within period 3 is calculated to be compared across the periods (P).

$$Recovery\ ratio\ =\ (mean\ P3\ -\ mean\ P2)/(mean\ P1\ -\ mean\ P2)$$

**3.2 Statistical approach**

Each proxy is compared across the three periods, and due to the survey data changes and small samples, an observed difference might arise by chance. To judge whether a difference is real - each comparison includes a p-value that is the probability of observing a difference when in reality there is no difference at all. A p-value below 0.05 is treated as statistically significant as a difference this large would occur by chance less than 5%. A p-value measures how surprising the data would be if there were no real differences in data.

Differences between period averages uses Welche's test, this does not assume that the periods are equally variable. To check whether the gap differs from zero, one sample t-test is used. Each slope's p-value is calculated by dividing the slope from its standard of error. Differences between two periods' slopes are tested with $z\ =\ (b_1\ -\ b_2)\ /\ \sqrt{(SE_1{}^2\ +\ SE_2{}^2)}$. All the calculations were carried out on excel and organised in tables.

It is also important to consider that with small sample sizes, p-value below 0.05 shows that the data is too limited to confirm no changes occurred. Statistical significance only indicates whether there is a real pattern or a coincidence.

**3.3 Limitations**

This methodology carries inherent limitations to be considered. The inflation attitudes survey is quarterly so fewer observations. Period 3 is especially small meaning the tests will have lower statistical power.

The UK and euro area have different survey horizons. The UK IAS question asks about inflation in the next five years. The ECB consumer expectations survey's museum term measure is three years ahead. The ECB added a five year ahead question in August of 2022 which can be used for comparisons. However, the ECB 3 year series is used for other periods which is a mismatch for results comparisions.

It's also important to consider the lack of consistency due to breaks in the survey series. In May of 2020, the IAS moved to online survey, which is a methodological break. The survey also switched from kantar to Ipsos in February 2022 which is a change in provider. To address this, the period 1 baseline is in two ways, 2015 Q1 - 2021 Q2 fully and 2020 Q2 - 2021 Q2 as a part representing online only. This is to show that conclusions don't depend on the break.

**4. Data Results - Has credibility recovered**

This section is to report the three credibility proxies across the pre-surge period, the surge period, and the recovery period.

### 4.1 Gap from target

The gap quantifies how far inflation expectations are from the target. If BoE has credibility, then more people will believe the bank will be able to hit the target and the gap between inflation expectations and target will be small. Large gap indicated households expecting the bank to miss the set target, implying lower public trust in BoE.

| Period averages | | | | | |
|---|---|---|---|---|---|
| Period | Quarters (n) | an expectation ( | Mean gap (pp) | SD of gap | Gap > 0 (p) |
| P1: 2015 Q1–2021 Q2 | 26 | 3.17 | 1.17 | 0.34 | < 0.001 |
| P2: 2021 Q3–2024 Q2 | 12 | 3.13 | 1.13 | 0.17 | < 0.001 |
| P3: 2024 Q3–2026 Q3 | 9 | 3.54 | 1.54 | 0.29 | < 0.001 |
| | | | | | |
| Differences within periods | | | | | |
| Comparison | Difference (pp) | p | | | |
| P2 – P1 | −0.04 | 0.66 | | | |
| P3 – P1 | 0.38 | 0.006 | | | |
| P3 – P2 | 0.41 | 0.003 | | | |
| | | | | | |
| Trend within P3 | | | | | |
| Sample | Slope (pp per quarter) | SE | p | | |
| All quarters | 0.017 | -0.04 | 0.69 | | |
| Excluding Q3 2026 | 0.082 | -0.016 | 0.002 | | |

Table 1. Household five year inflation expectations and the gap from 2% target by period.
Source- Bank of England Inflation Attitudes Survey. "Gap > 0" is a one-sample t-test. 'Differences within periods' uses Welch's t-test. 'Trend within P3'is an OLS regression of the gap on quarter number.

Figure 1 shows five year household expectations rose well above 2% during the inflation surge. The average gap of expectations from inflation target between P1 and P2 remained quite similar at 1.17 and 1.13. This shows that the anchor held during the crisis even with CPI reaching about 11%, five year expectations were no further above the target than pre 2021. However, the average gap increased for period 3 to 1.54 points, significantly further from the target and earlier periods. This indicates the damage to inflation expectations weakened confidence in bank post inflation surge. This can be explained through recency bias where agents tend to overvalue underperformance and recent performance which had an effect even after the inflation came down to the target. Within P3, expectations rose steadily from 3.2% in Q3 2024 to 3.9% in Q2 2026 then dropped to 3.0% in the recent quarter 2026 Q3. Inflation expectations in P3 till Q2 2026 increased by 0.08 points per quarter, highly unlikely to reflect a coincidence since p = 0.002. The inflation expectation of Q3 2026 used the inflation expectation from only July and August, and considering it is the most recent movement, it should be read with caution. The drop in expectations can indicate real improvement in inflation expectations or change of method, an answer that is inconclusive due to current limited data.

### 4.2 Sensitivity

| Slopes by period | | | | | | |
|---|---|---|---|---|---|---|
| **Period** | **n** | **CPI range (%)** | **Slope** | **SE** | **p** | **$R^2$** |
| P1 | 26 | 0.0–3.0 | 0.24 | -0.05 | < 0.001 | 0.49 |
| P2 | 12 | 2.1–10.7 | 0.01 | -0.02 | 0.45 | 0.06 |
| P3 | 9 | 2.0–3.8 | 0.27 | -0.17 | 0.17 | 0.25 |
| | | | | | | |
| **Differences in slopes** | | | | | | |
| **Comparison** | **Difference** | **z** | **p** | | | |
| P2 − P1 | −0.22 | −4.33 | < 0.001 | | | |
| P3 − P2 | 0.26 | 1.47 | 0.14 | | | |
| P3 − P1 | 0.03 | 0.19 | 0.85 | | | |

Table 2. Sensitivity of five-year inflation expectations to current CPI inflation
Notes- $R^2$ is the share of the variation in expectations that CPI explains.
Sources- Bank of England Inflation Attitudes Survey; ONS. Robustness: excluding quarters after the May 2020 survey change, the P1 slope is 0.22 (SE 0.04).

This measure uses slope to quantify the effect of current inflation on households' confidence and five-year expectations. A slope of 0.24 in P1 means for every 1 point rise in current CPI, the five year expectations rose by 0.24 points. A slope near 0 is the sign of a credible central bank.

The results show that during the inflation surge in P2, households looked past the high inflation with slope being 0.01. Because the CPI moved over such a wide range, this result is not an accident caused by a small data set. This is also supported by $R^2$ which refers to how closely CPI and inflation expectations move together. While CPI explained about half the movement in inflation expectations in P1, inflation only explained 6% of the movement in the long term expectations during the shock. Hence, this measure is consistent with credibility returning to normal evidenced by households not extrapolating high inflation to expectations and lack of lasting damage implied by results.

Since 2024 in P3, the sensitivity is back to normal at 0.27 which is almost identical to pre-2021 with p=0.85. However, it is essential to account for uncertainty in P3 considering only 9 quarters of data and CPI moving between 2% and 4%, the estimate has a wide margin of error. This evidence is consistent with sensitivity returned to normal, but cannot confirm it. Restricting P1 to quarters before the 2020 online survey changes to assess the impact of survey changes also gives a similar slope of 0.22.

### 4.3 Public trust

Table 3. Net satisfaction with the Bank of England.
Notes- Differences use Welch's t-test.
Sources- Source: Bank of England Inflation Attitudes Survey.

| Period averages | | | |
|---|---|---|---|
| **Period** | **n** | **Mean net satisfaction** | **SD** |
| P1 (full) | 26 | 28.2 | 4.24 |
| P1 (online only) | 21 | 22.2 | 2.59 |
| P2 | 12 | −3.75 | 11.5 |
| P3 | 9 | 1.78 | 2.82 |
| | | | |
| **Recovery and differences** | | | |
| **Measure** | **Value** | **p** | |
| P3 − P2 | 5.53 | < 0.001 | |
| Effect size (Cohen's d), P3 vs P2 | 0.62 | n/a | |
| P3 − P1 | −26.4 | < 0.001 | |
| Recovery ratio (full baseline) | 17.30% | n/a | |
| Recovery ratio (online baseline) | 21.30% | n/a | |
| | | | |
| **Trend within P3** | | | |
| **Slope** | **SE** | **p** | |
| −0.07 | -0.39 | 0.87 | |

The survey asks directly whether households are satisfied with the BoE's performance in using interest rates to control inflation. Net satisfaction was collected from IAS, where a negative number means more people are unhappy with the performance than happy. Unlike the previous measures being used to infer credibility, this question asks directly about it.

Before 2021, the mean net satisfaction over the period is 28.2 with a clear majority being satisfied. P2 mean is -3.75, more households dissatisfied during the price shock. Since 2024, the households seem to be more trusting as the mean net satisfaction increased to about 1.78, barely positive but higher than P2. That recovers only 17% of the lost trust measured against full P1 baseline or 21% against the online only baseline. The weak recovery cannot be explained by the change in survey method considering both the baselines give similar answers. P3 remains significantly below P1 net satisfaction. Within P3, the net satisfaction shows no trend (Slope is -0.07 points per quarter, p=0.87) which suggests that the recovery has stalled after 2024.

### 4.5 Summary

| | P1 | P2 | P3 | P3 vs P1 |
|---|---|---|---|---|
| Gap from target (points) | 1.17 | 1.13 | 1.54 | Higher (p = 0.006) |
| Sensitivity (slope) | 0.24 | 0.01 | 0.27 | No difference (p = 0.85) |
| Net satisfaction | 28.2 | −3.75 | 1.78 | Lower (p = 0.87) |

Together, the proxies show that damage done to credibility was delayed rather than immediate. During the price surge, long-term expectations held their level and ignored current inflation even as trust in BoE collapsed. Since 2024, the sensitivity has returned to normal while expectations sit significantly higher than the target and trust has recovered by a fifth of what was

lost. This pattern rules out full recovery and is more consistent with credibility that is partly restored but fragile in comparison to before.

**5. Discussion and conclusion**

The three results show that inflation targeting mechanism still works as sensitivity is back to pre 2021 level and households treat inflation spikes as transitory. This is the condition of expectations that enables the back to even consider the action of looking through inflation. Considering 5 year expectations are around 3.5%, the wages and prices will start from a higher perceived and assumed rate of inflation. However, the net satisfaction being close to zero brings into cushion whether public trust will regain in the future after barely recovering during the last inflation surge. In real time, the anchored but fragile credibility describes public trust that still expects inflation spikes as temporary but at a higher baseline.

If credibility is fragile, the minority's opinion in MPC is not about energy shock requiring higher rates but the bank can no longer assume that households will give it time to look through without impacting long term expectations and public confidence. BoE's history is also quite relevant, the bank judged rising inflation to be transitory during the 2021-2023 surge and the cost is visible in net satisfaction data. Looking through a second time draws public skepticism in accepting the same judgement.

The majority's position is supported by the data results supporting long term household expectations not responding to current inflation. Even during the worst inflation outcome, the households did not extrapolate high inflation with long term inflation expectations. The Q3 2026 reading also shows that inflation expectations fell to 3.0% in the latest quarter, if this trend continues even after the change in survey provider then the inflation expectations may already be revising.

Two results out of three evidence $H_1$. In this context, the action of looking through is weaker and riskier than in 2021, but is not completely gone. This supports a conditional look through where BoE holds rates while sensitivity remains low, but acts immediately to expectations or wage settlements start changing. However, it is imperative to consider the delay in transmission of monetary policy as rise in interest rates will not have immediate effect.

Has the Bank of England regained credibility lost in 2021-2023? The answer proves to be only partly. Households treat inflation spikes temporarily, long term expectations sit further above the 2% target than during or before the surge, and the public trust has recovered only to some extent.

**Works Cited**


"Inflation Attitudes Survey." Bank of England, 2015–2026

Office for National Statistics. "CPI ANNUAL RATE 00: ALL ITEMS 2015=100 - Office for National Statistics." Ons.Gov.Uk, 16 Apr. 2025.

Barro, Robert J. "Reputation in a Model of Monetary Policy with Incomplete Information." Journal of Monetary Economics, vol. 17, no. 1, 1986, pp. 3–20.

Bems, Rudolfs, et al. Expectations' Anchoring and Inflation Persistence. IMF Working Paper No. WP/18/280, International Monetary Fund, Nov. 2018.

Bernanke, Ben S. Forecasting for Monetary Policy Making and Communication at the Bank of England: A Review. Bank of England, 12 Apr. 2024.

Carvalho, Carlos, et al. "Anchored Inflation Expectations." American Economic Journal: Macroeconomics, vol. 15, no. 1, 2023, pp. 1–47.

Coibion, Olivier, and Yuriy Gorodnichenko. "Is the Phillips Curve Alive and Well after All? Inflation Expectations and the Missing Disinflation." American Economic Journal: Macroeconomics, vol. 7, no. 1, 2015, pp. 197–232.

Gürkaynak, Refet S., Andrew Levin, and Eric Swanson. "Does Inflation Targeting Anchor Long-Run Inflation Expectations? Evidence from the U.S., UK, and Sweden." Journal of the European Economic Association, vol. 8, no. 6, 2010, pp. 1208–42.

Haldane, Andrew, and Michael McMahon. "Central Bank Communications and the General Public." AEA Papers and Proceedings, vol. 108, 2018, pp. 578–83.

Kydland, Finn E., and Edward C. Prescott. "Rules Rather than Discretion: The Inconsistency of Optimal Plans." Journal of Political Economy, vol. 85, no. 3, 1977, pp. 473–92.

Reis, Ricardo. Losing the Inflation Anchor. Conference Draft, 2021.

Weber, Michael, et al. The Subjective Inflation Expectations of Households and Firms: Measurement, Determinants, and Implications. NBER Working Paper No. 30046, National Bureau of Economic Research, May 2022

Aoki, Kosuke. "Optimal Monetary Policy Responses to Relative-Price Changes." Journal of Monetary Economics, vol. 48, no. 1, 2001, pp. 55–80.

Bank of England. "Bank Rate Maintained at 3.75% – September 2026 Monetary Policy Summary and Minutes." Bank of England, 17 Sept. 2026.

"Rules, Discretion and Reputation in a Model of Monetary Policy." Journal of Monetary Economics, vol. 12, no. 1, 1983, pp. 101–21.

Christelis, Dimitris, et al. "Trust in the Central Bank and Inflation Expectations." International Journal of Central Banking, vol. 16, no. 6, 2020, pp. 1–37.

D'Acunto, Francesco, et al. "Exposure to Grocery Prices and Inflation Expectations." Journal of Political Economy, vol. 129, no. 5, 2021, pp. 1615–39.

Malmendier, Ulrike, and Stefan Nagel. "Learning from Inflation Experiences." Quarterly Journal of Economics, vol. 131, no. 1, 2016, pp. 53–87.

**Examining Factors Affect Online Review Helpfulness By Elise Duka**


## Abstract

This study examines which characteristics of Amazon book reviews are associated with review helpfulness. The analysis focuses on four factors: review rating, review text length, verified purchase status, and image inclusion. Amazon book review data was analyzed using correlation analysis and simple linear regression to examine the relationships between these variables and the number of helpful votes received by each review. The findings show that the image inclusion has a positive relationship with helpfulness, while rating, review text length, and verified purchase status have weaker relationships. Higher ratings and longer reviews are associated with more helpful votes, but these variables explain only a small amount of the variation in helpfulness. Verified purchase also shows a small positive relationship. Overall, the results suggest that review helpfulness is influenced by multiple characteristics rather than one single factor. The findings provide insight into how different features of online book reviews may affect how useful consumers find them.


## Introduction

Online reviews play a role in what people decide to buy. Online reviews help consumers evaluate product quality and reduce uncertainty before making a purchase decision (Chatterjee 2025). Before you purchase a product many people want to learn about the experiences people had with the product before buying it for themselves. The reviews can influence consumer opinions on the products, build trust, and increase sales. Thus, it's important to understand what makes a review helpful. In this paper, I aim to investigate the specific features such as the rating, review text, verified purchase, and images that affect how useful a review of the product appears. This study mainly focuses on identifying what factors and how they influence the helpfulness of online book reviews. To examine the relationships, this study uses correlation analysis and simple linear regression to analyze Amazon book review data. By focusing on online book reviews, this study creates a better understanding of which review characteristics are most closely associated with review helpfulness.

## Literature Review

Previous research has shown that there are several factors that influence how helpful the review is. A recent review of the literature concluded that review helpfulness is influenced by review content, credibility, reviewer characteristics, and platform-related factors (Pocchiari). Recent research reviews suggest that helpfulness is influenced by the different characteristics of the online reviews, including the quality of the information, credibility, and the text presentation (Pocchiari,Proserpio, and Dover 2025). Reviews with very high or low ratings tend to maintain more attention, but it's not always the most helpful. Positive reviews might seem less trustworthy, because the buyer might seem less trustworthy because they would think the review could be biased or fake.

Another factor of the review is the length or the detail the review has. Research shows that review content characteristics, including the amount and quality of information provided, can influence whether consumers would consider a review helpful ( Wang, Karmaka, Wang, and Pei 2025). Studies suggest that long reviews are often seen as more helpful because they provide more information, including both positive, and negative points.  However, long or repetitive reviews might lose the reader's attention.

Credibility also plays an important role in reviews. Verified purchase labels can improve trust because they can provide consumers with additional evidence that the reviewer has experience with the product ( Choi and Leon 2025) Reviews marked as verified purchases are often seen as more trustworthy, because they can confirm if the reviewer actually bought the product or not (add citations).  A recent study found that verified reviews are generally seen as more helpful because consumers use a different verification with a greater credibility, and authenticity. (Journal of Retailing and Consumer Services) These reviews might receive helpful votes, but the verification alone does not guarantee the use.

In addition, some reviews include images.Visual information can imrve review helpfulness when the images provide meaningful evidence that supports the written description of the product (Chu, Liu, and Lie 2025). The images could help show what the product looks like, or how it’s used, which could increase the helpfulness of the product.  However, if the images do not add good information, the buyer's impact could be limited.

Overall, past research suggests that helpfulness is influenced by multiple factors, including the rating, review text, credibility, and images. However, no single factor fully explains why some of the reviews are more helpful than others.

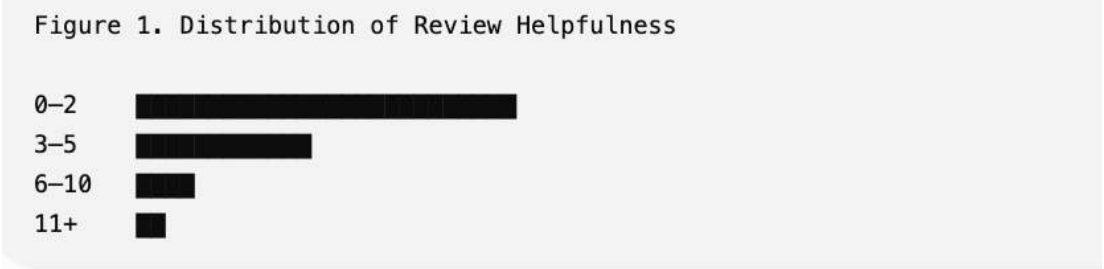
## Data and Method

### *Data*

The data that is used in the study comes from the Amazon book reviews. The samples were selected with books that were highly reviewed during a specific time period. . The data used in this study comes from Amazon book reviews because Amazon is a widely used online marketplace with a large number of customer reviews published in 2025.

The dependent variable is review helpfulness, shown by the number of helpful votes that each review had received. Figure one showed the distribution of review helpfulness across different helpful vote ranges. Most reviews received between 0, and 2 helpful votes, while fewer reviews received higher numbers of helpful votes..

Figure 1. Distribution of Review Helpfulness



The independent variables are the review text, rating, verified purchase, and images in the review. Review text length and rating are continuous variables and are summarized by using the mean, standard deviation, minimum, and maximum values. Verified purchase and image are dummy variables coded as 0 or 1, and are summarized by the % of reviews in each category. inclusion (0= no image, 1= Image included). Rating is measured on a scale from 1 to 5. Review text length is measured by the number of words in each review. Verified purchase is coded as 1 if the review is verified, and 0 if not. Image inclusion is coded as 1 if the review includes an image, and if it does not.

I use a bar chart to show the distribution of ratings. The chart shows that higher ratings are more common in the data set.

I used a bar chart to show verified purchases. This shows how many reviews were verified versus the ones not verified

I used a bar chart to show the image inclusion in reviews. This shows how many reviews included images compared to the reviews that did not.

The results show that there is a positive relationship between the image, and the helpfulness of the online review

***Method: Correlation Analysis***

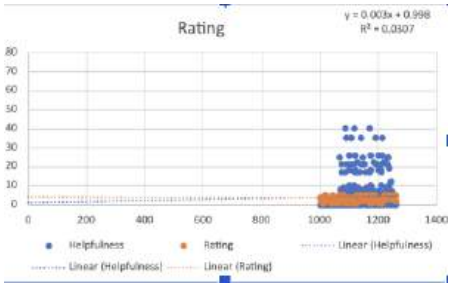
To show between variables and helpfulness,I used correlation analysis with simple linear regression. The dependent variable is helpfulness (y). The independent variables are review text length,rating,verify purchase, and image inclusion.

The regression equation used is: Y=bX+a

The R-squared ( ) value was also used to measure how much each variable explained the variation in helpfulness. Higher values indicate a stronger relationship.

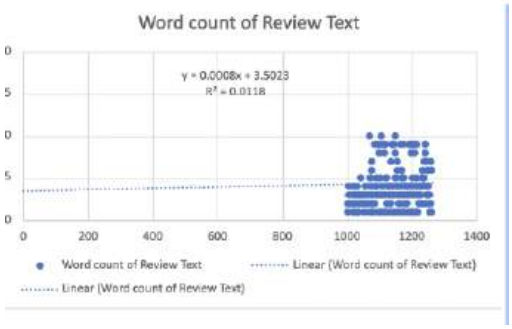
**Results and Discussion**

**Rating and Helpfulness:** Rating shows a weak relationship with helpfulness. The positive regression coefficient indicated that higher ratings are associated with slightly more helpful votes. Changes in rating slightly affect helpful votes, but the low R-squared suggests that rating does not strongly explain the helpfulness.



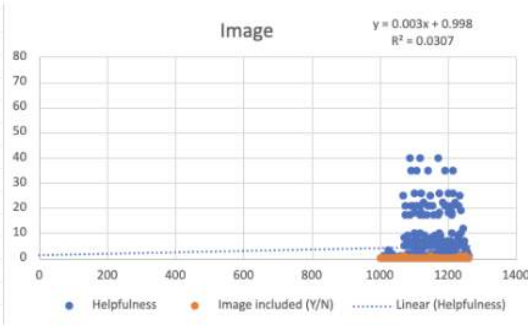
I created a scatterplot to analyze the relationship between the review text length, and the helpfulness.

**Review Text and Helpfulness:** Longer reviews tend to receive slightly more helpful votes, suggesting that detailed reviews may provide more useful information to consumers. However, the small effect and low R-squared indicate that text length alone does not strongly predict review helpfulness. only explains only a small portion of helpfulness.




**Verified Purchase and Helpfulness:** Verified Purchase has a small positive effect on helpfulness. Reviews marketed as verified tend to receive slightly more helpful votes, but the relationship overall is still weak.

**Image Inclusion and Helpfulness:** Image inclusion (0=no image, 1=image included) does not strongly affect the helpfulness. The results suggest that written content is more important than the images that determine the helpfulness for book reviews.



Verified purchases are common, while image inclusion had a mean of 0.59, with both the median and mode equal to one. Helpfulness goes across reviews, with most receiving a small number of helpful votes, and a few receiving more.

The results show that there is a positive relationship between the image, and the helpfulness of the online review. It implies that when an image is included in an online product review, it will increase the helpfulness of this online review.

**Conclusion**

Overall, the four variables offer valuable insights into review helpfulness and help explain some of the differences in how much consumers use the reviews. This research suggests that helpfulness is influenced more by factors such as clarity, writing quality, and detail. This research shows that the review helpfulness is positively associated with the inclusion of images, while rating, review length, and verified purchase status have weaker relationships with helpfulness.

Helpfulness is a complex thing, and cannot be explained by a single variable. Multiple factors interact to influence how users perceive a review.

This study is limited by its focus on one product category and a relatively small sample size. Future research could include larger samples and additional variables such as the sentiment, tone, or specificity of writing to better understand what drives helpfulness in online reviews.

**Works Cited**


Chatterjee, Swagato. “Effect of Construal Level on the Drivers of Online-Review-Helpfulness.” Electronic Commerce Research, vol. 25, no. 2, 2025, pp. 1115–1143. Springer. https://doi.org/10.1007/s10660-023-09716-2.

Choi, Hoon S., and Steven Leon. “When Trust Cues Help Helpfulness: Investigating the Effect of Trust Cues on Online Review Helpfulness Using Big Data Survey Based on the Amazon Platform.” Electronic Commerce Research, vol. 25, no. 3, 2025, pp. 1983–2018. https://doi.org/10.1007/s10660-023-09726-0.

Chu, Yongjie, Xinru Liu, and Cengceng Liu. “The Role of Visual Cues in Online Reviews: How Image Complexity Shapes Review Helpfulness.” Journal of Theoretical and Applied Electronic Commerce Research, vol. 20, no. 3, 2025, article 181. https://doi.org/10.3390/jtaer20030181.

Hong, Hong, Di Xu, G. Alan Wang, and Weiguo Fan. “Understanding the Determinants of Online Review Helpfulness: A Meta-Analytic Investigation.” Decision Support Systems, vol. 102, 2017, pp. 1–11. https://doi.org/10.1016/j.dss.2017.08.001.

Huang, Albert H., Kuanchin Chen, David C. Yen, and Trang P. Tran. “A Study of Factors That Contribute to Online Review Helpfulness.” Computers in Human Behavior, vol. 48, 2015, pp. 17–27. https://doi.org/10.1016/j.chb.2015.01.010.

Lutz, Bernhard, Nicolas Pröllochs, and Dirk Neumann. “The Longer the Better? The Interplay Between Review Length and Line of Argumentation in Online Consumer Reviews.” arXiv, 2019. https://arxiv.org/abs/1909.05192.

PowerReviews. The Ever-Growing Power of Reviews. 2023. https://www.powerreviews.com/power-of-reviews-2023/.

Ulu, Sevincgul, Can Uslay, and Sengun Yeniyurt. “Is This Review Helpful? How Product Gender, Verification, and Valence Shape Perceived Helpfulness.” Journal of Retailing and Consumer Services, vol. 87, 2025, article 104356. https://doi.org/10.1016/j.jretconser.2025.104356.

Wang, Shan, Shubhashis Karmakar, Fang Wang, and Yanli Pei. “Content Dissimilarity and Online Review Helpfulness: Contextual Insights.” Journal of Business Research, vol. 187, 2025, article 115250. https://doi.org/10.1016/j.jbusres.2024.115250.

# Why Small Food Businesses Remain: Community and Economic Endurance in the Mid Mon Valley By Owen Ryan


## Abstract

Many hospitality studies emphasize the need for a large consumer base and a thriving area as location is essential. However, in economically challenged areas like the Mid Mon Valley, where population numbers are dwindling and poverty and unemployment run rampant, it seems illogical that so many food businesses choose to remain in these areas. This study attempts to pinpoint what causes these local food businesses to stay in the Mid Mon Valley. The study utilized an attempted census of all food businesses in the Mid Mon Valley from Clairton to Brownsville, following all ZIP Codes that touched the river and were in between or equal to these boundaries. The study found statistically significant evidence that a business's decision to remain in the Mid Mon Valley correlated best with the reported importance of their community and their community's support of said business through a Spearman's Rank Correlation Test that resulted in a $r_s$ of 0.439 and a p-value of 0.0068 on an α of 0.01. In addition, financial benefits were observed to be important but did not have statistically significant backing from a Spearman's Rank Correlation Test giving a p-value of 0.222 on an α of 0.01. Through this study's findings, it is evident that local governments and economic development organizations must work to leverage this community engagement whilst fostering nonlocal support and communication with the businesses to mediate the gaps in incentives for staying in the Mid Mon Valley.


## Introduction

The Monongahela Valley (Mon Valley) is a fluvial landform that begins in Fairmont, West Virginia, and extends to Pittsburgh, Pennsylvania. This study focuses primarily on the middle portion of the Pennsylvania section of the Mon Valley, often referred to simply as the Mid Mon Valley.

The Mon Valley—a part of what is known as the Rust Belt—has a deep industrial legacy rooted in the coal, coke, glass, and steel industries. These boomed in the mid-to-late nineteenth century through to the end of the twentieth century. The economy of the Mon Valley was therefore closely tied to these industries, especially the steel industry. When the need for these goods dropped, the sustainability of maintaining expensive factories in the Mon Valley became a burden on many companies. With their livelihoods in the balance, management, owners, and unions failed to cooperate with one another, and the industry completely collapsed in the 1980s.

As such, the Mon Valley's economy was devastated, and this was reflected in the unemployment rates and population statistics. Pittsburgh was able to recover by pivoting from steel to healthcare, education, and even technology after steel collapsed. The same cannot be said for the middle and southern parts of the Mon Valley. While the local economy of these regions has seen some recovery, it is nowhere near what it once was during the steel boom. Many towns

remain very low in population and new commerce in the Mon Valley, displaying stagnant growth and severely limited clientele.

Such dreary circumstances and the limited customer base pose an intrinsic contradiction for business owners, namely those in the hospitality industry for food businesses. A food business's profitability and growth potential correlate strongly with its access to customers, and it would not make as much sense at face value to maintain, let alone start, a food business in these areas due to their overall economic health.

This raises an interesting and understudied research question: why do food businesses in areas like the Mid Mon Valley choose to remain? This study aims to explore this question by surveying food business owners in the Mid Mon Valley and utilizing census data to determine which factors appear to most directly correlate with food business ownership in the Mid Mon Valley.

**Literature Review**

The Mon Valley's story is one of many highs and lows. What began as a set of river communities would quickly flourish into the center for the steel and coal industries known as the Mon Valley. Although a major contributor to the iron, coal, and crucible steel industries, Pittsburgh and the Mon Valley did not see the major boom in industry they were known for until the area's first steel mill opened in 1875, starting a chain that would last over a century (Kobus 101). This was the eleventh mill in the United States to use the Bessemer Process, which was a steel-making process that allowed for cheap, mass-produced steel that did not take nearly as long as the older crucible methods (Kobus 101). This mill took advantage of the perfect location. The Pittsburgh area had both extremely easy access to coal and easy transportation along the river (Kobus 35, 46). In fact, coal could be found so easily in the Pittsburgh area that Colonials reportedly mined it as early as the mid-1700s (Kobus 35). In addition to the meeting of the three rivers of the Allegheny, Monongahela, and Ohio, the Pittsburgh area also had two other waterways in the Youghiogheny River and Conemaugh River, so the area had ample travel by water that led into the Mississippi River and to the ocean (Kobus 46). The Erie Canal was also accessible by traveling from the northern part of the Allegheny River on land to Buffalo, New York (Kobus 46).

With all of these benefits, it is no wonder that Pittsburgh and, by extension, the Mon Valley became the "Steel Capital of the World." The Mon Valley steel mills alone would reach a peak of employment around 120,000 (Skrabec, Jr. 109). The Mon Valley even produced the combined steel production of Germany, Japan, and Italy during World War II (Skrabec, Jr. 90). It was an area of wealth and determination. The lower-class immigrants often saw themselves or their children grow out of poverty and into skilled trades or management (Skrabec, Jr. 111). The prosperity could not last forever, however. By 1980, the steel mills of Pittsburgh and the Mon Valley were simply not competitive enough on a global scale (Hoerr 15).

The year 1982 was the year the collapse took root. De-industrialization and production cutbacks began to take full effect. There were many reasons for the collapse. Labor costs were a

major one. American companies paid out significantly more in labor per hour than international competitors like Japan (Skrabec, Jr. 119). In addition, recession, imports, and environmental regulation all took a toll on the American companies (Skrabec, Jr. 114). It was unsustainable. When the companies came to the unions and begged for concessions or wage cuts, the unions would turn them down because of chronic distrust on both sides (Skrabec, Jr. 112). The mills were operating at 43.8 percent capacity, and the economy was in a recession (Hoerr 16). The two sides needed to cooperate, but they could not. The issue was best summed up by a quote from a United Steel Workers union leader in 1983: "To the average union man, they're always crying wolf. And the wolf finally came" (Hoerr 23).

This wolf was ruthless. The end of 1981 saw the loss of over 200,000 steel jobs with 140,000 layoffs and 200 steel mills closing nationally (Skrabec, Jr. 85). Within only three months, 1982 saw more than 111,500 steelworkers laid off (Skrabec 89). The steel companies warned unions that it would get worse if something did not change. While at a union meeting at the end of July in the extravagant William Penn Hotel, a union leader named Lloyd McBride called a vote on concessions the steel companies made that he called "more than reasonable" (Hoerr 18). The union cried to reject it and began to cheer (Hoerr 18). But, after the cheers, McBride stood and reportedly said in relation to the union and the companies, "I say to you, the decision we have made today is not cause for celebration. The problems are mutual, and they have to be solved. This is a mutual failure" (Hoerr 19). He was correct.

By its end, the devastation reached "depression levels" of unemployment in the steel industry of 25 percent (Skrabec, Jr. 89). The damage was "monumental" with 1.2 million workers losing their jobs and 2,700 shutdowns throughout the nation (Skrabec, Jr. 87). 600 of these factory shutdowns were permanent, resulting in 250,000 jobs guaranteed to never come back (Skrabec, Jr. 87). The majority of those 1.2 million would never set foot on their factory or mill again.

The rapid collapse of the steel industry led to immense economic deterioration in Pittsburgh and the Mon Valley, resulting in a mass exodus of unemployed workers. The Mon Valley, however, lacked the resources and fluidity of Pittsburgh's urban sprawl. This forced an immediate and catastrophic change to a post-industrial reality, whereas Pittsburgh could cushion the fall of industry with the rise of new fields.

Now, Mon Valley towns that once held tens of thousands of residents often only hold a few thousand—even a few hundred. While Pittsburgh would recover by branching into the medical, technological, and financial fields, the Mon Valley could never hope to follow a major urban center's moves (Hoerr 13). It is stuck, lagging behind in the past with unemployment, deteriorating infrastructure, poor educational attainment, and a declining, aging population.

Despite this, the Mon Valley has seen a resilient and robust food scene. This brings the discussion back to the research question at hand: "why do food businesses in areas like the Mid Mon Valley choose to remain?" To this end, it makes sense to investigate the entrepreneurial mindset and the challenges faced by small businesses and hospitality.

Entrepreneurship is studied across several distinct fields and as such has many definitions. The type of entrepreneurship discussed here aligns with the Global Entrepreneurship Monitor's definition as "the act of starting or running a new business" (Hill et al. 20). Such entrepreneurs share several distinct personality characteristics that have been studied. Among them, high need for achievement, high self-confidence, opportunity obsession, moderate risk-taking, flexibility, and future orientation point to influences that could land someone in the Mon Valley (Principles of Business : Entrepreneurship 241–242). If someone thought the Mon Valley's situation might soon turn around, they may choose to locate therein to reap later profits. Perhaps they saw the Mon Valley as a larger challenge and thus a higher achievement if conquered.

Whether entrepreneur or not, everyone surveyed in this study has the shared trait of leading and representing a business in the hospitality industry, and all businesses share one thing: businesses want to maximize profits (Stringam 1; Schiller and Gebhardt 47). Consequently, managing and acquiring enough money to finance and maintain a business is always a challenge (Alsaati 18). Naturally, businesses in the hospitality industry need a stable and substantial customer base to remain competitive and in business. As such, "Location is an important element in the success of a hospitality business," and "Physical proximity to the origin markets of consumers creates a location-bound effect as well" (Stringam 69). A business without customers is destined to fail.

Labor issues also plague the hospitality industry (Stringam 70). Without access to labor, the business will struggle to keep up with the high turnover of employees that often manifests in the hospitality industry (Stringam 71).

One other consideration is that of competition. It is entirely possible that certain areas or towns may simply lack any other food-based businesses, resulting in the current business acting as a sort of local or spatial monopoly, whereby there is only enough customers for one business or the operations costs to add a second business are too great (Rutherford 501; Rutherford 786).

From this and beyond mere personality traits, several potential areas where the Mon Valley may be appealing can be identified. Perhaps costs or taxes are extremely low, which, therefore, increases profits. Maybe there is a reliable local or commuter community that supports the business—both options being integral to business, especially restaurant, success (Stringam 45; Stringam 69). Perhaps there is even a sturdy supply of cheap labor that enables the business to be more competitive and efficient. But maybe some would leave the Mon Valley but are locked after starting or taking over their business?

Starting a restaurant is not cheap. An article from Auguste Escoffier School of Culinary Arts places the median cost of opening a restaurant at $375,000 with the upper and lower quartiles being $175,500 and $750,500 respectively (“How Much Can It Cost to Start a Restaurant?”). Therefore, it is not an option for many businesses to simply move locations because of their previous choices. This aligns with what is known as path dependence, where "early decisions restrict what is achievable in the future" (Samuelsson et al. 322). Although applicable to all, this notion strongly follows what many owners who may have inherited or

purchased an existing establishment may encounter. If the previous owner decided to open their business in the Mon Valley because of past prosperity, current owners may lack resources or the capacity to change locations because of path dependence.

Similarly, what are known as sunk costs, "costs that were incurred in the past and cannot be recovered," may be affecting the ability of businesses to move on or relocate (Greenlaw et al. 33). These costs, however, should be ignored in present decisions to avoid falling for the sunk cost fallacy, where these past and unrecoverable costs play into the production and economic decisions of the business leaders (Greenlaw et al. 33; Rutherford 810).

Social capital may also be a key driver in why businesses chose to stay in the Mon Valley. Social capital is simply the networks and trust groups and individuals share in communities (Tripp et al. 1). When a business builds strong networks within communities, it gains advantages over competitors in both knowledge and customer basis among other things (Johnson 3-4). In other words, a business profits off of its relationships, and many of those relationships would be destroyed upon relocating.

Following this, it is entirely possible that the owner's connections to the Mon Valley are too emotionally difficult to part ways with, and as such, the owner chooses to stay in the Mon Valley from a sentimental standpoint or one of owing something to the community. This aligns with what is called place attachment, which is "one’s strong emotional bond with a specific location" (Debenedetti et al. 1). Furthermore, this place attachment may also affect the likelihood of locals spending money at local businesses because of their own attachment. A study conducted by Debenedetti et al. corroborates this as their data "suggest that places that consumers feel strongly attached to provide an experience that goes beyond what consumers believe that the market usually offers," which would positively aid local businesses (6).

Despite all of these potential reasons for why the Mon Valley contains so many food-based businesses, there is little that studies which of these actually contribute to the owners' decisions to stay. That is the void this study aims to address.

**Methodology**

For the purposes of this study, the Mid Mon Valley was defined as all ZIP codes touching the Monongahela River Valley including and in between the towns of Clairton (15025) and Brownsville (15417). All businesses not in ZIP code regions touching the river valley or geographically not in or between the designated regions were not considered in the results.

A thirty-two-question survey was made in SurveyMonkey Premier that asked businesses about their motives to stay in the Mid Mon Valley and how profitable said location was. The survey contained 17 Likert Scale questions for later analysis along with other numerical and business-related questions.

All nonnational or franchised restaurants and food-serving businesses with an online presence in the areas with valid ZIP codes touching the river valley were contacted through online means, most commonly email, Instagram, and Facebook Messenger, to see if they would be willing to participate in the study. To find these contacts, Google searches, public business

directories, and other online means were used to include as many businesses as possible. Furthermore, the Mon Valley Chamber of Commerce, Brownsville Area Revitalization Corporation, Washington County Chamber of Commerce, Mon Valley Alliance, and Mon Valley Independent were contacted to attempt to aid in passing around the survey even more.

In addition to this, the researcher drove around each of the valid ZIP codes extensively and gave each food-based business a short flyer with a QR code connected to the online survey to account for those businesses without an online presence and to drive up more responses.

All participants, who were current owners or managers of separate, active, and valid businesses, were briefly informed on the goals and general purpose of the study as well as the date it was due. In addition, all participants were to read an informed consent statement on what answering the survey meant, their privacy, the survey being completely voluntary, the fact that the survey could be stopped at any time, the fact that any question could be skipped, and the guarantee that they had for anonymity.

Once the day of required completion passed, the results were analyzed alongside census data of the towns and ZIP code regions. Analysis consisted of summary statistics, group comparisons, Cronbach's alpha, and Spearman's Rank Correlation Tests along with other graphs and number assessments.

**Results**

The final sample size was 31 businesses (n = 31). Note that the following tables and figures may display percentages that do not add up to 100% due to rounding, the ability to select multiple answers, or the data not being mutually exclusive.

| Business Metric | Median | Mean | Standard Deviation |
|---|---|---|---|
| Business Age (Years) | 14.5 | 27 | 29.475 |
| Current Ownership Tenure (Years) | 9 | 18.451 | 19.601 |
| Number of Employees | 9 | 17.29 | 19.566 |
| Percent of Business from Takeout/Delivery | 20% | 28.694% | 33.813% |

Table 1: General Summary Statistics of Business Metrics

| Business Fact | n | % |
|---|---|---|
| Family Owned/Generational | 10 | 32.26 |
| Bought (From Founder/Owner) | 9 | 29.03 |
| Founded | 11 | 35.48 |
| Owner 5 Minutes or Less to Business | 16 | 51.61 |
| Multiple Locations (Within the Valley) | 6 | 19.35 |
| Only Place to Dine in Their Area | 2 | 6.45 |
| Improved Profitability (From 3 Years Prior) | 10 | 32.26 |
| Worse Profitability (From 3 Years Prior) | 4 | 12.90 |
| Land and Building Ownership | 20 | 64.52 |

| | | |
|---|---|---|
| Renting or Leasing | 8 | 25.81 |
| Would Stay In the Mon Valley If Restarting | 17 | 54.84 |
| Would Not Stay In the Mon Valley If Restarting | 5 | 16.13 |

Table 2: Key Characteristics of Businesses Surveyed

| Type of Business | n | % |
|---|---|---|
| Family Restaurant (Eat-In) | 14 | 45.16 |
| Fine Dining | 2 | 6.45 |
| Bar | 1 | 3.23 |
| Coffee Shop | 3 | 9.68 |
| Food Truck | 3 | 9.68 |
| Bakery | 3 | 9.68 |
| Other | 5 | 16.13 |

Table 3: Types of Businesses Surveyed

| Type of Customer | n | % |
|---|---|---|
| Local/Resident | 27 | 87.10 |
| Commuter | 6 | 19.35 |
| Nonlocal/Highway Traffic | 5 | 16.13 |

Table 4: Customer Types Identified as Vital

The following 5 figures shown will display the results of the 17 Likert Scale questions asked in the survey. They are grouped into conceptual categories to make the data easier to follow. Each bar graph within these figures is ranked from the most agreed with (top) to the least agreed with (bottom) within the given group; however, the conceptual categories appear in no particular order. This ranking is done by converting answers into numbers, so, for instance, a “Strongly Agree” on one question is a 5 while a “Strongly Disagree” is a 1.

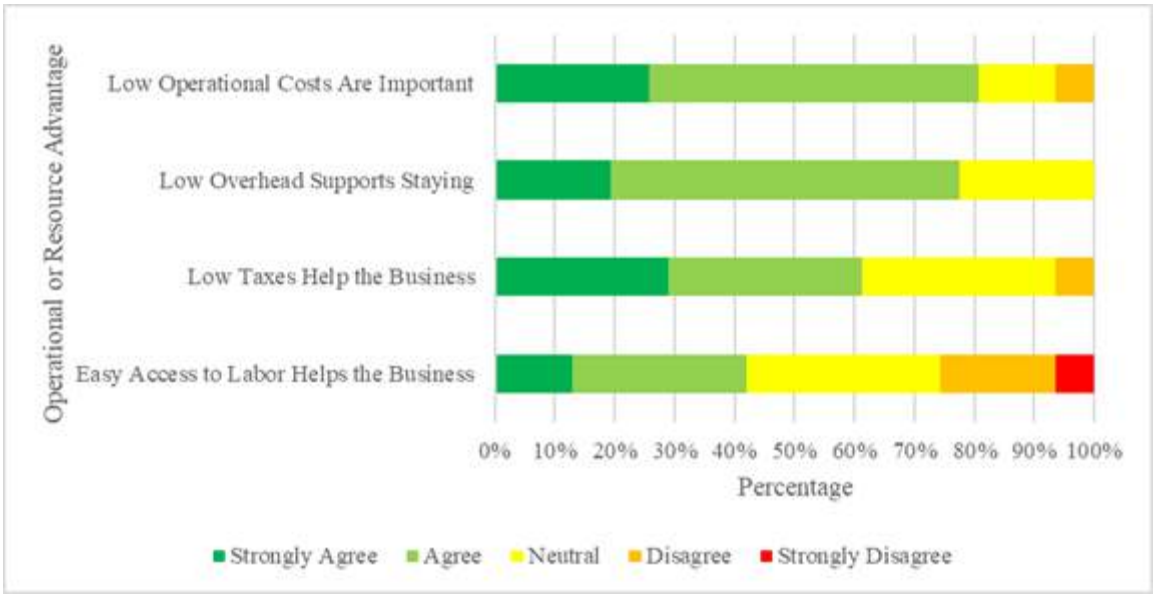



Figure 1: Costs and Resources

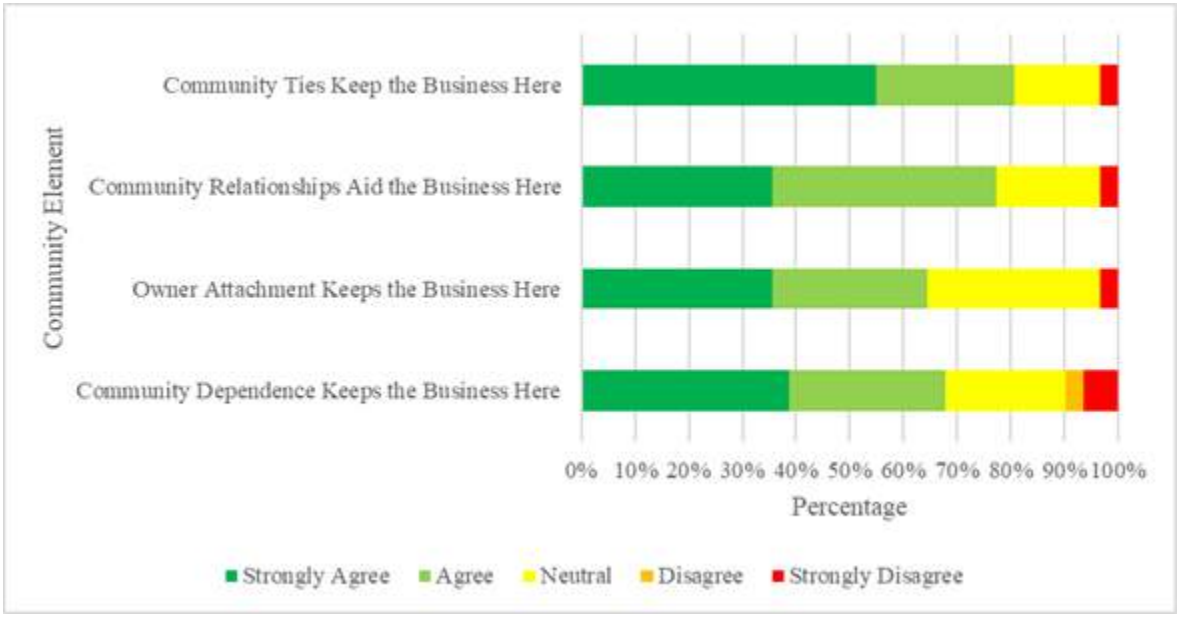



Figure 2: Community Engagement and Attachment

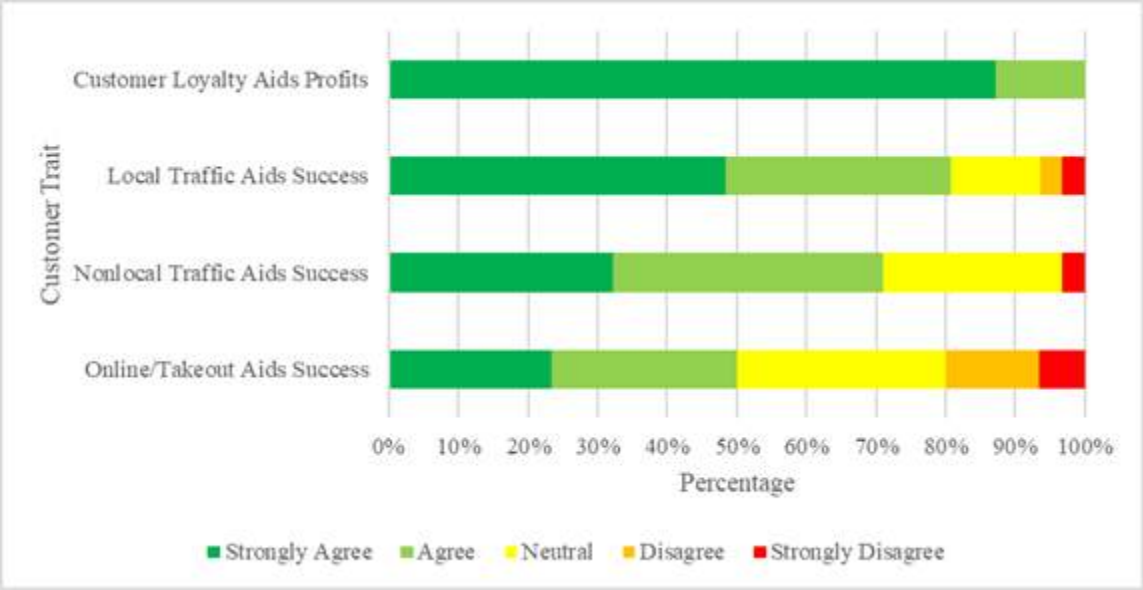



Figure 3: Customer Traits

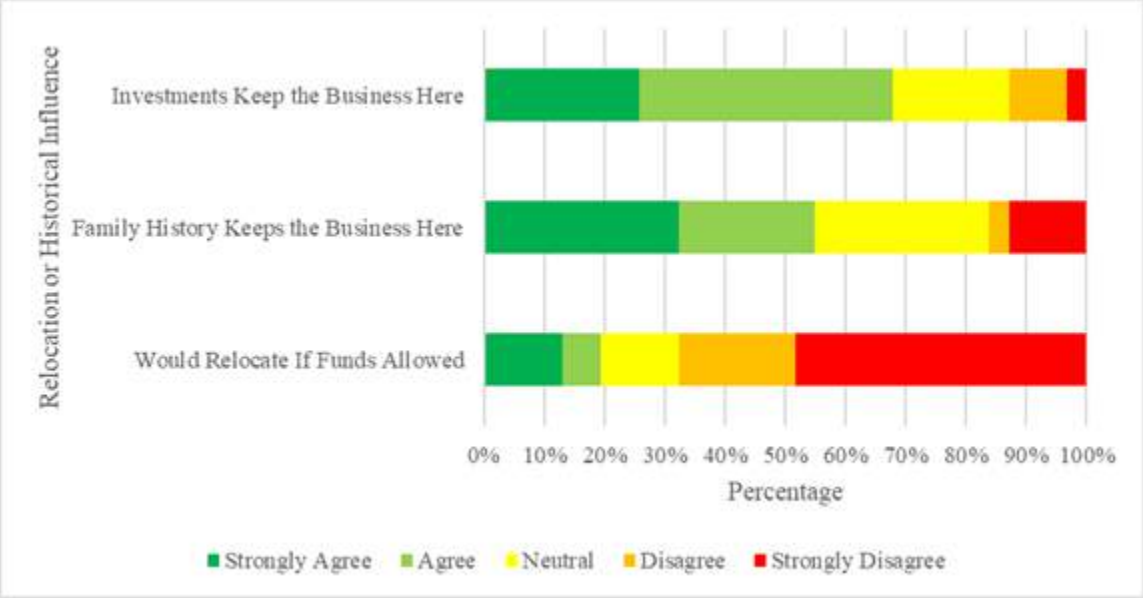



Figure 4: History and Relocation

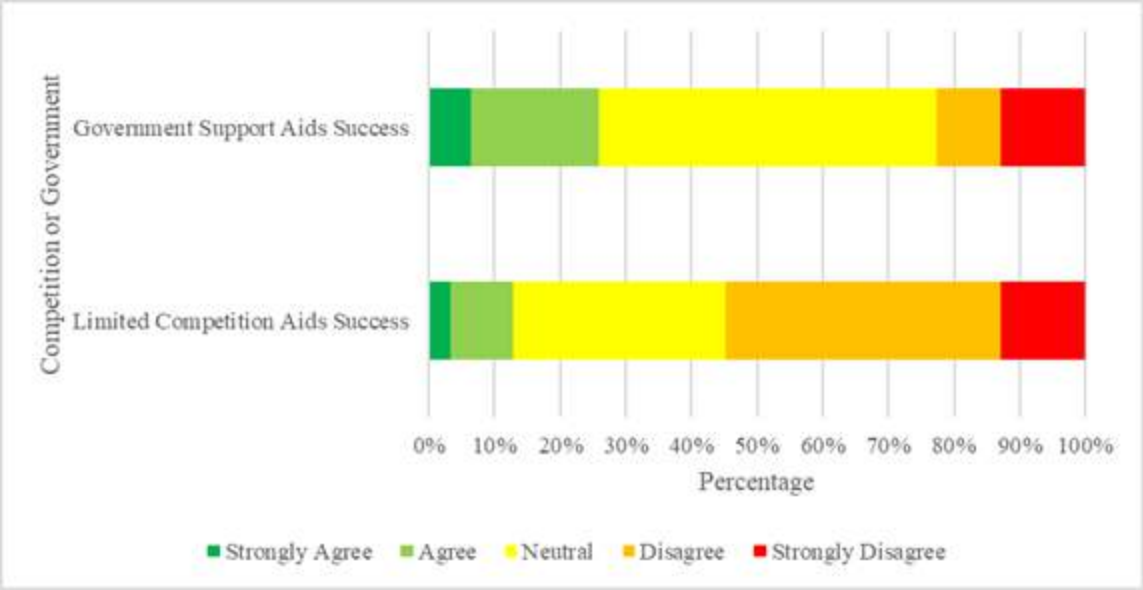



Figure 5: Government and Competition

| ZIP Code | Number of Participants | Population | Median Age (Years) | Percent Below Poverty Level | Unemployment Rate (%) |
|---|---|---|---|---|---|
| 15012 | 9 | 15756 | 47.1 | 8.6 | 5 |
| 15025 | 1 | 17118 | 41.4 | 6 | 6.4 |
| 15033 | 1 | 4677 | 42.8 | 17.4 | 7.4 |
| 15037 | 2 | 10519 | 48.5 | 3.2 | 6.1 |
| 15038 | 2 | 291 | 49.8 | >18.6* | 5.7 |
| 15062 | 3 | 6880 | 48.1 | 13.4 | 5.5 |
| 15063 | 2 | 11003 | 48.2 | 6.2 | 4.4 |
| 15332 | 2 | 8058 | 49.8 | 5.2 | 3.9 |
| 15412 | 1 | 257 | 50.7 | 4.8 | 4.9 |
| 15417 | 6 | 7973 | 42.1 | 22.6 | 12.3 |
| 15483 | 2 | 421 | 50.1 | 9.9 | 1.5 |
| AVG | 2.82 | 7541.18 | 47.15 | 10.25 | 5.74 |

Table 5: Census Data by Represented ZIP Codes

*No direct estimate was available, so the margin of error was used as an upper-bound estimate.
Source: U.S. Census Bureau, 2020 Decennial Census DP1 and 2020 American Community Survey 5-Year Estimates DP03

All of Table 1's metrics tended to vary greatly with standard deviations corroborating such. In fact, the age of the businesses surveyed had a maximum of 106 years. Looking at Table 2, the ways in which the current business owners came to ownership tended to be very evenly spread, including buying the business, inheriting it, and founding it. The majority of business owners were 5 minutes or less away from their business with the further distances becoming rarer the further away they were, and the most common form of land usage or rights was complete land and building ownership with the majority of those owners being fully paid off. In

Table 3, family eat-in-style restaurants were easily the most common form of business surveyed, although a fine mix was gathered with bakeries and coffee shops getting representation as well. Local residents or community members were most commonly labeled as of high importance to the business in Table 4.

Figure 1 sees easy access to labor as the least agreed with Likert question in said figure for why a business stays in the Mon Valley, being far below its counterpart questions in Figure 1. Figure 2 presents a very even spread of agreement among businesses, while Figure 3 is far more varied. Figure 3 has the most agreed with question in the survey, claiming that customer loyalty is important for profits, and a relatively low agreement question about the importance of online, takeout, or delivery to the business. Figures 4 and 5 see far less agreement than the previous figures with Figure 4's question claiming, provided it had the funds, the businesses would relocate being the least agreed with question in the survey.

Table 5 presents a varied picture of population, poverty, and unemployment among the businesses' areas, although the median age is relatively consistent, being mostly in the late forties with only three areas falling in the early forties and two in the early fifties.

**Analysis**

From the given data, a general picture for what keeps the food businesses of the Mid Mon Valley alive economically can be formed. The most telling factor is the customer base. 87.10% of the businesses claimed that locals were crucial to their business's success. That points to a very hyperlocal dependence, which is furthered by the age of these businesses and their tenures. With medians of 14.5 and 9 years respectively, the age and ownership tenure of these longstanding businesses builds community trust and therefore generates significant social capital. As such, a main factor for why these food businesses remain in the Mid Mon Valley appears to be this community trust and hyperlocal, almost niche, customer base.

These conjectures are corroborated by the Likert questions on the importance of customer loyalty and the local customer base. The question that asserted that customer loyalty significantly aided profits was by far the most agreed-with question with no business selecting the neutral or disagreement options. In fact, 87.10% selected the "Strongly Agree" option with the other 12.90% selecting the "Agree" option. This unanimity in agreement with customer loyalty's importance empirically demonstrates the influences of social capital on these businesses.

While less agreed with than the customer loyalty question, the question asserting that local traffic significantly aided profits saw 80.65% of businesses show some level of agreement, mirroring the 87.10% from the initial question. This works extremely well with the findings of Debenedetti et al. that suggest that consumers often report better experiences with personally meaningful places (6). As such, place attachment of the customers becomes another area in which businesses rely on and attract their hyperlocal clientele.

Of course, these businesses are not only leveraging the local support to remain in business. Another major factor appears to be landownership. 64.52% of the businesses owned the building and land they were based on, and of those businesses who owned the land, 65% were

fully paid off with no outstanding debt. This ownership appears to lower overhead by slashing fixed costs, allowing owners to keep larger cuts of their hyperlocal revenue.

Mirroring this claim, the Likert question asserting directly that low overhead was a notable help in the business's profits saw no disagreement with 19.35% selecting "Strongly Agree," 58.06% selecting "Agree," and 22.58% selecting "Neutral." This reinforces the idea that low overhead plays a major role in a food business's survival in the Mid Mon Valley.

In addition, low operational costs appear to substantially aid these businesses. For the question claiming that low operational costs were crucial to the business's success, 80.65% showed some level of agreement, and of those who agreed, 32% of them selected "Strongly Agree," and 68% selected "Agree." Complementing the effects of landownership and low overhead, these low operational costs allow these businesses to retain a greater slice of their revenue by cutting down variable costs.

This conjecture that low operational costs influence these businesses' survivals could be merged with their age and length of tenure, both of which allow for greater networking and potential relationships among suppliers of food, drinks, hardware, or any other items. These relationships could lead to certain deals or offers that younger businesses may not receive, which, in turn, lowers operational costs further.

Competition is an interesting factor to consider. Although it is evident in the two questions pertaining to a lack of competition that no spatial monopolies have been formed, it is entirely plausible that although these businesses may not be the only place to get food in their area, they may very well be the only place to offer their specific type of food. Two food businesses could coexist in one small town without necessarily competing.

For example, an Italian restaurant and a Chinese restaurant may both be competing on a large scale, but both are also fulfilling different roles and niches, which may, in turn, mean that certain areas are empowered by a lack of competition in their specific area of expertise. This is what is known as product differentiation, which is "the process of distinguishing a product from competitors' products in a way that makes it more attractive to a target costumer, consumer or audience" ("Product Differentiation and Place Branding | Indian Affairs"). Therefore, this product differentiation could serve to create a small or micro version of a spatial monopoly. This is likely not a main driving force behind businesses and their survivals in the Mid Mon Valley, but it should be taken into consideration.

In summary, food businesses appear to survive in the Mid Mon Valley off of a hyperlocal customer base, low overhead, and low operational costs. For those many businesses with long tenures and ages, relationships, social capital, and land and building ownership appear to further aid in their survival and profitability.

There are several factors that do not seem to significantly affect these businesses and their survival. Government aid, for instance, was quite low in agreement as only 25.81% of businesses selected any form of agreement on the Likert question claiming that the local government significantly aided the business. Easy access to labor was also not considered to be a major factor keeping the businesses where they were with only 41.94% selecting either

agreement level. Similarly, online, takeout, and delivery sales do not appear to be as crucial to the businesses' survivals as it was the least agreed with of the customer questions, coming in at 48.39% selecting any level of agreement. Additionally, the median for the percentage estimate for how much business came from delivery or takeout was only 20%, communicating a lack of importance to many businesses currently.

Now it is known how these businesses and restaurants survive, but the research question has yet to be answered: why do these businesses choose to remain in the Mid Mon Valley? To answer this, a more in-depth analysis is required, starting with the Likert-scale questions.

The following table will show all Likert questions converted into numerical scores. Each option will be converted into a number using the following: "Strongly Agree" is a 5, "Agree" is a 4, "Neutral" is a 3, "Disagree" is a 2, and "Strongly Disagree" is a 1. All questions are ranked from most agreed with (top) to least agreed with (bottom) using the score means.

| Factor that Aids Success/Keeps the Business Here | Mean | Median | Standard Deviation |
|---|---|---|---|
| Customer Loyalty | 4.87 | 5 | 0.34 |
| Community Ties | 4.29 | 5 | 0.97 |
| Traffic from Locals | 4.19 | 4 | 1.01 |
| Community Relationships for Profitability | 4.06 | 4 | 0.93 |
| Low Operational Costs | 4.00 | 4 | 0.82 |
| Low Overhead* | 3.97 | 4 | 0.66 |
| Traffic from Nonlocals* | 3.97 | 4 | 0.95 |
| Leader's Attachment to Community | 3.94 | 4 | 1.00 |
| Community's Dependence on the Business | 3.90 | 4 | 1.16 |
| Low Taxes | 3.84 | 4 | 0.93 |
| Deep Investments Already Made | 3.77 | 4 | 1.06 |
| Family/Generational Business Status | 3.58 | 4 | 1.34 |
| Online/Takeout/Delivery Sales | 3.39 | 3 | 1.27 |
| Easy Access to Labor | 3.23 | 3 | 1.12 |
| Local Government Support | 2.97 | 3 | 1.05 |
| Lack of Competition | 2.48 | 2 | 0.96 |
| Would Leave If Funds Allowed | 2.16 | 2 | 1.42 |

Table 6: Individual Likert Analysis

*These questions are tied in mean scores. The lower standard deviation is ranked higher.

From this table, the higher influence Likert factors can be determined. These are customer loyalty, community ties, traffic from locals, community relationships (that help profitability), and low operational costs. These are evidently the most important factors, as ranked by the businesses surveyed, that aid their profits and are reasons for their location. In contrast, access to labor, government support, and a lack of competition appear to be the least important.

Following this analysis, it would be useful to group these among related items. This would allow for general concepts to be tested rather than individual markers. To do this, each of the Likert questions will be separated into one of four groups. From Table 6, Group 1 consists of

customer loyalty, traffic from locals, community ties, community dependence, and community relationships. Group 2 consists of low taxes, low overhead, and low operational costs. Group 3 consists of family/generational business status, leader's attachment to the community, deep investments already made, and would leave if funds allowed. Group 4 consists of traffic from nonlocals, easy access to labor, sales from online/takeout/delivery, lack of competition, and local government support. For these groups, a composite mean score will be calculated using the same metrics as the Table 6's scores except for "would leave if funds allowed" as it will be reverse scored so that "Strongly Disagree" is a 5 and so on, and each group will be ranked according to said score.

| Group # | Mean | Standard Deviation | Cronbach's Alpha |
|---|---|---|---|
| 1 | 4.26 | 0.67 | 0.764 |
| 2 | 3.94 | 0.63 | 0.685 |
| 3 | 3.36 | 0.60 | 0.669 |
| 4 | 3.21 | 0.55 | 0.292 |

Table 7: Grouped Likert Scores

Group 1 shows the highest overall agreement of any group. Although it could be broadly named something like "community support," this group was specifically designed with social capital and consumer place attachment in mind. All the questions within this group test the agreement of the businesses surveyed in relation to how much the community and its local clientele matter to its profits. The α, or minimum acceptance level, for many studies utilizing Cronbach's alpha varies depending on the study. An article from the University of Virginia states, "many methodologists recommend a minimum α coefficient between 0.65 and 0.8" (Goforth). Because this study is more exploratory, an α of 0.65 shall be used.

Cronbach's alpha for Group 1 easily surpassed the α of 0.65 with a 0.764. As such, the questions in this group vary together at an acceptable level. What a business said for how important local customers were to their profits is a good indicator of how important they claimed their community ties were in their decision to remain in the Mid Mon Valley. The comparatively high mean of 4.26 to the other groups indicates that this is a considerably more important group of factors than the other Likert measures. Thus, social capital and the locals' place attachment appear to be highly valuable to the food businesses of the Mid Mon Valley.

Group 2 tested the financial benefits of the Mid Mon Valley in relation to their agreement. Its mean of 3.94 places it significantly above Groups 3 and 4 but below Group 1 in importance. Its Cronbach's alpha of 0.685 also surpasses the minimum α, indicating that this group also varies together but to a lesser extent than Group 1. This slightly less similar variance may stem from the question on overhead being included as those renting would reasonably have higher overhead than those who are owners; however, all businesses would be equally likely to take advantage of low taxes and operational costs. It is also true that rent is not the only item in overhead; therefore, even renters may still mark low overhead as an agreeable statement but may

be more likely to select "Agree" rather than "Strongly Agree." This group's mean of 3.94 indicates that it is more agreed with than its counterparts outside of Group 1 on average.

Group 3 tests the owner's place attachment and sunk costs. This group saw a major drop in mean scores at 3.36, indicating that this group is not nearly as important to why these businesses remain in the Mid Mon Valley when compared to the influences of social capital, consumer place attachment, and financial benefits. Group 3's Cronbach's alpha did barely surpass the minimum α of 0.65 with a 0.669, showing a degree of equivalent variance, albeit less than Group 2 and much less than Group 1. As such, sunk costs and owner place attachment do not appear to be driving factors behind the businesses' decisions to stay in the Mid Mon Valley.

Group 4 was the weakest grouping. It tested the external market influences that aided the business. The issue, however, is that these influences span many areas. It encompasses customer bases, access to labor, competition, and local government support. Conceptually, the grouping makes sense; these are all influences that are not purely financial that may aid a business. The issue, however, is that a business may feel greatly supported by its nonlocal customer base while also not feeling supported by the government or feeling substantial competition from other businesses. In addition, the question claiming that sales from online, takeout, or delivery were significant streams of revenue may exclude certain businesses that do not offer any of those options, causing great variance among businesses and their agreement levels. It therefore follows that this group had the lowest mean of 3.21 and a very low Cronbach's alpha of 0.292, well below the α of 0.65. That does not mean that these factors are not important in these businesses' decisions to stay in the Mid Mon Valley. Rather, it indicates that unlike the last groups, these values should not be examined in a group and should be considered individually. In fact, looking at Table 6, it is apparent that traffic from nonlocals is important to an extent as that question had a mean of 3.97, far above this group's mean of 3.21.

In summary, the most important factors and strongest grouping was Group 1, which tested social capital and the local customer base's place attachment. Group 2 followed closely behind, which tested the financial benefits of the Mid Mon Valley. Meanwhile, Groups 3 and 4 did not appear to demonstrate meaningful factors in these businesses' decisions to remain in the Mid Mon Valley. Group 3, however, was a valid grouping for this exploratory study, while Group 4 should not be considered valid and should be analyzed in separate questions.

Breaking apart the invalid Group 4, there is access to labor, traffic from nonlocals, sales from online/takeout/delivery, competition, and local government support. Examining Table 6, it is evident that support from local government and a lack of competition were not important factors in these businesses' decisions to remain in the Mid Mon Valley. Both had means below 3 (2.97 and 2.48 respectively), indicating that they are not benefits these businesses claim are highly important. Easy access to labor scored quite low as well with a mean of 3.23. This may be counterintuitive when looking at the unemployment rates of many of these businesses' areas with some reaching well above 10 percent, which would suggest that many are attempting to find work (Table 5). However, labor being available does not necessarily indicate qualified nor beneficial labor being available. Furthermore, many of these food businesses are not running

expansive operations with the median number of employees being 9 (Table 1), so it makes sense that these businesses do not necessarily consider access to labor as a vital influence keeping them in the Mid Mon Valley.

The importance of sales from online, takeout, or delivery was also a relatively low category in agreement at a mean of 3.39. Interestingly, Table 1 illustrates very high standard deviation in sales from delivery or takeout with a median of 20 percent but a standard deviation of 33.813 percent. This surfaces an interesting point. Several businesses noted that they were completely delivery or takeout based while others, still, noted that they did not offer any takeout, delivery, or online ordering. As such, this question's results are likely skewed heavily on either side. The standard deviation of 1.27 for this question mirrors this conjecture as that is the second highest standard deviation for one Likert question, being only behind the family or generational status question. The high age of many of these businesses and their tenures may be furthering some of these businesses' trailing progression into the digital age as one owner might not want to change or may not see the reason to do so because of the aging population (Table 1; Table 5).

Traffic from nonlocals is an anomalous question. Table 4 illustrates that only 16.13% of businesses claimed that nonlocals were vital to their business. But, the Likert version of this question saw a 3.97 mean score and 70.97% agreement. The given results are aberrant but likely stem from a few areas. This could be because of question wording or because the Likert question offered a lower level of agreement, a sort of middle ground, while the initial question simply asked for if nonlocals were important, a more all-or-nothing question. In addition, the initial question differentiated between commuters (19.35%) and nonlocal and highway traffic (16.13%). This wording difference may have caused some of the disparity, but even when combined these percentages result in a 35.48% nonlocal importance, which is barely half of the agreement of the Likert question. Also, the initial question was a multi-select question, which could have confused some participants, partially aiding in this discrepancy. With both results in mind, the target customer base is evidently locals for these businesses, but nonlocals do clearly influence the businesses to an extent.

Based on this analysis thus far, it has been determined that social capital and the locals' place attachment (Group 1) was the most agreed with influence for why businesses choose to remain in the Mid Mon Valley. To test this, a Spearman's Rank Correlation Test was conducted. The independent variable was the businesses' scores for Group 1, and the dependent variable was the reverse-scored answers to the Likert question asserting the business would leave if it had the funds. As stated, this Likert question being measured was reverse scored so that a "Strongly Disagree" (that they would not leave) is equivalent to a 5 as that shows strong willingness to stay. The null hypothesis was that there is no positive correlation between the level of agreement on the answers to the reverse-scored Likert question claiming they would leave if funds allowed and their Group 1 scores ($\rho \leq 0$). The alternative hypothesis was that there is a statistically significant positive correlation between the level of agreement on the answers to the reverse-scored Likert question claiming they would leave if funds allowed and their Group 1 scores ($\rho > 0$). A one-tailed Spearman's Rank Correlation Test was conducted at $\alpha$ of 0.01.

The results of this test granted a $r_s$ of 0.439, a t-statistic of 2.631, and a $p$-value of 0.0068. Because the p-value is less than the α, the null hypothesis is rejected. There is a statistically significant positive correlation between the level of agreement on the answers to the reverse-scored Likert question claiming they would leave if funds allowed and their Group 1 scores. As such, the conjecture made that Group 1 (social capital and consumer place attachment) was a significant and, in this case, the most agreed with factor for why food businesses choose to stay in the Mid Mon Valley is supported. Social capital and consumer place attachment correlate with a business's willingness to stay.

Because this last test was statistically significant, it follows that another test for Group 2 should be run as this was the second-most-agreed-with grouping with a significant Cronbach's alpha. As such, another Spearman's Rank Correlation Test was conducted. The independent variable was the businesses' scores for Group 2, and the dependent variable was the reverse-scored answers to the Likert question asserting the business would leave if it had the funds, using the same reverse-scoring method as the last test ("Strongly Disagree" is a 5, etc.). The null hypothesis was that there is no positive correlation between the level of agreement on the answers to the reverse-scored Likert question claiming they would leave if funds allowed and their Group 2 scores ($\rho \leq 0$). The alternative hypothesis was that there is a statistically significant positive correlation between the level of agreement on the answers to the reverse-scored Likert question claiming they would leave if funds allowed and their Group 2 scores ($\rho > 0$). A one-tailed Spearman's Rank Correlation Test was conducted at α of 0.01.

The results of this test granted a $r_s$ of -0.143, a t-statistic of 0.775, and a $p$-value of 0.222. Because the p-value is greater than α, the null hypothesis has failed to be rejected. There is not statistically significant evidence for a positive correlation between the level of agreement on the answers to the Likert question claiming they would leave if funds allowed and the businesses' Group 2 scores.

The results of this test are rather interesting. Countering the alternative hypothesis completely, the extremely weak correlation found was determined to be negative, indicating that oddly, businesses claiming to be benefitting more from the Mid Mon Valley financially were more likely to want to leave than those benefiting less financially. The extremely weak correlation, however, makes this counterintuitive conjecture invalid as a $r_s$ this low cannot reasonably be used to make a valid claim of correlation.

From this, the study's main research question has largely been answered. Group 1's emphasis on social capital and consumer place attachment has statistically significant evidence to assert its moderate positive correlation with a business's willingness to stay in the Mid Mon Valley and is far beyond any other meaningful grouping of measures presented in this study.

**Practical Implication**

Through this study, some very meaningful findings have been observed for areas like the Mid Mon Valley. Evidently, these food businesses' desires to stay in the economically challenged area of the Mid Mon Valley correlate with social capital and consumer place attachment. From

that, economic development organizations and movements should foster even greater community involvement and pride to allow for more bonding and to build further attachment among residents. However, it is dangerous for these areas to rely purely off of local clientele and their sentiments. As such, it would be advisable for organizations and local governments to try to attract greater nonlocal traffic and incentivize people to move to this area.

The Likert question that claimed that the local government's support was a major help to the business was one of the least agree with of all Likert questions. As such, it would follow that the government should attempt to mediate this gap in agreement by collaborating with local businesses on how to drive in more traffic and retain more revenue.

As shown by the Spearman's Rank Correlation Test, the economic benefits of the Mid Mon Valley, although attractive, are evidently not enough to encourage food businesses to stay in the Mid Mon Valley alone. Although attractive and helpful, reliance on these could result in a further decline in hospitality growth in this already struggling area.

Finally, local organizations and governments should work with businesses to implement online, takeout, and delivery systems to expand the reach of these businesses, make them more competitive, drive in a greater market, and even appeal to the elderly and less mobile by offering delivery in the aging population of the Mid Mon Valley. This may remedy the lack of importance of the given income stream as ranked by the owners.

**Limitations and Further Research**

Although this study was conducted with the utmost care, there are several limitations that must be acknowledged. The most obvious of these limitations is the study's sample size of 31. This is a low sample size; however, it likely represents an estimated 25-30% of the population for the area, meaning that a substantial share of the population was present. Regardless, further studies with larger budgets and more researchers should be conducted to try to expand this sample size so that statistical tests have greater power. This may reveal that items like financial benefits are actually statistically significant and correlate with a food business's decision to stay in the Mid Mon Valley. On top of this, further research should be conducted to involve other types of businesses to allow for economic recovery organizations to work across industries rather than in only one.

In addition to the sample size, this study contained nonresponse bias and may have been affected by response bias via question wording. These may have slightly resulted in skewed or altered results.

Finally, voluntary response bias affects the results of the present study. Despite the estimated 25-30% response rate, the study was inherently voluntary, and participants did not have to complete the survey. This may have made the sample less representative of certain groups or business types, which plausibly affected the results.

**Conclusion**

The Mid Mon Valley is an area marked by many historical ups and downs. Once a part of the steel capital of the world, it represents a spearheading example of the post-industrial landscape. The food businesses of this area, an industry marked by a need for a substantial consumer base, are aberrant in that they choose to reside in an area yet to recover economically to its peak whilst losing its population in the process. As such, this study inquired into what factors are keeping these businesses in this economically challenged area. Of all factors tested, those pertaining to financial and community incentives appeared to be most agreed with by businesses for why they chose to remain in the Mid Mon Valley. On the other hand, government support, a lack of competition, and easy access to labor were the least agreed with, marking a gap in support or incentives to stay. After statistical testing was conducted, it was found that the relationships with the local community (social capital) and said community's attachment to the area (place attachment) were statistically significant and moderately, positively correlated with the business's willingness to stay in the Mid Mon Valley, while the financial reasons were not statistically significant in their correlation.

As such, food businesses in the Mid Mon Valley who express a desire to stay appear to do so more often, as the correlation shows, because of their community rather than any economic or environmental benefits. This indicates that economic development organizations and local governments should leverage this advantage whilst attempting to mediate the gaps elsewhere. Community events, government collaboration, and efforts to generate greater nonlocal traffic should be implemented to do so. This study's main limitation of a small sample size ($n = 31$) should prompt further researchers to attempt to gather a larger sample size for greater statistical power. Throughout the story of the Mid Mon Valley, one key element is present: community. Whether it was coming together after losing a coworker to a steel accident or banding together in a union to end injustice, the Mid Mon Valley has always been a place of community. Clearly, that is what is carrying it through these economically dire times. From the people the researcher has talked to and the support they have received, it is abundantly clear that that exact advantage is what will see this humble Valley through these challenging times.

**Works Cited**


Alsaati, Mahmood. Small Business Management. Toronto Academic Press, 2024.

Debenedetti, Alain, et al. “Place Attachment in Commercial Settings: A Gift Economy Perspective.” Journal of Consumer Research, vol. 40, no. 5, Feb. 2014, pp. 904–23, https://doi.org/10.1086/673469.

Goforth, Chelsea. “Using and Interpreting Cronbach’s Alpha | UVA Library.” Library.Virginia.Edu, 2015, https://library.virginia.edu/data/articles/using-and-interpreting-cronbachs-alpha.

Greenlaw, Steven A, et al. Principles of Microeconomics. 3rd ed., Openstax, Rice University, 2022.

Hill, Stephen, et al. Global Entrepreneurship Monitor 2024/2025 Global Report: Entrepreneurship Reality Check. Global Entrepreneurship Research Association (GERA), 2025.

Hoerr, John P. And the Wolf Finally Came : The Decline of the American Steel Industry. Univ. Press, 1988.

“How Much Can It Cost to Start a Restaurant?” Escoffier, Auguste Escoffier School of Culinary Arts, 25 Jan. 2023, https://www.escoffier.edu/blog/food-entrepreneurship/how-much-does-it-cost-to-start-a-restaurant/.

Johnson, Cecil Douglas. Social Capital : Global Perspectives, Management Strategies and Effectiveness. Nova Publishers, 2015.

Kobus, Ken. City of Steel. Rowman & Littlefield Publishers, 2015.

Principles of Business : Entrepreneurship. Salem Press, 2017.

“Product Differentiation and Place Branding | Indian Affairs.” Www.Bia.Gov, https://www.bia.gov/service/running-business/product-differentiation-and-place-branding. Accessed 24 Sept. 2026.

Rutherford, Donald. Routledge Dictionary of Economics. Routledge, 2013.

Samuelsson, Mikael, et al. “Path Dependence in New Ventures’ Capital Structures.” Entrepreneurship Theory and Practice, vol. 45, no. 2, Feb. 2020, p. 104225872090171, https://doi.org/10.1177/1042258720901717.

Schiller, Bradley R, and Karen Gebhardt. The Essentials of Economics. McGraw Hill LLC, 2025.

Skrabec, Quentin R. Jr. The Fall of an American Rome. Algora Publishing, 2014.

Stringam, Betsy Bender. Profile of the Hospitality Industry. Business Expert Press, 2021.

Tripp, Gregory, et al. Social Capital. Nova Science Publ, 2009.

United States Census Bureau. 2020: ACS 5-Year Estimates Data Profiles, DP03: Selected Economic Characteristics. U.S. Department of Commerce, 2020.

United States Census Bureau. 2020 Census Demographic and Housing Characteristics File, DP1: Profile of General Population and Housing Characteristics. U.S. Department of Commerce, 2020.

**Consequence Framing and Academic Dishonesty: A Behavioral Economics Approach**
**By Evan Jin Rossmeisl**

**Abstract**

Academic dishonesty is often viewed as a decision based on the benefits of cheating and the potential consequences of being caught. However, behavioral economics suggests that the way these consequences are framed may also influence decision-making. This study examines how different consequence framings affect students' intentions to engage in academic dishonesty. Participants were randomly assigned to one of four hypothetical scenarios: a control condition with no stated consequences, a learning-loss condition emphasizing the potential loss of future learning and academic competence, a formal-punishment condition emphasizing academic consequences and detection, or a social-shame condition emphasizing embarrassment and loss of trust or respect. Participants then reported their reactions to the opportunity to cheat. The results suggest that formal punishment and social consequences were associated with lower stated cheating intentions, while emphasizing the potential loss of learning did not produce the same deterrent effect. These findings suggest that students may respond differently depending on how the consequences of academic dishonesty are presented, highlighting the importance of behavioral and social factors in academic decision-making.

**Introduction**

Academic dishonesty is often explained as a decision in which students weigh the potential benefits of cheating against its possible costs. From a traditional economic perspective, a student may compare the benefit of receiving a higher grade with the probability of being caught and the punishment that could follow. This approach assumes that students respond to incentives in a relatively rational and consistent way. However, students do not always make decisions based only on these objective costs and benefits. Emotions, social pressure, personal values, and the way consequences are presented can also influence how people respond to the same choice.

Behavioral economics provides a framework for understanding these differences. One relevant concept is loss aversion, which suggests that people may respond more strongly to losses than to equivalent gains. Applied to academic dishonesty, this raises the possibility that students could be discouraged from cheating when they focus on what they might lose rather than only on what they might gain. However, research on loss framing and academic honesty has produced mixed findings. Some studies suggest that emphasizing potential losses can increase honest behavior, while others find little or no effect. This suggests that the type of loss and the context in which it is presented may be important.

Social consequences provide another possible influence on dishonest decisions. Students may consider how their behavior will affect the way they are viewed by classmates, teachers, or other people around them. Research suggests that peer behavior and social norms can influence whether students view cheating as acceptable, while fear of judgment or shame may discourage

dishonest behavior. At the same time, social pressure does not affect everyone equally. If students believe that cheating is common or socially accepted, concern about reputation may become less important. Social consequences therefore represent a potentially important but inconsistent form of deterrence.

Academic consequences may operate through a different mechanism. Formal punishment makes the potential cost of cheating more concrete by connecting dishonesty to academic standing, future opportunities, or other external consequences. This is closely related to the traditional economic view that individuals respond to the expected costs of being caught. In contrast, a learning-loss perspective focuses on the internal consequences of cheating. Instead of emphasizing punishment, this approach suggests that cheating could prevent students from developing the knowledge and skills they need in future classes. Research on self-concept and achievement goals suggests that students may care about how their academic choices reflect their identity and whether they are genuinely developing competence.

Despite research on punishment, social influence, and loss framing, these mechanisms have often been examined separately. It remains unclear whether students respond more strongly to the possibility of formal academic consequences, social judgment, or the loss of future learning and competence. This distinction is important because these messages represent different reasons for acting honestly. A student might avoid cheating because they fear being caught, because they do not want others to lose respect for them, or because they value what they can learn from completing the work themselves.

This study addresses this gap by comparing three consequence framings within a single experimental design. Participants were randomly assigned to one of four hypothetical situations: a control condition with no stated consequences, a learning-loss condition in which no one finds out but cheating may harm future learning, a formal-punishment condition in which others find out and academic consequences may follow, and a social-shame condition in which others find out and the student may experience embarrassment or loss of trust and respect. By comparing these conditions, the study examines whether different ways of presenting consequences produce different levels of stated cheating intention.

The study therefore asks whether students are more discouraged from academic dishonesty when consequences are framed as a loss of learning, a formal academic punishment, or a social cost. More broadly, it examines whether academic decision-making is influenced not only by the size of a potential consequence, but also by what that consequence represents to the individual. Understanding these differences may provide insight into how behavioral-economic mechanisms such as loss aversion, social judgment, and responses to external incentives shape decisions about academic honesty.

**Literature Review**

Economic models of academic dishonesty typically begin with a simple cost-benefit calculation: students weigh the expected benefit of a better grade against the probability of detection and the severity of punishment (Ganuthula and Kumar). This rational-choice

framework predicts that cheating rises when perceived detection risk falls or when penalties are weak, and it has historically shaped how schools design deterrence policy — more proctors, harsher sanctions, clearer honor codes. But this model rests on an assumption worth questioning: that students process risk and reward the same way regardless of how the choice is framed to them. Behavioral economics suggests otherwise, and it is this gap between rational-deterrence theory and actual student behavior that motivates the present study.

A first behavioral complication comes from loss aversion — the tendency to weigh losses more heavily than equivalent gains. If this holds for academic integrity, then framing honesty around what students stand to lose, rather than what they might gain, should more strongly shift behavior than a gain-framed appeal of equal magnitude (Kristensen and Dyvik). The empirical picture, however, is inconsistent. Some studies find that loss framing does increase honest behavior, as people act to avoid a concrete, felt loss rather than pursue an abstract gain (Zeif and Yechiam), while other research finds little or no difference between gain- and loss-framed messages (Leisge et al.). This inconsistency does not necessarily mean loss aversion fails as a mechanism in academic settings; rather, framing effects appear to be moderated by task stakes, reward conditions, and individual differences in risk tolerance (Grolleau et al.; Cannito et al.). Taken together, this body of work suggests that loss framing is a real behavioral lever, but that what kind of loss is being framed — a grade, a reputation, a skill — may matter as much as the basic loss-versus-gain structure of the message itself. This is a distinction the existing literature has not tested directly, and it is central to the design of this study.

A second set of influences on cheating operates through social rather than purely individual channels. Peer behavior shapes how students interpret the norms around them: when students believe cheating is common among their classmates, they tend to see it as more socially acceptable, which in turn weakens the perceived weight of formal penalties (Ehrmann et al.; Zhao et al.). Shame operates through a related but distinct pathway. Concern about how one would be judged by peers, teachers, or oneself can independently discourage dishonest behavior, even in contexts where peer norms around cheating are relatively permissive (Cochran et al.). Yet shame's deterrent power is not uniform across students. Messages built around social judgment do not land the same way for everyone, and broader research on framing and social-norm interventions shows meaningfully different effects depending on the population studied and the context in which the message is delivered (Martuza et al.). Peer evaluation adds a further layer: the emotional response students have to being evaluated by classmates can shape how they think about their own honesty in ways that are somewhat independent of explicit social norms or formal rules (Lee and Yuan-Hsien). Together, these findings suggest that social-channel interventions — whether norm-based or shame-based — are powerful but unpredictable, working well for some students and poorly for others depending on factors the current literature has not fully mapped.

The most theoretically distinct thread in this literature concerns self-image and identity. A growing body of research suggests that people attend closely to how a given choice reflects on their sense of self, not only to its external rewards or costs (Goyal; Chen et al.). This idea closely

parallels Mazar, Amir, and Ariely's influential self-concept maintenance theory, which argues that people will cheat only up to the point where they can still view themselves as fundamentally honest — meaning dishonesty is bounded not just by fear of punishment, but by the psychological cost of updating one's self-image downward. This line of reasoning opens up a third possible lever for interventions, distinct from both loss aversion and social shame: framing honesty as identity-protective, tied to being a competent and capable learner, rather than as simply loss-avoidant or judgment-avoidant. This framing connects naturally to achievement-goal theory's well-established distinction between mastery goals — learning valued for its own sake — and performance goals, in which the priority is protecting a grade or an external image of competence (Dweck; Ames). If students primarily driven by mastery goals respond more strongly to a learning-focused message than to a punishment- or shame-focused one, this would suggest that motivational orientation, not just message content, shapes how effectively different framings deter dishonesty.

Prior research has examined punishment-based deterrence, peer and social norms, and gain-loss framing largely as separate literatures, and findings within each strand remain genuinely mixed (Bećirović et al.; Levitt et al.; Friedman et al.). What has not yet been tested directly is how these three message types — shame, formal punishment, and protection of long-term learning and competence — compare against one another within a single experimental design. This comparison matters because each message targets a functionally different psychological mechanism: shame targets social image, punishment targets calculated risk assessment, and a learning-loss frame targets self-concept and mastery motivation. This study addresses that gap by comparing all three framings under assessment pressure, examining both extensive cheating and smaller, more ambiguous acts of dishonesty (Mukherjee et al.), and testing whether students' choices more closely track a desire to master the material or a concern about grades and external judgment (Mukherjee et al.; Fritz et al.).

### *Design*

This study used a between-subjects experiment to examine whether the way consequences are described affects students' intentions to cheat. Participants were randomly assigned to one of four hypothetical scenarios and then answered the same questionnaire. Each participant saw only one scenario, so they would be less likely to compare the messages and tailor their answers. This design helps test whether the framing of a consequence matters, not just the consequence itself.

### *Participants*

Participants were university-aged students and young adults who could reasonably relate to the scenario. They were recruited through convenience and snowball sampling: the survey link was shared in classes, on social media, and through peer networks. Taking part was voluntary and unpaid. In total, 87 participants completed the survey, and 87 remained after exclusions, leaving approximately 21.75 participants per condition.

Participants first reported their age category. Those under 18 were asked to confirm that they had permission from a parent or guardian and were excluded if they could not. The introduction explained that participation was voluntary and anonymous and that participants could withdraw at any time. The survey did not collect names, email addresses, or other identifying information.

***Materials and Experimental Conditions***

In each scenario, participants imagined that they were a university student taking an important course. During an exam, they had a chance to access unauthorized answers that could substantially improve their grade. The scenarios differed in the consequences described. Each condition represented a mechanism discussed in the literature review:

1. Control: No consequences. Participants were told that no one would find out and that their choice would not affect their academic future or relationships. This condition provided a baseline for cheating intentions when no costs were described.
2. Treatment 1: Academic consequences, nobody finds out (learning-loss framing). Participants were told that no one would find out, but that they might struggle in future classes because they had not learned the material. This message focused on the possible loss of learning and competence, without detection or external punishment.
3. Treatment 2: Academic consequences, people find out (formal-punishment framing). Participants were told that others would find out and that their academic standing or future opportunities might be affected. This message focused on external consequences and the risks of detection.
4. Treatment 3: Social consequences, people find out (shame framing). Participants were told that others would find out and that they might feel embarrassed or ashamed, or lose others' trust and respect. This message focused on social image and reputation.

The comparisons were chosen to clarify how students respond to the different risk frames. Comparing each treatment with the control tests whether describing consequences reduces cheating intentions. The key comparison is Treatment 2 versus Treatment 3: because both scenarios involve others finding out, this comparison helps show whether formal academic consequences or social judgment is more discouraging. Comparing Treatment 1 with Treatment 2 tests whether academic consequences have different effects when they involve undetected learning loss rather than externally enforced consequences.

***Measures***

After reading their assigned scenario, participants rated 14 statements on a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree). The items covered three areas.
Cheating intention assessed participants' willingness or likelihood to cheat (for example, "I would be willing to cheat in this situation"). Decision and motivation assessed how participants weighed the potential grade benefit against the risks, including whether they would put grades

ahead of consequences or be more likely to cheat if they were not detected. Risk perception assessed how serious and discouraging participants found the consequences. These items were reverse-scored where needed so that higher scores consistently represented more of the construct being measured.

Three composite scores were calculated by averaging relevant items. The primary outcome was cheating intention . Secondary outcomes were motivation and perceived risk. Item 6 asked whether the consequences would affect participants' decisions. Because it more directly measured perceived deterrence than willingness to cheat, it was examined separately. Cronbach's alpha was used to assess the internal consistency of each composite ([α values]).

***Procedure***

The survey was delivered online through Qualtrics, which allowed participants to be randomly assigned to a condition. The introduction described the study broadly as an investigation of how different ways of presenting consequences may affect decision-making, without explaining the specific hypotheses. Participants were told the survey would take about 3–5 minutes, that their responses were anonymous, and that there were no right or wrong answers. They were asked to respond based on what they would personally do in the situation.

After the age and consent questions, Qualtrics randomly assigned each participant to one of the four scenarios. Participants then answered the questionnaire items, completed an attention check, and provided demographic information. Demographic questions appeared at the end so they would be less likely to influence responses to the scenario.

***Data Preparation and Analysis***

Responses were exported from Qualtrics and screened before analysis. Participants were excluded if they did not complete the survey, failed the attention check, or did not meet the age and consent requirements. Reverse-scored items were recorded, and composite scores were calculated as described above.

For each composite, differences across the four conditions were tested using a one-way analysis of variance. When the overall test was significant, Tukey's HSD was used to compare pairs of conditions. Because data based on Likert scales may not meet the assumptions of normality and equal variance, these assumptions were checked. Where they were not met, the Kruskal–Wallis test and Dunn's post-hoc comparisons were used instead. Effect sizes are reported with significant results. The significance threshold was $\alpha = .05$.

***Ethical Considerations***

Because academic dishonesty is a sensitive topic, the study uses hypothetical scenarios and asks about intended choices rather than asking participants to cheat or disclose specific misconduct. Participation was voluntary and unpaid, and participants were told that their responses were anonymous and that they could withdraw at any time. No direct identifying information was collected. Participants under 18 were asked to confirm permission from a parent

or guardian and were excluded if they could not. These safeguards help reduce pressure to participate and protect participants' privacy.

### *Limitations of the Design*

The study measures stated intentions in a hypothetical scenario, which may not predict what participants would do in a real exam. Responses may also be influenced by social desirability, even though the survey was anonymous. The conditions differ on more than one dimension: Treatment 1 and Treatment 2 vary in both whether others find out and the kind of academic consequence described. For that reason, the comparison between Treatment 2 and Treatment 3 provides the clearest test of how formal and social consequences differ when detection is held constant. Finally, convenience sampling limits how widely the findings can be applied beyond the recruited participants.

## Results and Data Analysis

A total of 87 participants completed the survey and remained in the dataset after the screening process. The sample was predominantly under 18, with 72 participants (88%) identifying as under 18 and 10 (12%) identifying as 18 or older in the initial age question. Among the participants who answered the later demographic questions, 45 (85%) were under 18, six (11%) were between 18 and 24, and two (4%) preferred not to report their age. Gender was relatively balanced, with 26 participants (49%) identifying as male and 25 (47%) as female. The sample was predominantly Asian, with 47 participants (90%) selecting Asian as their ethnicity.

The primary analysis examined participants' stated intention to engage in academic dishonesty after exposure to one of four consequence framings: the control condition, learning-loss framing, formal academic-punishment framing, and social-shame framing. Participants responded on a five-point Likert scale, with 1 representing "Strongly disagree" and 5 representing "Strongly agree." Therefore, higher scores indicated stronger stated willingness or intention to engage in cheating-related behavior.

Three reported items were used to examine cheating intention: whether participants would consider accessing unauthorized answers, whether they would be willing to cheat, and whether they would be likely to cheat if given the opportunity. Although the methodology specified a broader composite measure, the available survey export reports these three items separately by condition. The results therefore focus on these three available indicators rather than assuming values for the unreported items.

### *Descriptive Results*

The control condition produced the highest response for the first measure, "I would consider accessing the unauthorized answers in this situation," with a mean of 3.29 (SD = 1.44, n = 14). The mean for willingness to cheat was 2.50 (SD = 1.50), while the mean for likelihood of cheating if given the opportunity was 2.43 (SD = 1.40). These results provide the baseline against which the three consequence framings can be compared.

The learning-loss condition, in which participants were told that others would not find out but that failing to learn the material could create difficulties in future classes, produced a mean of 3.08 (SD = 1.44) for considering unauthorized answers. The willingness-to-cheat item had a mean of 2.83 (SD = 1.40), while the likelihood-of-cheating item also had a mean of 2.83 (SD = 1.40). These results were relatively similar to, and in some cases higher than, the control condition. In particular, willingness to cheat increased from 2.50 in the control group to 2.83 in the learning-loss condition.

The formal-punishment condition produced substantially lower scores. Participants reported a mean of 1.83 (SD = 0.69) for considering unauthorized answers, 1.75 (SD = 0.92) for willingness to cheat, and 1.92 (SD = 1.04) for likelihood of cheating if given the opportunity. The mean for considering unauthorized answers was therefore 1.46 points lower than the corresponding control-group mean of 3.29. This was the largest descriptive difference observed between the control condition and any treatment condition on that measure.

The social-shame condition also produced relatively low responses. The mean for considering unauthorized answers was 2.14 (SD = 1.25), while willingness to cheat had a mean of 1.57 (SD = 0.73). Likelihood of cheating if given the opportunity had a mean of 1.92 (SD = 1.21). The willingness-to-cheat score of 1.57 was the lowest mean observed across all four conditions for any of the three reported measures.

Overall, the descriptive pattern was consistent across the three indicators. The control and learning-loss conditions generally produced higher cheating-intention scores, while the formal-punishment and social-shame conditions produced lower scores. The approximate average across the three reported measures was 2.74 for the control condition, 2.91 for the learning-loss condition, 1.83 for the formal-punishment condition, and 1.88 for the social-shame condition. These averages should be interpreted as descriptive summaries of the three reported items rather than as the full primary composite specified in the methodology.

Inferential Statistical Analysis

To determine whether the differences between conditions were larger than would be expected from sampling variation alone, one-way ANOVAs were calculated for each of the three reported cheating-intention measures. The results differed depending on the outcome.

For consideration of accessing unauthorized answers, the one-way ANOVA was statistically significant, indicating that the four experimental conditions differed in their mean responses. The strongest descriptive difference was between the control condition (M = 3.29) and the formal-punishment condition (M = 1.83). The formal-punishment condition showed substantially lower consideration of accessing unauthorized answers than the control condition, suggesting that emphasizing potential academic consequences may have reduced stated cheating intentions.

The ANOVA for willingness to cheat was also statistically significant, indicating differences among the four experimental conditions. The lowest mean was observed in the social-shame condition (M = 1.57), followed by the formal-punishment condition (M = 1.75), while the control and learning-loss conditions were higher at 2.50 and 2.83, respectively. This

pattern suggests that consequences involving academic punishment or social judgment were associated with lower stated willingness to cheat.

In contrast, the ANOVA for likelihood of cheating if given the opportunity was not statistically significant. Although the descriptive means followed a similar pattern, with the control and learning-loss conditions higher than the formal-punishment and social-shame conditions, the evidence was insufficient to conclude that the four conditions differed significantly on this measure. Overall, the results suggest that consequence framing was associated with differences in some measures of stated cheating intention, but the effect was not consistent across all three measures.

Pairwise comparisons further indicated that the clearest difference occurred between the control and formal-punishment conditions for consideration of accessing unauthorized answers. The difference was statistically significant in the unadjusted comparison, with $p \approx .003$, and remained significant after a conservative Bonferroni adjustment ($p \approx .019$). Other pairwise differences were weaker and did not consistently remain significant after correcting for multiple comparisons. This suggests that the overall ANOVA effects were driven primarily by the substantially lower responses in the formal-punishment condition rather than by uniform differences between every pair of conditions.

***Relationship to the Hypotheses***

The results provide partial support for the theoretical expectations developed in the literature review. The formal-punishment condition produced substantially lower cheating-intention scores than the control condition, consistent with rational-choice models of academic dishonesty. When participants were told that their behavior could be detected and affect their academic standing or future opportunities, their reported willingness to engage in cheating decreased. This pattern is particularly strong for considering unauthorized answers and willingness to cheat.

The results also provide an important distinction between the learning-loss and formal-punishment framings. The learning-loss condition did not produce the expected reduction in cheating intention relative to the control. Its approximate three-item mean was actually slightly higher than the control condition. This finding suggests that simply emphasizing the possibility of losing future learning or competence may not be sufficient to deter cheating in this sample. It therefore provides a more complicated result for the self-concept and mastery-based mechanism described in the literature review.

The social-shame condition produced low cheating-intention scores and was broadly similar to the formal-punishment condition. This provides some support for the argument that social image and concern about others' judgments can discourage dishonest behavior. However, because the formal-punishment and social-shame conditions both involved others finding out, their relatively similar results do not establish that shame is more effective than formal punishment. The social-shame condition had a slightly higher mean for considering unauthorized answers (2.14 versus 1.83) but a lower mean for willingness to cheat (1.57 versus 1.75).

Overall, the findings suggest that consequence framing mattered, but its effect depended on the specific aspect of cheating intention being measured. The strongest evidence appeared for considering unauthorized answers and willingness to cheat, both of which showed statistically significant differences across the four conditions. The absence of a significant effect for likelihood of cheating if given the opportunity indicates that participants' general stated willingness and their more direct prediction of future behavior may not respond identically to consequence framing. The findings therefore support a more nuanced behavioral-economic interpretation: formal and social consequences appear capable of reducing some forms of stated cheating intention, while a learning-loss frame alone did not produce the same deterrent effect. Given the relatively small number of participants in each condition and the hypothetical nature of the scenarios, these results should be interpreted as evidence of an association between consequence framing and stated cheating intentions rather than definitive evidence that these framings would change actual cheating behavior.

**Discussion**

The results suggest that the way consequences are framed can influence students' stated intentions toward academic dishonesty, but the effect is not the same for every type of consequence or every measure of cheating intention. Across the three reported indicators, the control and learning-loss conditions generally produced higher scores, while the formal-punishment and social-shame conditions produced lower scores. The statistically significant differences for considering unauthorized answers and willingness to cheat provide evidence that consequence framing can affect how participants respond to an opportunity to cheat. However, the nonsignificant result for likelihood of cheating suggests that these effects should not be interpreted as evidence that framing alone determines actual behavior.

The strongest finding was the effect of the formal-punishment framing. Participants who were told that others would find out and that their academic standing or future opportunities could be affected reported substantially lower cheating intentions than participants in the control condition. The difference was particularly large for consideration of accessing unauthorized answers, where the formal-punishment group had a mean of 1.83 compared with 3.29 in the control group. The overall ANOVA was significant, and the control-versus-formal-punishment comparison remained significant after a Bonferroni adjustment. This finding is consistent with economic models of academic dishonesty that describe cheating as a decision involving expected benefits, detection risk, and punishment. When the scenario made the potential external cost more immediate, participants appeared less willing to engage in the dishonest behavior.

This result also supports the idea that students may not respond to an opportunity to cheat only by considering the immediate academic benefit. The formal-punishment condition introduced a clear future cost that could affect academic standing and opportunities. From a behavioral-economic perspective, this may have changed how participants evaluated the tradeoff between the short-term benefit of a better grade and the potential consequences of being caught.

The finding is therefore consistent with previous research suggesting that perceived detection and punishment can play an important role in discouraging academic dishonesty.

The social-shame condition produced a similar overall pattern. Participants reported relatively low cheating intentions when they were told that others would find out and that they could experience embarrassment, shame, or loss of trust and respect. This provides some support for research suggesting that social image and peer judgment can influence dishonest behavior. However, the results do not show that shame was more effective than formal punishment. The two conditions were relatively similar overall, and neither consistently produced lower scores across all three measures. In fact, the social-shame condition produced a slightly higher mean than formal punishment for considering unauthorized answers, while producing a slightly lower mean for willingness to cheat. Because the formal-punishment and social-shame conditions both involve other people finding out, their similarity suggests that the possibility of social or academic consequences may matter more than the specific type of external consequence.

The learning-loss condition produced the most unexpected result. Participants in this condition were told that nobody would find out, but that failing to learn the material could create difficulties in future classes. If students value mastery and their academic development, emphasizing the loss of learning might have been expected to discourage cheating. Instead, the learning-loss condition produced results that were similar to or slightly higher than the control condition. Its approximate average across the three reported measures was 2.91, compared with 2.74 in the control group. This suggests that emphasizing the long-term loss of learning was not sufficient to reduce stated cheating intentions in this sample.

This finding is important because it complicates a simple application of loss aversion to academic dishonesty. Although behavioral economics suggests that losses can receive greater psychological weight than equivalent gains, the type of loss appears to matter. A future loss of knowledge or competence may feel less immediate than the possibility of being caught or judged. In the hypothetical exam situation, participants may have placed greater weight on the immediate benefit of improving their grade than on the possibility that they would struggle with the material later. This could help explain why the learning-loss framing did not produce the same deterrent effect as the formal-punishment or social-shame framings.

The finding also relates to the distinction between mastery and performance concerns discussed in the literature review. A student who is strongly motivated by mastery may view learning as valuable in itself, making the learning-loss message more meaningful. However, if a student is primarily focused on grades, academic standing, or external evaluation, the possibility of missing some learning may not be strong enough to outweigh the immediate benefit of cheating. The results cannot directly determine which motivation participants held because the available analysis does not separate participants according to mastery or performance orientation. Nevertheless, the pattern suggests that emphasizing learning alone may not automatically change dishonest decision-making.

Another important finding is that the significant effects were not consistent across all three measures. The ANOVA for consideration of accessing unauthorized answers and the

ANOVA for willingness to cheat were significant, while the ANOVA for likelihood of cheating if given the opportunity was not. This difference suggests that participants may distinguish between what they would consider, what they would be willing to do, and what they believe they would actually do. Consequence framing may therefore influence attitudes toward cheating without producing an equally strong change in participants' predictions of their own future behavior. This is particularly important because the study measured hypothetical intentions rather than observed cheating behavior.

Taken together, the results provide partial support for the study's original expectations. The findings support the importance of formal and social consequences, but they do not support the expectation that a learning-loss message would necessarily be more effective than punishment-based messages. Instead, the strongest evidence points toward consequences that are immediate and externally meaningful. This does not mean that learning and academic growth are unimportant. Rather, it suggests that simply reminding students that cheating could interfere with future learning may be less persuasive under assessment pressure than making the possibility of academic or social consequences more salient.

The study also contributes to the research gap identified in the literature review by examining these different mechanisms within the same experimental design. Previous research has separately considered punishment, social norms, shame, and gain-loss framing. By placing formal punishment, social shame, and learning loss into a four-condition comparison, this study provides a more direct view of how different consequence frames may influence cheating intentions. The results indicate that these mechanisms should not necessarily be treated as equally powerful. Formal punishment and social shame produced lower stated cheating intentions, whereas the learning-loss frame did not.

Several limitations should be considered when interpreting these findings. First, the study measured hypothetical intentions rather than actual cheating behavior. Participants may respond differently when facing a real exam, where stress, time pressure, academic stakes, and the immediate opportunity to cheat are present. Second, social desirability may have influenced responses even though the survey was anonymous. Third, the sample was relatively small, with approximately 22 participants per condition, which limits the precision and generalizability of the findings. Finally, the conditions were not identical in every respect. In particular, the learning-loss and formal-punishment conditions differed in whether other people found out as well as in the type of academic consequence described. Therefore, the comparison between formal punishment and social shame provides the clearest test of whether academic consequences or social judgment are more discouraging when detection is held constant.

Future research could address these limitations by using a larger and more diverse sample and by separating the different mechanisms more precisely. For example, future studies could independently manipulate detection, academic punishment, social judgment, and learning loss rather than combining some of these features within conditions. Researchers could also measure participants' mastery and performance goals to determine whether learning-loss messages are more effective for students who place greater value on mastery. Finally, studies using behavioral

measures rather than hypothetical scenarios could determine whether the patterns observed here translate into actual decisions when participants face meaningful incentives.

Overall, the findings suggest that students' responses to academic dishonesty are influenced by more than the potential academic benefit of cheating. Formal consequences and social judgment appeared to reduce several measures of stated cheating intention, while a learning-loss message alone did not produce the expected deterrent effect. The results therefore support a behavioral-economic view in which the framing and psychological salience of consequences can influence decision-making, while also showing that different forms of loss are not necessarily equally persuasive.

**Conclusion**

This study examined whether different ways of framing the consequences of academic dishonesty affect students' stated intentions to cheat. By comparing a control condition with learning-loss, formal-punishment, and social-shame framings, the study tested whether students respond differently to academic, social, and personal consequences.

The results provide partial support for the study's expectations. Formal punishment and social shame were associated with lower cheating-intention scores, with the strongest evidence appearing in the formal-punishment condition. In contrast, the learning-loss framing did not reduce cheating intentions compared with the control condition. The findings therefore suggest that consequences that are immediate and externally meaningful may have a stronger deterrent effect than simply emphasizing the possibility of losing future learning opportunities.

At the same time, the lack of a significant difference for likelihood of cheating if given the opportunity shows that the effects of framing were not consistent across every measure. Because the study measured hypothetical intentions rather than actual behavior and used a relatively small convenience sample, the findings should not be interpreted as proof that these messages would prevent cheating in real academic settings.

Despite these limitations, the study demonstrates the value of applying behavioral-economic ideas to academic dishonesty. Students may respond not only to the potential benefits of cheating, but also to how different consequences are presented and how psychologically important those consequences seem. Understanding these differences may help educators develop approaches to academic integrity that go beyond punishment alone and consider the roles of social judgment, academic consequences, and students' perceptions of learning.

**Work Cited**


Bećirović, Damir, et al. “Using Loss Aversion and Framing to Nudge Students’ Classroom Performance.” Management, vol. 27, no. 2, July 2022, pp. 5–17, https://doi.org/10.30924/mjcmi.27.2.2.

Cannito, Loreta, et al. “Measure for Measure: Effects of Money Exposure, Reward Size and Loss Aversion on Cheating.” Current Research in Behavioral Sciences, vol. 4, Jan. 2023, p. 100110, https://doi.org/10.1016/j.crbeha.2023.100110.

Chen, Guan-Zhao, et al. “Beyond Monetary Value: How Reward Type Drives Cheating in a Gender-Judgment Task.” Frontiers in Psychology, vol. 15, May 2024, p. 1290793, https://doi.org/10.3389/fpsyg.2024.1290793.

Cochran, John K., et al. “Shame, Embarrassment, and Formal Sanction Threats: Extending the Deterrence/Rational Choice Model to Academic Dishonesty.” Sociological Inquiry, vol. 69, no. 1, Jan. 1999, pp. 91–105, https://doi.org/10.1111/j.1475-682x.1999.tb00491.x.

Czibor, Eszter, et al. “Loss-Framed Incentives and Employee (Mis-)Behavior.” Management Science, vol. 68, no. 10, Feb. 2022, pp. 7518–37, https://doi.org/10.1287/mnsc.2021.4280.

Dimant, Eugen, et al. “Requiem for a Nudge: Framing Effects in Nudging Honesty.” Journal of Economic Behavior & Organization, vol. 172, Mar. 2020, pp. 247–66, https://doi.org/10.1016/j.jebo.2020.02.015.

“Requiem for a Nudge: Framing Effects in Nudging Honesty.” SSRN Electronic Journal, RELX Group (Netherlands), 1 Jan. 2019, https://doi.org/10.2139/ssrn.3416399.

Ehrmann, Thomas, et al. “Does Character Matter When Everyone Cheats? Peer Influence and Environmental Drivers of Academic Misconduct.” Kyklos, vol. 79, no. 1, Sept. 2025, pp. 112–28, https://doi.org/10.1111/kykl.70017.

Friedman, Adi, et al. “Cheating and Feeling Honest: Committing and Punishing Analog versus Digital Academic Dishonesty Behaviors in Higher Education.” Interdisciplinary Journal of E-Skills and Lifelong Learning, vol. 12, Jan. 2016, pp. 193–205, https://doi.org/10.28945/3629.

Fritz, Tanja, et al. “Elucidating the Associations Between Achievement Goals and Academic Dishonesty: A Meta-Analysis.” Educational Psychology Review, vol. 35, no. 1, Mar. 2023, https://doi.org/10.1007/s10648-023-09753-1.

Ganuthula, Venkat Ram Reddy, and Singh Kumar Manish. “Strategic Interactions in Academic Dishonesty: A Game-Theoretic Analysis of the Exam Script Swapping Mechanism.” ArXiv.Org, 17 Oct. 2025, https://doi.org/10.48550/arxiv.2510.15307.

Goyal, Saanvi. “Understanding Dishonesty: Integrating Rationality, Morality, and Social Factors.” International Journal of Science and Research (IJSR), vol. 12, no. 8, Aug. 2023, pp. 899–903, https://doi.org/10.21275/mr23807191147.

Grolleau, Gilles, et al. “Cheating and Loss Aversion: Do People Cheat More to Avoid a

Loss?" Management Science, vol. 62, no. 12, Feb. 2016, pp. 3428–38, https://doi.org/10.1287/mnsc.2015.2313.
Hua, Jocelyn. "Beyond Exams: Investigating AI Tool Impact on Student Attitudes, Ethical Awareness, and Academic Dishonesty in Online College Assessments." International Journal of Educational Management and Development Studies, vol. 4, no. 4, Dec. 2023, pp. 160–85, https://doi.org/10.53378/353030.
Kristensen, Ivar, and Per Haarr Dyvik. "The Effect of Loss Aversion on Cheating Behavior: An Experimental Study." Duo Research Archive (University of Oslo), 2019, http://hdl.handle.net/11250/2616582. University of Oslo.
Lavecchia, Adam, et al. "Behavioral Economics of Education: Progress and Possibilities." SSRN Electronic Journal, Jan. 2015, https://doi.org/10.2139/ssrn.2568069.
Lee, Wing Shing, and Chuang Yuan-Hsien. "The Impact of Moral Character Evaluations on Cheating, Emotions, and Self-Evaluation." International Journal of Research in Business and Social Science (2147-4478), vol. 11, no. 7, Nov. 2022, pp. 340–49, https://doi.org/10.20525/ijrbs.v11i7.2039.
Leisge, Kai, et al. "Framing Effects and Social Context as Determinants of Dishonest Behavior." Discover Psychology, vol. 5, no. 1, Dec. 2025, https://doi.org/10.1007/s44202-025-00494-6.
Levitt, Steven D., et al. "The Behavioralist Goes to School: Leveraging Behavioral Economics to Improve Educational Performance." American Economic Journal Economic Policy, vol. 8, no. 4, Nov. 2016, pp. 183–219, https://doi.org/10.1257/pol.20130358.
"The Behavioralist Goes to School: Leveraging Behavioral Economics to Improve Educational Performance." National Bureau of Economic Research, 1 June 2012, https://doi.org/10.3386/w18165.
Martuza, Jareef, et al. "Do Honesty‐nudges Really Work? A Large‐scale Field Experiment in an Insurance Context." Journal of Consumer Behaviour, vol. 21, no. 4, Apr. 2022, pp. 927–51, https://doi.org/10.1002/cb.2049.
Maux, Benoît Le, and Sarah Necker. "Honesty Nudges: Effect Varies with Content but Not with Timing." Journal of Economic Behavior & Organization, vol. 207, Feb. 2023, pp. 433–56, https://doi.org/10.1016/j.jebo.2023.01.011.
Mukherjee, Suvadeep, et al. "Can Theory-Informed Message Framing Drive Honest and Motivated Performance with Better Assessment Experiences in a Remote Assessment?" arXiv (Cornell University), Cornell University, 4 Feb. 2026, https://doi.org/10.48550/arxiv.2602.04450.
"Can Theory-Informed Message Framing Drive Honest and Motivated Performance with Better Assessment Experiences in a Remote Assessment?" ArXiv.Org, Feb. 2026, http://arxiv.org/abs/2602.04450.
Murdock, Tamera B., and Eric M. Anderman. "Motivational Perspectives on Student

Cheating: Toward an Integrated Model of Academic Dishonesty." Educational Psychologist, vol. 41, no. 3, Aug. 2006, pp. 129–45, https://doi.org/10.1207/s15326985ep4103_1.

Pulfrey, Caroline, et al. "Under Pressure to Achieve? The Impact of Type and Style of Task Instructions on Student Cheating." Frontiers in Psychology, vol. 10, Aug. 2019, p. 1624, https://doi.org/10.3389/fpsyg.2019.01624.

Wu, Yu-Wei, et al. "Can Priming Legal Consequences and the Concept of Honesty Decrease Cheating

Zeif, Dana, and Eldad Yechiam. "The Effect of Gain and Loss Framing on Cheating: Evidence for the During Examinations?" Frontiers in Psychology, vol. 10, Jan. 2020, p. 2887, https://doi.org/10.3389/fpsyg.2019.02887.Null." Judgment and Decision Making, vol. 20, Jan. 2025, https://doi.org/10.1017/jdm.2025.10016.

Zhao, Li, et al. "Academic Dishonesty and Its Relations to Peer Cheating and Culture: A Meta-Analysis of the Perceived Peer Cheating Effect." Educational Research Review, vol. 36, May 2022, p. 100455, https://doi.org/10.1016/j.edurev.2022.100455.

# The NBA's Anti-Tanking Reform: History, Current Problems, and the 3-2-1 Draft Lottery

By MingCheng Yuan

## Abstract

Tanking has existed in the NBA for decades because the draft rewards weak records with better access to young talent. The issue became unusually serious in the 2025–26 season, when a strong draft class, several very poor records, and the fines given to the Utah Jazz and Indiana Pacers brought new attention to competitive integrity. In May 2026, the NBA approved the 3-2-1 Draft Lottery for the 2027–2029 drafts. The new system lowers the lottery chances of the three worst teams, includes 16 teams, draws all 16 lottery positions, and adds restrictions on repeated high picks. This paper reviews the history of NBA tanking, explains the situation in 2026, and considers the strengths and weaknesses of the reform. It concludes that the policy should reduce extreme tanking, but may move the problem toward the middle lottery and create new difficulties for genuinely weak teams.

## Introduction

For many years, tanking has been one of the NBA's most difficult problems. In simple terms, tanking happens when a team makes decisions that lower its chance of winning because a worse record can lead to a better draft pick. Players are rarely asked to miss shots on purpose. More often, the decision comes from the front office: experienced players are traded, healthy stars are rested, young lineups receive heavy minutes, or a team stops trying to improve its roster late in the season.

For a team, tanking can seem reasonable. One great young player can change a franchise, and a high draft pick is cheaper and easier to control than an established star. This matters especially to small-market teams that struggle to attract major free agents. However, the same choice can hurt the league. Fans pay for competitive games, broadcasters need meaningful contests, and other teams expect serious effort. The team keeps the possible future benefit, while fans, sponsors, and the rest of the NBA share the cost.

The 2025–26 season pushed this old debate back to the center of the league. Commissioner Adam Silver said tanking was worse than it had been in recent memory, and the NBA moved from warnings and fines to a major change in the draft lottery (Reuters). On May 28, 2026, the Board of Governors approved the 3-2-1 Draft Lottery, which will be used for the 2027, 2028, and 2029 drafts (NBA Communications, "NBA Board of Governors Approves New"). The reform is important, but it is not a complete answer. It reduces the reward for finishing in the bottom three, while leaving other reasons for strategic losing in place.

## How Tanking Became Part of the NBA

The NBA did not always use a draft lottery. For many years, the first pick was decided by a coin flip between the worst teams in the two divisions. The lottery began in 1985, when seven non-playoff teams received equal chances. In 1990, the league introduced weighted odds, giving

worse teams a better chance. The worst team's odds rose to 25 percent in 1994. Another major change was approved in 2017 and began with the 2019 draft: the three worst teams each received a 14 percent chance of the first pick, and the lottery decided the first four selections (NBA Communications, "NBA Board of Governors Approves Changes"; Reynolds).

Each version tried to balance two goals. The draft should help weak teams become competitive, but it should not make losing more valuable than winning. These goals are difficult to combine. When the worst record carried a very strong chance of the first pick, teams had a clear reason to chase last place. When the odds became flatter, truly weak teams lost some protection and rebuilding became more dependent on luck.

Research suggests that teams respond to these rules. Taylor and Trogdon found that NBA performance changed when the reward for losing changed, while Price et al. also connected league policy with late-season results. Gong et al. ("Exploring Tanking Strategies") found that teams eliminated from playoff contention rested more healthy players, especially when competition for draft position was stronger. These studies support a basic point from the course materials: front offices react to incentives even when the result looks bad to fans.

The Philadelphia 76ers' "Trust the Process" period made tanking part of everyday NBA discussion. Philadelphia accepted several losing seasons, collected picks and young players, and created salary flexibility. Later results were mixed, but recent research found improvements in wins and operating income, with the financial result appearing stronger than the overall sporting return (Minuci et al.). The case showed that long-term losing can produce useful assets, but it also showed that high picks do not guarantee a championship. Draft mistakes, injuries, and poor development can waste the advantage.

**Why the Problem Became More Serious in 2026**

The 2025–26 season created a "perfect storm." The draft class was viewed as unusually strong, with prospects such as AJ Dybantsa, Darryn Peterson, Cameron Boozer, and Caleb Wilson giving weak teams a powerful reason to value lottery position. By early March, the seven worst teams had combined for only 20 wins and 87 losses since the beginning of February, although this record alone does not prove that every loss was intentional (Ingle). At the All-Star break, five teams—Brooklyn, Indiana, New Orleans, Washington, and Sacramento—had winning percentages below .290 (Reuters).

The clearest cases involved the Utah Jazz and Indiana Pacers. The NBA fined Utah $500,000 after Lauri Markkanen and Jaren Jackson Jr. were removed before the fourth quarter of two close games even though they were able to continue. Indiana received a $100,000 fine after the league concluded that Pascal Siakam and two other starters could have played, at least for reduced minutes (NBA Official). These decisions were important because they focused on team management rather than accusing players of trying to lose.

The reason teams take this risk is not complicated. A losing season can create salary space, more minutes for young players, and a chance to select a future star. It can also give owners a clear story to sell to disappointed fans: the team is suffering now because it is building

something better. For a franchise with almost no playoff chance, one extra win may feel less valuable than protecting draft position. If several weak teams think this way, none wants to be the only one still winning meaningless games and falling behind in the lottery race.

Still, what helps one franchise can hurt the NBA as a whole. Fans may stop attending games or trusting injury reports. A weak late-season schedule can also affect playoff races because some contenders face opponents using full-strength lineups while others do not. Studies find that fans generally dislike perceived tanking and that public discussion of tanking can be connected with attendance (Aiken et al.; Gong et al., "Do Consumer Perceptions"). The problem is therefore not only about draft fairness. It is also about the quality and credibility of the league's product.

**What the 3-2-1 Draft Lottery Changes**

The name "3-2-1" comes from the number of lottery balls given to different groups of teams. The lottery expands from 14 to 16 teams. The seven teams that miss both the playoffs and the Play-In Tournament, but do not finish in the bottom three, receive three balls each. The three worst teams receive only two balls, meaning they are "draft relegated." The four teams that enter the Play-In as ninth or tenth seeds also receive two balls. The two teams that lose the seventh-versus-eighth Play-In games receive one ball each (NBA Communications, "NBA Board of Governors Approves New").

Under the new distribution, the seven three-ball teams each have about an 8.1 percent chance of receiving the first pick. The bottom three and the ninth- or tenth-seeded Play-In teams have about a 5.4 percent chance, while the one-ball teams have about a 2.7 percent chance. This is a major change from the current system, in which the three worst teams each have a 14 percent chance. All 16 lottery positions will be drawn, although a bottom-three team cannot fall below the twelfth pick (NBA Communications, "NBA Board of Governors Approves New"; Reynolds). The reform also limits repeated high selections. A team's own pick cannot be first in two consecutive drafts or top five in three consecutive drafts. This rule follows the pick even if it has been traded to another team. The league will also have stronger power to reduce lottery odds, change draft position, and issue large fines when it finds clear tanking behavior. The system is temporary for three drafts, giving the NBA a chance to examine the results before deciding what should happen from 2030 onward.

**Strengths and Weaknesses of the Reform**

The strongest part of the new rule is easy to see: finishing third-worst instead of fourth-worst will no longer improve a team's chance at the first pick. It will reduce that chance from about 8.1 percent to 5.4 percent. A team near the bottom therefore has a reason to avoid falling into the last three places. Drawing all 16 positions also makes an extremely poor record less safe. These changes should reduce the most obvious race to the bottom and make some late-season games more competitive.

The rule may also help coaches and players. Coaches face conflict when front offices value draft position while players still want to win. Lower rewards for last place make it easier to develop young players without making every loss useful. More competitive games could protect ticket sales, television interest, and fan trust. The lottery does not only distribute talent; it also helps the league show that its competition is legitimate (Soebbing and Mason).

However, the reform may move tanking rather than end it. Teams finishing fourth through tenth from the bottom receive better first-pick odds than the bottom three. A team could try to stay inside that group instead of falling lower. The Play-In rules create another difficult area. The ninth- and tenth-seeded Play-In teams retain lottery access even if they later qualify for the playoffs. Most players and coaches will still try to win, but front offices may find new ways to manage rosters around these boundaries.

The second problem is fairness to genuinely weak teams. A bad record is not always proof of tanking. Injuries, unsuccessful draft choices, young rosters, and small-market limits can all produce real losing. These teams depend heavily on the draft because free agency may not provide another path to star talent. Giving them worse odds may stop deliberate losing, but it can also keep an honestly bad team near the bottom for longer. The NBA must be careful not to punish weakness as if it were cheating.

The repeated-pick rules also affect traded draft choices. Because the restriction follows a team's own pick even after a trade, another franchise may receive a weaker asset because of results it could not control. Teams made some of these trades before the new policy existed. The league should explain clearly how older trades will be treated and should consider limited protection when the rule changes the expected value of an existing deal.

Finally, enforcement will remain difficult. A team can have good basketball reasons to rest an injured player, reduce a veteran's minutes, or give young players a larger role. The Jazz and Pacers cases included specific evidence that available players could have continued, but many cases will be less clear. The NBA needs written standards, medical review, and consistent penalties. Otherwise, teams and fans may see enforcement as selective rather than fair.

**Rebuilding Without Years of Losing**

The new lottery makes extreme long-term tanking less attractive but does not remove the need to rebuild. Young players can receive meaningful minutes while experienced players keep the team organized. A rebuilding team may still lose often, but losing while developing a real system is different from making choices mainly to damage the record.

Teams should not treat the first pick as their only plan. Scouting later picks, international prospects, and undrafted players can create value, while the G League offers another path for development. Flexible contracts and careful trades can also create opportunities that do not depend on lottery luck. A draft pick is an asset that can be used on a player, saved for the future, or traded when the roster is ready for a major move.

The 2026 champion New York Knicks offer one example. The Knicks had a 17–65 season in 2018–19 but did not receive the first pick. They selected RJ Barrett third, and Barrett

later became part of the trade for OG Anunoby. Jalen Brunson arrived in free agency, while Karl-Anthony Towns, Mikal Bridges, Anunoby, and Josh Hart arrived through trades (Silverman). The lesson is not that every team can copy New York. The Knicks benefit from a large market and strong free-agent appeal. The useful lesson is that a disappointing lottery result does not have to end a rebuild. Picks and young players can become parts of a longer roster plan. The league also has work to do. It should publish clear signs that may trigger an investigation, such as repeatedly removing healthy leading players from close games or making unexplained late-season changes in availability. It should monitor not only the bottom three teams, but also teams near the middle-lottery and Play-In boundaries. During the 2027–2029 trial, the NBA should study player availability, late-season competitiveness, attendance, fan response, small-market recovery, and the value of traded picks. If tanking simply moves to a new part of the standings, the rule should be adjusted.

**An Economic Perspective**

From an economic perspective, tanking shows how rules change the value of winning and losing. A weak team may benefit from one more loss through a better draft position, while fans, broadcasters, sponsors, and other teams share the cost. A choice that seems reasonable for one franchise can therefore harm the league as a whole. The 3-2-1 system lowers the reward for finishing last, but teams may shift their strategy toward the middle of the lottery. Its success should be judged not only by the records of the worst teams, but also by late-season competition, fan interest, and rebuilding opportunities.

**Conclusion**

The lottery's history shows that no set of odds can remove tanking completely. High picks can change a team's future, while the league needs every game to remain believable. In 2025–26, a strong draft class, poor late-season records, and the Jazz and Pacers fines made this conflict unusually visible and convinced the NBA that the 2019 system was no longer enough. The 3-2-1 Lottery lowers the reward for finishing in the bottom three, increases uncertainty, restricts repeated high picks, and strengthens penalties. It should reduce extreme tanking, but it may create new pressure around the middle lottery and Play-In, make rebuilding harder for truly weak teams, and change the value of traded picks.

The reform should therefore be treated as a three-year test. Teams need several ways to find and develop talent, while the NBA needs clear enforcement and honest evaluation of where losing incentives move. The best policy will not make every team equally strong. It will give weak teams a fair path to improve without intentionally losing the most attractive choice.

**Works Cited**


Aiken, K. D., et al. "Investigating Sport Fan Attitudes toward Tanking: The Role of Values-Based Connections." Sport Marketing Quarterly, vol. 29, no. 4, 2020, pp. 256–68. https://doi.org/10.32731/SMQ.294.122020.02.

Gong, Hua, et al. "Do Consumer Perceptions of Tanking Impact Attendance at National Basketball Association Games? A Sentiment Analysis Approach." Journal of Sport Management, vol. 35, no. 3, 2021, pp. 254–65. https://doi.org/10.1123/jsm.2020-0274.

Gong, Hua, et al. "Exploring Tanking Strategies in the NBA: An Empirical Analysis of Resting Healthy Players." Sport Management Review, vol. 25, no. 3, 2022, pp. 546–66. https://doi.org/10.1080/14413523.2021.1970972.

Ingle, Sean. "NBA's Bizarre 'Tanking' Problem Has Spewed Theories but No Solutions." The Guardian, 10 Mar. 2026, https://www.theguardian.com/sport/blog/2026/mar/10/nba-tanking-problem-theories-no-solutions.

Minuci, Eduardo G., et al. "Should More Teams 'Trust the Process' of Tanking?" Contemporary Economic Policy, vol. 44, no. 2, 2026, pp. 408–25. https://doi.org/10.1111/coep.12705.

NBA Communications. "NBA Board of Governors Approves Changes to Draft Lottery System." NBA Communications, 28 Sept. 2017, https://pr.nba.com/nba-draft-lottery-reform-player-rest-rules/.

NBA Communications. "NBA Board of Governors Approves New Draft Lottery System to Address Tanking." NBA Communications, 28 May 2026, https://pr.nba.com/nba-board-of-governors-approves-new-draft-lottery-system-to-address-tanking/.

NBA Official. "NBA Fines Utah Jazz and Indiana Pacers." NBA Official, 12 Feb. 2026, https://official.nba.com/nba-fines-utah-jazz-and-indiana-pacers/.

Price, J., et al. "Tournament Incentives, League Policy, and NBA Team Performance Revisited." Journal of Sports Economics, vol. 11, no. 2, 2010, pp. 117–35. https://doi.org/10.1177/1527002510363103.

Reuters. "Reports: NBA Plans to Add Anti-Tanking Rules Next Season." Reuters, 20 Feb. 2026, https://www.reuters.com/sports/reports-nba-plans-add-anti-tanking-rules-next-season--flm-2026-02-20/.

Reynolds, Tim. "A History of the NBA Draft Lottery and Its Changes over Time." NBA.com, 28 May 2026, https://www.nba.com/news/nba-draft-lottery-changes-over-the-years.

Silverman, Blake. "What All Tanking NBA Teams Should Take Away from the Knicks' Title." Sports Illustrated, 17 June 2026, https://www.si.com/nba/what-tanking-nba-teams-should-take-away-from-knicks-title.

Soebbing, B. P., and D. S. Mason. "Managing Legitimacy and Uncertainty in Professional Team Sport: The NBA's Draft Lottery." Team Performance Management, vol. 15, nos. 3–4, 2009, pp. 141–57. https://doi.org/10.1108/13527590910964928.

Taylor, Beck A., and Justin G. Trogdon. "Losing to Win: Tournament Incentives in the National Basketball Association." Journal of Labor Economics, vol. 20, no. 1, 2002, pp. 23–41. https://doi.org/10.1086/323930.

# Canva's Critical Success Factors: The Study of Visual Design Tools and Platform Business Models By Weizi Liu


## Abstract

Canva's emerging presence within visual design industries has led businesses to increasingly heightened interest in the mechanisms that have led to such high customer loyalty. Throughout the landscape of visual design platforms, some have expanded beyond providing simple templates for posters and flyers, stretching into platform businesses with freemium models and platform ecosystems. These new implementations are what allow platforms to differentiate themselves from their competitors, while others continue to refine the features they already provide to retain their loyal customers with evolving updates. This dynamic creates a critical research gap regarding whether Canva's competitive advantage stems primarily from its foundational identity as a visual design tool or its evolution into a platform business model. Using a quantitative 7-point Likert Scale survey, this study gathered empirical data from targeted respondents who actively utilize Canva, as well as several others within the website development field, measuring four independent variables, User-Friendliness, Perceived Usefulness, AI Features, and Continuous Innovation, against Brand Loyalty. A multiple linear regression analysis found that these variables explained 66.6% of the variance in Brand Loyalty ($R^2 = 0.666$, $F(4, 125) = 62.31$, $p < 0.001$), with Continuous Innovation emerging as the strongest predictor ($\beta = 0.562$, $p < 0.001$), while Perceived Usefulness and AI Features were also significant and User-Friendliness was not. These findings indicate that Canva's competitive advantage stems primarily from its evolution into a platform business rather than its visual design identity alone, contributing to extant research on digital and hybrid businesses focused on long-term retention.


## 1. Introduction

In a constantly changing generation where visual design platforms shape the way individuals and businesses engage with the world, the demand for accessible and powerful design platforms has continued to grow. As digital content becomes the centerpiece of modern marketing and consumer engagement, the platforms enabling such creation have progressively become an incredibly prominent and consequential force in the technology field. Customer loyalty and retention are essential components of how digital platform businesses succeed, yet despite continuous inquiry, companies that try to explore the relationship between technological innovation and loyalty find it difficult to uncover the factors associated with company loyalty (Pereira et al., 2025), leaving the question of what truly compels a user to serve as both a strategic and academic challenge worth studying. As digital technologies continue to evolve, many software applications have expanded beyond functioning solely as standalone products into interconnected platform businesses that facilitate collaboration, third-party integrations, and ecosystem growth. Rather than operating exclusively as software products, these organizations increasingly combine the characteristics of both a traditional product and a digital platform,

resulting in what many scholars describe as hybrid digital businesses (Hein et al., 2020). In these hybrid models, long-term success depends not only on delivering an effective product, but also on cultivating an ecosystem that continuously creates value for multiple groups of users

Among the most renowned of these design platforms is Canva, a hybrid digital business that integrates the functionality of a visual design tool with the characteristics of a platform ecosystem. With its first official launch in 2013, Canva's foundational creation was designed to make graphic design obtainable by offering free, drag-and-drop tools alongside thousands of templates, graphics, and fonts. However, since then Canva has evolved far beyond a traditional design application through the incorporation of collaborative workspaces, artificial intelligence (AI), educational resources, external applications, and template variety, illustrating its transition toward a hybrid platform model. Many businesses, aspiring to achieve such impactful milestones as Canva are unable to find the key differences that separate sustained success from lack of growth. Numerous platforms enter the industry with comparable tools, from varying templates, pricing structures, and user-friendly services, yet no competing platform has managed to challenge Canva's dominance yet. As Hein et al. (2020) argue , digital platforms dominate our modern landscape, disrupting traditional businesses by revolutionizing how audiences access and deliver services. Additionally, in a platform ecosystem, firms that make strategic differentiation are more likely to survive within the constantly changing platform markets (Kretschmer et al., 2022).

While some companies remain solely software products and others operate primarily as platform businesses, an increasing number of organizations now occupy the intersection of both (Helfat & Raubitschek, 2018). This evolution introduced an important theoretical question regarding hybrid businesses: are their competitive advantages primarily driven by the excellence of the product itself, the strategic advantages of the platform ecosystem, or the middle ground between both? Although existing literature has extensively examined digital platforms and design software independently, few research studies investigate organizations that simultaneously operate as both. Current studies frequently emphasize either platform ecosystems and governance or usability and interface design, but rarely integrate these perspectives as a whole to explain how hybrid digital businesses achieve successful cultivation of consumer attraction and retention (Kretschmer et al. 2022). This study seeks to answer the following research questions: What are Canva's critical success factors that contribute to its ability to attract and maintain consumers? Are Canva's critical success factors primarily derived from its visual design tools or its platform business model?

## 2. Literature Review

### 2.1 Canva's Emergence Within the Visual Design Tool Industry

Founded by Melanie Perkins, Canva's foundational steps were undertaken in 2007 while teaching university design students who struggled with complex, expensive software like Adobe Photoshop (Canva, 2023). Initially developed as an online visual design application, Canva was

created to simplify the graphic design process by providing users with an intuitive drag-and-drop interface, customizable templates, and accessible design resources. Unlike traditional professional design software that often required extensive technical expertise, Canva allowed individuals and organizations across occupations to efficiently create visual content including presentations, social media graphics, posters, and marketing materials. As a result of Perkins' rigorous efforts, by 2018 the company achieved a $1 billion valuation and had expanded across many regions such as Mainland China (Canva, 2023). What had started as a simple inconvenience evolved into a globally recognized visual design tool. Visual design software has expanded substantially over the past decade, with numerous companies competing to simplify digital content creation for individuals and organizations.

Established companies such as Adobe, as well as newer competitors including Figma, VistaCreate, and Adobe Express, each provide users with accessible design capabilities while continuously expanding their digital offerings. Although these platforms share similar objectives in facilitating visual communication like Canva, they differ considerably in their target audiences, technological capabilities, collaborative functions, and overall business strategies. Adobe, one of Canva's primary competitors, represents a more traditional software-centered approach to visual design, with its Creative Cloud ecosystem centered on specialized professional design applications. Canva, in contrast, has expanded beyond its original visual design tools by incorporating platform-based features that encourage broader collaboration and ecosystem participation. Canva has differentiated itself by transitioning into a hybrid platform. This comparison between Canva and Adobe highlights the distinction between traditional visual design software and hybrid platform businesses, providing a foundation for understanding how the visual design tool market has evolved.

**2.2 Visual Design Tool Markets**

Within the growing design software industry, major players dominate the market. Companies like Adobe and Figma experience consistent growth, suggesting the presence of specific success factors that many competitors fail to see. Adobe leads the visual tool market, simplifying creation for non-experts using accessible elements, with AI features that power professional-grade creativity with deep technical precision through Photoshop for high-end media production. At the same time, demand for visual design tools has expanded rapidly because businesses, educators, entrepreneurs, and content creators increasingly depend on digital content for communication and marketing (Mitrović et al., 2020). As organizations shift toward online branding and social media engagement, visual design software has become an essential business tool rather than simply a resource for professional designers.

Individuals are exposed to design tools in everyday life, and actively participate in the design process through slide presentations or building personal websites. These tools are also designed to meet the evolving needs of professional graphic and user interface designers (Tanner, 2019). Many underestimate the centrality of tools like Adobe and Canva; design tools serve as an essential component of how individuals complete their daily tasks. Businesses increasingly rely

on social media marketing, or the ability to create engaging visual content to attract and retain audiences (Mitrović et al., 2020). The demand for high-quality content drives engagement on social media, demonstrating the growing importance of effective visual communication in digital marketing. With design tools covering a broad scope, various occupations, from entrepreneurs to business owners, utilize design tool features as a primary strategy for engagement and overall brand reputation. For many, these tools have become vital to supporting their interactions and outreach.

While many applications remain standalone design tools whose primary value comes from their software functionality, others have gradually evolved into platform businesses by supporting collaboration, integrations, third-party applications, templates created by users, AI services, educational communities, and ecosystem participation. Utilizing multi-sided platform businesses strategically is a significant business innovation model for various industries, since they allow companies to have direct interactions between two or more distinct types of affiliated customers(Evans, Hagiu, & Schmalensee, 2006). The most successful contemporary companies operate under this new business model, one predicated on collaboration, emerging technologies, externally driven innovation, and vibrant ecosystems, which we call a platform (Simon, 2011).

**2.3 Evolution of Platform Businesses**

Platform businesses are a separate category of industries that create value through direct production and exchange between firms and consumers. These systems are operated by facilitating interactions among multiple groups of users rather than relying solely on internal resources; these businesses create value by coordinating ecosystems where consumers, developers, creators, and external partners can contribute to the growth and refinement of the platform. Through these interconnected networks, platform businesses transform from being simple providers of products or services into facilitators of collaboration, innovation, and value exchange.

Digital platforms are visible, everyday platforms that challenge incumbents by changing how we consume and provide digital products and services; they serve as technical infrastructures, and their ecosystems continue to change entire industries and strategic roadmaps that oversee and orchestrate software platforms (Tiwana, 2013). Digital platforms combine and deploy technologies in new ways to effectively coordinate an ecosystem of supply and demand (Hein et al., 2019). Platform ecosystems consist of multiple participants whose activities must be coordinated to maintain stability and efficiency across communication and development. Platform leaders coordinate these relationships among participating firms and other ecosystem actors, allowing different participants to contribute resources, knowledge, and innovations that support the platform's continued development and competitive advantage (Adner, 2017; Jacobides et al., 2018; Kapoor, 2018).

Platform businesses exist across numerous industries, demonstrating their ability to transform traditional markets through digital coordination. Software-as-a-Service (SaaS) platforms represent one prominent example, where companies provide users with cloud-based

software solutions while continuously improving functionality through subscriptions, integrations, and user feedback. Beyond software, companies such as Airbnb and Uber demonstrate platform models by connecting independent suppliers with consumers through digital infrastructures that facilitate transactions, trust, and ongoing participation. Although these industries differ significantly, they share the same fundamental principle: creating value by enabling interactions among multiple stakeholders rather than solely delivering a standalone product.

According to the theory of two-sided markets, platform industries dominate as a value stream that aims to help firms overcome changes and grow sustainably. Their ability to open innovative assets maintains sustainable company growth through open innovation (Kim et al., 2019). Platform businesses are typically divided into three types: Supplier, Tailor, and Facilitator. Supplier businesses deliver resources to consumers through the platform, tailored to be consumer-centered; they are often referred to as "tailor-made," as users are in charge of how they use the platform. Facilitator businesses allow individuals to act as "prosumers," meaning both parties produce and consume products, with their services going straight to the platform (Kim et al., 2019).

Although they may share the same concept and features, platform businesses and digital businesses are not entirely the same. There are different functions that distinguish digital businesses as they create their own authentic value based on how they facilitate and maintain brand identity and consumer loyalty. As platform businesses continue to evolve, many firms have moved beyond operating under a single platform model, instead opting to diversify themselves to gain exposure to new potential during growth. Hybrid models have become a transitional milestone for many platforms to serve as gold mines toward enriching opportunities of progression. Renowned companies, including Amazon, transitioned from traditional value chain companies to hybrid multi-sided platforms (Trabucchi et al., 2021).

Well-established companies that have already built the foundation, have assets, pre-existing relationships, and levels of expertise, and customer loyalty is given the upper hand amid the transition into a hybrid multi-sided platform to unlock new value flows and innovation opportunities (Dell'Era et al., 2021). Industry platform progression requires firms to continuously manage innovation and adapt to technological and market disruptions as their ecosystems evolve. Gawer and Cusumano (2014) define internal platforms as an organizational asset from which a company can develop streams of production, whereas external platforms consist of coordinating external business ecosystems that independently develop technologies and services. Platform businesses are rarely released in their full form. Instead, they are constantly designed, managed, and altered as their ecosystems and conditions change, evolving to accomplish complex tasks in uncertain markets (Helfat & Raubitschek, 2018). Platform leaders operate in a world of market failure characterized by imperfect information about relevant choices and the outcomes of making them, in a setting that calls for highly interdependent decisions due to the frequent indirect variables that may hinder their ability to adjust and adapt in the platform business field (Helfat & Raubitschek, 2018).

A major determinant of platform success is the ability to adapt to changing consumer needs rather than relying solely on existing product functionality. Parker et al. (2016) explain that platforms can generate greater value as interactions between producers and consumers increase, while Tiwana (2013) emphasizes the importance of managing platform architecture and innovation over time. These factors allow platforms to continuously expand their capabilities and remain relevant as consumer expectations change. For Canva, this suggests that continuously expanding its design tools may play an important role in its ability to attract and maintain consumers.

In the rapidly expanding network of platform users, complementary applications capture entire markets, with many industries utilizing differentiation strategies based on distinctive positioning; a platform's performance improves only when a platform system is highly distinctive relative to its rivals (Cennamo & Santaló, 2013). Platform differentiation allows for applications to gain an upper hand in retaining their customers. While Canva positions itself between both a visual design tool and a platform business, many platforms share the same middle ground, navigating through complex strategic trade-offs rather than relying on a single tactic. As a direct competitor to Canva, Adobe pairs simplified, web-based design templates with the power of professional imaging and generative AI. Conversely, Visme is heavily focused on data visualization and business communications; it provides a hybrid drag-and-drop editor and asset marketplace that allows users to create and embed interactive infographics and presentations. VistaCreate is similar to Visme:  a budget-friendly design tool and platform that features thousands of templates, stock assets, and built-in brand kits for social media and marketing.  Given that Canva operates within a highly competitive visual design software market rather than as an isolated platform, the true inquiry remains: despite their similar functions, how can Canva continue to reshape its techniques and strategies to remain competitive in the market?

## 3. Research Methodology

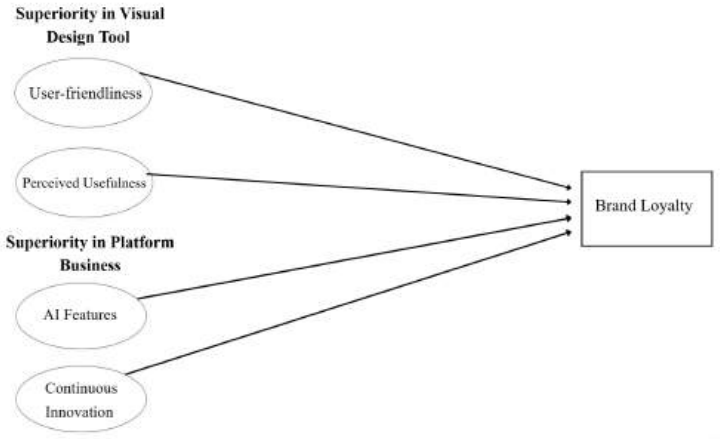



Fig 1: Research Model

### 3.1 Hypothesis Development

Figure 1 is the research model used to observe how brand loyalty is influenced by either visual design tool functionalities or platform business features. It divides four independent variables into two categories: the superiority in Visual Design Tools and Platform Businesses. The independent variables within the visual design tools consist of user-friendliness and perceived usefulness, whereas the platform business section has AI features and continuous innovation. All of these variables have arrows pointing to Brand Loyalty, the dependent variable.

Brand loyalty is the main outcome connecting all of these variables and providing the foundation for the given hypotheses. Brand loyalty is what encourages and retains the consumer to continue to use the product or service the business provides (Chaudhuri & Holbrook, 2001). It serves as a go-to business that customers can always rely on despite new, rising competitors. Chaudhuri and Holbrook (2001) show that many factors measure brand loyalty, from determining how much customers are willing to pay for services or products with brand trust and loyalty resulting in a higher relative price for the brand. These outcomes, which in turn drive brand profitability, depend on distinct dimensions of behavioral and attitudinal loyalty. Brand-loyal consumers have a greater likelihood of making greater investments into a company due to their perception of the business, believing they have unique values that few alternatives can compete with (Chaudhuri & Holbrook, 2021). Driven by the perceived company value, it forms greater trust in the reliability of the business, which creates a cycle of repeated purchases and constant investments, developing a more prominent image. Therefore, evaluating how visual design features and platform business characteristics influence brand loyalty is of high theoretical and practical significance.

This research seeks to assess whether User-Friendliness, Perceived Usefulness, AI Features, and Continuous Innovation are associated with Brand Loyalty. The analyses measure the relationships between these four independent variables and Brand Loyalty, while also comparing their relative strength within the Visual Design Tool and Platform Business categories.

**Hypothesis 1. User-friendliness is highly associated with Brand Loyalty.**

User-friendliness measures how effortless, convenient, and accessible the online platform's digital functions are for users. Davis (1989) states that user-friendliness, also considered to be perceived ease of use, is the extent to which users believe that utilizing the system requires minimal effort to dissect and comprehend. When implementing Canva into this context, user-friendliness, drag-and-drop features, and pre-made templates eliminate technical barriers, allowing individuals to navigate the platform effortlessly regardless of their design experience. As a result, minimizing user frustration and simplifying digital creation fosters a positive overall user experience, increasing user retention and encouraging long-term platform commitment over more complex alternatives. Thus, having such a close-knit relationship towards relevance and overall importance, having a convenient interface significantly increases customer loyalty and long-term relationships.

**Hypothesis 2. Perceived usefulness is highly associated with Brand Loyalty.**

Perceived usefulness is the ability to have straightforward features that increase confidence and efficiency during usage. Davis (1989) defines perceived usefulness as individuals finding the system they utilize to increase their job quality. Furthermore, terms such as "relevant" and "valuable" are also associated with user-friendliness. Thus, value parallels perceived usefulness, as both ease of use and effectiveness influence user preferences among alternative communication technologies (Davis, 1989). Perceived usefulness is demonstrated through its ability to assist users in creating professional-grade visual content rapidly, facilitate smooth communication amidst workflows, and produce high-impact assets without requiring formal training or background knowledge. When users perceive that Canva significantly elevates their work quality and efficiency, Canva as a design tool becomes indispensable, whether it be in personal or professional routines. This high utility creates dependence and appreciation, driving continuous tool use and discouraging users from migrating to alternative design software.

**Hypothesis 3. Continuous innovation is highly associated with Brand Loyalty.**

To test whether continuous innovation positively influences user brand loyalty toward Canva, identifying continuous innovation is necessary to link to this correlation. Continuous innovation refers to the business's ability to constantly evolve based on consumer needs while maintaining its reputation within the competitive industry. Huang and Ichikohji (2023) argue that business model innovation (BMI) serves as a vital complement to traditional evolutionary methods. They explain that utilizing BMI allows management to navigate the inherent trade-offs between innovation costs and returns, ultimately stabilizing the firm's strategic value. Within Canva, innovation goes beyond simple platform updates; rather, Canva refines various areas of its offers, from the template libraries to its freemium models. Through continuous platform capabilities and endless consumer adaptability, Canva's diverse platform ecosystem prevents stagnation in the stream of value towards usability.

**Hypothesis 4. AI features are highly associated with Brand Loyalty.**

AI features are a prevalent monopoly within business growth; they offer quick-response tools from automated bots that are designed to serve as passive tools for assistance. Based on prior studies, AI can serve as a tangible asset if combined with human and internal resources. When used properly, this combination captures a positive relationship with both organizational creativity and overall performance abilities (Mikalef & Gupta, 2021). Building on this prior work, AI features may allow for upgraded work capabilities, therefore building stronger customer loyalty through enhancements. By constantly fine-tuning these features and strengthening their connection to performance, platforms can improve usability by training AI to identify user needs through personalized algorithms. As a result, this can make users less likely to abandon the platform for competitors that lack comparable AI integration.

### 3.2 Secondary Research Questions: Which variables from the Design Tool or the Platform Business category have the greatest influence on Brand Loyalty?

Considering both categories of visual design tool and platform business, this question will measure which factors from which categories have stronger correlations with Brand Loyalty. The number of significant correlations from each category will be counted to see the relative importance of the two categories.

### 3.3 Research Design

In order to test hypotheses, the method consists of using a quantitative Likert Scale survey to assess the variables introduced. Through a cross-sectional survey design, this approach examines the correlations between User-Friendliness, Perceived Usefulness, AI Features, and Continuous Innovation and their relationship with Brand Loyalty. Questions were framed from a first-person perspective and assessed the perceived accessibility and benefits of each feature to an individual's daily routine (Refer to the questionnaire in Appendix 1)

The survey was sent to participants randomly sampled within the same geographical location. Data were gathered through a structured online survey that collected responses through Google Forms using a 7-point Likert scale, ranging from 1 (Strongly Disagree) to 7 (Strongly Agree). The questionnaire was structured into three modules. The first section introduced the purpose of the research to participants to clear any misunderstandings. In Section 2, Visual Design Tool Constructs (IV 1 & IV 2) measured traditional software functionality through specific 7-point Likert items. User-Friendliness (IV 1) consisted of three questions based on Davis (1989), assessing the ease of learning tools, management of the interface, and overall clarity of the features provided. Perceived Usefulness (IV 2) also included three questions based on Davis (1989), focusing on task completion efficiency, overall workflow enhancement, and utility for routine tasks.

Furthermore, Section 3 measured platform business constructs and Brand Loyalty. AI Features were assessed through three questions based on Huang and Ichikohji (2023), focusing on the perceived value added by automated features, the accuracy of generative design suggestions, and the time saved compared to manual design. Continuous Innovation was assessed through two questions based on Mikalef and Gupta (2021), focusing on the frequency of new feature releases and the platform's ability to adapt to market trends. Brand Loyalty was assessed through a single question measuring future reuse intent and preference over alternative design software.

## 4. Research Findings

The quantitative analysis evaluated N = 130 completed responses capturing user perceptions of Canva's platform attributes and overall brand loyalty. To evaluate the internal consistency of each multi-item scale prior to hypothesis testing, Cronbach's alpha was calculated. Table 1 presents the descriptive statistics along with the internal reliability scores for each construct.

| Construct | Items (k) | Mean (μ) | Standard Deviation (SD) | Cronbach's Alpha (α) |
|---|---|---|---|---|
| User-Friendliness | 3 | 5.68 | 0.98 | 0.882 |
| AI Features | 3 | 5.82 | 1.05 | 0.894 |
| Continuous Innovation | 3 | 4.62 | 1.29 | 0.835 |
| Perceived Usefulness | 2 | 5.48 | 1.18 | 0.861 |
| Brand Loyalty | 1 | 5.62 | 1.23 | — |

Table 1. Construct Reliability and Descriptive Statistics

As shown in Table 1, all multi-item constructs demonstrated strong internal consistency, with Cronbach's alpha values ranging from 0.84 to 0.89, well exceeding the standard reliability threshold of 0.70. A multiple linear regression analysis was conducted to determine how effectively the four independent variables predict Brand Loyalty. Table 2 summarizes the overall predictive power and fit of the linear regression model.

| Metric | Symbol/Formula | Computed Value |
|---|---|---|
| Multiple Correlation Coefficient | R | 0.816 |
| Coefficient of Determination | $R^2$ | 0.666 |
| Adjusted $R^2$ | Adj. $R^2$ | 0.655 |
| Standard Error of Estimate | $SE_e$ | 0.772 |
| Overall Model Significance | F(4,125) | 62.31 ($p < 0.001$) |

Table 2: Model Fit Summary

*Note.* $R$ = multiple correlation coefficient; $R^2$ = coefficient of determination; $SE_e$ = standard error of the estimate. $N$ = 130.

The model summary indicates a strong positive correlation ($r = 0.816$) between the combined independent variables and Brand Loyalty, explaining 66.6% of the total variance ($R^2 = 0.666$). To confirm whether this explanatory power is statistically significant, an Analysis of Variance (ANOVA) was performed, as presented in Table 3.

| Source of Variation | Sum of Squares | Degrees of Freedom (df) | Mean Square (MS) | F-value | p-value |
|---|---|---|---|---|---|
| Regression | 129.97 | 4 | 32.4925 | 62.31 | < 0.001 |
| Residual (Error) | 65.18 | 125 | 0.5214 | — | — |
| Total | 195.15 | 129 | — | — | — |

Table 3: ANOVA Test Table

| Independent Variable | Unstandardized Coefficient (B) | Standard Error (SE) | Standardized Beta | t-statistic | p-value | Significance Outcome |
|---|---|---|---|---|---|---|
| User-Friendliness | 0.075 | 0.091 | 0.060 | 0.824 | 0.411 | Not Significant |
| Perceived Usefulness | 0.211 | 0.085 | 0.180 | 2.482 | 0.014 | **Statistically Significant ($p < 0.05$)** |
| AI Features | 0.140 | 0.053 | 0.147 | 2.642 | 0.009 | **Statisically Significant ($p < 0.01$)** |
| Continuous Innovation | 0.586 | 0.076 | 0.562 | 7.711 | <0.001 | **Statistically Significant ($p < 0.01$)** |

Table 4: Multiple Regression Coefficients

The ANOVA results in Table 3 validate that the overall regression model is highly significant ($F(4, 125) = 62.31$, $p < 0.001$), proving that the predictors collectively account for a significant portion of user loyalty. To isolate the individual contribution and significance of each specific factor, Table 4 details the regression coefficients.

Based on the unstandardized coefficients (B) presented in Table 4, the mathematical relationship predicting Brand Loyalty is expressed in the following equation:

Brand Loyalty = 0.140 + 0.075 + 0.211 + 0.140 + 0.586

The empirical findings demonstrate that the combination of User-Friendliness, Perceived Usefulness, AI Features, and Continuous Innovation explains approximately 66% of the total variance in Canva's Brand Loyalty ($R^2 = 0.666$, $F(4, 125) = 62.31$, $p < 0.001$). Continuous

innovation emerged as the strongest predictor of brand loyalty ($\beta = 0.562$, $t = 7.711$, $p < 0.001$). Specifically, for every 1-unit increase on the 7-point Continuous Innovation scale, Brand Loyalty increases by 0.568 units when holding all other variables constant. Users view Canvaś proactive release of features and rapid adaptation of market trends as the defining reason to remain loyal to the ecosystem rather than migrating to competing platforms.

Perceived Usefulness. ($\beta = 0.180$, $t = 2.482$, $p = 0.014$) and AI Features ($\beta = 0.147$, $t = 2.642$, $p = 0.009$) both exhibited statistically significant positive relationships with brand Loyalty at the $p < 0.05$ and $p < 0.01$ levels, respectively. Practical workflow improvements directly impact user commitment. Concurrently, specialized AI additions such as Magic Studio and Background Remover act as high-value differentiators that enhance the platform's utility. User-friendliness did not reach statistical significance within the full regression model ($\beta = 0.060$, $t = 0.824$, $p = 0.411$). While descriptive statistics indicate high overall user satisfaction with usability ($\mu = 5.68$), the regression output reveals that ease of use functions as a foundational baseline. User-friendliness is necessary for adoption, but advanced utility and continuous feature rollout serve as the main driver of brand loyalty over time.

## 5. Conclusion

This study evaluated whether Canva's competitive advantage and high customer retention stem primarily from its identity as an intuitive visual design tool or its strategic evolution into a platform business model. Through multiple linear regression analysis ($N = 130$), empirical evidence confirms that while foundational product features establish initial utility, Canva's strategic position as a platform business serves as the primary engine driving long-term customer loyalty. The regression model accounted for 66.6% of the variance in Brand Loyalty, indicating that the selected factors collectively had a strong relationship with customer retention. Continuous Innovation emerged as the strongest factor associated with Brand Loyalty, while Perceived Usefulness and AI Features also showed meaningful relationships. In contrast, User-Friendliness appeared to function more as a baseline expectation than a key differentiator of retention.

These findings contribute directly to the literature on hybrid digital businesses (Hein et al., 2020; Helfat & Raubitschek, 2018), demonstrating that hybrid platforms cannot sustain market dominance through software usability alone. Canva's competitive advantage instead comes from positioning itself as an evolving platform ecosystem rather than a static visual editor. This also builds upon Parker et al. 's (2016) claim that platforms generate greater value as producer-consumer interactions increase, since Continuous Innovation's dominance as a predictor suggests that Canva's expanding ecosystem of features and interactions, rather than the interface alone, sustains loyalty over time. Moreover, the findings are also consistent with Davis's (1989) technology acceptance model, which treats ease of use as a central driver for continued adoption. However, given that User-Friendliness did not reach statistical significance in the study, it is possible that once a platform such as Canva becomes intuitive enough, usability may shift from being a differentiator of customer retention to the baseline.

For emerging software providers and hybrid digital businesses attempting to replicate Canva's market presence, these results offer clear strategic guidance. Prioritize Innovation Speed Over Interface Refinement. Allocating resources toward continuous feature rollouts and trend alignment yields higher customer retention than incrementally refining an already functional user interface. AI tools such as Magic Studio act as primary value drivers that elevate perceived usefulness rather than novelty additions. Treat Usability as a Baseline: High user-friendliness is mandatory for entry into the visual design space, but retention strategies must focus on continuous value creation to prevent user churn.

Several limitations should be acknowledged when interpreting these findings. To start, Brand Loyalty was measured using a single-item scale rather than a multi-item construct, which may, as a result, not fully capture the behavioral and attitudinal dimensions of loyalty described by Chaudhuri and Holbrook (2001). Moreover, the data collected were through convenience sampling within a single geographic location, limiting the extent to which these results can be generalized to Canva users in other regions or professional backgrounds. Future studies should use multi-item loyalty measures, broader sampling, and longitudinal designs to clarify how feature rollouts affect retention over time.

Beyond Canva specifically, these findings offer a broader lens for understanding brand loyalty within hybrid businesses generally. As software companies across industries increasingly evolve into platform-based models, identifying which factors are critical to sustain customer loyalty carries implications for the business beyond visual design. The findings may also suggest a distinction between attracting users and maintaining their loyalty. An accessible interface may function as the initial entry point that reduces barriers to adoption, but once usability becomes an expected industry standard, it may no longer be sufficient to distinguish one platform from another. Instead, continued feature development and the introduction of new capabilities may give users a reason to remain. In this sense, the design tool may serve as the hook that attracts consumers, while the evolving platform ecosystem becomes a stronger reason for them to stay. Hence, this creates an additional challenge for emerging competitors as replicating accessible design features alone may not be enough to recreate the level of Brand Loyalty associated with an established hybrid platform. Competitive advantage may therefore depend less on introducing a single successful feature and more on sustaining an ongoing cycle of adaptation.

**Works Cited**

"About Canva." Canva, 2022, www.canva.com/about.

Adner, Ron. "Ecosystem as Structure: An Actionable Construct for Strategy." Journal of Management vol. 43, no. 1, 2017, pp. 39–58. Crossref, https://doi.org/10.1177/014920 6316678451.

Adobe Creative Cloud. "Match VS." Match VS, 2026, match-vs.com/en/tool/adobe-creative-cloud.

Alt, Rainer, Witold Abramowicz, and Haluk Demirkan. "Service-Orientation in Electronic Markets." Electronic Markets, vol. 20, no. 3–4, 2010, pp. 177–80. Crossref, https://doi.org/10.1007/s12525-010-0047-6.

Cennamo, Carmelo, and Juan Santalo. "Platform Competition: Strategic Trade-offs in Platform Markets." Strategic Management Journal, vol. 34, no. 11, 2013, pp. 1331–50. Crossref, https://doi.org/10.1002/smj.2066.

Chaudhuri, Arjun, and Morris B. Holbrook. "The Chain of Effects from Brand Trust and Brand Affect to Brand Performance: The Role of Brand Loyalty." Journal of Marketing, vol. 65, no. 2, 2001, pp. 81–93. Crossref, https://doi.org/10.1509/jmkg.65.2.81.18255.

Davis, Fred D. "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology." MIS Quarterly, vol. 13, no. 3, 1989, pp. 319–40. JSTOR, https://doi.org/10.2307/249008.

Dell'Era, Claudio, Daniel Trabucchi, and Stefano Magistretti. "Exploiting Incumbents' Potentialities: From Linear Value Chains to Multi-Sided Platforms." Creativity and Innovation Management, vol. 30, no. 1, 2021, pp. 31–46. Crossref, https://doi.org/10.1111/caim.12413.

Evans, David S., Andrei Hagiu, and Richard Schmalensee. Invisible Engines. The MIT Press, 2006. Crossref, https://doi.org/10.7551/mitpress/3959.001.0001.

Gawer, Annabelle, and Michael A. Cusumano. "Industry Platforms and Ecosystem Innovation." Journal of Product Innovation Management, vol. 31, no. 3, 2014, pp. 417–33. Crossref, https://doi.org/10.1111/jpim.12105.

Hein, Andreas, Maximilian Schreieck, Thomas Riasanow, et al. "Digital Platform Ecosystems." Electronic Markets, vol. 30, no. 1, 2020, pp. 87–98. Crossref, https://doi.org/10.1007/s12525-019-00377-4.

Hein, Andreas, Jörg Weking, Maximilian Schreieck, et al. "Value Co-creation Practices in Business-to-Business Platform Ecosystems." Electronic Markets, vol. 29, no. 3, 2019, pp. 503–18. Crossref, https://doi.org/10.1007/s12525-019-00337-y.

Helfat, Constance E., and Ruth S. Raubitschek. "Dynamic and Integrative Capabilities for Profiting from Innovation in Digital Platform-Based Ecosystems." Research Policy, vol. 47, no. 8, 2018, pp. 1391–99. Crossref, https://doi.org/10.1016/j.respol.2018.01.019.

Huang, Wei, and Tomomi Ichikohji. "A Review and Analysis of the Business Model Innovation Literature." Heliyon, vol. 9, no. 7, 2023, art. e17895. Crossref, https://doi.org/10.1016/j.heliyon.2023.e17895.

Jacobides, Michael G., Carmelo Cennamo, and Annabelle Gawer. "Towards a Theory of Ecosystems." Strategic Management Journal, vol. 39, no. 8, 2018, pp. 2255–76. Crossref, https://doi.org/10.1002/smj.2904.

Kapoor, Rahul. "Ecosystems: Broadening the Locus of Value Creation." Journal of Organization Design, vol. 7, no. 1, 2018, art. 12. Crossref, https://doi.org/10.1186/s41469-018-0035-4.

Kim, Jisun, and Jemin Min. "Supplier, Tailor, and Facilitator: Typology of Platform Business Models." Journal of Open Innovation: Technology, Market, and Complexity, vol. 5, no. 3, 2019, art. 57. Crossref, https://doi.org/10.3390/joitmc5030057.

Kretschmer, Tobias, Asta Leiponen, Melissa Schilling, and Gurneeta Vasudeva. "Platform Ecosystems as Meta-Organizations: Implications for Platform Strategies." Strategic Management Journal, vol. 43, no. 3, 2022, pp. 405–24. Crossref, https://doi.org/10.1002/smj.3250.

Lucas, Henry C., and Joseph M. Goh. "Disruptive Technology: How Kodak Missed the Digital Photography Revolution." The Journal of Strategic Information Systems, vol. 18, no. 1, 2009, pp. 46–55. Crossref, https://doi.org/10.1016/j.jsis.2009.01.002.

Mikalef, Patrick, and Manish Gupta. "Artificial Intelligence Capability: Conceptualization, Measurement, Calibration, and Empirical Study on Its Impact on Organizational Creativity and Firm Performance." Information & Management, vol. 58, no. 3, 2021, art. 103434. Crossref, https://doi.org/10.1016/j.im.2021.103434.

Mitrović, K., A. Jakšić, and J. Spajić. "The Analysis of Graphic Design Platforms Used in Social Media Marketing." Proceedings of the 10th International Symposium on Graphic Engineering and Design—GRID 2020, edited by Faculty of Technical Sciences, University of Novi Sad, 2020, pp. 651–57. Crossref, https://doi.org /10.24867/GRID-2020-p74.

Oestreicher-Singer, Gal, and Lior Zalmanson. "Content or Community? A Digital Business Strategy for Content Providers in the Social Age." MIS Quarterly, vol. 37, no. 2, 2013, pp. 591–616. Crossref, https://doi.org/10.25300/MISQ/2013/37.2.12.

Parker, Geoffrey G., Marshall W. Van Alstyne, and Sangeet Paul Choudary. Platform Revolution: How Networked Markets Are Transforming the Economy and How to Make Them Work for You. W. W. Norton & Company, 2016.

Pereira, M. de S., B. S. de Castro, B. A. Cordeiro, M.G. , M. Peixoto, E. C. M. da Silva, and M. C. Gonçalves. "Factors of Customer Loyalty and Retention in the Digital Environment." Journal of Theoretical and Applied Electronic Commerce Research, vol. 20, no. 2, 2025, art. 71. Crossref, https://doi.org/10.3390/jtaer20020071.

Simon, Phil. The Age of the Platform: How Amazon, Apple, Facebook, and Google Have Redefined Business. BookBaby, 2011.

Tanner, K. W. Visual Design Tools in Support of Novice Creativity. 2019. Stanford University, PhD dissertation. Stanford Digital Repository, purl.stanford.edu/vp479xs4758.

Thakur, T. "Figma Statistics 2025: Growth, AI, Global Use." SQ Magazine, 25 Aug. 2025, sqmagazine.co.uk/figma-statistics/.

Tiwana, Amrit. Platform Ecosystems: Aligning Architecture, Governance, and Strategy. Morgan Kaufmann, 2013.

Trabucchi, Daniel, and Tommaso Buganza. "Landlords with No Lands: A Systematic Literature Review on Hybrid Multi-Sided Platforms and Platform Thinking." European Journal of Innovation Management, vol. 25, no. 6, 2021, pp. 64–96. Crossref, https://doi.org/10.1108/EJIM-11-2020-0467.

**Japanese Governmental Policy Efficacy and Intention in Addressing the Gender Wealth Gap By Ava Gilardi**

## Introduction

Japan is often acknowledged as a country with a strong economy, constant technological advancements, and high education standards. Despite this lively economy, Japan still has one of the largest gender wealth inequality gaps in the world. Many believe this is due to a strong societal emphasis on gender roles and a lack of opportunities in roles that progress to leadership because of rigid beliefs in expectations and qualifications in the workplace. Japan is a relevant case study to this issue because many countries that have high gender wealth inequality gaps are also less developed, and are countries where there exist severe women's human rights crises. However, Japan is a considerably developed country that still has a high gender inequality gap even though the government has tried to intervene in multiple forms. The Japanese government has attempted to address this issue by implementing government policies to reduce inequality through increasing child care and setting rigid expectations to prevent workplace discrimination. The success of these policies would be shown through a decrease in the wage gap and a higher percentage of leadership roles.

This essay will examine the extent to which the Japanese government policies reduce the gender wealth gap seen through human capital theory. Human Capital theory assumes that, if women invest in their human capital, they will receive higher earnings in return. In this essay, the gender wealth gap is measured through the gender income gap, using the OECD wage gap and the female share of labor income. This essay will also examine the feminist perspective on this issue and how gender wealth inequality is not caused by a lack of policy or individual choices but due to structural discrimination in society as a whole. Feminist theory assumes that structural discrimination prevents people with equal qualifications from receiving equal pay. While Human Capital Theory assumes equal qualifications lead to equal pay, feminist theory expects that gender identity will lead to pay inequality for the same quality of labor. Japanese government policies from 2011 to 2025 had limited success in reducing gender wealth inequality because, while investment in women’s human capital increased workforce participation, the policies failed to address structural barriers that limited increases in leadership roles and did not address the gender income gap.

## Methodology

When evaluating gender wealth inequality in a state and its underlying causes, there cannot be solely one specific issue at the root of this problem. This essay evaluates both qualitative and quantitative data. This essay focuses on the period from 2011-2025 to examine how different prime ministers implemented policies regarding labor practices and gender discrimination. The qualitative data were obtained from Japanese government policies on the workplace and gender discrimination, as well as research papers on Japanese social practices and how they shape beliefs about gender roles. (Oi) One limitation of this approach is that social

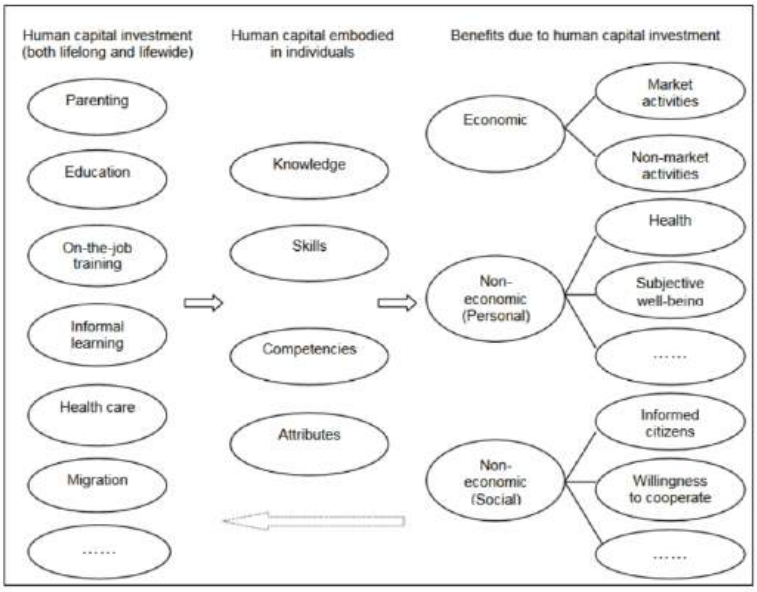
practices and expectations are both generalizations and evolving, as technology and social media are challenging historical practices, as people now have platforms for advocacy. (Council on Foreign Relations) Both measures can be used when evaluating the same evidence, to show where they converge and diverge and what this suggests about the case study.

The quantitative data was obtained through looking at the difference in average income by gender . (IMF) One limitation of this quantitative data is that it does not take into account women who are not working but instead are staying home to take care of the household and family responsibilities. Additionally, data were examined by looking at the implications of the wealth inequality gap and how that has started to shift practices and trends for women, specifically in Japan's declining birth rates, to evaluate how this gap is now affecting society as a whole and the government. Furthermore, I compared Japanese income distribution on a global scale to evaluate how this significant wealth gap affects Japan's economy as a whole when comparing it to similar countries with similar GDP or policy. (IMF) To determine the extent of policy effectiveness, I looked at the implementation, the enforcement, and the impact. Implementation considers how the policy is introduced into workplace practices and existing laws, and how long it has taken to fully implement. The enforcement of the policy considers how the policy is measured in the workplace and the strength of government intervention in aiding wealth inequality. Lastly, impact can be measured through wealth gap data, data on sustained increases in women in the workplace, and what roles they are taking on.

**Theoretical Framework**

This essay examines the theoretical framework through both feminist political theory and economic human capital theory to evaluate the political and economic lenses. Feminist theory focuses on “the exclusion of the experience of women from global politics as a discipline and political processes in public life around the world, as well as making significant contributions through exploration of patriarchy.” (Hull & Keely) Some key strands of feminist theory are related to women's role and presence in global politics. As a theoretical framework, feminist theory challenges patriarchy and how it has created male prominence in major political, economic, and social institutions. One key value of feminist theory is that it highlights the lack of work in diminishing issues solely categorized as women's and also allows structures in global politics to have inequality in systems where women are seen as second to men. One limitation to feminist theory is that it puts men in the power of being able to make change in society, which contradicts the theory within itself. (Hull & Keely) This is significant for Japan because these policies were created in a governance that is male-dominated

Human capital theory is an economic theory that states that human capital is “the knowledge, skills, competencies and attributes embodied in individuals that facilitate the creation of personal, social and economic well-being.” (OCED). This theory asserts that investment in skills can increase productivity. For Japan, productivity is the increase in wages.




The model highlights the reason behind why Japan implemented policy funding education and childcare. However, the model fails to highlight that even though parenting is a skill, in Japan this often hurts careers. Specifically to Japan, one characteristic of human capital is competence and attributes that are widely subjective and a common area for discrimination in evaluating people's human capital based on their identity and not their capital. One value of these investments is that they have different results and do not produce the same outcomes. It is important to see how, according to the Human Capital Theory, parenting embodies an investment that produces knowledge and skills. With that said, in Japan, pregnancy is looked down upon in the workplace and seen culturally as the end of one's professional life. This provides an interesting contradictory perspective in how parenting provides individuals with knowledge and skills, but also how knowledge and skills from the maternal perspective are devalued in the workplace. (OCED)

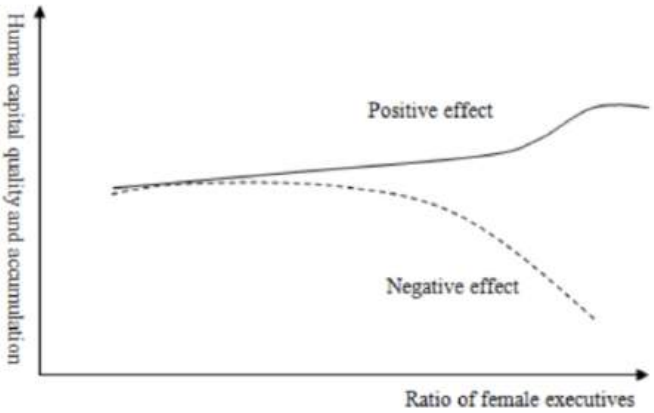



**Figure 3.** Role of female executives in human capital theory

The increasing sloping solid line represents the positive effect of increasing the ratio of female executives and the increasing level of human capital. The positive effect represents that as ratio increases and more perspectives are gained, overall human capital increases. The negative effect shows how structural barriers limit the use of female skills, which decreases human capital overall. Moreover, this shows how gender diversification is positive in increasing human capital. In Japan, women only represent 9% of executive roles, which would be trending towards the negative effect. (Zhang et al.) This graph helps illustrate that Japan is wasting its investment because if the human capital of women is lost when a woman becomes pregnant, then the country is wasting its own human capital, ultimately decreasing the workforce entirely.

Human Capital theory is a lens through which I can evaluate how different human attributes make candidates valuable to the workplace as a whole and how that can relate to a country's economic outcome. For different countries, the valuable attributes of a candidate can differ based on cultural perspectives and are not one set list of characteristics and skills. This essay will specifically weigh how attributes socially attributed to women and mothers are seen in the Japanese market and evaluate how that also affects a country's GDP. One limitation of human capital theory is that the theory does not take into account whether people's skills, attributes, and education are valued equally in the workforce due to systemic discrimination. (Oi) The combination of human capital and feminist theory would assume that Japan's current policies will increase participation in the workplace but not equal pay or job trajectory outcomes.

**Analysis of policy combating gender wealth inequality gap**

Japanese government policies have helped in starting to address the issue of gender wealth inequality in Japan by prioritizing women's role in the workplace, investing in human capital by providing government-funded programs such as subsidized daycare. Japan has a variety of laws in its constitution that have been added in the past 20 years regarding gender discrimination. The main laws that discuss this are the Basic Act for Gender Equality in the Workplace, the Equal Employment Opportunity Law, and most recently, the Promotion of Women’s Participation and Advancement Law, which was enacted in 2015. The Basic Act for Gender Equal Society states that “every citizen is able to fully display their individuality and ability regardless of gender.” (Constitution of Japan) From the human capital perspective, these policies should increase female earnings because of the removal of barriers to using skills. However, from the feminist perspective, laws that remove physical barriers do not address the workplace culture, limiting their effectiveness.

The Equal Employment Opportunity Law prevents the discrimination of men and women in all areas of employment. In this, the government emphasizes that indirect discrimination includes requirements that prevent someone from being employed there due to their gender identity. The act promoting women's participation and advancement law was put in place with the goal of recognizing gender inequality and giving opportunities to any woman who wants to participate or advance in her current employment by encouraging firms to hire more women in both private and public sectors. (Hiroshima University)

One value of the equal employment opportunity law is to help address issues with large disparities in economic participation and roles in making decisions, especially with executive roles. The sector where women are particularly left out in Japan is management, and many women feel that they can not move forward in their positions because they are not simply given the opportunity to. Another structural barrier that Japan is hoping to overcome is the long working hours associated with many jobs, and that women are seen to have the role as the sole caregivers in families, so without access to flexible work arrangements, this is not possible. (Pearson and Prichett). Additionally, since 2022, workplaces with over 100 employees have had to analyze their policies and practices so that they can fix any discrepancies and make it more inclusive. They also have to create and publish their action plan on their implementation; however, how this is being monitored is unclear. Publishing a plan does not guarantee advancement because plans have not been consistently implemented. In essence, this policy highlights the limitation of Human Capital Theory: reducing logistical barriers does not consistently lead to equal income.

**Origin: Driven by financial gains, not inclusion**

The inability to reduce gender wealth inequality through government policy is because the policy was driven by promoting the economy, not equality, to benefit women's well-being. Japan began to prioritize reducing gender wealth inequality because of a severe issue with a shrinking population and the ability to maintain a stable economy through a consistent labor force. Additionally, one motivator to implement these laws is that these current societal structures are hurting the country as a whole, and have led to declining birth rates as women are forced to choose between their career and having a child. Structural barriers, including the negative connotation in the workplace associated with being pregnant, having to take time off, and the societal expectation to be the sole caregiver of the family are often too hard for mothers to balance, so there has been a serious effort to make universal child care more sustainable so that this can happen. An alternative perspective is that the motive is unimportant if goals are similar. Prior to this, the Japanese daycare system was very complex in rules and regulations in order to get it subsidized, and there were many waitlists for this program of over 470000 kids because no daycare wanted to deal with these rules and restrictions (BoyCoff).

**Implementation: unclear guidelines & plans instead of action**

These policies were put in place to economically address issues that have systematically led to gender wealth inequality; these policies were ineffective in actually helping women have roles in leadership instead of part-time and low-paying jobs based on manual skills. When investing in Japanese policies, the government did not account for the fact that investing in human capital alone would not bridge the gender wealth gap because the value of human capital differs for men and women. From a feminist perspective, this is because women and men with equal qualifications do not lead to equal outcomes. This highlights how investing in human capital has allowed more women to join the workforce without reducing the wealth gap because

leadership roles remain highly concentrated. Additionally, when putting these practices into implementation, the government did not take into consideration the societal barriers that make it the most difficult for women to work and the shame that women's employment can bring to certain families due to gender stereotypes and beliefs, which is why they're failing to meet many of their goals and improving their gender wealth inequality. Changing societal norms and structures is very difficult and often takes a lot of time. However,, this does not mean they aren't evolving; with the presence of social media and globalization, as countries become more interconnected in different ways, these beliefs are shifting, but not at the rate these government policies hoped to enact.

Secondly, Japan's investment in human capital for women does not help in addressing the toxic workplace culture in higher-paid management positions. In Japan, having a child is seen as a weakness and a negative in the workplace, and women are often seen as less than capable because of this. The term "Matahara," which is a term for people bullying employed women for being pregnant, is a common occurrence within the workplace. (Matahara net) Because this is societally accepted, there are no repercussions for these behaviors. The policy put in place is not enough because factors including office behavior and traditional gender roles hurt Japanese economic policies. Women are also significantly disadvantaged because of harmful workplace culture because they are supposed to balance household expectations with workplace expectations, and if they are unable to balance these responsibilities, they are seen as weak. This is why many women in Japan do work, though for short periods of time, because they are forced to quit when they cannot fulfill both duties. This high turnover rate does not hurt the Japanese economy as a whole because Japan is a highly developed country with a large group of individuals who are willing and able to perform the same task, and as a result, women are typically unable to return to the job afterward. (Johnson)

**Impact: efficacy and limitations due to structural barriers**

The Japanese government policies have not been successful in targeting the root of the issue because, while more women have joined the workforce, their roles are not in long-term careers but more short-term and part-time work on account of systemic barriers preventing women from balancing long-term careers and motherhood.

One source states: "The government was only pushing to increase the number of female leaders, without offering any fundamental changes or support. The burden was on women to bear."(Oi) This quote can highlight how, while Japan's government is trying to make the effort to fix this wealth inequality, the way they are going about this is ineffective. Investing in human capital for women, the government forgot to take into account that, from a feminist perspective, the labor force rewards different genders inequitably for the same qualifications. For example, if Japan is investing in providing women ways to be able to get job training, education, and experience, if women with equal experience to men were qualified in this way, they would still be undermined by their chance in fertility and the chance they may need time away from the company. This can be shown through Japan's continued low rankings in the world economic

forums global gender gap report in 2021, where they ranked 120 out of 156 countries, which has decreased from 2016 but not to the level they were hoping to achieve.

In Japan, women have clear roles that no policy or effort has sought to truly dismantle or evolve. Often, men in Japanese culture are seen as making decisions, which correlates with leadership roles in the workplace, while women are seen as caregivers in the household. These societal expectations highlight a limitation to Human Capital Theory because of the assumption that all human capital is sought after in the same way. This case study clearly disproves this because women with the same skill level are hurt by their gender identity. T

When comparing Japan's relatively new policy that promotes equality in the workplace and also has a stable economy, in comparison to the real world, Japan is still falling behind significantly in comparison to countries with similar GDPs. From a global political perspective, the best way to track gender wealth inequality is the Organization for Economic Co-operation and Development. In 2022 Japan was ranked fourth out of 38 OECD countries with a gender wage Gap of 22%, so even though this has decreased, the rate at which it is dropping is not as significant as a rate of population decline. (OECD) Moreover, some believe that this gap can be attributed to the percentage of managerial positions, which is roughly 9%, and a high percentage of females only being able to get part-time or short-term jobs, which is roughly 50%. (Society) Japan wants to make this a priority for multiple reasons: one is because they're experiencing outrage for being a very economically stable and prosperous country, yet with very high gender wealth inequality gaps. Many argue that, to further prosper as a country, they need more women to participate in the labor markets. (Society)

Japan's female labor income share has increased since 2013, but is still significantly behind the other countries. This graph highlights how Japan, compared to other developed countries, is still significantly falling behind despite economic policies created in 2013, and how growth has started; the level of growth is nowhere near where it needs to be, which can show that until the female labor share of income can increase to be similar to other developed countries, the birth rate will still decline. While the gap in female labor share is decreasing, the rate of decrease is very slow as the population is decreasing rapidly. This suggests that the policies are not successful in raising birth rates.

**Conclusion**

Japanese government policies have reduced the gender wealth gap to a limited extent. While the policies created a strong framework for an ideal society with no discrimination, the implementation of these policies had a limit in reducing inequality. This is largely due to rigid societal norms and because the primary goals of helping bridge this gap were to stabilize the economy and stop the striking population issue, rather than addressing and condemning the large patriarchal workforce and its toxic culture. The creation of political policy in order to combat the gender wealth inequality gap has led to an increase in female participation in the workforce, but not in roles of leadership and often unstable incomes and part-time work due to family responsibility. This reality illustrates how investing in human capital can be effective in increasing participation in the workforce, rather than striving for equality within it.

The feminist perspective emphasized how the government's investment in human capital was effective in bringing more women to the workforce. Though, due to social inequality between females and males, the same qualifications do not ultimately yield the same results, which is why more women in Japan began working part-time, lower-level, and labor-intensive jobs. One limitation of the feminist theory, however, is that Japan also created this policy due to economic pressure to maintain a stable workplace. On the other hand, one value of the feminist perspective is that it illuminates the limitations of human capital theory, specifically that identical qualifications did not lead to the same output.

The perspective of Human Capital Theory was ultimately important in highlighting why the government needed to start by investing in human capital, and how women are unable to work if they don't have support in childcare and education. The main limitation of this lens is that it assumes that there is equal return for equal input. This limitation can be disproven when structural discrimination is considered as an assumption of the theory.

The inclusion of both theoretical frameworks illustrates that the investment in human capital increases the percentage of women in the workplace, without addressing the lack of opportunity for advancement into managerial positions due to structural barriers that allow women with the same qualifications to receive unequal return on investment. However, without a societal shift in customs and greater emphasis on anti-discrimination laws in the workplace, Japan is unlikely to truly and efficiently remediate the root cause issue.

## Works Cited


BoyCoff, Pamala . “6 Years after Japan Launched Its “Womenomics” Policy Is It Working?” NPR.org, 17 May 2019, www.npr.org/2019/05/23/726294619/6-years-after-japan-launched-its-womenomics-policy-is-it-working.

Brian, Keeley. OECD Insights Human Capital. Org. for Economic Cooperation & Development, 20 Feb. 2007.

Dept, International Monetary Fund Asia and Pacific. “Sustainable Path to Inclusive Growth in Japan: How to Tackle Income Inequality?” IMF Staff Country Reports, vol. 2024, no. 119, 13 May 2024, www.elibrary.imf.org/view/journals/002/2024/119/article-A003-en.xml, https://doi.org/10.5089/9798400276682.002.A003.

HULL, BEN., et al. Global Politics for the IB Diploma. Hodder Education, 26 Apr. 2024.

Johnson, Arianna C., "Abenomics’ Effect on Gender Inequality in Japanese Society and the Workplace" (2021). Honors College Theses. 583.https://digitalcommons.georgiasouthern.edu/honors-theses/583

Maringe, Felix. “Human Capital Theory - an Overview | ScienceDirect Topics.” Sciencedirect.com, 2023, www.sciencedirect.com/topics/social-sciences/human-capital-theory.

Matahara net. “WHAT IS “MATAHARA.”” マタハラNet, mataharanet.org/en/what-is-matahara/.

Nations., United. Guide on Measuring Human Capital. UNECE, 2016.

Oi, Mariko. “Why Japan Can’t Shake Sexism?” Www.bbc.com, BBC, 8 Apr. 2021, www.bbc.com/worklife/article/20210405-why-japan-cant-shake-sexism.

Pearson, Nina , and David Prichett. “Promoting Women’s Participation and Advancement in the Workplace in Japan - EDGE Certified Foundation.” EDGE Certified Foundation, 19 Feb. 2026, www.edge-cert.org/article/promoting-womens-participation-and-advancement-in-the-workplace-in-japan/. Accessed 30 Apr. 2026.

Society, Japan. “Japan Addresses the Wage Gap by Requiring Gender Pay Gap Disclosure.” Japan Society, 12 Sept. 2023, japansociety.org/news/japan-addresses-the-wage-gap-by-requiring-gender-pay-gap-disclosure/.

“Spotlight on Japan: Growing Economies through Gender Parity.” Council on Foriegn Relations, www.cfr.org/womens-participation-in-global-economy/case-studies/japan/.

Tragakes, Ellie. Economics for the IB Diploma Coursebook. 2009. 3rd ed., Cambridge University Press, 2020.

“Various Laws Relating to Gender Equality | Hiroshima University.” Www.hiroshima-U.ac.jp, www.hiroshima-u.ac.jp/en/gender/students/laws.

“World.” WID - World Inequality Database, wid.world/world/#spllin_p0p100_992_f/US.

Zhang, Chang-zheng , et al. “How Female Executives Affect Firm Performance? A Multi-Approach Perspective.” Advances in Economics and Business, vol. 4, no. 7, July 2016, pp. 351–365, https://doi.org/10.13189/aeb.2016.040705.

**Across the San Francisco Metropolitan Area By Lucas Rhee**


## Abstract

Access to essential destinations such as supermarkets, schools, and healthcare is a substantial determinant of socioeconomic inequality across various geographic environments, with the access provided by these destinations greatly depending on the quality of the offered resources and services. However, most existing accessibility research neglects quality in their assessments, instead considering every destination as providing equal access. This study develops two gravity-based accessibility indices to assess the San Francisco Metropolitan Area (SFMA): the first variation measures distance only, whereas the second weights each destination by a quality score computed by combining a Bayesian-adjusted rating with its log-scaled review volume. Libraries, hospitals, supermarkets, schools, and train stations were queried from 1,759 census tract centroids via Google Places API to compile comprehensive data on the locations, ratings, and review counts of essential destinations across the SFMA. Then, 5,177 census block groups were assessed using both accessibility indices. The two indices produced nearly identical rankings of census block group accessibility (Spearman $\rho = 0.990$), indicating that the quality adjustment contributed little additional depth to evaluating realistic accessibility. This result is likely because Google ratings are tightly clustered and were further compressed by the Bayesian adjustment. Sensitivity analyses demonstrated that rankings were stable across the quality assessment parameters and the number of nearest-neighbors evaluated in accessibility computation, but were slightly more sensitive to the distance-decay rate. These findings ultimately underscore the limits of crowd-sourced ratings as an applicable measure of quality.


## 1. Introduction

Defined as the ease with which people can reach valued or important nearby destinations, destination accessibility or destination access is a central concept and measure of socioeconomic inequality in the field of urban planning. As described classically by Hansen (1959), accessibility evaluates the spatial distribution of opportunities and the potential of interactions with those opportunities, which are often locations of services, resources, or employment. While *mobility* concerns specifically concerns the efficiency and means of travel, such as the speed and connectivity of public transportation, *accessibility* concerns outcomes of mobile living: whether people can reach or are reaching the necessary services and resources within reasonable time and cost (Levinson & Wu, 2020).

This distinction is important for urban socioeconomic equity analysis because transit connectivity and destination accessibility are distinctive measures, despite often correlation and conflation. One neighborhood may be well connected to a highway or transportation network, yet still far in distance and time from their desired destinations. Conversely, another neighborhood may be distant from means of transit, yet has many of their desired destinations in close proximity, perhaps even walking distance. These spatial inequalities in destination access are not distributed randomly across metropolitan areas—rather, decades of research have

documented strong connections between low-income neighborhoods consisting predominantly of people of color and poor access to key destinations. The consequences of this extend beyond just inconvenience: limited accessibility to important destinations often compounds poverty and reinforces socioeconomic segregation in areas already under-resourced.

These important destinations are defined as "essential destinations," types of locations that offer services and resources integral to the health and future success of individuals, such as supermarkets and schools. This study focuses on access to these essential destinations, although other research has explored disparate accessibility to other opportunities like employment, which have found that public transit accessibility is strongly correlated with employment opportunities and economic inequality in Beijing and the unreasonable reachability of employment opportunities strongly impacts economic development across cities in the U.S. (Zhang, Chun, and Joyce Man, 2015; OECD, 2018). Additionally, while accessibility is a measure of proximity primarily, the essential destinations that individuals choose to access are not necessarily the closest—in fact, many do not opt for their nearest option. Other characteristics of destinations may inform these decisions: the connectivity of a train station, the literature variety of a library, or the produce quality of a supermarket.

The San Francisco Metropolitan Area (also referred to colloquially as the Bay Area) exemplifies this complex decision involved in accessing preferable essential destinations. Encompassing nine counties of California (Alameda, Contra Costa, Marin, Napa, San Francisco, San Mateo, Santa Clara, Solano, and Sonoma), the region houses residents who endure some of the longest average commute times in the country. In the past couple decades, they have fluctuated but ultimately increased from 25 minutes in 2010 to 33 minutes in 2019 and now 30 minutes in 2024 (Metropolitan Transportation Commission, 2026). These high average employment commute times may signal a willingness to travel long distances for essential destinations as well. Additionally, the region's atypical geographic layout forces many residents to cross bridges that span miles over the bay via cars to access their desired destinations, contributing heavily to regional carbon emissions.

This research aims to quantify and compare the accessibility of essential destinations across the San Francisco Metropolitan Area, ultimately highlighting areas within the region that lack quality, accessible services and resources. Insufficient access to these necessities often afflicts poorer individuals and subsequently reinforces poverty within their neighborhoods. Therefore, the novel data and insights of quality-based accessibility presented in this paper may hopefully inform future analysis of regional inequality.

## 2. Literature Review

### 2.1 Taxonomy of Measures

There exist four primary families of accessibility measures common in existing accessibility literature. Each differs in their data requirements, facet of realistic access they attempt to emulate, and ease of interpretation for policymaking audiences.

*Nearest single-destination measures* record an origin's distance or travel time to its nearest facility of a given type, such as the closest hospital or supermarket. Nearest single-destination measures are typically the most computationally straightforward and easily communicable of the accessibility measures, and are therefore a more practical option for many studies. However, the measure is more susceptible to influences from outlier nearby destinations since it only evaluates a single destination, and thus is less realistic to the true access of a community whose residents likely access multiple destination options in their proximity instead of just one. Nearest single-destination measures are more applicable to rural deserts where there is often only one option nearby, but less appropriate for urban environments where multiple facilities are within a reachable distance.

*Cumulative opportunity measures,* or isochrone measures, sum the number of destinations of a given type reachable within a given travel time threshold. These are the most widely adopted approaches in access equity research because of their interpretability. For instance, an access evaluation like "12 grocery stores reachable within 30 minutes by transit" succinctly and comprehensibly communicates accessibility to a variety of audiences (El-Geneidy & Levinson, 2022). However, cumulative opportunity measures are subject to threshold sensitivity: the binary cutoff of its fixed travel time threshold causes drastic changes in accessibility assessments with minor adjustments in the threshold. Namely, a destination just beyond the given threshold contributes nothing to the composite accessibility score, even though it is nearly as accessible as a destination just within it. Xi et al. (2018) empirically demonstrate that such varying of the travel time threshold possibly shifts geographic patterns of inequality in access substantially, highlighting a major shortfall of cumulative opportunity measures.

*Gravity-based or potential measures* address this threshold problem by instead weighting destinations using a distance-decay function. They prioritize the nearest destinations and progressively penalize the farther locations, rather than grouping destinations by a binary cutoff. The applied decay function (negative exponential, Gaussian, log-logistic) is specified by a researcher and has significant effects on results. Gravity-based measures more realistically capture the spatial proximity component of accessibility, but are less intuitive to communicate, and the specifications of their decay function introduce an additional source of subjectivity.

Finally, *utility-based or random utility measures* are evaluated from discrete choice models involving individual travel behavior. They measure the welfare surplus individuals gain from access to destinations given their preferences, constraints, and alternatives. While, theoretically, they are the most in-depth measures, they are also the most data-intensive, requiring detailed travel behavior surveys. Thus, they are uncommon in existing destination access literature, where researchers have typically prioritized comparing accessibility across demographic groups rather than predicting behavior (Geurs & van Wee, 2004; Levinson & Wu, 2020).

**2.2 Comparing Cumulative and Gravity-Based Measures**

The convergences and trade-offs between cumulative measures and gravity-based measures have been closely examined in practice. Santana Palacios and El-Geneidy (2022)

compare the two families of measures for transit employment access in Montreal. When the cumulative threshold is set to the mean transit commute time for both low-wage and non-low-wage employment, they found correlations above 0.9. This strong correlation indicates that the decision of accessibility measure is relatively inconsequential, thereby supporting cumulative measures as the methodologically simpler option.

Kapatsila et al. (2023) replicate this comparison for eight Canadian metropolitan regions using both transit and car modes, and find similarly high (>0.9 coefficient) correlations between the two measurement types across the eight regions when the travel time threshold is set around the regional mean travel time and when the gravity measure is evaluated using a Log-Logistic or Gaussian decay function. These high coefficients suggest that the correlation between the cumulative opportunity measures and gravity-based measures is robust across diverse urban environments, indicating the same correlation would remain true for other North American metropolitan regions including the geographically unique San Francisco Metropolitan Area.

However, when applying cumulative opportunity, gravity-based, and hybrid measures to the evaluation of a bus rapid transit project, Klar et al. (2023) found that whether a project ultimately produced net positive or net negative access depended on the chosen measure and parameter. This indicates that despite such high correlations found by Santana Palacios and El-Geneidy (2022) and Kapatsila et al. (2023), the accessibility measurement can still significantly impact assessment conclusions, suggesting the importance of sensitivity analyses for parameters involved.

**2.3 Spatial Metrics and GIS Tools**

Beyond selecting a family of accessibility measures, accessibility research has significantly depended on the geospatial networks and computational tools used to estimate travel times. Mavoa et al. (2012) created a GIS-based system for measuring destination accessibility via public transit and walking in Auckland, New Zealand. They developed a combined public transit and walking accessibility index (PTWAI) that considers all stages of one transit journey: walk access, wait time, in-vehicle time, and egress. In their research, they establish that both transportation and pedestrian components of travel should be integrated to more accurately assess accessibility.

Blanchard and Waddell (2017) later extended this integration methodology into an open-source tool called UrbanAccess. It combines GTFS transit schedule data with OpenStreetMap pedestrian network data using the Pandana network analysis library to compute multimodal accessibility at a range of spatial scales. They demonstrated this tool in a San Francisco Bay Area case study, suggesting the region's validity for other future accessibility research.

Liu et al. (2021) also contribute an open-data framework using OpenStreetMap and other public sources to evaluate walkability and pedestrian destination access indicators. They found consistent inequalities in walkability and accessibility across cities, where lower access scores were concentrated in peripheral low-income regions.

### 2.4 Gaps in Existing Research

While numerous studies have tested various methods of accessibility measurement, there exists little research that considers the quality of destinations. Rather, accessibility research often weighs every destination equally, assuming that proximity alone accounts for an individual's access. However, true access is as much a measure of service quality as proximity—as mentioned in the introduction, quality plays a major role in which destinations are sought. This research therefore intends to bridge that gap in existing literature by quantifying destination quality and incorporating those measures into ultimate accessibility assessments.

## 3. Methods

### 3.1 Area of Study

This research focuses on the San Francisco Metropolitan Area (SFMA) of California as its geographic region of study, selected for its unique geography, high average commute times and distances, and proximity to the author of this paper (see the Introduction). Like any region in the United States, it is divided into census tracts and census block groups (CBG), holding 1,200 –8,000 and 600–3,000 residents respectively. For this study, census tracts and census block groups composed entirely of geographic water were removed from the datasets before their geographic centroids were extracted via QGIS, an open-source Geographic Information System. After the removal, 1,759 census tracts and 5,177 census block groups remained and were used in the data collection and analysis processes.

### 3.2 Units of Data Collection and Analysis

Census tract centroids were used as origin points for API queries, while census block groups were chosen as the spatial statistical units of analysis for accessibility scores across the SFMA. The centroids of census tracts were selected for data collection because the tracts' relatively consistent populations are served by a similar number of services and businesses regardless of their geographic size: small urban tracts are packed densely with establishments, while large rural tracts are sparsely scattered with them. Therefore, queries that originate at the centroids of these census tracts would return a more comprehensive set of places than if their origin points were evenly dispersed across the geographic region, equidistant from adjacent points. As more granular geographic divisions than census tracts, census block groups across the same region were selected as a unit of analysis for greater precision in measuring individual communities' access to the services and businesses around them.

Five "essential destination" types were selected and synthesized to measure composite accessibility: libraries, hospitals, supermarkets, schools, and train stations. These five types were selected for their significant influence on quality of living: libraries offer educational resources, hospitals grant healthcare, supermarkets supply food, schools provide education, and train stations afford access to further destinations.

### 3.3 Data Source and Method

To compile a comprehensive dataset of every essential destination across the San Francisco Metropolitan Area, the extensive location data—including the names, assigned place IDs, destination types, geographic coordinates (latitude and longitude), ratings, and total reviews count of each individual destination—publicly accessible through the Google Places API database was used. Via the programming language R in the Integrated Development Environment (IDE) RStudio, Places API queries were called using the "Googleway" R package. These queries were conducted for each of the five essential destination types as the query's "place_type" argument, with proximity prioritized for returned results (by indicating "distance" for the query's "rankby" argument) to avoid popularity bias in thorough data collection. As mentioned above, the centroids of geographic census tracts were used as the origin point of the queries. Each query returns at most 20 Google places, and these results were then appended to a dataframe containing the place data of all previous queries. Unnecessary fields in the results returned were removed, leaving the following fields: name, place ID, Google rating (which will be referred to as "rating" from here on), total reviews count (referred to as "reviews" or "review count" from here on), the index of the census tract the place was called from, its latitude, longitude, and destination type.

Once the results were collected in five dataframes (one for each essential destination type), any duplicate places (identified by their place ID) were removed. Then, the library, hospital, and supermarket dataframes were filtered by removing any places with fewer than 10 reviews to create a more realistic database of essential destinations, as many of the places with under 10 reviews were either incorrectly labeled or did not resemble the type of place people would realistically visit when attempting to access a library, hospital, or supermarket. Several places in the initial returned results were cross-referenced with their actual Google Maps entry or official website at random, and were found ineligible as a true library, hospital, or supermarket. For instance, many of the libraries in the initial dataframe were Little Free Libraries, small boxes throughout neighborhoods with which people exchange books, but not establishments considered a library. None of these Little Free Libraries had received 10 or more reviews while a majority of true libraries had, and therefore any places under 10 reviews were removed. Hospitals and supermarkets similarly contained unconventional establishments in their initial dataframes, and were filtered likewise.

However, many genuine schools and train stations had under 10 reviews as those destinations typically receive very few reviews through Google, so those two dataframes were not filtered. For the schools and train stations that did not have any rating or reviews data, their rating was replaced by the average rating of all destinations for the purposes of later analysis, while their reviews were set as 0 as the place likely received few reviews if no data was returned.

### 3.4 Destination Quality Assessment

To more accurately measure true accessibility to essential destinations, the "quality" of each essential destination was assessed and incorporated into the second variation of the final gravity-based spatial accessibility model. The calculation of quality involved a weighted sum of normalized log-review count and normalized Bayesian-adjusted ratings, with the formulas as follows:

$$R_{bayes} = \frac{n}{n+m} \cdot R + \frac{m}{n+m} \cdot \overline{R}$$

| Variable | Meaning |
|---|---|
| $R_{bayes}$ | Bayesian-adjusted rating |
| $R$ | raw rating from Google Places API, 1–5 scale |
| $\overline{R}$ | mean rating across all places in filtered dataset |
| $n$ | review count |
| $m$ | = 20, confidence threshold |

Next, the reviews count (plus 1 to account for places with 0 reviews) was adjusted to a natural log scale:

$$L = ln(1 + n)$$

| Variable | Meaning |
|---|---|
| $L$ | log-adjusted review volume |
| $n$ | review count |

Then, the Bayesian-adjusted rating and log-review count were normalized to a 0 to 1 scale:

$$R_{norm} = \frac{R_{bayes}}{5}, \quad L_{norm} = \frac{L}{L_{max}}$$

| Variable | Meaning |
|---|---|
| $L_{max}$ | maximum log-reviews value across the dataset |

Finally, the composite quality score $Q$ was calculated as the sum of the rating and log-review count, normalized and weighted:

$$Q = w_R \cdot R_{norm} + w_L \cdot L_{norm}$$

| Variable | Meaning |
|---|---|
| $Q$ | quality score |
| $w_R$ | = 0.4, weight on rating |

$w_L$ = 0.6, weight on review count

The Bayesian-adjusted rating addresses small sample size-biased ratings by pulling the original rating closer to the mean for places with a review count under the confidence threshold $m$—the minimum number of reviews a place needs to deem its rating trustworthy. The value of $m$ was selected as 20 after a sensitivity analysis using Spearman rank correlation found that ultimate quality scores are robust to changes in $m$. Spearman rank correlation was preferred to raw quality score comparisons as accessibility scores—calculated using quality scores (see Section 3.5)—exist on a relative scale and are interpreted comparatively (to identify which census block groups are relatively more or less accessible), rather than used as absolute values.

**Table 1**

*Spearman Rank Correlation Matrix for 5 values of $m$*

| | $m$ = **5** | $m$ = **10** | $m$ = **20** | $m$ = **50** | $m$ = **100** |
|---|---|---|---|---|---|
| $m$ = **5** | 1.0000 | 0.9958 | 0.9945 | 0.9920 | 0.9901 |
| $m$ = **10** | 0.9958 | 1.0000 | 0.9995 | 0.9979 | 0.9964 |
| $m$ = **20** | 0.9945 | 0.9995 | 1.0000 | 0.9993 | 0.9983 |
| $m$ = **50** | 0.9920 | 0.9979 | 0.9993 | 1.0000 | 0.9998 |
| $m$ = **100** | 0.9901 | 0.9964 | 0.9983 | 0.9998 | 1.0000 |

*Note.* Spearman rank correlation coefficients ρ between quality scores computed at each value of $m$. All pairwise ρ values exceed 0.99, indicating rankings are extremely stable across the tested range.

Spearman rank correlation matrices measure the strength of a monotonic relationship between two variables. In this case, destinations are ranked by their quality score for values of $m$ varied across five values (5, 10, 20, 50, and 100) spanning a 20-fold range, while holding the rating and review weights constant at $w_R = 0.4$ and $w_L = 0.6$. Then, each ranking was paired with every other and compared for similarity in ordering, resulting in the matrix above. Table 1 indicates that every pair of rankings is ordered almost identically, with even the pairing of range extremes ($m$ = 5 and $m$ = 100) exceeding 0.99 in their coefficient value ρ. Therefore, any value of $m$ in the tested range would produce a nearly identical ranking and has very little influence on quality scores relative to other scores.

The values of $w_R$ and $w_L$ were chosen because review count was believed to more accurately represent a place's value than its rating. Since people more often opt for favorable destinations, which subsequently receive more reviews, a review count indicates how many people the destination serves. These values were confirmed after another Spearman rank correlation analysis conducted for five pairs of review-rating weights:

$$w_R = 0.3, w_L = 0.7$$

$w_R = 0.4, w_L = 0.6$
$w_R = 0.5, w_L = 0.5$
$w_R = 0.6, w_L = 0.4$
$w_R = 0.7, w_L = 0.3$

**Table 2**

*Spearman Rank Correlation Matrix for 5 review-rating weight pair values of* $w_R$ and $w_L$

| | $w_R$ = **0.3,** $w_L$ = **0.7** | $w_R$ = **0.4,** $w_L$ = **0.6** | $w_R$ = **0.5,** $w_L$ = **0.5** | $w_R$ = **0.6,** $w_L$ = **0.4** | $w_R$ = **0.7,** $w_L$ = **0.3** |
|---|---|---|---|---|---|
| $w_R$ = **0.3,** $w_L$ = **0.7** | 1.0000 | 0.9994 | 0.9961 | 0.9797 | 0.9299 |
| $w_R$ = **0.4,** $w_L$ = **0.6** | 0.9994 | 1.0000 | 0.9985 | 0.9857 | 0.9411 |
| $w_R$ = **0.5,** $w_L$ = **0.5** | 0.9961 | 0.9985 | 1.0000 | 0.9924 | 0.9559 |
| $w_R$ = **0.6,** $w_L$ = **0.4** | 0.9797 | 0.9857 | 0.9924 | 1.0000 | 0.9756 |
| $w_R$ = **0.7,** $w_L$ = **0.3** | 0.9299 | 0.9411 | 0.9559 | 0.9756 | 1.0000 |

*Note.* Spearman rank correlation coefficients ρ between quality scores computed at each combination of review-rating weight pairs. All pairwise ρ values exceed 0.9, indicating rankings are stable across the tested range.

Similar to Table 1, Table 2 demonstrates that the values of the review and rating weights in destination quality assessment have little effect on the ranking of destinations by quality, with the lowest ρ values still exceeding 0.9, albeit lower on average than the ρ values for $m$. These high ρ values indicate the selection of weights is largely insignificant to composite accessibility scores, and the original weight values of $w_R = 0.4$ and $w_L = 0.6$ were kept.

### 3.5 Gravity-based Spatial Accessibility Modeling

To accurately calculate Euclidean distance, the coordinates of all destinations and census block group centroids were transformed between coordinate reference systems (CRS) from WGS 84 (EPSG:4326) to the CONUS Albers planar projection (EPSG:5070). Then a k-nearest-neighbor search was conducted, identifying $k = 5$ nearest destinations of each type for every CBG centroid.

For each destination type, each of the five nearest-neighbor destinations was assigned a distance-decay weight whose value was determined using the distance to its corresponding CBG centroid, and calculated by the following negative exponential decay function:

$$D_{ij} = e^{-\beta \cdot d_{ij}}$$

**Variable** **Meaning**

| | |
|---|---|
| $d_{ij}$ | Euclidean distance (km) between CBG centroid $i$ and destination $j$ |
| β | = 1.2. Decay rate parameter |
| $D_{ij}$ | distance-decay weight (0–1) |

Using these distance-decay weights, two variations of the accessibility score were then modeled: one involving strictly the distance-decay weights summed (where all quality scores are assigned a value of 1 and thereby disregarded), and another summing the destination quality scores (assessed in Section 3.4), each discounted by their corresponding distance-decay weight.

**Variation 1.** $$A_{i,t} = \sum_{j=1}^{k} D_{ij}$$

**Variation 2.** $$A_{i,t} = \sum_{j=1}^{k} D_{ij} \cdot Q_j$$

| Variable | Meaning |
|---|---|
| $k$ | number of nearest-neighbor destinations evaluated per centroid |
| $Q_j$ | Quality score of the $j$-th nearest destination |
| $A_{i,t}$ | accessibility score of CBG $i$ for destination type $t$ |

Third and fourth sensitivity analyses were conducted for 4 values of $k$ and 5 of β.

**Table 3**

*Spearman Rank Correlation Matrix for 4 values of k*

| | $k$ = **3** | $k$ = **5** | $k$ = **10** | $k$ = **15** |
|---|---|---|---|---|
| $k$ = **3** | 1.0000 | 0.9936 | 0.9756 | 0.9632 |
| $k$ = **5** | 0.9936 | 1.0000 | 0.9914 | 0.9821 |
| $k$ = **10** | 0.9756 | 0.9914 | 1.0000 | 0.9974 |
| $k$ = **15** | 0.9632 | 0.9821 | 0.9974 | 1.0000 |

*Note.* Spearman rank correlation coefficients ρ between composite accessibility scores computed at each combination of $k$. All pairwise ρ values exceed 0.95, indicating rankings are highly stable across the tested range.

**Table 4**

*Spearman Rank Correlation Matrix for 5 values of* β

| | β = **0.5** | β = **1** | β = **1.2** | β = **2** | β = **3** |
|---|---|---|---|---|---|
| β = **0.5** | 1.0000 | 0.9807 | 0.9692 | 0.9233 | 0.8760 |
| β = **1** | 0.9807 | 1.0000 | 0.9984 | 0.9753 | 0.9383 |

| | | | | | |
|---|---|---|---|---|---|
| β = **1.2** | 0.9692 | 0.9984 | 1.0000 | 0.9857 | 0.9541 |
| β = **2** | 0.9233 | 0.9753 | 0.9857 | 1.0000 | 0.9897 |
| β = **3** | 0.8760 | 0.9383 | 0.9541 | 0.9897 | 1.0000 |

*Note.* Spearman rank correlation coefficients ρ between composite accessibility scores computed at each combination of β. All pairwise ρ values exceed 0.9 except β = 0.5 and β = 3, indicating rankings are somewhat stable across the tested range.

Table 3 shows composite accessibility scores (explained in Section 3.6) for all pairwise values of $k$, all exceeding 0.95, which confirm that the number of nearest-neighbor destinations assessed has little impact on composite accessibility scores. Table 4 instead suggests that β (the decay rate parameter) has a more significant effect. This Spearman Rank Correlation contains the lowest correlation coefficients of any conducted in this research (0.8760 for β = 0.5 and β = 3), indicating that the selection of β is the most influential to ultimate accessibility scores among input parameters.

### 3.6 Composite Accessibility Index

Finally, the per-destination type accessibility scores were summed to create a single composite score on a 0–5 scale. Note that while destination types may be weighted, they were not for the purposes of this study as it is difficult to quantify a type of destination's contribution to an individual's overall access, and challenging to assess how much more important one destination type is than another. For instance, hospitals are seldom visited by the average person yet offer life-saving medical care, while schools offer a service inessential to short-term survival but significantly influential to an individual's access and well-being, especially as their life progresses.

$$A_i = \sum_{t=1}^{5} w_t \cdot A_{i,t}$$

| Variable | Meaning |
|---|---|
| $t$ | essential destination type (library, hospital, supermarket, school, train station) |
| $w_t$ | weight for destination type $t$ (all kept equal at a value of 1 for this study) |
| $A_{i,t}$ | accessibility score of CBG $i$ for destination type $t$ |
| $A_i$ | composite accessibility score for CBG $i$ |

The composite accessibility scores were joined with the shapefile geometries of their corresponding census block group by their GEOIDs, and visualized via QGIS.

## 4. Results

**Table 5**

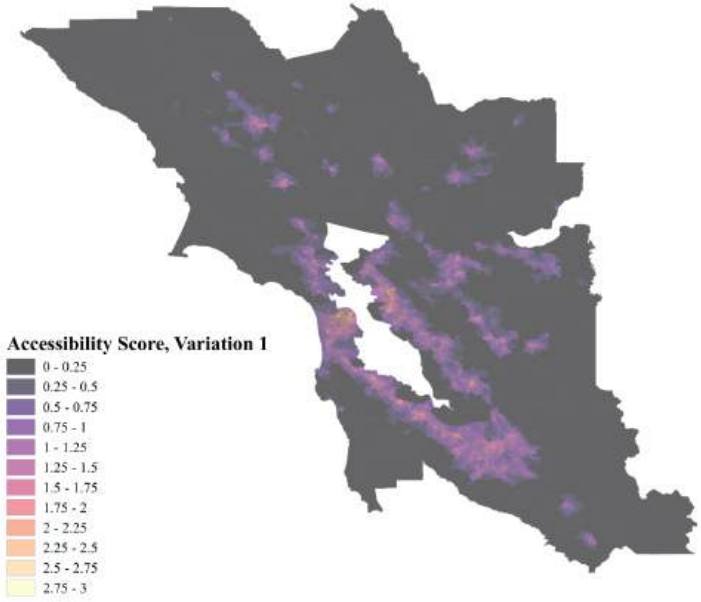
*Summary Statistics of Results*

| | Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max. | Var. |
|---|---|---|---|---|---|---|---|
| **Variation 1 Access Scores** | 0.000 | 0.652 | 0.949 | 0.982 | 1.259 | 2.862 | 0.277 |
| **Variation 2 Access Scores** | 0.000 | 0.334 | 0.515 | 0.543 | 0.703 | 1.756 | 0.099 |

*Note.* Minimum (Min.), First Quartile (1st Qu.), Median, Mean, Third Quartile (3rd Qu.), Maximum (Max.), and Variance (Var.) of Variation 1 and Variation 2 Accessibility Scores.

**Figure 1**

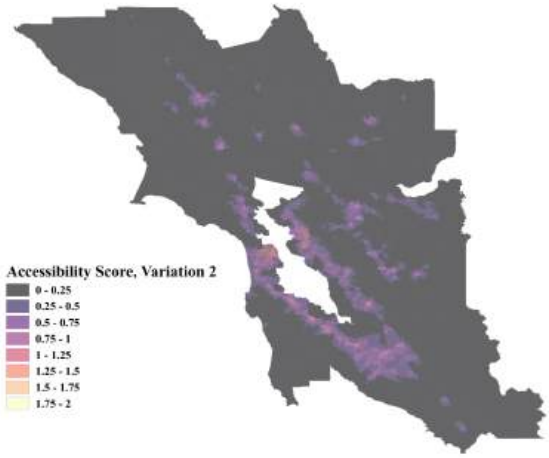
*Accessibility Scores, Variation 1 (Distance Only)*




*Note.* Visualization of Variation 1 Accessibility Scores across the entire San Francisco Metropolitan Area. Scores range from 0.00 to 2.86 in value. Legend scales from 0.00 to 3.00.

**Figure 2**

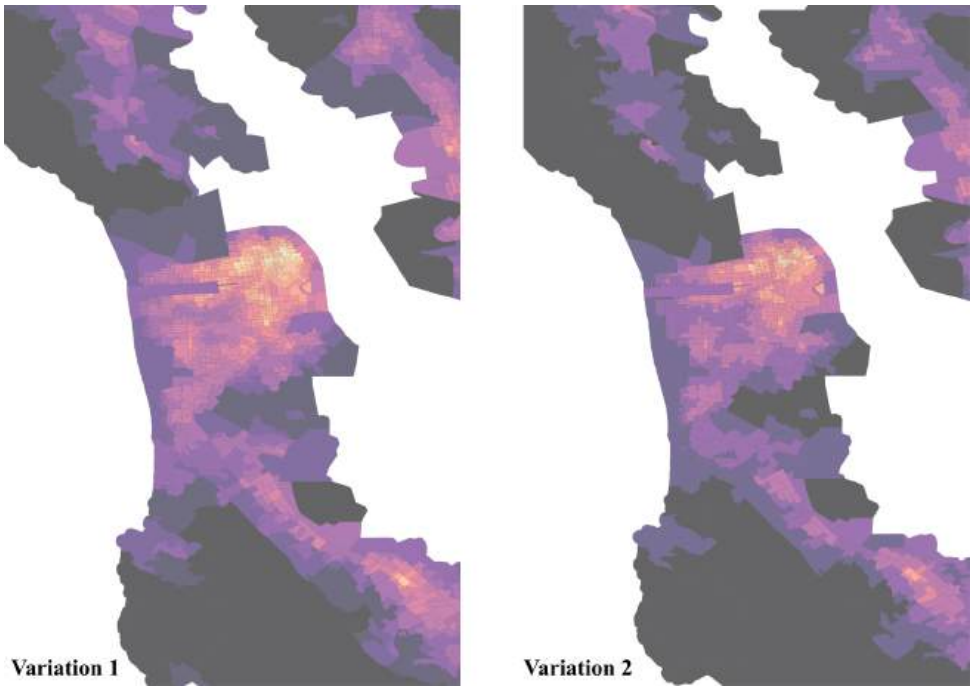
*Accessibility Scores, Variation 2 (Distance and Quality)*




*Note.* Visualization of Variation 2 Accessibility Scores across the entire San Francisco Metropolitan Area. Scores range from 0.00 to 1.76 in value. Legend scales from 0.00 to 2.00.

**Figure 3**

*Comparison between Variation 1 and Variation 2*




*Note.* Side-by-side comparative visualization of Variation 1 and Variation 2 Accessibility Scores, focused on the San Francisco region.

## 5. Discussion

### 5.1 Analysis of Results

The near-identical results between Variation 1 and Variation 2 Accessibility Scores can be visually identified in Figure 3 (note that Figure 1 and Figure 2 use different legend scales) and are confirmed by a Spearman Rank correlation coefficient of 0.9902 between the variations. A correlation value exceeding 0.99 indicates that corresponding accessibility scores between the

two variations would be practically identical if both were normalized to a consistent scale, and reveals that the incorporation of quality scores provided little additional insight into a more realistic quantification of accessibility.

The most significant contributing factor to this strong similarity is likely the lack of variance (around 0.019) in quality scores caused by an existing compression of Google ratings (which, normalized, have an extremely small variance of 0.002 and standard deviation of 0.047). This compression is also easily understood by the ratings' interquartile range of 0.200 (between a Q1 of 4.600 and a Q3 of 4.800) as well as the calculated Bayesian ratings' significantly smaller interquartile range of 0.019 (from 4.588 to 4.607). The Bayesian average evaluation shrank an already narrow middle 50% of rating values into a range of less than 0.02 stars, collapsing the distribution of values and diminishing variance of ultimate Variation 2 Accessibility Scores. This severe compression suggests the adjustment or possibly removal of the Bayesian average may be necessary to calculating meaningful quality scores using Google or other ratings with preexisting compression (which is a common occurrence for many typical rating systems).

Furthermore, the normalized ratings and normalized reviews had a negative correlation of -0.402, which highlights a common phenomenon of places with a higher volume of reviews receiving more negative ratings. This negative correlation suggests that the ratings and reviews are partially canceling each other out when calculating Bayesian ratings, explaining the resulting compression of values. These compressed Bayesian ratings contributed to a decrease in variance from 0.099 in Variation 2 (which involved quality scores assessed using Bayesian ratings) to 0.277 in Variation 1 (Table 5).

**5.2 Data Collection and Filtration Selection Biases**

The collection process of essential destinations involved many areas of possible selection bias. Selecting a spread of query origin points across the region (census tract centroids) was necessary to assemble a full dataset of destinations since Google Places API only returns a maximum of 20 results per query. However, this roundabout collection method may have missed destinations that lie on the periphery of census tracts because destinations nearest to each query origin were prioritized for returned results. With destinations missing from the assembled dataset, accessibility to some census block groups' nearest essential destinations may have been underestimated. This phenomenon is more likely for schools and supermarkets because they had the highest volume of destinations of the five destination types, while the initial datasets of train stations, libraries, and hospitals contained many duplicate entries that were removed during data filtration, indicating their datasets are likely more comprehensive.

Data filtration also involved several decisions that introduced additional selection bias. Determining whether a destination from a Google Places API query is a valid establishment of a given type is a subjective and tedious process, and the solution chosen—to remove any library, hospital, or supermarket with fewer than 10 reviews—very likely removed valid destinations while retaining invalid ones. Filtering only libraries, hospitals, and supermarkets while leaving the returned lists of schools and train stations untouched was deemed the best compromise to

avoid removing facilities of those destination types that were verified as valid, but of course contributed another dimension of subjectivity. With the data provided by Google Places API, it is impossible to determine with complete certainty that a "hospital" is in fact a hospital or instead a clinic or other healthcare facility, and impractical to determine without introducing even more selection bias that would likely invalidate subsequent analysis involving this data.

This determination of destination type validity also underlines errors in Google Places API's type classifications. After investigating several destinations via Google Maps searches, many of them were not exact matches with their respective type classification: the dataset of "hospitals" contained chiropractic clinics and "schools" contained tutoring services, for instance. Therefore, destinations for each type were likely overcounted and accessibility scores overestimated in regard to individuals' access to destinations they would realistically seek. Additionally, Google Places API—while generally offering an extensive catalog of stores and services to casual users—does not provide a completely exhaustive database of every establishment in service, and the accessibility scores of census block groups that are less documented by Google (likely more rural areas) are thereby underestimated.

### 5.3 Modeling and Methodological Limitations

The assessment of quality for essential destinations was calculated as a Bayesian average Google rating and reviews count, which favored reviews count over rating ($w_L = 0.6$ to $w_R = 0.4$) in their weightings. These weighting selections, while proven largely insignificant to accessibility scores by a Spearman Rank Correlation test (Table 2), were not informed by existing literature on rating adjustments by a Bayesian average, of which there exists unfortunately little. There are arguments to be made for different weighting pairs, which would cause minor but possibly meaningful adjustments to accessibility scores and overall research outcomes.

The selection of $k$ and β in the spatial accessibility model (Section 3.5) also possibly influenced accessibility outcomes. Accessibility scores proved particularly elastic to changes in β , whose correlation coefficient between its two range extremes was 0.8760, falling under 0.9 (see Table 4). β determined the contribution of access that destinations farther from the census block group centroid provide to the census block group's cumulative accessibility score. The value of β in a distance decay model is typically determined through empirical calibration, collecting real-world origin-destination travel data and applying the data to a regressed linear-form gravity equation. This data collection and calibration process was infeasible for the scope of this project; however, it could be conducted for future extensions of this work in an attempt to develop a more realistic accessibility assessment model. Additionally, while both Santana Palacios and El-Geneidy (2022) and Kapatsila et al. (2023) determined that the cumulative-threshold and gravity-based measures for geospatial modeling were highly correlated and led to similar outcomes for metropolitan regions across Canada, a cumulative-threshold measure may have been worth modeling to identify general model limitations for this gravity-based San Francisco Metropolitan Area study.

In addition to aggregate proximity measures, alternative distance measures could have been explored. While this model used Euclidean (straight-line) distance after projecting only the destinations and census block group centroids to a planar CRS, these origin-destination distances could have instead been calculated via a road or transportation network to more accurately measure travel distance and, consequently, composite accessibility scores. Involving a travel network for gravity-based models introduces many difficulties in model development; therefore, Euclidean distances were chosen for simplicity. Travel mode could also be explored in research extensions, introducing mobility and its influence on overall access.

## 6. Conclusion

This study attempted to more realistically measure access to essential destinations across the San Francisco Metropolitan Area by incorporating destination quality, a dimension overlooked by most accessibility research. Using Google Places data on five essential destination types, the accessibility of 5,177 census block groups was modeled with two gravity-based measures, one using distance only and another that also incorporated a quality score that synthesized Bayesian-adjusted ratings and review volume.

The primary finding is that, in the way it was implemented for this study, the addition of quality changed very little. The two measures produced nearly identical rankings of CBGs (Spearman $\rho = 0.990$), and both displayed a concentration of access in the densest neighborhoods of San Francisco, in the East Bay, and along the Peninsula stretching down to San Jose. This outcome likely reflects a narrow distribution of Google ratings, which Bayesian shrinkage further compressed, as well as a negative relationship between rating and review volume that further reduced quality score variance. The results suggest that these rating data alone may not capture destination quality well enough to meaningfully change accessibility estimates, so quality parameters should be chosen carefully.

Several methodological limitations influenced these results as well. Compiling destinations in individual queries from census tract centroids and filtering the libraries, hospitals, and supermarkets likely missed some valid establishments. Google's destination type classifications also encompassed a wider variety of destinations than what would generally be considered an essential destination. Euclidean distance omits real travel networks and mode considerations, and parameters like the decay rate $\beta$ and weights on quality components were not empirically calibrated or informed by prior literature.

Future research should address these limitations by using road or transit networks, calibrating parameters, and exploring alternative quality indicators beyond Google reviews in assessment of accessibility, like facility characteristics or performance data, although comprehensive datasets including that information for every destination are currently sparse and difficult to compile manually.

## Works Cited


Blanchard, Samuel D., and Paul Waddell. "UrbanAccess: Generalized Methodology for Measuring Regional Accessibility with an Integrated Pedestrian and Transit Network." Transportation Research Record, vol. 2653, no. 1, 2017, pp. 35–44, https://doi.org/10.3141/2653-05.

El-Geneidy, Ahmed, and David Levinson. "Making Accessibility Work in Practice." Transport Reviews, vol. 42, no. 2, 2022, pp. 129–33, https://doi.org/10.1080/01441647.2021.1975954.

Gazzé, Ludovica. "Inequality in Job Accessibility via Transit in US Cities." Divided Cities: Understanding Intra-urban Inequalities, OECD Publishing, 2018, https://doi.org/10.1787/9789264300385-en.

Geurs, Karst T., and Bert van Wee. "Accessibility Evaluation of Land-Use and Transport Strategies: Review and Research Directions." Journal of Transport Geography, vol. 12, no. 2, 2004, pp. 127-40, https://doi.org/10.1016/j.jtrangeo.2003.10.005.

Hansen, Walter G. "How Accessibility Shapes Land Use." Journal of the American Institute of Planners, vol. 25, no. 2, 1959, pp. 73–76, https://doi.org/10.1080/01944365908978307.

Kapatsila, Bogdan, et al. "Resolving the Accessibility Dilemma: Comparing Cumulative and Gravity-Based Measures of Accessibility in Eight Canadian Cities." Journal of Transport Geography, vol. 107, 2023, article 103530, https://doi.org/10.1016/j.jtrangeo.2023.103530.

Klar, Ben, et al. "The Impacts of Accessibility Measure Choice on Public Transit Project Evaluation: A Comparative Study of Cumulative, Gravity-Based, and Hybrid Approaches." Journal of Transport Geography, vol. 106, 2023, article 103508, https://doi.org/10.1016/j.jtrangeo.2022.103508.

Levinson, David M., and Hao Wu. "Towards a General Theory of Access." Journal of Transport and Land Use, vol. 13, no. 1, 2020, pp. 129–58, https://doi.org/10.5198/jtlu.2020.1660.

Liu, Shiqin, et al. "A Generalized Framework for Measuring Pedestrian Accessibility around the World Using Open Data." Geographical Analysis, vol. 54, no. 3, 2022, pp. 559–82, https://doi.org/10.1111/gean.12290.

Mavoa, Suzanne, et al. "GIS Based Destination Accessibility via Public Transit and Walking in Auckland, New Zealand." Journal of Transport Geography, vol. 20, no. 1, 2012, pp. 15–22, https://doi.org/10.1016/j.jtrangeo.2011.10.001.

Metropolitan Transportation Commission. Vital Signs: Commute Time - Region. 2026, data.bayareametro.gov/dataset/Vital-Signs-Commute-Time-Region/nmef-9x2v/about_data.

Santana Palacios, Manuel, and Ahmed El-Geneidy. "Cumulative versus Gravity-Based Accessibility Measures: Which One to Use?" Findings, 10 Feb. 2022, https://doi.org/10.32866/001c.32444.

Xi, Y. L., et al. "Exploring the Impact of Different Cut-off Times on Isochrone Measurements of Accessibility." Transportation Research Record, vol. 2672, no. 49, 2018, pp. 113-24, https://doi.org/10.1177/0361198118783113.

Zhang, Chun, and Joyce Man. "Examining Job Accessibility of the Urban Poor by Urban Metro and Bus: A Case Study of Beijing." Urban Rail Transit, vol. 1, no. 4, 2015, pp. 183–93, https://doi.org/10.1007/s40864-015-0026-5.

# AI-Powered Personal Finance in Kazakhstan and the United States: A Comparative Review of Digital Tools for Budgeting, Saving, and Financial Decision-Making By Damily Murat


## Abstract

Artificial intelligence and digital financial technologies are changing how individuals monitor spending, manage budgets, save money, and make everyday financial decisions. However, the availability and usefulness of these tools differ across national financial ecosystems. This comparative narrative review examines personal finance technologies in Kazakhstan and the United States, focusing on budgeting, automated saving, personalized financial recommendations, and the relationship between technological convenience and financial behavior. Kazakhstan provides an example of a highly integrated digital banking environment in which consumers can access payments, transfers, shopping, and other services through multifunctional applications. The United States offers a more diverse ecosystem of banks, independent financial technology providers, budgeting applications, and automated investment services. Drawing on academic research and institutional evidence, the review distinguishes conventional financial automation from machine learning and generative AI. It argues that digital financial tools can reduce the effort required to manage money, increase awareness of financial activity, and support more personalized services. Nevertheless, convenience and personalization do not necessarily produce better financial outcomes. Their effectiveness depends on financial literacy, consumer trust, data quality, transparency, and the design of the tools themselves. The comparison suggests that successful AI-powered personal finance requires not only advanced technology but also accessible services, appropriate consumer protections, and meaningful support for informed financial decision-making.



**Keywords** artificial intelligence, personal finance, financial technology, Kazakhstan, United States, budgeting, savings, digital banking, financial decision-making


## 1. Introduction

Managing personal finances involves a series of decisions that individuals make throughout their lives. People must decide how much of their income to spend, how much to save, which financial products to use, and how to prepare for unexpected expenses. These decisions can be difficult because consumers face limited time, incomplete information, competing financial priorities, and uncertainty about future income or expenses. Traditional personal finance methods, such as written budgets and spreadsheets, require individuals to collect information and monitor their own behavior. Digital financial technologies increasingly perform parts of this work automatically.

Mobile banking applications and financial technology platforms can record transactions, organize spending information, provide account notifications, and help users establish savings goals. More advanced services may apply machine learning to identify spending patterns,

estimate future expenses, or generate personalized recommendations. Generative AI can also support conversational interfaces through which users ask questions about financial information. These developments create opportunities to make financial management more accessible, but they also introduce concerns about accuracy, privacy, transparency, and excessive reliance on automated advice.

Kazakhstan and the United States provide useful cases for comparison because their digital financial ecosystems have developed through different market structures. Kazakhstan is notable for integrated mobile banking and payment services, with Kaspi.kz representing an important example of a platform combining several financial and commercial functions. The United States has a larger and more diverse market of financial institutions, banking applications, specialized budgeting services, and automated financial advice. These differences may influence how easily consumers access financial information and how personal finance tools fit into everyday spending and saving behavior.

This paper examines the research question: How do AI-powered and automated personal finance tools in Kazakhstan and the United States differ in their approaches to budgeting, saving, and financial decision-making, and what opportunities and limitations do these differences create for consumers? The central argument is that digital technology can make financial management more convenient and personalized, but the quality of financial decisions depends on more than the sophistication of the technology. Financial literacy, consumer protection, trust, and the design of financial services remain essential.

**2. Research Approach and Comparative Framework**

This paper uses a comparative narrative literature review rather than an original experiment or statistical survey. The purpose is to connect academic findings on personal financial behavior, financial technology, and automated advice with institutional evidence about the digital financial environments of Kazakhstan and the United States. This approach is appropriate because the two countries do not necessarily collect identical statistics or classify financial technology services in the same way. A narrative comparison can identify meaningful differences without assuming that every measure is directly comparable.

The analysis considers three main areas of personal financial management: budgeting, saving, and financial decision-making. Budgeting includes transaction monitoring, expense categorization, spending alerts, and financial planning. Saving includes automated transfers, goal-setting features, and tools designed to encourage regular contributions. Financial decision-making includes personalized recommendations, digital financial guidance, and automated investment advice. The review also considers risks involving privacy, algorithmic accuracy, financial literacy, and consumer trust.

An important methodological distinction concerns the meaning of AI-powered finance. Not every digital feature uses artificial intelligence. A scheduled transfer or a fixed spending alert can operate through conventional software rules. By contrast, a system that learns from transaction patterns or produces personalized predictions may involve machine learning. A

conversational financial assistant may use generative AI, but its capabilities and reliability depend on how it accesses and verifies financial information. The review therefore avoids treating all digital banking functions as evidence of AI adoption.

**3. Two Different Personal Finance Ecosystems**

Kazakhstan's digital finance environment is characterized by the important role of multifunctional banking applications. Kaspi.kz is a prominent example because its ecosystem connects payments, transfers, marketplace activity, and financial services within a single consumer platform. This integration can reduce the need to switch between separate applications for everyday transactions. For consumers, having several services available through one interface may simplify financial activity and make digital payments more convenient. However, integration should not automatically be interpreted as evidence that every service uses AI or that consumers achieve better financial outcomes.

The United States has a more varied personal finance market. Consumers may use a bank's mobile application for deposits and transfers, another service for budgeting, a separate investment platform, and additional tools for payments or financial planning. This environment creates opportunities for specialized financial products designed around particular consumer needs. It can also require users to connect information across several services. The availability of many alternatives may encourage innovation, but greater choice can make it harder for consumers to understand which tools are appropriate for their financial circumstances.

The difference between these ecosystems is important for personal finance management. An integrated platform can provide a convenient starting point for monitoring everyday transactions, while specialized applications may offer more detailed budgeting, savings, or investment functions. Neither structure guarantees better financial behavior. A consumer may have access to an excellent spending dashboard but rarely examine it. Another person may use a simple banking application consistently and maintain strong savings habits. The relationship between technology and behavior therefore depends partly on how consumers use the tools available to them.

Financial inclusion provides a broader perspective. Demirgüç-Kunt et al. (2022), in the World Bank's Global Findex Database 2021, reported that 76% of adults worldwide had an account at a bank, another financial institution, or with a mobile money provider in 2021, compared with 68% in 2017. This illustrates the expanding infrastructure through which digital financial services can reach consumers. However, account ownership is not the same as effective personal financial management. Having access to digital finance creates an opportunity to budget and save, but it does not demonstrate that a consumer possesses the knowledge or habits needed to use that opportunity effectively.

**4. Understanding Spending: Digital Budgeting Tools**

Budgeting is one of the most direct applications of digital personal finance technology. Traditional budgeting requires individuals to record income and expenses, classify purchases, and compare actual spending with financial plans. Digital applications can simplify this process

by displaying transaction histories, identifying recurring payments, and organizing expenses into categories. Some systems also provide notifications when spending approaches a predetermined limit. These features reduce the effort involved in collecting financial information and may make spending patterns easier to recognize.

More advanced systems can attempt to identify patterns that are not immediately obvious to users. Machine learning may be applied to transaction classification, recurring-expense recognition, or spending predictions. For example, a system could identify that certain expenses regularly increase near the end of the month or that several subscriptions account for a significant share of discretionary spending. The potential benefit is not simply automation but the ability to transform transaction records into information that consumers can use. Nevertheless, these capabilities must be distinguished from basic rule-based categorization, which does not necessarily require AI.

Budgeting tools also interact with behavioral factors. Financial decisions are not always based on careful calculations. Consumers may underestimate small repeated purchases, prioritize immediate rewards, or fail to monitor spending consistently. Lusardi and Mitchell (2014) explain that financial literacy is closely connected with economic decision-making, while Fernandes et al. (2014) demonstrate that financial education and financial behavior have a complicated relationship. Digital tools may help users notice spending patterns, but providing information does not guarantee that they will change their behavior.

The Kazakhstan-U.S. comparison is useful here because budgeting functions may be delivered through different types of platforms. In Kazakhstan, spending information can be accessed through major banking applications that consumers already use for everyday transactions. In the United States, consumers can use banking applications as well as independent budgeting services. The important question is not which country has more applications, but whether the available tools help consumers understand spending, establish realistic limits, and make decisions consistent with their financial goals.

**5. From Saving Intentions to Saving Habits**

Saving money is often difficult because consumers must balance future financial goals against immediate spending needs. Even when individuals intend to save, they may postpone transferring money or spend funds that were originally reserved for another purpose. Digital financial services can address some of these difficulties through automatic transfers, recurring contributions, savings goals, and reminders. These features can reduce the number of decisions a person must make each month and make saving more consistent.

Behavioral research provides an important foundation for understanding why automation may help. Thaler and Benartzi (2004) examined the Save More Tomorrow program, which encouraged employees to commit in advance to increasing retirement savings when their salaries rose. In the program's first implementation, participants' average saving rates increased from 3.5% to 13.6% over 40 months. This finding concerns retirement savings rather than modern

mobile banking applications, but it illustrates a broader principle: financial behavior can improve when saving is connected to well-designed commitments and automatic processes.

Digital applications can apply related ideas through recurring transfers and goal-based savings. A consumer may establish an emergency fund and arrange for a fixed amount to be transferred after each salary payment. More advanced tools may attempt to identify opportunities to save based on income and spending patterns. Such systems could potentially recommend smaller contributions during expensive periods and larger ones when the consumer has more available funds. However, the reliability of these recommendations depends on whether the system accurately understands the user's financial obligations and circumstances.

The two countries may offer different pathways to these functions. In Kazakhstan, automated saving can be integrated into banking services that consumers already use for payments and deposits. In the United States, saving tools may be provided by banks, investment platforms, and independent financial technology companies. These differences affect convenience and choice, but they do not establish which country produces better saving habits. The strongest evidence concerns the potential benefits of automation and commitment mechanisms, while direct comparative evidence about AI-driven savings outcomes in Kazakhstan and the United States remains limited.

**6. Financial Decisions and the Role of AI Advice**

Personal financial decisions extend beyond tracking expenses and transferring money into savings. Consumers must also decide how to manage debt, compare financial products, prepare for emergencies, and invest for long-term goals. These decisions can involve unfamiliar terminology and uncertainty about future outcomes. AI-powered services may help by organizing information, explaining financial concepts, comparing options, or generating recommendations based on user-provided information. Their usefulness, however, depends on the quality of the information and the appropriateness of the advice.

Robo-advisors provide an established example of automated financial decision support. These services typically collect information about an investor's goals, time horizon, and risk tolerance before recommending or managing an investment portfolio. D'Acunto et al. (2019) examined robo-advising in retail investment and found that automated advice can influence portfolio decisions, particularly among investors who previously held poorly diversified portfolios. This demonstrates that financial technology can affect actual decisions rather than simply making information easier to access. However, investment management is only one part of personal finance, and findings about robo-advisors should not automatically be applied to budgeting applications.

Generative AI introduces a different type of interaction because users may ask financial questions in conversational language. A person might request an explanation of compound interest, ask for help understanding a budget, or compare different saving strategies. Such interfaces could make financial information more accessible to people who find traditional financial terminology difficult. At the same time, generative AI can produce inaccurate or

incomplete explanations. A response that sounds confident may still overlook important facts about a user's income, obligations, financial goals, or risk tolerance. Recent academic research continues to examine the relationship between AI-based financial advice, financial knowledge, and consumer behavior. Olajide et al. (2026), writing in Financial Planning Review, investigate AI use for financial advice and the role of financial knowledge. This research direction highlights an important distinction between access to advice and the ability to evaluate it. A financially knowledgeable consumer may use AI as an additional information source, while a less knowledgeable user may have difficulty recognizing inappropriate recommendations. AI can therefore support financial decision-making, but it should not replace consumer judgment or appropriate professional advice in consequential situations.

**7. Kazakhstan versus the United States: Comparative Analysis**

The comparison between Kazakhstan and the United States reveals two different approaches to the organization of digital financial services. Kazakhstan's market illustrates the importance of integrated banking ecosystems, while the United States provides a wider range of specialized personal finance products. These differences influence how consumers encounter financial technology. In an integrated ecosystem, users may access spending information and financial products through an application they already use frequently. In a more fragmented environment, consumers may choose specialized services that offer additional functionality but require connections between separate accounts and platforms.

The distinction also affects the potential development of AI-powered personal finance. Integrated platforms may be well positioned to provide financial insights based on activity occurring within their ecosystems. Specialized platforms may focus more narrowly on budgeting, automated saving, financial planning, or investment advice. However, access to transaction data does not automatically imply that a platform uses machine learning or provides personalized financial recommendations. Product capabilities must be established through provider documentation or independent evaluations rather than inferred from the popularity of an application.

The countries also differ in the institutional context surrounding digital finance. The United States has an established system of consumer financial regulation, financial reporting, and research on household financial behavior. For example, the Federal Reserve's Survey of Household Economics and Decisionmaking provides recurring evidence about financial well-being and the ability of households to manage unexpected expenses. Kazakhstan's National Bank and financial regulators provide information about payment systems, financial institutions, and the development of digital financial services. These sources are useful for understanding national markets, although their indicators should not be treated as directly comparable without examining their definitions and collection methods.

The main comparative finding is therefore structural rather than a claim that one country is technologically superior. Kazakhstan demonstrates how integrated financial services can make everyday digital transactions convenient, while the United States illustrates how specialized

providers can offer consumers different tools for managing specific financial needs. Both models create opportunities for AI-based personalization. Yet neither guarantees improved budgeting, saving, or financial decisions. The effectiveness of personal finance technology ultimately depends on whether consumers understand the information provided and use it to make appropriate choices.

**8. When Financial Convenience Becomes a Risk**

The first major risk associated with AI-powered personal finance is inaccurate or unsuitable advice. A budgeting system may incorrectly categorize a transaction, while a predictive model may fail to recognize an unusual expense. A conversational AI assistant may misunderstand a question or generate an explanation that appears reliable but contains factual errors. These problems become more serious when users treat automated outputs as instructions rather than suggestions. Financial decisions can have lasting consequences, making verification especially important.

Privacy and data protection create another challenge. Personal finance applications may process sensitive information about income, spending, debts, investments, and other aspects of an individual's financial circumstances. Integrating several services can make financial management more convenient, but it can also concentrate information within a platform or increase the number of organizations with access to personal data. Consumers may not fully understand how their information is collected, analyzed, shared, or used to personalize financial products. These concerns apply to both integrated banking ecosystems and independent fintech services.

Algorithmic personalization can also create fairness and transparency problems. An AI system may generate different recommendations for different users, but consumers may not understand why those recommendations were produced. If a model relies on incomplete or unrepresentative information, its outputs may be less suitable for certain groups. This is especially important when financial technology influences decisions involving credit, investment risk, or other consequential products. The problem is not simply whether an algorithm is technically sophisticated, but whether its recommendations are understandable, appropriate, and open to meaningful evaluation.

Financial literacy remains a central limitation. Lusardi and Mitchell (2014) establish the importance of financial knowledge for economic decision-making, while Fernandes et al. (2014) show that improving financial behavior requires more than the transfer of information. Digital tools may simplify complicated tasks, but they can also encourage passive dependence if users stop questioning recommendations. Effective personal finance technology should therefore help people understand their decisions rather than merely automate them. Consumer education, transparent explanations, and opportunities for human review are important safeguards.

**9. Discussion: Do Smarter Tools Create Better Financial Habits?**

The reviewed evidence suggests that digital financial technologies can improve access to information and reduce the effort required for everyday money management. Transaction histories, automatic categorization, recurring transfers, and personalized notifications can make budgeting and saving easier to organize. Behavioral research also supports the idea that appropriately designed automatic mechanisms can encourage consistent saving. However, these benefits should not be confused with proof that every AI-powered application improves financial outcomes. The relationship between technology and behavior remains dependent on individual circumstances and product design.

One important distinction is between convenience and financial improvement. A mobile application can make spending easier without necessarily encouraging responsible spending. Similarly, an automated savings feature may be useful for someone with stable income but less suitable for a person whose income varies significantly. Personalized financial recommendations can be valuable when they reflect accurate information and realistic goals, but inappropriate recommendations can create additional risk. The effectiveness of financial technology must therefore be evaluated through consumer outcomes rather than convenience alone.

The Kazakhstan-U.S. comparison reinforces this point. An integrated financial ecosystem can reduce friction between financial activities, while a market of specialized applications can offer more tailored tools. Neither structure is automatically preferable. The important question is whether consumers can access useful information, understand the services they use, and make choices consistent with their financial circumstances. Differences in financial literacy, trust, income, regulation, and access to technology may matter as much as differences in application design.

AI adds another layer to this relationship. Machine learning can identify patterns and support prediction, while generative AI can make financial information easier to discuss through conversational interfaces. Yet greater personalization does not eliminate uncertainty or replace financial judgment. Consumers still need to understand basic financial principles, recognize limitations in automated recommendations, and consider their own goals. The most promising role for AI is therefore as a decision-support technology rather than an unquestioned decision-maker.

Future research would benefit from direct empirical comparisons of personal finance users in Kazakhstan and the United States. Such studies could examine whether integrated banking platforms or specialized applications are associated with stronger budgeting habits, more consistent saving, or greater financial confidence. They could also distinguish ordinary automation from genuinely AI-powered features and measure outcomes rather than relying only on reported satisfaction or adoption. Until such evidence becomes available, conclusions about the superiority of either national ecosystem should remain cautious.

## 10. Conclusion

AI-powered and automated personal finance tools are expanding the ways individuals can monitor spending, organize budgets, save money, and obtain financial information. Kazakhstan

and the United States illustrate different digital financial environments, with integrated banking platforms playing an important role in Kazakhstan and a broader range of specialized financial technology services available in the United States. Both models can support more convenient financial management and create opportunities for personalized assistance.

However, technological convenience does not automatically produce better financial habits or more informed decisions. The usefulness of personal finance tools depends on accurate information, appropriate recommendations, financial literacy, consumer trust, and effective safeguards. AI is most valuable when it helps consumers understand their finances and make decisions that reflect their own circumstances rather than encouraging unquestioning reliance on automation. The development of responsible personal finance technology should therefore focus on improving consumer outcomes, not simply increasing the number or complexity of digital features.

## Works Cited


D’Acunto, F., Prabhala, N., & Rossi, A. G. (2019). The promises and pitfalls of robo-advising. The Review of Financial Studies, 32(5), 1983–2020. https\://doi.org/10.1093/rfs/hhz014

Demirgüç-Kunt, A., Klapper, L., Singer, D., & Ansar, S. (2022). The Global Findex Database 2021: Financial inclusion, digital payments, and resilience in the age of COVID-19. World Bank. https\://www\.worldbank.org/en/publication/globalfindex

Fernandes, D., Lynch, J. G., Jr., & Netemeyer, R. G. (2014). The effect of financial literacy and financial education on downstream financial behaviors. Management Science, 60(8), 1861–1883. https\://doi.org/10.1287/mnsc.2013.1849

Lusardi, A., & Mitchell, O. S. (2014). The economic importance of financial literacy: Theory and evidence. Journal of Economic Literature, 52(1), 5–44. https\://doi.org/10.1257/jel.52.1.5

Olajide, O., Chimbane, T., & Lu, Y. (2026). Artificial intelligence used for financial advice and financial behaviors: The role of financial knowledge. Financial Planning Review. https\://doi.org/10.1002/cfp2.70045

Thaler, R. H., & Benartzi, S. (2004). Save More Tomorrow™: Using behavioral economics to increase employee saving. Journal of Political Economy, 112(S1), S164–S187. https\://doi.org/10.1086/380085

Board of Governors of the Federal Reserve System. (2025). Report on the economic well-being of U.S. households in 2024. https\://www\.federalreserve.gov/consumerscommunities/shed.htm

Kaspi.kz. (n.d.). Corporate information and investor relations. https\://ir.kaspi.kz/

National Bank of Kazakhstan. (n.d.). Official statistics and information on payment systems. https\://nationalbank.kz/en

Consumer Financial Protection Bureau. (n.d.). Consumer financial protection and financial well-being resources. https\://www\.consumerfinance.gov/

**Trust in AI and Human Financial Advisors Richard Zhang**


## Abstract

As Artificial Intelligence (AI) is increasingly integrated into financial decision-making, understanding when and why people trust AI advice over humans' becomes critical. This paper examines how the perceived urgency and importance of financial decisions affect preference for AI or human financial advisors, while also reviewing existing research on the neuroscience behind human trust, the distinct traits of AI and human advisors, and how honest communication of confidence or uncertainty shapes trust. A survey was conducted where 10 participants assessed 10 financial scenarios, and selected between two AI-generated recommendations that were labeled as AI or human generated. Then, participants rated each scenario's perceived urgency and importance. Participants selected AI over human-labeled advice 54% of the time. The mean perceived importance was similar for AI and human-labeled selections, while mean urgency was higher for AI rather than for human-labeled selections. This discrepancy suggests that the urgency, rather than importance of a financial scenario, may be more closely related to advisor preference. These findings are relevant for individuals seeking financial advice, because understanding the driving factors of trust in AI can help with making informed choices about relying on AI advice. However, these results should be interpreted with caution due to the limited sample size and findings should be regarded as preliminary.


## Introduction

Trust plays a major role in decision-making when people face uncertainty and must rely on someone else despite the risk that comes with doing so (Pauls et al., 2016). Finance is an especially important area for studying trust because investment decisions involve uncertainty, risk, and large amounts of money. At the same time, artificial intelligence has become more common in finance through tools such as robo-advisors, automated investment recommendations, and other AI financial services (Zhang et al., 2021). As these tools spread, people must decide whether to rely on automated systems, traditional human experts, or some combination of the two. Current research has examined trust in human financial advisors and in AI systems such as robo-advisors separately. However, one important question remains less clear: whether people's trust changes depending on how quickly an investment decision must be made and how much money is at stake. This study therefore asks, how do the urgency and scale of investments influence human trust in AI advisors versus human financial advisors?

## What neural networks are involved in processing risk and reward when making decisions?

Trust is a complicated process that uses many regions of the brain to help people evaluate risk, predict outcomes, and understand the intentions of others (Pauls et al., 2016). Trust does not rely on a single brain region, and is instead formed by neural systems, or networks of nerves, that also shape social reasoning, risk-reward processing, and learning from previous experiences.

During situations of uncertainty, these neural systems help people determine the potential rewards and risks of trusting other people, and figure out what their true intentions are (Pauls et al., 2016).

Research in neuroscience has identified multiple brain regions involved in these processes. In an fMRI study of social interactions, Krueger et al. (2007) found that building a trusting relationship activated the paracingulate cortex, which drives humans' social decision making, while conditional trust activated the ventral tegmental area, which is responsible for the body's dopamine production. These findings suggest that trust depends not only on evaluating another person's intentions but also on weighing the good and bad outcomes of trusting them. These neural systems allow people to evaluate trust as they make decisions in uncertain situations.

**How is trust gained or lost in terms of neuroscience?**

It is widely known that trust is not fixed, and can change when people gain new knowledge or experiences with others (Krueger et al., 2007). Positive interactions strengthen trust because they reinforce expectations that another person is reliable, while negative interactions cause individuals to reevaluate those expectations, which in turn decreases trust levels between people (Pauls et al., 2016). Neuroscience shows that this process happens through learning mechanisms that compare expected outcomes with actual outcomes, allowing people to make future decisions based on their previous experiences (Krueger et al., 2007). As a result, trust tends to develop gradually rather than appearing immediately.

The amount of risk in a decision also affects how easily trust is established (Pauls et al., 2016). Decisions involving more uncertainty or larger potential losses usually require higher trust levels before individuals are willing to rely on others (Pauls et al., 2016). Understanding how the brain builds and updates trust is foundational when analyzing and comparing trust levels between people but also with AI.

**What does AI have to build trust that humans don't, and vice versa?**

Artificial intelligence has created a modern problem: people now must decide whether to trust recommendations coming from a computer more than another person. Although AI systems do not have emotions or personal experiences, they still have many characteristics that increase users' trust. AI can process large amounts of information quickly, give consistent recommendations, and avoid emotional biases that would influence a human's judgment (Mehrotra et al., 2024). These characteristics make AI appear more objective, especially when decisions depend on analyzing complex data. However, AI programs lack human social interactions, meaning that in some situations, they may not effectively account for human emotions (Zhang et al., 2021). This makes AI advisors much faster and unbiased, but sometimes emotionally incompetent.

Human advisors, on the other hand, offer qualities that AI cannot replicate. Human advisors can respond to their client's feelings and personal circumstances when giving advice (Pauls et al., 2016; G P & Dsa, 2024).

Because of this, trust in AI might be decided by whether users prefer consistency or interpersonal understanding, which shows that both human advisors and AI systems each have strengths that may be trusted differently depending on the situation.

**How effective is anthropomorphism?**

Researchers have also examined whether making AI seem more human can increase trust. Anthropomorphism refers to giving AI systems human-like traits, such as conversational language, names, or human-like voices and personalities. Large language models, or LLMs, naturally communicate using human language, making interactions with them feel similar to conversations with another person. Research has found that providing explanations, and expressing confidence or uncertainty in a decision can increase users' trust by making AI recommendations easier to understand and seem more human (Mehrotra et al., 2024). However, these same features sometimes cause people to put too much trust in AI, even when its recommendations should be questioned. So, researchers suggest that AI systems should promote appropriate trust rather than maximizing trust (Mehrotra et al., 2024).

**How do user age, AI literacy, and risk tolerance affect trust levels?**

Trust in AI also depends on characteristics of the user. Younger individuals are more familiar with AI technology because they have grown up using digital tools, while older adults may be less comfortable using AI for important decisions (Sidoti & Gottfried, 2025). Additionally, people with better AI literacy often have a better understanding of AI's capabilities and its limitations, allowing them to make more informed decisions about when to rely on AI recommendations (Mehrotra et al., 2024).

Individual attitudes toward risk also influence trust in AI (Aini). People who are more comfortable making decisions under uncertainty evaluate AI recommendations differently than individuals who tend to avoid risk (Aini, 2025). This evidence suggests that trust in AI depends not only on the characteristics and ability of the technology but also on the personality and traits of the person using it. These fast-changing factors become especially important in financial decision-making, where the consequences of trusting AI can lead to gaining or losing large amounts of money.

**How does trust between humans affect financial decisions?**

Trust plays an important role in financial decision-making because investors depend on others to explain information and help them make choices involving risk. Some examples of financial decisions are choosing investments, planning for retirement, or deciding how big of a

risk to take. Investors must decide whether they trust an advisor to be knowledgeable and reliable when providing guidance (G P & Dsa, 2024).

Trust in financial advice does not only depend on the skills of the advisor. A person can trust an advisor personally even if they don't trust the company they are working for or the financial advice industry in general (Pauls et al. 2016). Because of this, an investor's willingness to follow advice can depend on trust in both the individual and the organization they work for (Pauls et al., 2016).

Trust also develops through experience. Positive experiences can strengthen the belief that an advisor is reliable, while negative experiences cause an investor to doubt future recommendations (Fareri, 2019). As discussed previously, trust is not permanent and can change as people compare what they expected to happen with what actually happened. Investors continuously evaluate their advisor's judgment and past performance to decide how much to trust them in the future (Pauls et al., 2016).

**What are investors trusting advisors to do?**

Investors trust financial advisors to do more than choose investments that might increase in value. Research on the role of financial advisors shows that these professionals can shape investment decisions by helping investors organize information and understand the possible consequences of different choices (G P & Dsa, 2024). Investors are also trusting advisors to understand them personally. Two people may have similar incomes but completely different financial responsibilities, goals, and levels of comfort with risk. A human advisor can ask questions about these differences and change a recommendation based on information that is not obvious from just financial data.

Another important responsibility is explaining uncertainty. Advisors cannot guarantee success for an investment, so they have to help their client understand the possible benefits and risks of their recommendations. So, an advisor can gain trust not by claiming to be completely certain, but by clearly explaining what is known, what remains uncertain, and why they made a recommendation.

**What is AI's role in finance today?**

Investors rely on AI advisors because they can process large amounts of information and produce recommendations both quickly and consistently. Unlike humans, AI systems do not become tired or let emotions influence the judgment and calculations they make. These characteristics make AI advisors seem like the better option when a decision depends more heavily on comparing large amounts of data and possible outcomes (Mehrotra et al., 2024).

Robo-advisors, AI chatbots that give generalized advice, can make financial guidance more accessible. Instead of meeting directly with a human advisor, users can tell the program about their goals and risk preferences through an online system. The system can then use that

information to give recommendations. The information may include the investor's goals, investment timeline, financial circumstances, and tolerance for risk.

AI can also support human financial advisors instead of replacing them. A human advisor can use AI to organize information, or identify patterns in data before discussing the results with a client. In this situation, AI is involved in part of the analysis, while the human advisor interprets the information and communicates with their client. This creates a hybrid approach that takes advantage of both the analytical power of AI and the interpersonal skills of humans.

**How do people choose between human and AI advice?**

Perception of certain institutions can also influence whether someone prefers human or AI advice. Bosompim et al. (2026) suggests that trust in an advisor may be influenced not only by the quality of their recommendations, but also by the people or institutions surrounding the person giving that advice. For example, if a client has had a negative experience with human or AI advisors, that may cause them to shy away from one of the options. Similarly, if a client had a positive experience or has seen promotional advertisements, they will be drawn towards one of the options.

These many factors, whether it be the urgency of a situation or pre-existing beliefs in an institution, make it difficult to always pick the right advisor for a given situation. On top of that, with humans and AI having their clear advantages and disadvantages, human and AI advisors may be more effective when used together. However, the balance between the two depends on the scale of the decision being made and what the most important traits are to the investor.

**Present Study**

The context of a financial decision can also affect whether someone prefers advice from a human or AI advisor. However, it is unclear how different factors of a financial situation, like urgency or importance, affect preference for human or AI advisors. To find out, I conducted a survey with ten separate financial situations. For each situation, participants chose between two pieces of AI advice. However, for each situation, one piece of advice was labeled as AI advice and the other as human advice. Participants were then asked to rate each situation based on both how urgent and how important the decision felt, on a scale of one to five. I hypothesized that in general, participants would choose AI advice in situations they considered more urgent, while they would choose human advice in situations they considered more important.

**Methods**

***Participants***

Ten participants were recruited through family and friends for this survey, and they all responded to the same 10 financial scenarios.

***Procedures***

For each scenario, participants were shown two pieces of advice, with one labeled as advice from an AI financial Advisor and the other labeled as advice from a human advisor. Participants selected which recommendation they preferred and then rated how urgent and important the financial scenarios felt on a scale from one to five, giving a separate rating for each one.

***Analysis***

After all the data was collected, the average urgency and importance ratings for AI and human selections were calculated, and these results were assessed to determine whether perceived urgency or importance was related to preference for human or AI financial advice.

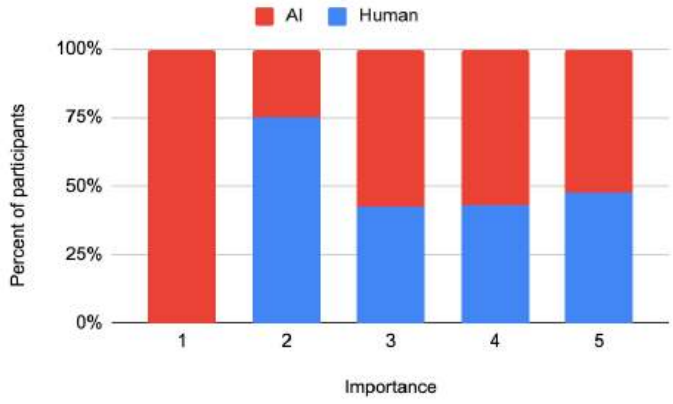
## Results

***How does the importance of the decision affect trust levels in advice from humans and AI?***

In this survey, participants chose AI over human advice 54% of the time. Out of these responses, the average importance value when choosing AI advice is 3.65, versus 3.59 when choosing human advice. This suggests that the perceived importance of a decision does not greatly affect preference for human or AI advice. This is highlighted when looking at responses to individual questions. (**see Figure 1**).

**Figure 1.**

*Human vs. AI advisor selections based on importance.*




*Note.* Total N = 10; 1 is the lowest perceived importance, and 5 is the highest perceived importance.

For example, in question three, where a recently graduated college student chooses between a paid job over the summer and an auspicious unpaid-internship (3.9/5 perceived Importance), seven out of the 10 participants selected the AI response. On the other hand,

participants averaged an importance value of 3.6 out of five in question eight, where the participant chooses between a temporary transportation solution by paying a monthly train fee, or buying an affordable used car. All 10 participants chose the human advice. Additionally, out of all scenarios where the perceived importance was rated as 1 out of 5, which is the lowest perceived importance, 100% of participants chose AI advice. This shows that in situations with especially low importance levels, there may be a stronger preference for AI advisors. These results suggest that although there is no significant trend between perceived importance and advisor preference, and situations with similar levels of importance can create preferences for both human and AI advice, importance may still become a factor at extremely low levels of importance.

***How does the urgency of the decision affect trust levels in advice from humans and AI?***

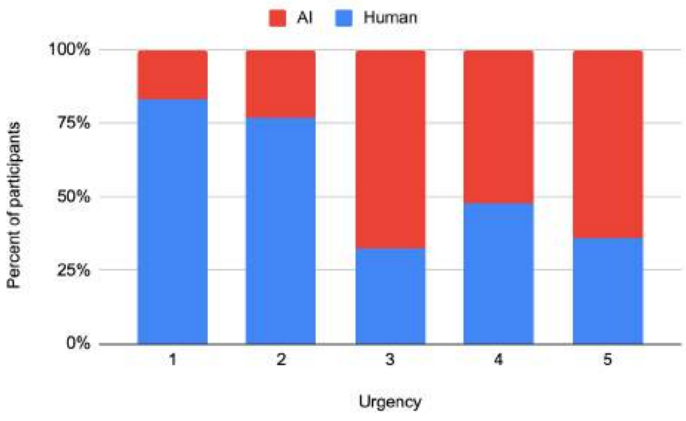
The survey results show a clearer relationship between urgency and preference for AI advice, with AI advice appearing more frequently in situations that respondents viewed as more urgent.

For example, in question 10, where a college student chooses between repairing their laptop or buying a new one, the average urgency score was 3.5. In this scenario, eight out of 10 participants selected AI advice.

Additionally, similar to the average importance value, the average urgency value for all AI advice selections was 3.46. However, the average urgency value dropped to 3.04 when calculated with the human advice selections. This data suggests that participants choose AI advice over human advice in situations they deem more urgent.

**Figure 2.**

*Human vs. AI advisor selections based on urgency.*




*Note.* Total N = 10. 1 is the lowest perceived urgency, and 5 is the highest perceived urgency.

## Discussion

This study looked at how the urgency and importance of a financial decision affect human trust in AI and human financial advisors. To study this, a survey was conducted where 10 participants responded to 10 financial situations, chose between advice labeled as AI or human generated, and rated each situation's urgency and importance on a scale of one to five. Overall, participants chose AI advice 54% of the time. The average importance values were similar between AI and human selections, while the average urgency value was higher when participants chose AI advice. These results suggest that urgency may play a larger role than importance in determining preference for human or AI financial advice.

One reason could be that one of AI's main advantages is its ability to provide information quickly. When someone needs to make a decision in a shorter amount of time, receiving an immediate recommendation feels more assuring.

However, not all situations that leaned towards AI advice were considered urgent by respondents, as some situations that were considered relatively urgent still resulted in a preference for human advice.

While this survey does suggest the impact of urgency on the preference of human or AI advisors, it also shows that there are factors other than just perceived importance or urgency that alter trust levels in financial advisors. These findings relate to previous research, and show that AI is trusted to quickly process large amounts of information, while human advisors are valued more for interpersonal understanding. The preference for AI in urgent situations reflects this idea, since faster, analytical recommendations become more reliable when there is less time to make a decision. On the other hand, the lack of a clear relationship between perceived importance and advisor preference suggests that when choosing between human and AI advice, people prioritize the type of support they want over how important the decision feels.

## Future Directions and Limitations

Another area to be explored by future research is studying how the way AI communicates its recommendations affects trust in financial advice. Research has found that explanations that show confidence or uncertainty can increase trust by making AI advice easier for users to understand (Mehrotra et al., 2024). However, these same characteristics can sometimes cause users to place too much trust in AI, even when its recommendations should still be questioned. A future experiment could build on this idea by changing how confidently an AI advisor presents the same recommendation across financial situations with different levels of urgency and importance. This could help researchers find out whether explaining uncertainty makes financial advice more or less convincing.

There were also several limitations to the survey conducted in this study. First, the survey only recorded results from ten participants who each responded to ten financial situations, making the amount of analyzable data rather small. Because of this small sample size, the findings should be treated as preliminary trends rather than conclusive findings. In a study with

more resources, a power analysis could be conducted to determine an optimal sample size before recruiting respondents.

Another limitation was the diversity of the participant pool. Participants were primarily recruited through family and friends, meaning the survey did not have a broad range of backgrounds and experiences. A future experiment could distribute the survey to a wider group of people to see whether the results remain consistent across different populations.

Finally, the survey was limited to ten questions to keep the questionnaire manageable for respondents, since they were not being compensated for their time. However, with additional funding, future researchers could compensate participants and include more scenarios, which would allow for a wider range of financial scenarios included in the survey.

Together, these changes would create a more complete follow-up experiment that looks into not only whether people prefer human or AI financial advice, but also how confidence or uncertainty demonstrated by advisors can influence that preference. Existing research already shows that trust in AI depends partly on how clearly and honestly explanations are communicated, making this a key area for further research as AI becomes more involved in financial decision-making (Mehrotra et al., 2024).

**Conclusion**

This paper looks at how the urgency and importance of a financial decision affects human trust in AI and human financial advisors. Our discussion was centered around human trust, which is built through neural processes that calculate risk, reward, and past experiences. Human trust is greatly shaped by personal factors such as AI literacy and risk tolerance. For this reason, financial trust doesn't only depend on the advisor's expertise, but also on their ability to clearly and accurately communicate risk and uncertainty. AI offers quick decisions and the power to process large amounts of data, while humans offer interpersonal understanding. Offering a hybrid approach to get the benefits of both types of advisors could be useful. In the survey conducted, it was observed that importance didn't play as big of a factor as urgency when determining trust levels, with urgency making people lean towards AI advisors more. However, this survey was limited by its small sample size and lack of diversity in its participants, which future studies can work on. Looking ahead, emerging research can also look into how different levels of demonstrated confidence or uncertainty can affect trust. These findings allow financial advisors and AI developers better understand when people are most likely to rely on AI, and would help them create tools that support more appropriate trust.

**Works Cited**


Aini, Sukma. "Trusting AI in Finance: The Role of Psychological Traits and AI Literacy in Shaping User Behaviour Toward Robo-Advisors." The South East Asian Journal of Management, vol. 19, no. 2, Oct. 2025. DOI.org (Crossref), https://doi.org/10.7454/seam.v19i2.1929.

Bosompim, Prince, et al. "In Whom We Trust? Social Networks and Trust in AI vs. Human Experts Financial Advice." Finance Research Letters, vol. 91, 2026, article 109437. https://doi.org/10.1016/j.frl.2025.109437.

Chatterji, Aaron, et al. How People Use ChatGPT. National Bureau of Economic Research, 2025. NBER Working Paper, no. 34255. https://doi.org/10.3386/w34255.

Fareri, Dominic S. "Neurobehavioral Mechanisms Supporting Trust and Reciprocity." Frontiers in Human Neuroscience, vol. 13, 2019, article 271. https://doi.org/10.3389/fnhum.2019.00271.

G P, Pallavi, and Kusum Thantry Dsa. "Literature Review on the Role of Financial Advisors in Shaping Investment Decisions." Journal of Emerging Technologies and Innovative Research, vol. 11, no. 1, 2024. https://doi.org/10.5281/zenodo.10475591.

Kavitha, M., et al. "The Role of Artificial Intelligence in Financial Decision-Making." Journal of Marketing & Social Research, vol. 2, no. 2, 2025, pp. 189–198.

Krueger, Frank, et al. "Neural Correlates of Trust." Proceedings of the National Academy of Sciences of the United States of America, vol. 104, no. 50, 2007, pp. 20084–20089. https://doi.org/10.1073/pnas.0710103104.

Mehrotra, Siddharth, et al. "A Systematic Review on Fostering Appropriate Trust in Human-AI Interaction: Trends, Opportunities and Challenges." ACM Journal on Responsible Computing, vol. 1, no. 4, 2024, article 26, pp. 1–45. https://doi.org/10.1145/3696449.

Pauls, Thomas, et al. Broad-Scope Trust and Financial Advice. Social Science Research Network, 2016. https://doi.org/10.2139/ssrn.2670484.

Sidoti, Olivia, et al. "About a Quarter of U.S. Teens Have Used ChatGPT for Schoolwork—Double the Share in 2023." Pew Research Center, 15 Jan. 2025, www.pewresearch.org/short-reads/2025/01/15/about-a-quarter-of-us-teens-have-used-chatgpt-for-schoolwork-double-the-share-in-2023/.

Zhang, Lixuan, et al. "Who Do You Choose? Comparing Perceptions of Human vs. Robo-Advisor in the Context of Financial Services." Journal of Services Marketing, vol. 35, no. 5, 2021, pp. 634–646. https://doi.org/10.1108/JSM-05-2020-0162.

**The Role of Ephemeral Content in Snapchat's Growth and Success By Brandon Chen**

**Abstract**

Snapchat has numerous factors that have contributed to its growth and success as a digital media company. This research specifically explores the role of ephemeral content in Snapchat's growth and success. The paper focuses on how Snapchat's ephemeral content features played a significant role in the company's growth and success by driving innovation and differentiation, increasing user engagement, and accelerating revenue growth. Online research was conducted mainly through Snapchat's annual reports and resources, but also through outside sources explaining Snapchat's history, features, user engagement, and business performance. The findings suggest that ephemeral content helped Snapchat differentiate itself from competing social media platforms while encouraging frequent communication and repeat visits, ultimately driving growth in user engagement and revenue. According to Snapchat's annual reports from 2017 to 2025, daily active users tripled, and revenue increased from approximately $404 million in 2016 to $5.93 billion in 2025. Ultimately, this research demonstrates that ephemeral content played a significant role in Snapchat's growth and success through innovation and uniqueness, user engagement growth, and revenue generation. Snapchat's long-term success will rely on its ability to continue innovating, competing with other social media platforms, and improving its financial stability.

**Keywords** Snapchat, ephemeral content, user engagement, social media, disappearing messages, disappearing content, digital media, Streaks, Snapchat Stories

**The Role of Ephemeral Content in Snapchat's Growth and Success**

Snapchat is one of the world's leading social media platforms and visual messaging applications (Snap Inc., 2026). Brian O'Connell (2020) explores the early history and development of Snapchat. Founded in 2011 by Stanford University students Evan Spiegel, Reggie Brown, and Bobby Murphy, Brown and Spiegel imagined an app focused on disappearing content, so in July 2011, the students released an app called "Picaboo," the predecessor of Snapchat, but Brown was ultimately removed from the company. Spiegel and Murphy relaunched Picaboo as Snapchat two months after its release due to a business affair. The chief executive officer, Spiegel, and Murphy continued developing Snapchat (O'Connell, 2020). Snapchat's early growth was also supported by venture capital funding, an investment used to help startups with high scaling potential. Like many other social media companies before going public, Snapchat "raised $2.26 billion in funding from investors like Yahoo, Lightspeed Venture Partners, and Benchmark" (Tracxn, 2026).

Snapchat integrates several examples of ephemeral, or disappearing, content features, including Stories, chats, and visual messaging. Snap Inc. (2026) describes Stories as a way for users to engage with their friends and community by sending photos and videos that are shown in chronological order. According to Snap Inc. (n.d.-b), Stories normally last 24 hours by default

before disappearing, unless the user changes to a different setting. Snap Inc. (n.d.-b) also notes that the app includes duo and group conversations and that chats are deleted 24 hours after being viewed or 31 days after they are sent, unless settings are changed. Snapchat's system also deletes all duo and group conversation Snaps after every member of the chat views them (Snap Inc., n.d.-b). In this context, "Snaps" refers to visual messaging through Snapchat.

Before exploring how ephemeral content contributed to Snapchat's growth, it is necessary to define business success. Success cannot be measured by one factor, especially in the social media industry. Instead, a company's success can be determined by innovation and uniqueness, user growth and engagement, and revenue growth. Although there are countless ways to define success, this paper will focus on the three factors listed here, as they provide a framework for exploring Snapchat's ephemeral content and its role in the platform's growth and success. Snapchat's ephemeral content has contributed to its success by differentiating it from other social media platforms, shaping a large user base, and ultimately helping Snapchat build a large, revenue-generating company.

**From Innovation to User Engagement**

Snapchat's uniqueness and innovation drove its growth and success by releasing a social media platform that differentiated itself from other existing platforms. When Snapchat launched in 2011, platforms such as Instagram were centered around preserving content rather than ephemeral content. Alison Eldridge (2026) describes one of two ways a user can post on Instagram. The first allows users to post content to a permanent feed. When Snapchat first launched, Snapchat differentiated itself by building a platform around ephemeral content, rather than taking inspiration from existing platforms that focused on preserving content. Snapchat's ephemeral content became so popular that Instagram, another popular social media platform, drew "inspiration" from its features. Instagram launched Stories, "a special selection where content remains for 24 hours before disappearing (unless specifically archived)" (Eldridge, 2026). Harry Boulton states that "Instagram 'stole' Snapchat's most iconic feature three years after they introduced it to the world," highlighting Snapchat's uniqueness and innovation at launch (Boulton, 2025). Eventually, Snapchat's features became so iconic that other platforms began taking inspiration from them.

What started as a technical innovation evolved into an alternative to some of social media's more problematic side effects. Because Instagram is so public and visible to other users, many individuals try to be picture-perfect, often to the point of self-consciousness about their appearances or day-to-day activities, and trying to fit in can create a lot of stress. Common Sense Media (2024, p. 7) found that "51% of [American] teens feel pressure to look their best or present themselves in a certain way (e.g., follow trends, have a certain body type)." Snapchat, however, took a different approach, emphasizing ephemeral communication over permanent content. Malik (2024) says that "users tend to post more casual, unfiltered content, knowing it won't be permanently archived." This means users are influenced to connect and engage more casually and authentically, rather than trying to create an idealized profile of themselves.

Because users can communicate more casually and authentically on Snapchat, it can help relieve the stress of needing to post picture-perfect content on social media. Although content on Snapchat will always leave a digital footprint, its ephemerality provides users with a sense of casual security, knowing that what they post will not be widely searchable. Since Snapchat differentiated itself from other social media platforms at its launch, its ephemeral content features have helped it build a massive user base.

Along with the uniqueness that has helped Snapchat gain a massive following, its ephemeral content has also helped the platform build a large user base by giving users reasons to communicate frequently and constantly return to the app. Snapchat combines ephemeral content with other features, such as Streaks and its home screen, to encourage frequent user engagement on the platform. Streaks form when friends Snap each other at least once a day, and a fire emoji will appear on the chat screen after friends have formed a Streak (Snap Inc., n.d.-a). Since friends need to Snap each other every day to maintain a Streak, users are motivated to return to the app daily, which contributes to Snapchat's high daily active user count (DAU).

Anizoba (2025) explains that Streaks work due to psychological principles such as reward systems and dopamine hits combined with Snapchat's sense of urgency and quickness. She says that when users continue a Streak, they get a dopamine hit that encourages repeated behavior and keeps them coming back. The author observed that quick Snap Streaks are also easier to maintain than other routines because they take little to no time to send, and Streak benchmarks such as 100, 200, 300, etc., can feel rewarding for the user. Streaks consistently boost user engagement, and they also pair exceptionally well with Snapchat's ephemeral content. The author also observed that Streaks disappear after one day, and Snaps eventually disappear, so users feel more urgency to respond and continue their Streaks. Snapchat also uses its home screen feature to keep users engaged. The author discovered that whenever the Snapchat application is opened, the camera automatically appears as the home screen. Morris (2025) states that the "front-facing camera allows users to quickly capture their reaction, current expression, or immediate surroundings from their own perspective." Snapchat already encourages sharing casual and genuine moments, and this feature enhances that experience by making it easier and quicker for users to capture their current selves and surroundings.

The impact of Snapchat's user growth strategies over time can be reflected in its daily active user (DAU) statistics in its annual reports. From the fourth quarter of 2016 to the fourth quarter of 2025, Snapchat's DAUs have tripled, from 158 to 474 million, reflecting consistent growth driven by user engagement strategies (Snap Inc., 2018, 2026). This shows that Snapchat's strategies have been effective in gathering a large user base. Except for 2017 to 2018, with a 1 million decrease in DAU year over year, Snapchat's DAU count has consistently increased year-over-year all the way until the fourth quarter of 2025, which was the latest annual report (Snap Inc., 2018, 2019, 2020, 2021, 2022, 2023, 2024, 2025, 2026). Furthermore, according to the latest annual report, Snapchat reached its peak with 474 million daily active users in the fourth quarter of 2025, the highest it has ever reported, showing growth in user engagement (Snap Inc., 2026).

**Monetizing Snapchat's User Base**

Snap Inc. (2026) classifies its revenue into two categories, advertising and other, and states that advertising accounts for the majority of its revenue. Snap Inc. (2026) notes that its advertising payouts are defined by legal agreements, usually based on ad engagement or a fixed fee over a set period of time. They also reported that the advertisements are shown on Snapchat, mainly through Snap Ads and Augmented Reality (AR) Ads. Snap Inc. (2026) reports that other revenue comes from subscriptions such as Snapchat+, Lens+, and Snapchat Platinum, which provide access to features that are unavailable to regular Snapchat users. Additionally, Snap Inc. (2026) reported that other revenue consists of storage plans, partnerships, and physical products. Snapchat earned $5.93 billion in revenue in 2025, but approximately $5.2 billion of that came from advertising, meaning the vast majority of its revenue came from advertising deals (Snap Inc., 2026). In the fourth quarter of 2025, Snapchat had 474 million daily active users, a 5% year-over-year increase, and $3.62 average revenue per user, a 5.23% year-over-year increase (Snap Inc., 2026).

Snapchat's large user base plays a major role in maximizing advertising revenue. With 474 million daily users, Snapchat gives advertisers consistent access to a large, engaged user base and audience (Snap Inc., 2026). Snapchat mainly generates revenue through advertising, so expanding that user base is key to ultimately increasing revenue generation. Due to its large user base, it allows businesses and advertisers to promote their products and services to users active on the platform (Snap Inc., 2026). Snapchat advertising deals can vary based on advertising engagement, so more users ultimately bring in more revenue. Besides advertising revenue, Snapchat also generates revenue from subscriptions, partnerships, storage plans, and physical products (Snap Inc., 2026). As Snapchat continued to expand its user base, its revenue increased from $404 million in 2016 to $5.93 billion in 2025 (Snap Inc., 2018, 2026), showing how user engagement can drive revenue generation. This growth illustrates Snapchat's ability to build and sustain a large audience, which has helped the company commercialize its platform and drive revenue growth.

**Snapchat's Competition and Future Growth**

Snapchat originally stood out for its innovative ephemeral content features, but its competitive advantage has been difficult to maintain as other platforms have adopted similar features. As other social media platforms began to take off, they started to take notice of Snapchat's iconic features, such as Stories. This produced plausible challenges for Snapchat because other platforms with larger user bases and greater financial resources could offer the same features. Snap Inc. (2026) also acknowledges that it competes with platforms like YouTube, TikTok, and Meta, but states that it "invests heavily in constantly improving and expanding" its products. As a result, they have continued developing features such as AR Lenses, Snap Map, and other products to differentiate themselves from competitors (Snap Inc., 2026). The company's ability to grow revenue and expand its user base will likely depend on whether it can continue innovating while improving its ability to attract advertisers and generate

revenue. Additionally, Snapchat has faced net losses and has yet to become profitable despite its revenue earnings, so maintaining its business while striving to become profitable will be Snapchat's challenge on its path to becoming a fully successful company (Snap Inc., 2026). Hern (2024) suggests that it is too early to predict if Snapchat's attempts to compete with other platforms will ultimately succeed, making its long-term future unpredictable. However, if Snapchat continues developing unique products, maintaining its large user base, and strengthening its promotional business, the company has potential for further long-term growth and success.

**Conclusion**

Overall, this research demonstrates that Snapchat's ephemeral content features have played a significant role in the platform's growth and success. Ephemeral content has driven Snapchat's competitive advantage over other social media platforms, as its uniqueness and innovation have attracted many users. While many other social media platforms focused on preserving content, Snapchat emphasized ephemeral content, which encouraged more casual communication (Malik, 2024). With the home screen camera, Streaks, and ephemeral content, Snapchat also kept users engaged while growing its user base (Anizoba, 2025; Morris, 2025). Ultimately, Snapchat's uniqueness and user engagement formed a valuable audience that it could monetize through advertising. (Snap Inc., 2018, 2026). Therefore, although ephemeral content significantly contributed to Snapchat's success and growth, maintaining long-term success will depend on the company's ability to continue innovating, competing, and improving its financial stability.

**Limitations of Research**

However, this research contains several limitations. First, the majority of the research relied on information from Snap Inc., including annual reports and other company resources. Although these resources report accurate, official statistics on Snapchat's financial performance, user growth, and platform features, they may present the information primarily from Snapchat's perspective. Along with the positive effects of Snapchat's ephemerality, there are also possible negative effects such as a false sense of security, risky behavior, and possible mental health impacts that this research does not explore due to the focus of the paper. Finally, Snapchat's growth and success have not solely relied on its ephemeral content features and the tools supporting it, as other factors such as other features, strategies, and business decisions may have also supported its growth. Therefore, although this research demonstrates that ephemeral content has played a significant role in Snapchat's growth and success, it is difficult to determine how much of that success can be attributed to it.

**Future Research Opportunities**

Future research could further examine Snapchat's future regarding growth and success. Because Snapchat is not a perfect or fully successful company, future research can determine

whether it can become consistently profitable while competing with other social media companies. Additionally, future research could investigate how Snapchat's newer features, including augmented reality and subscription features, have influenced user engagement and revenue. As Snapchat continues to develop, future research will be crucial to determining whether the company can maintain its large active user base while achieving long-term financial success.

**Works Cited**


Anizoba, P. (2025, April 14). Designing for user retention: The psychology behind streaks. Bootcamp. Medium. https://medium.com/design-bootcamp/designing-for-user-retention-the-psychology-behind-streaks-cf0fd84b8ff9

Boulton, H. (2025, March 28). Snapchat CEO reveals how he reacted to finding out Instagram was going to copy the iconic "stories" feature. UNILAD Tech. https://www.uniladtech.com/social-media/snapchat/snapchat-ceo-reacted-finding-out-instagram-copied-stories-189799-20250326

Common Sense Media. (2024). Unpacking grind culture in American teens: Pressure, burnout, and the role of social media. https://www.commonsensemedia.org/sites/default/files/research/report/2024-unpacking-grind-culture-in-american-teens_final-release-updated-for-web.pdf

Eldridge, A. (2026, September 1). Instagram. Britannica Money. https://www.britannica.com/money/Instagram

Hern, A. (2024, May 7). TechScape: How Snapchat is saving itself – and keeping up with Silicon Valley giants. The Guardian. https://www.theguardian.com/technology/article/2024/may/07/techscape-newsletter-snapchat

Malik, I. (2024, November 12). Snapchat's ephemeral revolution: Transforming digital behavior and mental health. Our Mental Health. https://www.ourmental.health/screen-time-sanity/snapchats-ephemeral-revolution-transforming-digital-behavior-and-mental-health

Morris, O. (2025, December 1). Why does Snapchat open with a front camera? Aeanet.Org. https://www.aeanet.org/why-does-snapchat-open-with-front-camera/

O'Connell, B. (2020, February 28). History of Snapchat: Timeline and facts. TheStreet. https://www.thestreet.com/technology/history-of-snapchat

Snap Inc. (n.d.-a). How do streaks work and when do they expire? Snapchat Support. Retrieved September 7, 2026, from https://help.snapchat.com/hc/en-us/articles/7012394193684-How-do-Streaks-work-and-when-do-they-expire

Snap Inc. (n.d.-b). When does Snapchat delete snaps and chats? Snapchat Support. Retrieved September 6, 2026, from https://help.snapchat.com/hc/en-us/articles/7012334940948-When-does-Snapchat-delete-Snaps-and-Chats

Snap Inc. (2018). Annual report pursuant to section 13 or 15(d) of the Securities Exchange Act of 1934. https://d18rn0p25nwr6d.cloudfront.net/CIK-0001564408/fa22ca71-844f-4e4b-891f-9795e5e8a44e.pdf

Snap Inc. (2019). Annual report pursuant to section 13 or 15(d) of the Securities Exchange Act of 1934. https://d18rn0p25nwr6d.cloudfront.net/CIK-0001564408/2e4d8346-d986-4ca9-9225-29b911d5bc11.pdf

Snap Inc. (2020). Annual report pursuant to section 13 or 15(d) of the Securities Exchange Act of 1934. https://d18rn0p25nwr6d.cloudfront.net/CIK-0001564408/0cfebc98-816e-44ac-8351-5067b4f88f0c.pdf

Snap Inc. (2021). Annual report pursuant to section 13 or 15(d) of the Securities Exchange Act of 1934. https://d18rn0p25nwr6d.cloudfront.net/CIK-0001564408/799dba00-c613-4d87-b858-916aff0d832a.pdf

Snap Inc. (2022). Annual report pursuant to section 13 or 15(d) of the Securities Exchange Act of 1934. https://d18rn0p25nwr6d.cloudfront.net/CIK-0001564408/da8288aa-d492-4fd1-b6ce-62dce8206e9f.pdf

Snap Inc. (2023). Annual report pursuant to section 13 or 15(d) of the Securities Exchange Act of 1934. https://d18rn0p25nwr6d.cloudfront.net/CIK-0001564408/c22ae9bd-7418-456e-82d4-48129de1df54.pdf

Snap Inc. (2024). Annual report pursuant to section 13 or 15(d) of the Securities Exchange Act of 1934. https://d18rn0p25nwr6d.cloudfront.net/CIK-0001564408/0dadfe63-2f5e-41e4-8029-6374d3e6e620.pdf

Snap Inc. (2025). Annual report pursuant to section 13 or 15(d) of the Securities Exchange Act of 1934. https://d18rn0p25nwr6d.cloudfront.net/CIK-0001564408/5c4563f0-758b-428b-9274-1e58f51f6095.pdf

Snap Inc. (2026). 2025 Annual report. https://s25.q4cdn.com/442043304/files/doc_financials/2025/q4/2025-Annual-Report.pdf

Tracxn. (2026, September 1). Snapchat - company profile. https://tracxn.com/d/companies/snapchat/__CXkhrdaPPHmlMwJop_X05L5lJcCtoRFeHUGVdeNcdu0#about-the-company

# How Common Are Five Behavioural Biases in the Investment Decisions of Indian Respondents, and How Do They Vary by Age, Gender, and Self-Reported Confidence?

By Karthik Paruchuri

## Abstract

Traditional economic models treat people as rational, but behavioural finance shows they routinely are not. This study asked how common five behavioural biases (loss aversion, present bias, herd behaviour, time inconsistency, and recency bias) appear in the investment decisions of 234 Indian respondents, and how they differ by age, gender, and self-reported confidence. Respondents completed an online questionnaire of fifteen forced-choice scenarios, each pairing a biased option against a rational one, plus a confidence rating and a decision-driver question. Each answer was coded as biased (1) or rational (0), and results were analysed descriptively. Respondents chose the biased option on 24.7% of scenarios overall. Present bias (32.9%), time inconsistency (32.1%), and loss aversion (29.3%) formed the highest group, recency bias (19.1%) sat in the middle, and herd behaviour (10.0%) was clearly the lowest. Measured bias fell at every step of the confidence scale, from 38.1% at the lowest level to 19.6% at the highest, which runs opposite to a simple overconfidence account. Women reported markedly lower confidence than men (3.41 versus 3.92 out of 5) despite only a small difference in actual choices, pointing to a gap between how capable women feel and how capable they are. Respondents who said they rely on their own research were the least biased (21.2%), while those following social media were the most biased (37.1%). The five biases were only weakly related across individuals, suggesting they are separate tendencies rather than one general trait. The findings locate this sample's departures from rationality mainly in self-control, and suggest that self-reported confidence carries real diagnostic information about financial decision quality.

**Keywords** behavioural finance, behavioural biases, financial confidence, retail investors, India

## Introduction

Traditional economic theory assumes people are rational: that they use all available information to make the choice that benefits them most. This assumption sits behind the Efficient Market Hypothesis, the idea that asset prices already reflect all known information (Fama, 1970), and the expected-utility framework it rests on (von Neumann & Morgenstern, 1944). These models were tidy, but they fit real behaviour poorly. An early challenge argued that people are only boundedly rational, limited by time and information, so they settle for good-enough rather than perfect choices (Simon, 1955). The larger break came with prospect theory, which showed that people feel losses more sharply than equal gains (Kahneman & Tversky, 1979). Later work carried these ideas into everyday spending and saving (Thaler, 1985). What emerged was a set of predictable biases that shape whether people save, panic in a downturn, or misjudge risk.

A recurring theme in this literature is gender. Women tend to be more cautious with financial risk and less likely to hold risky assets (Croson & Gneezy, 2009), which carries a real cost, because

holding safer, lower-growth assets over a lifetime leaves less accumulated wealth. The difference seems to stem less from ability than from confidence. Cross-country evidence shows that self-rated confidence predicts risky-asset ownership even after actual financial knowledge is accounted for (Cupák et al., 2021), while men tend towards overconfidence, trading more and earning lower net returns (Barber & Odean, 2001). The barrier women face is often about how capable they feel, not how capable they are. Because this study measures both demonstrated bias and self-reported confidence, it can examine that gap directly.

India is a timely setting. The number of dematerialised (demat) accounts rose from roughly 36 million in 2019 to over 210 million by late 2025 (Securities and Exchange Board of India [SEBI], 2025), pulling in millions of first-time investors, and roughly 80% of retail equity investors now come through app-based brokers (India Brand Equity Foundation [IBEF], 2026). Yet financial literacy remains low; the National Centre for Financial Education (2019) found only 27% of Indian adults financially literate. Regulators estimate that most retail derivatives traders lose money: over 91% of individual traders in the equity derivatives segment lost money in the 2025 financial year, with net losses near 1.06 trillion rupees (SEBI, 2025), and these losses fall hardest on the young and lower-income. Access has, in effect, outpaced capability, and that gap is exactly where behavioural biases prove most costly. The gender divide is visible too: women held only about 19% of demat accounts in early 2024 (National Stock Exchange of India, 2025).

Most existing evidence, though, comes from experienced investors, from single biases studied on their own, or from Western populations (Barber & Odean, 2001; Croson & Gneezy, 2009). Far less is known about how a broad, mostly inexperienced public shows several biases at once, or how those biases differ across ordinary groups such as age and gender. Two questions in particular remain open. First, the five biases here are rarely measured side by side in one sample, so their relative prevalence is unclear. Second, it is not well established whether self-reported financial confidence reflects actual decision quality or misleads. This study addresses both. The central research question is: how common are five behavioural biases in the investment decisions of Indian respondents, and how do they vary by age, gender, and self-reported confidence?

## Materials and Methods

### *Participants and Recruitment*

Data came from 234 respondents who completed a self-administered online questionnaire built in Google Forms. The link was shared through the researcher's personal network, including family, friends, and their wider contacts, who were invited to pass it on. Participation was voluntary and anonymous. Because respondents were reached through existing contacts rather than drawn at random from a defined population, the sample is a non-probability convenience sample. There were 133 female and 101 male respondents. By age, the sample skewed older: 14 respondents were aged 16 to 18, 15 were 19 to 25, 59 were 26 to 35, 53 were 36 to 45, and 93 were 46 and above.

### *Instrument*

The questionnaire contained 19 items of three kinds. Two demographic items captured age group and gender. Fifteen forced-choice scenario items (Q1 to Q15) each described a realistic financial situation with two options, one bias-consistent and one rational, with three items measuring each of the five biases. Two self-report items captured a financial confidence rating on a 1-to-5 scale and a single decision-driver question with four options (own research, friends or family, past experience, and social media). The scenarios covered situations such as choosing between a guaranteed payout and a favourable gamble, taking a reward now versus a larger reward later, and reacting to what others were doing. The full wording of all items appears in the Appendix.

### *Coding and Analysis*

Every scenario answer was coded 1 if the respondent chose the biased option and 0 if they chose the rational option; in this instrument the biased option was always labelled A and the rational option B. For each respondent, an overall bias score was calculated as the mean of their 15 coded answers, giving the proportion of scenarios on which they made the biased choice. Bias rates were then averaged for each item, for each of the five biases (the mean of its three items), and overall. Groups were compared by averaging the person-level bias score within each gender, age group, confidence level, and decision-driver category. The analysis used descriptive statistics (means and proportions) rather than inferential tests, so all group differences are reported as patterns observed in this sample rather than as statistically established effects. There were no missing values across the 234 responses.

## Results

### *Overall Prevalence of the Five Biases*

Across all 15 scenarios, respondents chose the biased option 24.7% of the time, so they made the rational choice most of the time. The five biases fell into three clear levels (Table 1 and Figure 1). Present bias (32.9%), time inconsistency (32.1%), and loss aversion (29.3%) sat close together at the top; recency bias (19.1%) was intermediate; and herd behaviour (10.0%) was clearly the lowest. Because the analysis is descriptive, the three highest biases are treated as a group that stood out rather than as proven different from one another.

**Table 1.** Overall bias rate for each of the five constructs.

| Behavioural bias | Survey items | Mean bias rate |
|---|---|---|
| Present bias | Q4–Q6 | 32.9% |
| Time inconsistency | Q10–Q12 | 32.1% |
| Loss aversion | Q1–Q3 | 29.3% |
| Recency bias | Q13–Q15 | 19.1% |
| Herd behaviour | Q7–Q9 | 10.0% |
| Overall (15 items) | Q1–Q15 | 24.7% |

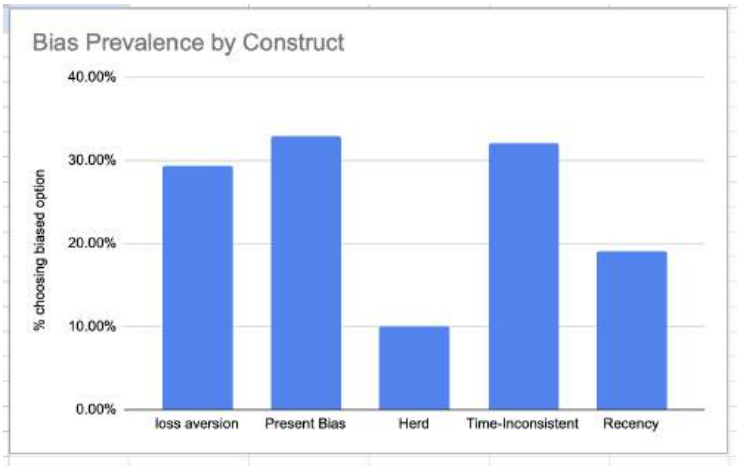



Figure 1. Bias prevalence by construct. Each bar shows the mean percentage of respondents (n = 234) who chose the biased option across the three items measuring that bias. Present bias (32.9%), time inconsistency (32.1%), and loss aversion (29.3%) cluster at the top, recency bias (19.1%) sits in the middle, and herd behaviour (10.0%) is clearly the lowest.

***Variation Within Constructs***

Averaging the three items in each bias hid large differences between individual scenarios. Within loss aversion, 62% chose the biased option on Q1, where the choice was between a guaranteed amount and a riskier gamble worth more on average, but only 14% and 12% did so on the other two items. Within present bias, one item (Q5) reached 53%, well above the construct average of 32.9%; this was the item that placed an immediate reward directly against a larger delayed one. Time inconsistency showed a similar spread, from 17% to 42% across its items. Herd behaviour stayed low across all three of its items (6% to 15%), while recency bias was comparatively even (16% to 22%).

***Bias by Age and Gender***

Average bias tended to fall as age rose (Figure 3): 31.4% for those aged 16 to 18, 27.1% for 19 to 25, 28.8% for 26 to 35, 21.9% for 36 to 45, and 22.2% for those 46 and above. The two youngest groups were small (14 and 15 respondents), so their figures are read with caution. By gender, women chose the biased option slightly more often than men overall (26.4% versus 22.4%; Figure 4), but this gap was small and concentrated in two biases, time inconsistency and recency, rather than spread across all five (Table 2).

**Table 2.** Bias across demographic and self-report groups.

| **Comparison** | **Lower bias** | **Higher bias** |
|---|---|---|
| Gender | Men 22.4% | Women 26.4% |
| Age | 46+ 22.2% | 16–18 31.4% |
| Confidence | Level 5 19.6% | Level 1 38.1% |
| Decision driver | Own research 21.2% | Social media 37.1% |

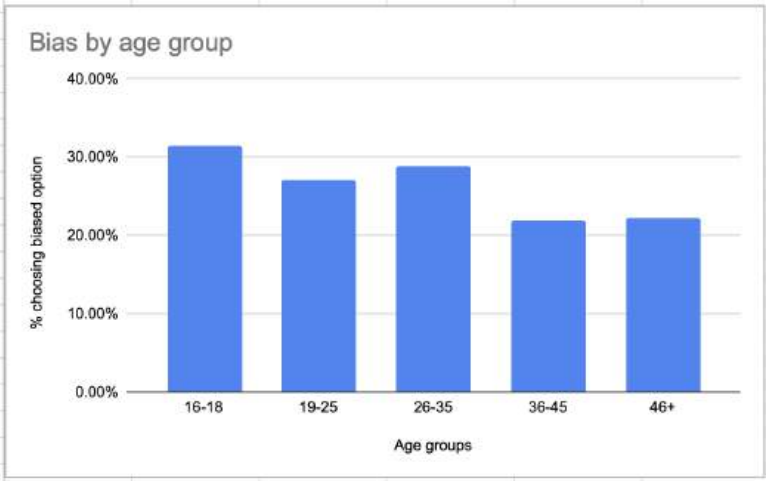



Figure 3. Bias by age group. Bias tends to be higher among younger respondents (31.4% for ages 16 to 18) and lower among older ones (22.2% for 46 and above). The two youngest groups were small (n = 14 and n = 15), so their bars are read with caution.

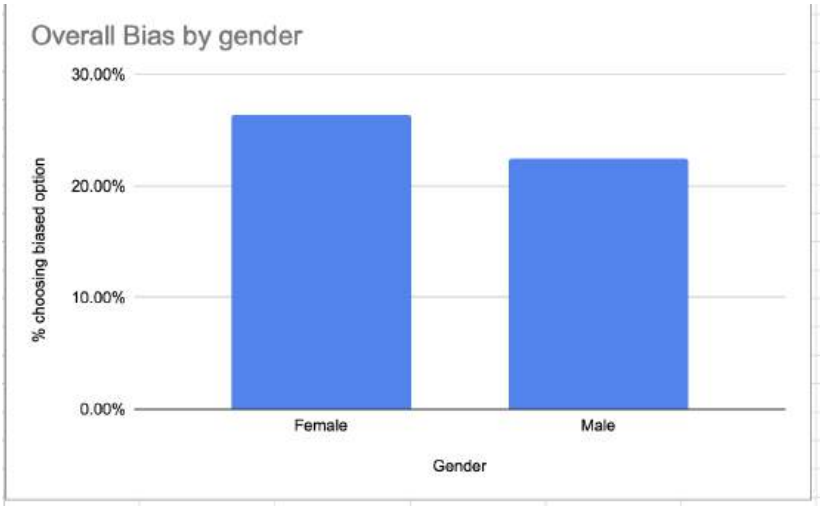



Figure 4. Overall bias by gender. Women chose the biased option slightly more often than men (26.4% versus 22.4%), but the gap is small and came mainly from two biases (time inconsistency and recency) rather than a general tendency across all five.

***Confidence and Bias***

The clearest pattern involved confidence (Figure 2). Measured bias fell at every step of the 1-to-5 confidence scale: 38.1% at level 1, 30.7% at level 2, 26.8% at level 3, 23.2% at level 4, and 19.6% at level 5. Because bias dropped at every level rather than only at the extremes, the downward pattern is a strong sign of a real relationship in the sample, in which more confident respondents made better financial choices. Separately, women rated their confidence lower than men on average (3.41 versus 3.92 out of 5). These two findings are kept distinct: confidence was linked to lower

bias, and women reported lower confidence even though their actual choices were only slightly more biased than men's.

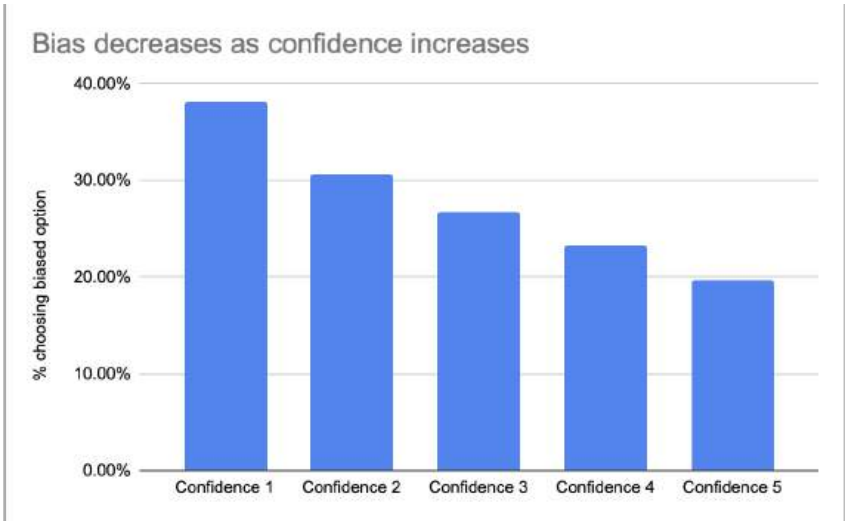



Figure 2. Bias decreases as confidence increases. Bias falls at every step of the scale, from 38.1% at level 1 to 19.6% at level 5. Because the decline is consistent across all five levels, it points to a genuine negative relationship, the opposite of a simple overconfidence pattern.

***Decision Drivers and the Relationships Among Biases***

Respondents' stated decision drivers matched their measured behaviour. Those who relied on their own research were the least biased (21.2%), followed by past experience (27.4%), then friends or family (34.4%), with social media the most biased (37.1%). Only seven respondents chose social media, so that figure is treated with caution. Finally, being high on one bias did not mean being high on the others; a respondent prone to one bias was not necessarily prone to the rest. This suggests the five biases are fairly separate tendencies rather than one general habit, a pattern visible only because all five were measured in the same people.

**Discussion**

The central question was how common five behavioural biases are in the investment decisions of Indian respondents, and how they vary by age, gender, and confidence. The short answer: biased choices were the exception rather than the rule (24.7% overall), and the biases that did appear clustered around self-control. Present bias and time inconsistency, both rooted in the pull of the immediate over the delayed, were the two strongest, and the single most bias-inducing scenario (Q5, at 53%) was the one that pitted a reward now against a larger reward later. This locates the sample's departures from rationality mainly in self-control rather than in social conformity or information processing (Kahneman & Tversky, 1979; Simon, 1955), and it matters for investing, where building wealth depends on the willingness to wait.

Loss aversion complicates the picture. It ranked among the three strongest biases on average, but that average is misleading: 62% chose the biased option on the uncertain-gain item, against

roughly 12% to 14% on the other two. The three items meant to capture one bias behaved very differently, which suggests that what gets grouped as loss aversion may really be several tendencies, avoiding risk on a possible gain, sticking with sunk costs, and anchoring to an original price, that do not always travel together. This fits the mixed record in the literature, where loss aversion sometimes drives decisions strongly and sometimes barely at all (Aigbovo & Ilaboya, 2019; Suriyanti et al., 2024). It also carries a measurement lesson: a few scenario items are a blunt tool for a construct this internally varied.

Herd behaviour was the weakest bias at 10.0%, which on its face suggests people rarely follow the crowd. That reading deserves caution. Because the scenarios carried an obviously sensible option, respondents may have been reluctant to cast themselves as followers, a form of social-desirability bias well documented in self-report financial research, where people give more biased answers about themselves than about a hypothetical other. The paper's own data support this worry: herd behaviour looked low in the direct items, yet the respondents who said social media drives their decisions were the most biased group of all. Social influence may work in ways people do not notice or admit, consistent with evidence that influencers and peer cues now shape investing more than formal sources (Setyadi et al., 2025).

The gender and confidence results are best read together. The overall gender gap in bias was small and concentrated in two biases, yet women rated their financial confidence clearly lower than men (3.41 versus 3.92). Put side by side, these point to a gap between confidence and competence: women made only slightly more biased choices, but judged their own ability much lower. This mirrors a central finding in the literature, where much of the apparent gender gap in finance reflects confidence rather than ability, and where lower test scores for women partly reflect anxiety rather than knowledge (Barber & Odean, 2001; Cupák et al., 2021; Tinghög et al., 2021).
The confidence result is the study's strongest, and its direction is what makes it interesting. Bias fell at every step of the confidence scale, and more confident respondents made fewer biased choices, not more. That is the opposite of a naive Dunning-Kruger story, in which the least able are the most sure of themselves. In this sample, confidence was broadly calibrated to decision quality rather than inflated. Set against the gender finding, a fuller reading emerges: confidence tracked ability within the sample, but women appear to under-rate themselves relative to that ability, a kind of reverse miscalibration.

**Conclusion**

Returning to the research question, the evidence gives a clear answer: in this sample, behavioural bias was neither pervasive nor uniform, but concentrated in self-control, and self-reported confidence was a surprisingly honest signal of decision quality. First, present bias and time inconsistency led the field, which shows that the sample's main financial weakness was resisting immediate temptation rather than following the crowd or misreading information. Second, bias fell steadily as confidence rose, which shows that confidence here reflected genuine decision quality rather than empty bravado. Third, women reported much lower confidence than men despite making similar choices, which shows that the gender story in this sample is about perception as

much as ability. Together these results suggest that behavioural bias is not one general trait but a set of distinct, situation-triggered tendencies, and that how capable people feel is worth measuring alongside how capable they are.

Several limitations bound these conclusions. The analysis used descriptive statistics only, so every group difference is a pattern in this sample, not a proven effect. The sample is a non-probability convenience sample of 234, reached through one network and skewed older and more female. All answers were self-reported responses to imaginary scenarios rather than choices over real money, which invites both social-desirability and hypothetical bias, and likely understates herd behaviour. The design was cross-sectional, so it cannot separate the effects of age from those of growing up in a different financial era. Future work with larger, random samples, incentivised tasks, more items per bias, and longitudinal designs could test whether the confidence-calibration pattern holds under real stakes, and whether closing the confidence gap changes women's investing behaviour.

**Acknowledgements**

The author thanks his research mentor, a doctoral researcher at Cornell University, for guidance throughout this project. The author received no external funding and declares no conflicts of interest.

## Works Cited


Aigbovo, O., & Ilaboya, O. J. (2019). Does behavioural biases influence individual investment decisions. Management Science Review, 10(1), 68–89.

Barber, B. M., & Odean, T. (2001). Boys will be boys: Gender, overconfidence, and common stock investment. The Quarterly Journal of Economics, 116(1), 261–292.

Croson, R., & Gneezy, U. (2009). Gender differences in preferences. Journal of Economic Literature, 47(2), 448–474.

Cupák, A., Fessler, P., & Schneebaum, A. (2021). Gender differences in risky asset behavior: The importance of self-confidence and financial literacy. Finance Research Letters, 42, 101880.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. The Journal of Finance, 25(2), 383–417.

India Brand Equity Foundation. (2026). The new Indian investor: Young, digital and confident.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. Econometrica, 47(2), 263–291.

National Centre for Financial Education. (2019). Financial literacy and inclusion in India: Final report.

National Stock Exchange of India. (2025). Registered investor data, FY25.

Securities and Exchange Board of India. (2025). Demat account and retail participation data.

Setyadi, V., Indarto, & Indriana, D. (2025). Behavioral biases, financial literacy, and female investors: The role of social media. Matrik, 19(1).

Simon, H. A. (1955). A behavioral model of rational choice. The Quarterly Journal of Economics, 69(1), 99–118.

Suriyanti, S., et al. (2024). Exploring financial behavior: A qualitative investigation into psychological factors influencing risk preferences and investment decisions. Golden Ratio of Finance Management, 4(2), 100–112.

Thaler, R. H. (1985). Mental accounting and consumer choice. Marketing Science, 4(3), 199–214.

Tinghög, G., et al. (2021). Gender differences in financial literacy: The role of stereotype threat. Journal of Economic Behavior & Organization, 192, 405–416.

von Neumann, J., & Morgenstern, O. (1944). Theory of games and economic behavior. Princeton University Press.

Zahera, S. A., & Bansal, R. (2018). Do investors exhibit behavioral biases in investment decision making? A systematic review. Qualitative Research in Financial Markets, 10(2), 210–251.

**Economic Crisis, Political Response: Comparing the USSR's Decline and America's Great Recession Recovery By Fedor Klopotovskii**

**Abstract**

In the past 100 years, the Soviet Union and the United States of America (USA) both experienced major economic crises. The economic decline of the Soviet Union began in the 1970s, while the USA faced the Great Recession in the 2000s. Although the two examples are separated in both time and space, they serve as a useful comparison since the USA and USSR were both economic superpowers. By collecting and analyzing evidence-based literature articles, I compare and contrast the USSR's and America's response to their economic crises. I conclude that the USA's response is much more effective, since its crisis lasted for a much shorter time than the one faced by the Soviet Union. More importantly, I argue that the state of the political system, that is, the degree of liberalization and the level of public morale, played a larger role in the differing effectiveness of response to crisis than any particular economic policy. Although this analysis focuses on two superpowers experiencing the same challenge, its implications extend beyond this case, demonstrating how a country's political state shapes its economic, religious, and societal state.

**The Soviet Union: From Prolonged Stagnation to Chaotic Transition**

It is often incorrectly claimed that the Union of Soviet Socialist Republics (USSR; Soviet Union), a communist regime, was always economically stagnant. Between the end of World War II (1945) and the 1970s, the USSR saw continuous economic growth on par with Western Europe. This situation changed in the 1970s, as the USSR's GDP largely stopped growing (Zhuravskaya, Guriev, and Markevich 6). By the 1980s, Mikhail Gorbachev was tasked with fixing a worsening situation. He attempted to liberalize the economy by providing limited freedom to industrial complexes, giving them an opportunity to make their own financial and strategic decisions. As economists Christine Lang and Shlomo Weber write, however, the Gorbachev regime's decision to decentralize the economy was "internally contradictory. They decentralized economic decision-making without establishing the market institutions, price signals, or property rights that would make decentralized decisions rational." (Lang and Weber 185).

The Gorbachev regime also had to contend with a highly traditional and conservative Soviet bureaucracy. The bureaucrats and future economists of the Soviet Union, who were expected to enhance the economy, appeared to lack confidence and willingness to help the country escape from exhausting economic stagnation. As described by Paul Roderick Gregory, a professor of economics at the University of Houston, surveys exposed a common fear that pushed bureaucrats away from economic reforms — the fear of losing their jobs (Gregory 860).

After the fall of the Soviet Union, when the communist dictatorship was overthrown and Boris Yeltsin's administration took a course of democratization and decentralization, a new challenge shook Russia. Following the 1993 coup, as new administrators were trying to establish

a new order, multiple mafias and gangs were born on the streets of Russia. These organizations significantly worsened the political instability in the country. As Lang and Weber observe, economic and political progress in this period was "hampered by the poor and criminalized state of business and political spheres" (Lang and Weber 179). The criminalization had direct consequences for Russia's economic situation - the privatization program proceeded without the international financial support that had been promised, and that would have been necessary to stabilize the macroeconomic environment (Lang and Weber 187). Russia undertook price liberalization, trade liberalization, and privatization simultaneously, without the foreign investments that reformers hoped would reach them. The financial "aid" that was supposed to support a newborn economy never happened, largely because of the political chaos that made Russia an unattractive destination for foreign investments.

As the criminal situation cooled down, the chaos transformed into corruption and was shifted to another level, from petty smuggling to corruption in the government and large corporations. Police crackdowns eliminated mafias and gangs in cities and rural areas, while the new government gifted previously government-owned companies to friends and acquaintances of the administration. Russia gained most of its income from selling natural resources, which could give the new-born country a variety of opportunities to develop its economic and societal state. However, as described by Thomas Friedman, Yeltsin "let business clans rip off Russia's gas, oil and minerals… this wealth has gone to enriching individuals rather than cushioning the society in transition" (Lang and Weber 189). Described as "privatization" and "decentralization", in reality this new approach inevitably helped corruption thrive, which paralyzed effective and constructive entrepreneurial activity.

Overall, as the Soviet and later Russian economy experienced partial, and later full liberalization and decentralization, it could not operate as economies did in Western countries. The political chaos and criminality that were brought with the military coup in 1993 made the situation worse - the economy had to proceed without the international financial support that was supposed to stabilize the macroeconomic environment of Russia. This way, Russia had to carve its own economic path and bring peace to the country before major changes would come. The military actions of the 2003 war in Iraq shook the oil trade in the Middle East, transferring the opportunity to sell oil to Russia. This economic instability granted the Russian economy new incomes from trade, which led to Russian economic prosperity by the late 2000s (Arvedlund).

**The United States of America: From Financial Fall to Rapid Recovery**

In the early 2000s, in the USA, prospective homeowners were taking loans from banks and buying properties without regard for their ability to pay the loans back. In 2006, the foreclosure crisis escalated as citizens of the USA. declared their inability to pay back countless mortgages. Economist John Tatom describes how "the broader financial crisis developed in stages from this starting point: by spring 2007, several hedge funds and mortgage bond insurers had either failed or had their credit seriously downgraded, while the insurance giant AIG required a federal bailout of approximately $180 billion to avoid complete collapse" (Tatom

175). As historian Adam Tooze argues, the crisis that materialized was not the sovereign debt or currency crisis that many analysts anticipated, but rather one triggered by the housing market that quickly threatened the stability of the world economy (Tooze 7, 14). It spread far beyond American borders, shaking the financial systems across the United Kingdom, East Asia, and Eastern Europe.

The response of the USA's government and the Federal Reserve was swift and historically unprecedented. "The Fed assumed the role of the liquidity provider of last resort not only to American institutions but to the global banking system, establishing the dollar swap lines with a hand-picked group of central banks that allowed them to issue dollar credits on demand" (Tooze 9). Inside the US, the Fed launched a series of so-called "quantitative easing programs (QE)", or QE1, QE2, and QE3, that were meant to stabilize the economic environment by feeding it tons of capital. "The scale of these interventions was extraordinary: reserve balances grew from approximately $10 billion before the crisis to over $2.400 billion" (Taylor 62). Even though economic and job growth during this period fell short of the Fed's own forecasts, and that long-term rates rose rather than fell following the launch of QE (Taylor 62), the effectiveness of QEs as a stabilizer of the economy and growth stimulus is unarguable.

The political factor also played some significant role in the USA's rapid recovery. Following the 2008 elections, the Democratic Party held unified control of American branches of government, which helped create the political conditions that were necessary for a proper economic/fiscal response. Political scientists Andrea Campbell and Michael Sances describe the pattern in Democrat-controlled areas: "personal income tax increases were more likely in states with Democratic legislatures, while midyear spending cuts were smaller in states with larger public sectors" (Campbell 252). More significantly, Campbell and Sances's research reveals what might be described as a "double-vectoring" effect of two-party governance during the recession. Democratic controlled states tended to rely on progressive tax systems, while republican governments were shown to react more quickly to unexpected economic situations and "make larger net adjustments to spending" (Campbell 258). The result was the political configuration of two different approaches, each effective in its own way, that worked simultaneously and formed the base of the United States economic response. This collaborative solution stands in direct comparison to the Soviet Union's earlier crisis, where neither of the aforementioned solutions was present. The communist regime was too traditional to reform, while the post-Soviet state was too chaotic to stabilize itself.

**Aftermath from Two Crises: Politics or Policy?**

The United States and Soviet Union both faced severe economic crises, and both attempted to respond with some combination of liberalization and direct governmental intervention. The difference in outcomes is shocking: the Soviet and Russian economies drowned for decades, while the U.S. recovered within a few years from its most dangerous crisis. This paper has argued that the effectiveness of economic policy is determined not only by the policies themselves, but by the political conditions in which they are implemented. The

Soviet bureaucracy feared reform, while the post-Soviet state lacked political stability, was wracked by criminality, and promoted corruption. The American government, in contrast, implemented the "two-vectored" approach to economic recovery and restored the American people's access to capital through the Federal Reserve's QE1, QE2, and QE3 measures. While the Soviet bureaucracy feared reform, the American government pushed it to occur. This great difference in political conditions, more than any particular fiscal policy, explains the difference in outcomes.

**Works Cited**

Arvedlund, Erin E. "Russian Growth Accelerates, Stoked by Oil." New York Times, 7 Jan. 2004, https://www.nytimes.com/2004/01/07/business/russian-growth-accelerates-stoked-by-oil.html.

Campbell, Andrea Louise and Michael W. Sances. “State Fiscal Policy during the Great Recession: Budgetary Impacts and Policy Responses.” The Annals of the American Academy of Political and Social Science, vol. 650, Nov. 2013, pp. 252–73. JSTOR, http://www.jstor.org/stable/24541685.

Gregory, Paul R. “The Impact of Perestroika on the Soviet Planned Economy: Results of a Survey of Moscow Economic Officials.” Soviet Studies, vol. 43, no. 5, 1991, pp. 859–73. JSTOR, http://www.jstor.org/stable/152453.

Lang, Christine, and Shlomo Weber. “Ten Years of Economic Reforms in Russia: Windows in a Wall.” The International Lawyer, vol. 34, no. 1, 2000, pp. 179–92. JSTOR, http://www.jstor.org/stable/40707515.

Tatom, John A. “Crises and the Great Recession.” Business Economics, vol. 48, no. 3, 2013, pp. 175–81. JSTOR, http://www.jstor.org/stable/23491699.

Taylor, John B. "The Role of Policy in the Great Recession and the Weak Recovery." The American Economic Review, vol. 104, no. 5, May 2014, pp. 61-66. American Economic Association Web, https://www.aeaweb.org/articles?id=10.1257/aer.104.5.61.

Tooze, Adam. Crashed: How a Decade of Financial Crises Changed the World. Penguin Books, 2018.

Zhuravskaya, Ekaterina, et al. "New Russian Economic History." Journal of Economic Literature, vol. 62, no. 1, Mar. 2014, pp. 47-114. Economic Association Web, https://www.aeaweb.org/articles?id=10.1257/jel.20221564.

# Nike's Shift from Traditional Advertising to Digital Marketing: Impact on Consumer Engagement and Brand Loyalty By Vivaan Daruka (student)[1]*


## Abstract

Social media is an internet-based technology, which allows users to share, exchange, and create information and content from anywhere. The growth of digital platforms has fundamentally transformed how brands communicate with consumers, shifting marketing strategies away from traditional platforms such as television, towards social media platforms. The aim of this paper is to examine how Nike's transition from traditional advertising to digital and social media marketing has influenced consumer engagement and brand loyalty. This research will use quantitative and qualitative research, to help understand Nike's shift from traditional advertising to digital and social media marketing has influenced consumer engagement and brand loyalty, from the lenses of consumers or potential consumers. The method of collecting data was a survey which adopted a non probability conceive sampling method. The findings indicated that the majority of respondents encountered Nike through social media platforms, and most frequently among teenagers, who rated the advertisement most effective. Influencer campaigns produced a positive but moderate effect on engagement, while peer influence emerged as a significant driver of brand interaction. In conclusion 86% of respondents credited social media in strengthening their loyalty to Nike. Hence the digital shift had a positive impact on both engagement and loyalty.



## Keywords

Marketing, Digital Marketing, Traditional Marketing, Nike, Consumer Engagement, Brand Loyalty


## Introduction

Social media is an internet-based technology, which allows users to share, exchange, and create information and content from anywhere in the world. It is used by more than 6 billion people around the world, representing more than 73% of the world's population (Kemp, 2024). In today's era, most users of social media tend to gain knowledge, information, and news from social media alone, with the average person spending more than 2 hours on social media scrolling through information (Williams, 2022). This ability to reach such a large audience in such a short time makes social media a prime target for marketing for brands and companies around the world. In addition to the widespread usage, social media has changed how consumers interact with information and brands. Studies show that users are no longer just receivers of information, but actively engage with the information, through likes, comments, shares, and even content creation (Lu et al., 2025).


1. Pathways School Noida, Sector 100, Noida, Uttar Pradesh, 201301, India
* Corresponding author email: darukavivaan@gmail.com

This results in increasing two way communication and feedback, becoming a fundamental part in brand creation and keeping up with competition. Hence, consumers play a role in shaping brands perception and influencing others' options. This creates social media as a dynamic environment, significantly impacting consumers and brand decision making process. In addition to the two way communication, digital marketing is much affordable and accessible for small and medium sized enterprises (Pros, 2024), which often work with financial constraints. Social media platforms are able to provide low cost or even free tools for promotion, allowing businesses to reach a larger audience without investing a significant amount of money. This enables more frequent marketing efforts. This leads to increasing brand recognition and visibility over time. Digital marketing uses algorithms, which allow users to target a specific group of audiences, based on demographics, interest, online behaviours, generating more returns. Businesses have been rapidly shifting from traditional advertisement methods, such as television commercials, print media, and billboards, which tend to reach a limited audience, are expensive, and do not provide instant feedback (Veleva & Tsvetanova, 2020). Social media, on the other hand, allows businesses to easily reach a wider audience and potential customers, is more affordable for smaller businesses, and provides immediate feedback on the goods and services they offer (Hayes, 2025). Traditional marketing methods are constrained by their inability to filter out specific target audiences, as advertisements are broadcast to a broad and general population, regardless of their interests or relevance (Excellent Publicity, 2024).

This causes inefficient use of funds and investment, as businesses may invest a large amount of money, but end up reaching consumers with limited interest in their products. High costs are associated with production and placement, such as televisions, print space, these methods are less accessible, particularly for smaller businesses. Another limitation of traditional marketing would be its one way communication nature, where consumers of the information or advertisement have little or no chance of providing feedback to business, or even interact with the brand, making it difficult for businesses to measure the effectiveness of their campaigns, or adapt to changing consumer preferences and keep up with competition. Hence, businesses have started overlooking traditional marketing due to it becoming less flexible and less responsive, while coming with higher costs. This results in switching to alternatives, in this case digital marketing.

Digital marketing is projected to grow to over 390 billion dollars by the end of 2026 (Ketan, 2025), due to social media becoming a widely adopted tool for marketing across a variety of businesses. This rapid growth is supported by the increasing accessibility of smartphones and internet connectivity, which has made digital platforms more integrated in daily life (Kemp, 2024). The leading consequence to the rise in digital marking is the development of brand loyalty, by enabling continuous interactions between brands and consumers. Social media allows brands and consumers to have regular communications, feedback, build trust and connection, developing brand loyalty. Tools offered by digital marketing, helps provide personalised recommendations, loyalty programs, interactive campaigns, encourage repeated engagement and purchases.

This makes consumers feel more valued and connected with brands through online interactions. This makes them more likely to be loyal and advocate towards a specific brand. In addition to the point, additional user generated content, and enabling interaction between consumer to consumer, enhances loyalty furthermore. This is influenced by others' opinions. This establishes a sense of community and creates a stronger attachment with the brand. This makes consumers less likely to switch to competitors.

Therefore, showcasing a massive opportunity for businesses to grow by leveraging the benefits of social media, making it important to study digital and social media marketing and its influence on consumer engagement and brand loyalty. It is within this context that this paper seeks to explore, through the lens of Nike, how the shift from traditional advertising to digital and social media marketing has influenced consumer engagement and brand loyalty.

**Literature Review**

Research on the shift from traditional to digital marketing has commonly documented a fundamental change in how brands communicate with consumers, traditional marketing approaches have consistently been known as one directional, with top-down communication and untargeted strategies. Which offer less personalised and interactive advertising (Awan et al., 2024). Awan et al. (2024) further state that traditional methods fail to harness data and analytics in a meaningful way, leaving brands unable to predict or respond to individual consumer needs. In contrast, digitally transformed brands have been found to make use of advanced data systems, artificial intelligence, and personalisation technologies that allow them to anticipate consumer behaviour and tailor their messaging accordingly. Separately, Tran et al. (2025) highlight that digital brands are significantly more effective due to two-way, real-time engagement through social media platforms and influencer partnerships.

The relationship between social media marketing and brand loyalty has been explored in considerable depth in the existing literature. Wilson et al. (2024) identify personalisation, high-quality content, and authentic influencer endorsements as the primary factors through which social media marketing enhances the emotional connection between a brand and its consumers, thereby building credibility and long-term loyalty. Wilson et al. (2024) also state that community management and a clearly defined brand identity play a major role in generating positive consumer experiences, suggesting that loyalty is not merely a transactional outcome but is cultivated through ongoing relational engagement in digital spaces. Furthermore, Wilson et al. (2024) find that when consumers feel that the content they encounter on social media is genuinely relevant to their interests and values, their attachment to the brand deepens in ways that converts into purchasing behaviour. The study thus establishes that the mechanisms of digital and social media marketing help build the emotional loyalty that traditional mass-market advertising has struggled to generate.

Other studies have examined in greater detail the specific mechanisms through which digital marketing strategies contribute to brand growth and consumer engagement. Wang (2024) emphasises the use of search engine marketing and video-driven content as tools through which

brands create strong emotional connections with consumers, noting that the integration of multiple digital channels allows for seamless and personalised consumer experiences underpinned by large-scale data analysis. Wang (2024) further discusses how corporate social responsibility is used effectively in digital environments to improve brand image, reflecting the broader finding in the literature that digital consumers increasingly expect brands to demonstrate authentic social values rather than purely commercial intent. Additionally, Wang (2024) highlights that the combination of digital channels allows brands to maintain a consistent and coherent identity across different consumer touchpoints, reinforcing brand recognition and trust in ways that fragmented traditional advertising campaigns often fail to achieve.

The use of experiential marketing, emotional marketing, and hunger marketing as components of an integrated digital strategy has also been documented in the literature. Li and Yu (2023) explain that brands combine these approaches to influence consumer demand, with limited product releases and exclusive campaigns used to create incentives and scarcity that drive both immediate engagement and longer-term brand desire. Li and Yu (2023) further state that the shift toward a direct-to-consumer model, enabled by digital infrastructure, allows brands to use data analytics and mobile applications to deliver personalized shopping experiences that traditional retail and advertising models could not support. Separately, Lim et al. (2025) identify the emergence of artificial intelligence and the metaverse as future opportunities for deepening consumer engagement through digital channels, while acknowledging that challenges around data privacy and consumer trust will need to be carefully managed as these technologies develop. Taken together, the studies suggest that the transition from traditional to digital marketing is not a simple change of platform at which brands promote themself, but a transition with logic and reasoning, using which brands thrive to deepen relationships with consumers. Despite the amount of research available on the shift from traditional to digital marketing and its effects on consumer engagement and brand loyalty there is a significant gap in the literature. While there are several studies dressing digital marketing strategies in general or specific strategies, there is a lack of studies that measure effects of the shift from traditional to digital on consumer engagement and brand loyalty and specifically from Nike's perspective. Studies focused on Nike’s digital strategies, such as Li and Yu (2023) and Wang (2024), provide detailed accounts of the mechanisms and tools Nike employs but do not offer a comparison with traditional advertising and its effectiveness or any qualitative analysis. Wilson et al. (2024), who focus on social media and brand loyalty, and Awan et al. (2024), who examine AI-driven personalisation effectiveness, remain general in nature and do not apply their findings to Nike as a specific case. This gap in the literature is what the study seeks to cover, by examining how Nike’s particular shift from traditional advertising to digital and social media marketing has influenced consumer engagement and brand loyalty.

**Methods**

This research will use both quantitative and qualitative research design, to help investigate how bikes shift from traditional advertising to digital and social media marketing has

influenced consumer engagement and brand loyalty. The survey method was selected due its practicality, cost effectiveness and ability to collect measurable data from a wider audience without any time frame (SmartSurvey, 2025). The questionnaire is made to gather both quantitative and qualitative data regarding participants perceptions, behaviours and response towards Nike's traditional vs digital marketing strategies.

The research primarily focused on understanding how consumers engage with nikes content across different platforms and whether digital marketing strategies influence loyalty and engagement. The survey includes questions relating to advertising exposure, social media engagement, influencer impact, brand perception and behavioural responses to Nike's marketing campaigns, and use of qualitative components with open ended questions allowing personal opinions and explaining how Nike's digital presence may have influenced their trust and loyalty toward the brand.

This study involved a non probability conceive sampling method, as participants were selected based on accessibility and whether they were willing to participate. The target population included individuals across different age groups who are familiar with Nike and have used digital platforms, such as social media.

1. Demographics – Questions relating to age
2. Traditional vs Digital Marketing – Questions where respondents most commonly encounter Nike advertisements and its effectiveness of digital marketing compared to traditional advertising methods.
3. Consumer Engagement – Questions with interaction with Nike's digital content, including likes, shares, comments, website visits, and responses to athlete or influencer collaborations.
4. Brand Loyalty – Questions investigating levels of loyalty towards Nike, willingness to recommend the brand, and how Nike's social media presence has influenced trust and loyalty.

The survey includes open-ended questions which helps enable respondents to provide more detailed explanations regarding their experiences and thoughts about Nike's marketing strategies, adding reasoning to the quantitative findings.

All survey responses were collected through Google Forms , which generated summary statistics and visual charts for each question automatically. Therefore helped organise and analyse patterns relating to consumer engagement and brand loyalty.

Demographic Data:

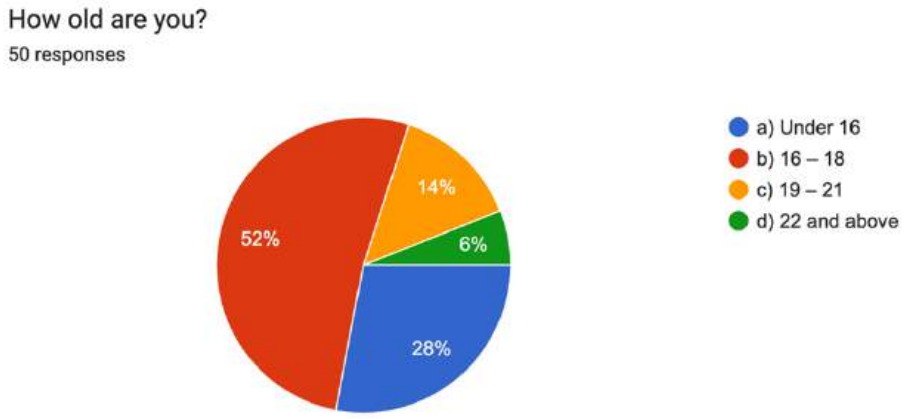



**Figure 1.**

Behavioural Data

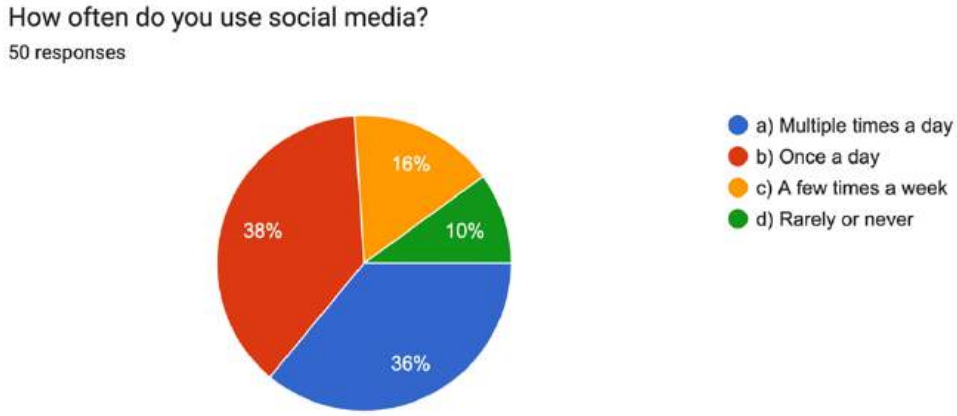


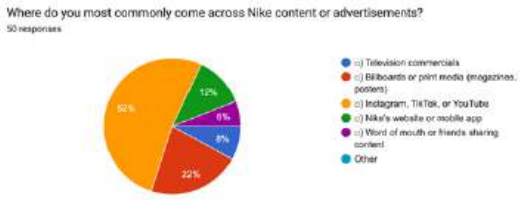
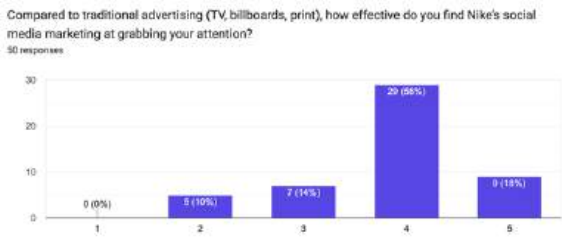
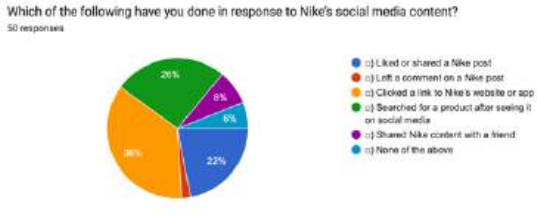

**Figure 2**.

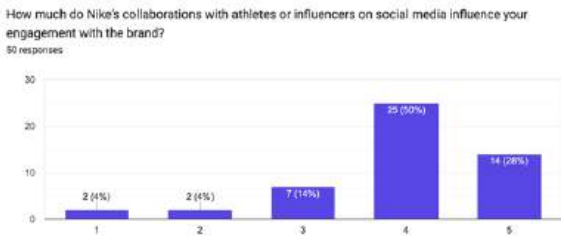
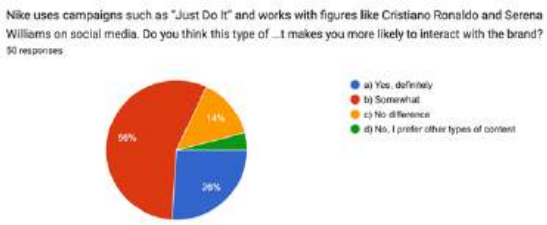
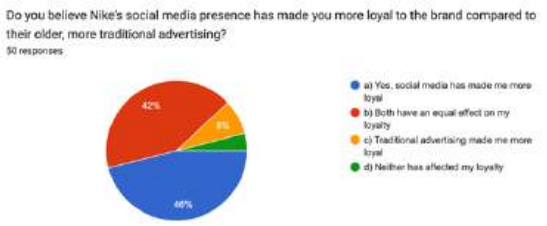
Every Likert-scale and multiple-choice question is analysed in a particular subheading. Frequency distributions and percentages were used to analyse responses. Furthermore graphs such as pie charts and bar graphs were used to find trends in the responses. Not all graphs and data are shown, the ones which have a connection with customer loyalty and engagement have been chosen.

## Results and Discussion

### *Quantitative Analysis*

**Table 1.** This table discusses the responses of respondents to quantitative questions in the survey and their analysis and findings in relation to changing marketing strategies and its impact on consumer engagement and brand loyalty.

| Subsection: | Graph and Interpretation |
|---|---|
| Change from traditional to digital advertisement |  <br>**Figure 3.**<br>Figure 3 shows that 52% of respondents come across Nike through social media platforms while 30% from traditional advertising methods. This indicates the shiviting advertisement from traditional to digital. |
| Perception of Nike's social media content in terms of attention grabbing. And frequency of the audience engaging with the content. |  <br>**Figure 4.**<br>Figure 4 shows that 58% of the responder give Nike's social media contents ability to grab attention a rating of ⅘. And 18% giving it a 5/5. 76 percent of responders believe that Nike is able to capture user attention through their ad campaigns in social media. This showcases that Nike's digital media content is engaging for social media users hence they tend to stop and view Nike's content, and are able to grab attention successfully. |
| |  <br>**Figure 5.**<br>In addition to attention grabbing content, figure 5 highlights that 94 percent of the respondents have interacted with Nike's social media content and only 6 percent have not interacted in any form. This shows that social media campaigns and advertisements are very engineering and attractive with the audience and provides the audience to be able to communicate with Nike through likes, shares, comments etc, unlike traditional advertising. |

| | |
|---|---|
| Views on digital marketing strategies adopted by Nike and their effectiveness on engagement and interaction. |  <br>**Figure 6.**<br>Figure 6 demonstrates Nike's marketing strategy of including influencers and athletes in their digital ad campaigns and whether the strategy is effective in engaging the audience or not. 50% responded that nike was effective in doing so ⅘ times will 28% said Nike's strategy was highly influential in increasing engagement. |
| |  <br>**Figure 7.**<br>Furthermore, figure 7 indicates that strategies such as “Just do it” or use of public figures helped make viewers interact with the brand. The graph shows that more than 72% of respondents had a change in interaction with the brand due to public figures and Nike campaigns such as “just do it” highlighting their effectiveness. In conclusion, the shift in focus towards digital media and marketing of Nike has helped elevate Nike's engagement and interactions. |
| Change in loyalty and trust from digital advertisement in comparison to traditional. |  <br>**Figure 8.**<br>Figure 8 indicates that the majority of responders either believe that social media and traditional advertising had equal affect on loyalty or either social media had more effect on loyalty. With 46% of respondents stating that social media had made me more loyal and 42% saying both had the same affect on loyalty. |

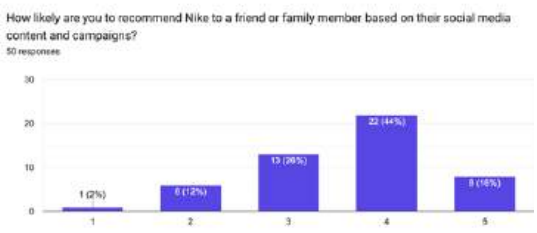



**Figure 9.**
In addition to the comparison between both advertising strategies, respondents are likely to recommend Nike to friends and family based on social media content, as from figure 9, 60% of responders gave it a likelihood rating of above 3. Demonstrates the loyalty responders have for Nike as they are willing to recommend it to their family and friends.

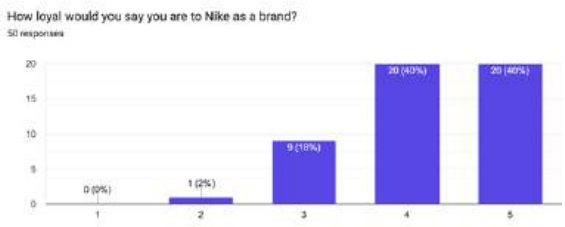



**Figure 10.**
Figure 10 shows that Nike's loyalty is also evident as 40 percent of responders rated their loyalty towards Nike a ⅘ and 40% gave a rating of 5/5.

***Qualitative Analysis***

Open ended questions, questions were grouped into themes based on similarity of meaning.

Question: In your own words, do you think Nike's shift to digital and social media marketing has strengthened or weakened your trust and loyalty towards the brand? Please explain why.

**Table 2.** This table discusses the responses of respondents to the open ended question "In your own words, do you think Nike's shift to digital and social media marketing has strengthened or weakened your trust and loyalty towards the brand? Please explain why." in the survey and their analysis and categorizing the outcomes deduced from responses into identified themes in relation to changing marketing strategies and its impact on consumer engagement and brand loyalty.

| Theme 1 - Strengthened trust and loyalty | Theme 2 - Weakened trust and loyalty |
|---|---|
| Most participants stated that Nike's presence in the digital world has helped it strengthen loyalty and trust. Giving the reasoning that using is able to deliver more personalised content to users, making the ads more engaging. Furthermore digital marketing allows Nike to make content users can engage with through likes, comments and shares, hence participants of the survey say, this makes them likely to purchase Nike products when they see Nike content which has been interacted with by known people. | Some participants of the survey have answered that Nike's digital shift has weakened their loyalty and trust. Stating that digital marketing does not provide them with the authenticity they got from traditional Nike ad campaigns through traditional marketing methods, the loss of authenticity have prompted them to switch to alternatives. |

**Table 3.** This table discusses the responses of respondents to the open ended question "In your opinion, what is the biggest advantage Nike gains by using social media over traditional advertising?" in the survey and their analysis in relation to changing marketing strategies and its impact on consumer engagement and brand loyalty.

| Question: | Responses: |
|---|---|
| In your opinion, what is the biggest advantage Nike gains by using social media over traditional advertising? | Survey takers stated that the advantages of Nike's use of social media compared to traditional advertising was the ability to get more audience engagement by promoting more personalized content, furthermore being able to link websites and products directly through the ad that gets more traffic through their website. Furthermore participants stated that Nike could make more interactive content and due to increased social media popularity, more people are likely to view the ad campaign, hence more audience presence. |

This study looked at how Nike is moving from traditional advertising to digital and social media marketing and how it has changed consumer engagement and brand loyalty. Figure 1 shows that 50 people responded to the survey which consisted of responses with different age groups. Showing how different age groups and people with different social media usage view nike and nikes shift to digital advertising. The discussion looks at four main criterions, where people see Nike's content the most, how age affects the responses and behaviour towards Nike's social media ad campaigns, and how influences play a role in advertising and brand engagement. And the overall effect on brand loyalty and engagement.

The first finding from Figure 2 where people view Nike and Nike's advertisements, 52% of respondents said they came across Nike's content on Instagram, youtube and other social media platforms. Only 8% said they viewed Nike's content through their television while 22% said billboard or print media. This showcases the shift from traditional to digital marketing, as Nike has minimised their presence in print, television, billboards, etc, and increased their presence in social media and other digital platforms, in response to increasing popularity of

social media platforms. The  Pew Research Center (2024) found that 9 out of 10 teenagers use YouTube, and around 6 out of 10 use TikTok and Instagram. This indicates that Nike is changing their marketing strategy based on where the audience is going towards or already is. Zhu (2024) also supports this, stating that social media puts brands in front of people when they are already engaged and scrolling, which is more effective than a TV ad interrupting something they are watching.

Different age groups had different opinions and views on Nike and their shift. People aged 16 to 18 gave Nike's social media marketing an average effectiveness score of 4.12 out of 5. Highest among other groups. But people aged 22 and above gave it only 2.00 on average. That is a huge gap. Under 16s averaged 3.79 and 19 to 21 year olds averaged 3.71. Scores based on age groups reveal a pattern, the younger the audience the more effective they find Nike's digital marketing. Nadanyiova and Sujanska (2023) found that Gen Z responds much better to social media marketing than traditional advertising. They connect more with content that feels authentic, personalised, and interactive. Older consumers are different. Nielsen (2025) found that older age groups still prefer traditional TV and established media. This explains why the 22 and above respondents in this study were less impressed by Nike's digital approach. Hence Nike's digital shift is the most effective on younger audiences compared to older audiences who found it less effective.

The survey also questioned the use of influencer collaborations and patternerships for the advertisements. For instance “Just do it” and collaborations with Cristiano Ronaldo and Serena Williams. Figure 7 shows that 54% of respondents said these make them somewhat more likely to interact with Nike. 24% said yes definitely. Only 4% said they prefer other content.  Therefore suggesting that Nike's collaborations with celebrities and influences were appealing and positive. However, a number of respondents said that influences had a somewhat impact on their behaviour, indicating that most people are not fully won over by celebrity campaigns. They notice them, but it is not a strong pull. Ahmed et al. (2024) found that influencer marketing works best when the audience feels a personal connection to the influencer and feel the content is relatable. So Nike's big name campaigns may be creating some engagement. Yet they may not be able to build a better connection.

One of the most honest and revealing findings came from the open ended questions from Table 2. Many respondents did not just talk about Nike's official content. They talked about their friends. One person wrote that seeing friends liking Nike products makes them want to buy too. Another said their trust in Nike grew because they keep seeing Nike videos being reshared and reposted. This is called electronic word of mouth, or eWOM. It is when consumers spread brand content to each other without the brand doing it directly. Gayle (2025) found that peer recommendations on social media are trusted more than any traditional advertisement. People trust their friends more than they trust a billboard or a TV commercial. This is a big deal for Nike. It means their strategy is not just working because of what Nike posts. It is working because their audience is spreading the content for them. Every share, every like, every repost is free marketing. And it reaches people in a more trusted and personal way.

There are some limitations to this study. The sample was 50 people, with most of them being between 16 and 18, and under 16. Only 3 respondents were 22 or above. So the findings about older age groups are based on very few responses and cannot be taken as fully representative of a larger population. The responses were also self reported, meaning people answered based on how they feel, not based on actual purchase records or data. Someone might say they feel loyal to Nike but not actually buy the products regularly and the sample also came from one school setting, so it does not reflect a wide range of backgrounds or locations. Future research should use a bigger and more diverse group of people to get a fuller picture.

**Conclusion**

This study explored how Nike's shift from traditional advertising to digital and social media marketing has influenced consumer engagement and brand loyalty. Based on survey responses from 50 participants. Most respondents encounter Nike through Instagram, TikTok, and YouTube rather than television or billboards. Younger respondents, particularly aged 16 to 18, found Nike's digital marketing the most effective, while older respondents remained more inclined towards traditional marketing. Influencer campaigns had a positive but moderate effect, and peer influence emerged as a strong reason for engagement, with many respondents saying they interact with Nike's content because friends like and share its content.

On brand loyalty, 86% of respondents credited social media as having at least some role in making them more loyal to Nike. A small minority were inclined towards traditional advertising hence Nike's digital shift has had a clear and positive impact on consumer engagement and brand loyalty. Traditional advertising still supports brand loyalty and engagement, but it is no longer the only major force especially among the youth. In conclusion, Nike's move to digital and social media marketing has not just changed where it advertises. It has changed how consumers relate to the brand.

**AI Acknowledgement**

The author declares that no artificial intelligence (AI) tools were used in the preparation of this work.

## Works Cited


Ahmed, S., Islam, T., & Ghaffar, A. (2024). Shaping brand loyalty through social media influencers: The mediating role of follower engagement and social attractiveness. SAGE Open, 14(1). https://doi.org/10.1177/21582440241242928

Awan, T. M., Hassan, S., Perveen, S., & Sarfraz, M. (2024). AI-driven personalization in digital marketing: Effectiveness and ethical considerations. ResearchGate. https://www.researchgate.net/publication/383137190

Veleva, S. S., & Tsvetanova, A. I. (2020). Characteristics of the digital marketing advantages and disadvantages. IOP Conference Series: Materials Science and Engineering, 940(1). https://doi.org/10.1088/1757-899X/940/1/012065

Gayle, L. (2025). The impact of social media on consumer decision making. In Regent Research Roundtables proceedings (pp. 115–127). Regent University School of Business & Leadership. https://www.regent.edu/wp-content/uploads/2025/01/Regent-Research-Roundtables-2025-Business-and-Society-Gayle.pdf

Hayes, A. (2025, June 2). What is social media marketing? Investopedia. https://www.investopedia.com/terms/s/social-media-marketing-smm.asp

Kemp, S. (2024). Digital 2024: Global overview report. DataReportal. https://datareportal.com/reports/digital-2024-global-overview-report

Ketan. (2025, December 1). Digital marketing statistics and facts. Market.us. https://market.us/statistics/digital-marketing/

Li, Z., & Yu, Y. (2023). The impact of Nike's social media marketing strategy on consumer demand. Advances in Economics, Management and Political Sciences, 38, 172–178. https://www.ewadirect.com/proceedings/aemps/article/view/6060

Lim, X. J., Cheah, J.-H., Ng, S. I., & Dwivedi, Y. K. (2025). The bright and dark side of metaverse marketing. Global Business and Organizational Excellence. https://doi.org/10.1002/joe.22271

Lu, X., Balakrishnan, K., Chan, T. J., & Na, M. (2025). The role of real-time engagement in shaping social media check-in behavior: Moderating effects of trust and peer influence. Brain and Behavior, 15(9). https://doi.org/10.1002/brb3.70887

Nadanyiova, M., & Sujanska, L. (2023). The consumer behavior of Generation Z: An impact in modern marketing innovations. International Journal of Research and Innovation in Social Science. https://rsisinternational.org/journals/ijriss/articles/the-consumer-behavior-of-generation-z-an-impact-in-modern-marketing-innovations/

Nielsen. (2025). Why marketers should not discount age in today's media landscape. https://www.nielsen.com/insights/2025/why-marketers-should-not-discount-age/

Pew Research Center. (2024). Teens, social media and technology 2024. https://www.pewresearch.org/internet/2024/12/12/teens-social-media-and-technology-2024/

SmartSurvey. (2025, August 8). Advantages of questionnaires in online research. https://www.smartsurvey.co.uk/blog/advantages-of-questionnaires-in-online-research

Tran, T. A., Nguyen, H. T., & Le, V. H. (2025). The impact of generative AI in shaping customer engagement: Humanising conversational marketing. In Advances in artificial intelligence and machine learning applications. IntechOpen. https://www.intechopen.com/online-first/1214874

Veleva, S. S., & Tsvetanova, A. I. (2020). Characteristics of the digital marketing advantages and disadvantages. IOP Conference Series: Materials Science and Engineering, 940(1). https://doi.org/10.1088/1757-899X/940/1/012065

Wang, B. (2024). The impact of integrated digital marketing on brand growth: Nike. Finance & Economics, 2(10). https://doi.org/10.61173/3j4b6v91

Williams, K. (2022, July 12). 5 excellent reasons why social media marketing is essential. Servon Design. https://www.servondesign.co.uk/5-reasons-you-need-social-media-marketing/

Wilson, G., Johnson, O., & Brown, W. (2024). The influence of social media marketing on brand loyalty. Preprints. https://doi.org/10.20944/preprints202408.0332.v1

Zhu, Y. (2024). Leveraging social media marketing: A case study of Starbucks digital success. EFMD Global. https://blog.efmdglobal.org/2025/11/05/marketing-social-media-branding

# Understanding Commonalities Across Six Financial Risk Episodes in the United States: A Behavioral Economics, Law and Economics, and Financial Media Approach

By Katya Kong

**Abstract**

This paper looks at six different financial episodes and tries to understand whether they have any common patterns. The cases are the collapse of Enron, Lehman Brothers, Madoff Ponzi, the GameStop trading surge, FTX and Alameda Research, and Silicon Valley Bank. This paper compares these cases using three perspectives: behavioral economics, law and economics, and financial media. These perspectives allow us to understand the role of human behavior, regulation, incentives, and information and media. This comparison shows that there is not one single cause that explains all six cases; however, some similar patterns appear in several of them. Herd behavior and overconfidence appeared in multiple cases, while other behavioral factors are specific to each case. The cases also show problems related to regulation, incentives, enforcement, conflict of interest, and information asymmetry. The role of financial media is different across all cases. In some cases, journalists helped to uncover important information, while in others, media coverage became stronger after the financial problem was already visible. Overall, the paper finds that financial risk cannot be explained by only one factor: human behavior, incentives, and information altogether influence how different financial issues develop.

**Keywords** Financial risk, behavioral economics, law and economics, financial media, information asymmetry.

## Introduction

Financial risks are often explained through their main or most visible effects or immediate causes: plummeting asset prices, extreme debt, and bank failures (Cooter and Ulen, 1998). However, these explanations do not always fully address all the factors that determine how financial risk really begins in the first place, and why extreme financial problems continue to persist. To better understand these events, it is useful to look at how people make decisions, how regulation and incentives work, and what information is available to investors and the public.

Individuals, institutions, regulations, and information systems that influence financial decisions are ultimately behind market events and movements. From corporate fraud to bank runs to speculative trading, financial instability can emerge when behavioral biases, weaknesses in regulation, and financial media intersect.

This paper examines six major episodes of financial instability: the collapse of Enron, the 2008 failure of Lehman Brothers, the Madoff Ponzi scheme, the GameStop trading phenomenon, the collapse of Silicon Valley Bank, and the failure of FTX and Alameda. Although these cases seem heterogeneous, each case demonstrates how financial crises cannot be truly understood with traditional economics alone and whether similar patterns could appear in different types of

financial events. These cases are very different; they include investment bank failure, bank failure, speculative stock trading, corporate fraud, ponzi scheme, and cryptocurrency failure. The causes of these crises extend beyond asset prices and rapid credit growth; they are further shaped by institutional incentives, irrational behavior, lack of regulatory structures, and the degree to which the media participates in their level of reporting. Examining these factors altogether provides a more complete explanation for why financial instability occurs and why in some cases, it is particularly harder to detect.

Behavioral economics helps to look at people's decisions, and biases may contribute to financial risk. The law and economics perspective helps examine regulation, enforcement, incentives, and information asymmetry. It supports understanding of whether regulation and oversight was present in each case. Financial media is important because investors depend on public information, so it is important to see whether journalists identify the risk before the event occurs or explain the issue after the fact.

Media coverage and journalism plays a significant role in this process, as financial news does not always appeal to the public in a neutral perspective (Schiffrin, 2015). Journalists and news organizations control which financial events and risks are emphasized to the public, how explanations to these events are presented, and whose perspectives are being highlighted (Schiffrin, 2015). Political perspectives further shape how the public interprets financial events. Ultimately, examining financial crises and events must not only utilize economic perspectives, but also how journalism presents this information.

The main research question of this paper is what behavioral, regulatory, and media related patterns appear across different financial risk episodes in the United States. And, how are these factors connected? The purpose is not to find one cause that explains every case, instead the paper looks for recurring patterns while also considering important differences between six episodes. Identifying the key shared characteristics between crises in the 21st century is significant because it may improve our understanding of systemic financial risk, pushing the media and policymakers to improve information output and help stabilize regulatory systems. This can lead to less panic and misinformation when encountering market instability in the future. The comparison shows recurring patterns involving behavioral biases, information asymmetry, issues with regulation and enforcement, and differences in roles of financial journalism. However, these factors do not appear in an identical manner in each case. This shows how financial risk is better understood by looking at the interactions between behavioral, institutional, and informational aspects.

The main contribution of this paper is comparing six different financial episodes using the same three lenses: behavioral economics, law and economics, and financial media. Comparing six different financial events is useful because it demonstrates whether or not common patterns and factors contribute to different financial events. This approach is different because this comparison shows patterns such as overconfidence, herd behavior, information asymmetry, and regulatory enforcement, however these patterns are not always identical in each case, and all six cases do not share one cause. These factors have significant impact on economic

instability, and additionally have influence on each other. Human behavior increases risk, regulation can influence humans' decisions and incentives, and the media can affect what information becomes available to both investors and the public. After the analysis of each financial episode, financial risk is not always caused by one problem and looking through the perspective of several factors together can allow us to better understand financial episodes.

The remainder of this paper is organized as follows: Section 2 explains the methodology used to compare 6 financial risk episodes. Section 3 reviews the existing literature of how scholars look at financial risk episodes, followed by the behavioral economics framework, the law and economics framework, and the financial media framework. Section 4 presents six case studies of each financial episode and will be provided along with analysis of each episode through three lenses. Section 5 concludes the paper.

**Methodology**

The paper uses a comparative case study approach to examine six different financial risk episodes: Enron, Lehman Brothers, Madoff, GameStop, FTX and Alameda Research, and Silicon Valley Bank. These cases were selected because they represent very different types of financial risk. This allows us to determine whether similar patterns could be present in a varied selection of financial risk episodes. They include investment bank failure, bank failure, speculative stock trading, corporate fraud, ponzi scheme, and cryptocurrency failure. Each episode has connections to the United States, although FTX's main international exchange was headquartered in the Bahamas (Knauth and Hals, 2023). Additionally, the effects and participants of these episodes extended beyond national borders. It is significant to select cases that vary because it allows us to determine whether a pattern is only specific to a certain type of financial risk episode, or if this pattern can be generalized to all.

Additionally, each case occurs at different times, under different markets, and different regulatory conditions. Each case is examined using the same three lenses–behavioral economics, law and economics, and media. Under behavioral economics, patterns of herd behavior, overconfidence, and loss aversion were studied. Under law and economics aspects like regulation, externalities, incentives, conflicts of interest, information asymmetry, and enforcement were studied. Under the financial media lens, the role of journalism before and after each episode was analyzed. For each case, the same process was applied in order to enable successful comparison between cases. In each case, not every factor is expected to appear in every case. Behavioral or regulatory factors are only included when the evidence from that cause supports it. Similarly, the role of financial media may vary from one case to another. The paper looks for similarities and differences across six cases, as well as considering recurring patterns. If a specific factor is only found in one case, it is specific to that certain type of case. The paper considered recurring patterns when it appeared several times, although it does not have to appear in all cases. If a factor appears in only one case, it cannot be properly identified as a recurring pattern. The purpose of this method is not to show that all six episodes have the same cause, and instead is to show that they have similar or related causes.

**How Scholars Look at 'Financial Risk Episodes'**

Scholars examine financial risk episodes by focusing on several different perspectives. These perspectives include individual and market behavior, regulatory frameworks, and the influence of information. Scholars examine how financial bubbles and risk-taking grow into systemic crises due to panic, herd behavior, contagion, and leverage (Brunnermeier and Oehmke, 2012). This intersection of financial behavior and financial crisis continues to act as the backbone behind these collapses. For example, scholars like Nicolas Barberis view financial bubbles as stemming from the mental tendency of some participants in the market to extrapolate past prices too far into the future (Barberis, 2013, 4) Scholars are informed by behavioral economics and law and economics to develop different approaches to financial regulation.

A second approach focuses on the role of regulation in preventing or limiting financial instability (Chaudary and Salvador-Adebayo, 2014). Scholars disagree over how active governments should be in regulating and intervening in financial markets, and how regulation should allow both market efficiency, systemic stability, and investor protection to equally coexist (IOSCO Report, 2003). Some scholars emphasize the ability of markets to resolve problems through incentives and competition, while other scholars argue that regulation is necessary to resolve information asymmetry, conflicts of interest, or excessive risk-taking that generates problems in the market (Chang, 2022). These conflicting approaches illustrate the weakness in regulatory frameworks; gaps and failures in regulatory structures contribute to financial crises.

**Behavioral Economics Framework**

The behavioral economics lens examines how individuals' behavior influences their decision-making, and how this can contribute to financial risk. This lens does not assume that investors and individuals always act rationally, and instead examines behavioral factors like panic, contagion, herd behavior, and overconfidence. Overconfidence is a significant behavioral bias that leads investors to trade too often, underestimating portfolio risk. Scholars like Daniel Hirshleifer theorize that overconfidence, specifically where people overestimate the precision of their forecasts, contributes largely to the formation of bubbles (Kent et al., 1997).

Ho et al. (2016) suggests that manager overconfidence causes significant heterogeneity in bank risk-taking behaviors during a boom and the performance of these banks following the financial crisis (Ho et al., 2016). This perspective is informed by the fact that chief executive officers are often the most influential in the decisions that financial institutions make, making the CEO's assessment of borrower prospects crucial to the stability of the bank (Ho et al., 2016). When CEOs are overconfident, they often think they have more accurate knowledge about future events than they actually have, believing they will experience more favorable outcomes than they actually will. This overconfidence can lead to an overestimation of their ability to generate higher returns, leading to overinvestment and ultimately vulnerability in the financial institution. Ho et al. (2016) examine overconfidence using stock option-based measures developed by Campbell et al. (2011). Overall, this result suggests that overconfident participants in the market showcase more aggressive and risk-taking behavior.

Furthermore, scholars view herd behavior as a main driver of financial crises, causing widespread systemic panic and contagion (Aljifri, 2024). This behavioral bias shows how uncertainty and pressure may lead investors to adopt the behavior of others. The main causes of herd behavior stem from information asymmetry, reputational risk, and emotional responses like fear of missing out. Information asymmetry occurs when one party acquires more financial information than the other. When this gap grows larger, it causes lenders, banks, and traders to panic and ultimately causes frozen credit markets.

Loss aversion, the tendency for individuals to weigh the pain of losses more heavily than the pleasure of gains of equal magnitude, is a cognitive bias with direct consequences of market instability. A key feature of prospect theory, formalized by Kahneman and Tversky (1979), loss aversion has been incorporated into asset-pricing models to explain abrupt price movements not accompanied by news reports and contagion between unrelated assets (Ouzan, (2020). This exemplifies how loss aversion creates market instability, where the cognitive bias does not destabilize prices itself, but the uncertainty it produces in the minds of other individuals on the market that are attempting to interpret trading behavior. This uncertainty causes huge price swings in response to news when investors are uncertain about what an asset is actually worth.

Overconfidence, herd behavior, and loss aversion can cause mispriced risk. Overconfident CEOs and traders become too confident in their own judgements and believe that their information is more accurate than it truly is, inflating asset prices during booms (Kent et al., 1997). Herding causes investors to mimic the observed behavior of others, replacing their own judgment under conditions of information asymmetry (Aljifri, 2024). Loss aversion demonstrates how loss-averse traders' reluctance to realize losses and their aggressive withdrawal from the market can cause sudden and volatile market dips (Ouzan, 2020). These three biases exemplify how financial episodes can become the product of interaction between these three behavioral biases.

**Law and Economics Framework**

The law and economics lens examines how federal regulatory frameworks support financial markets, and how the lack of them can lead to instability and discrepancies in financial markets.

Traditional law and economics often assumes that individuals are rational actors who have predictable responses. However, financial crises that happen before proper detection show how investors, executives, and financial institutions do not always act rationally.
Law and economics examines the incentives that laws create, and whether or not these incentives lead individuals and companies toward beneficial or harmful behavior.

The modern law and economics framework was developed through the work of scholars such as Ronald Coase and Richard Posner, who helped connect legal rules with economic principles. Coase (2013) additionally demonstrated that the economic consequences of laws and government regulation could be analyzed through the lens of incentives, costs, and externalities. Rather than only analyzing whether or not a law was ethically right or wrong, the law and

economics framework provided an additional perspective by challenging how different laws change people's incentives, which ultimately impact economic outcomes (Coase, 2013).

An important concept in law and economics is negative externalities. In financial markets, individuals may participate in risky behavior because while they receive benefits, the resulting consequences transfer to investors, employees, or the general economy (Kenton, 2026). This outcome can be addressed through regulation acknowledging and ensuring that firms internalize these consequences through penalties, liability, and other forms of stricter oversight. Laws that incentivise negative externalities often do so because regulators assume rational behavior from participants in the markets.

This issue is especially seen in the Enron scandal, as executives had strong financial incentives to maintain the appearance of successful growth and profit. Due to a lack of oversight, executives were allowed to misrepresent the company's true financial conditions and conceal debt, harming investors and clients (Segal, 2026). This was ultimately caused by an insufficient institutional structure that failed to create incentives to prevent risky behavior. Enron demonstrates how the existence of regulation alone is not sufficient to prevent financial crises; the enforcement and magnitude of legal consequences ultimately determine whether regulation actually changes economic incentives.

**Media Role Framework**

The media role framework analyzes how financial journalists report financial risk and whether they can warn the public before a crisis occurs. Schiffrin (2015) uses the housing bubble in 2008 to exemplify how financial media failed to warn viewers and identify abuse and risky practices within the financial market, causing a vulnerable outcome (Schiffrin, 2015). Because the media took on an overly optimistic view and relied on information from financial institutions and the government, they did not properly inform people of the instability within the housing market (Schiffrin, 2015). This is ultimately caused by the lack of investigation and challenging information within financial journalism. Additionally, during the housing bubble in 2008, the media failed to cover the long term impact, leading to increased uncertainty and the public not gaining a deeper understanding and lesson from this crisis (Schiffrin, 2015).

There is additionally a lack of financial incentive for journalists to be more proactive in investigating information, as those who are more skilled instead work in the financial sector to thrive financially (Stiglitz, 2015: 122). Gillian Doyle argues that financial journalists, in general, are good at analyzing and interpreting companies; however, they lack the initiative to investigate and challenge stories, which is due to the lack of time, opportunity, and training (Doyle, Gillian, 2015: 122). This connects to the regulatory framework applying to financial journalism, as financial news significantly influences market behavior, yet the system of journalism does not incentivize journalists to provide the public with more accurate and investigative reporting. Additionally, journalists often rely on CEOs and government officials; however, if they criticize and challenge these sources, they may risk losing access to crucial information that shapes their

reports (Stiglitz, 2015: 25). Additionally, journalists often work under limited time spans, resulting in less time to accurately investigate financial events.

This lens additionally focuses on information asymmetry, where a bank knows much more information than a journalist, putting them at a disadvantage. This can also lead to cognitive capture, leading journalists to stop asking more difficult and investigative questions. Stiglitz (2015) notes that some of the asymmetry can be reduced through improved training in economics and independent funding of investigative stories (Stiglitz, 2015).

Based on this understanding of how scholars look at these three perspectives and frameworks, this paper will now use a comparative and theory-driven approach, applying each of the three analytic lenses–behavioral economics, law and economics, and media–across a set of six modern financial risk episodes.

## Six Analytical Case Studies

### Silicon Valley Bank

Silicon Valley Bank was founded in 1983, becoming the largest bank by deposits in Silicon Valley, serving technology startups, venture capital firms, and life sciences companies (Shefrin, 2023). The majority of its clients were tech startups, creating a concentration of money from one sector, especially when the technology sector experienced unforeseen growth, causing the deposits at SVB to triple between 2019 and 2021 (Ramasastry, 2023). Because rising interest rates reduced the market value of its bond investments, these technology startups began withdrawing their deposits at SVB, as venture capital funding was hindered (Ramasastry, 2023). However, because SVB had invested that money, they were unable to meet these withdrawals. As a result, SVB sold securities it already owned on the open market, leading to increased panic among its customers (Shefrin, 2023). When the bank announced that it had sold approximately $21 billion of securities at a $1.8 billion loss and planned to raise additional capital, this only further caused widespread panic among depositors. When companies withdrew their funds, this led to a bank run, leading to the closure of the bank on March 10, 2023 (Shefrin, 2023).

#### *Behavioral Economics Lens*

SVB was exposed to duration risk because it had a large amount of long-term securities whose market would decline when interest rates increased. However, the failure of SVB illustrated how deposits can be a significant source of bank runs.

According to Choi et al. (2024), when venture capital firms began advising startups to withdraw their money, depositors, without independently evaluating SVB's financial condition, began to follow this behavior Panic spread through several social networks, where the actions of others became an incentive to withdraw money, deeming it as the safest option. Additionally, behavioral economics shows that people prefer avoiding losses over equivalent gains, aligning with how SVB customers experienced loss aversion. Because many SVB customers had deposits exceeding the FDIC insurance limit, depositors, despite having uncertainty about whether

deposits would actually be lost, preferred avoiding the possibility of losing this extreme amount of money. This cascading effect of withdrawals caused by herd behavior and fear ultimately led SVB to fail. Instead of following their own analysis, people assumed from early withdrawals that something must be wrong, believing that earlier depositors had comparatively better information. This caused information cascades, as people tried to imitate others instead of relying on their own assessments. Additionally, the 2008 financial crisis caused people to overestimate the likelihood that SVB would fail, causing them to all withdraw in panic. This panic was also further pushed when rumors about SVB's financial instability were spread, causing depositors to search for confirming information. Because positive information received less attention, news relating to bond losses and capital raising supported these existing beliefs, as contradicting evidence was ignored (Choi et al., 2024).

***Law and Economics Lens***

According to Casey (2025), Because SVB served venture capital firms and technology firms, it experienced significant growth, growing from $45 billion in 2017 to over $200 billion in March 2022. SVB did not keep these funds liquid, however, instead investing in long-term government securities with long maturities. When interest rates increased in 2022, bond values fell, venture capital funding slowed, causing SVB to face declining asset values and major deposit withdrawals.

The governmental response to this crisis was the Treasury, Federal Reserve, and FDIC invoking the Systemic Risk Exception, guaranteeing all deposits, and protecting even uninsured depositors. The government also stated that taxpayers would not bear the losses, as the Deposit Insurance Fund would cover it. FDIC only protects deposits up to the insured limit; however, the systemic risk exception allows regulators to bypass these limits if failing to do so will cause financial instability.

However, the main concern is moral hazard, as if depositors are protected, they will stop paying attention to whether or not banks are taking excessive risks. Additionally, it will cause banks to have less incentive to manage their risk cautiously (Casey, 2025). The failure of SVB additionally provided the first post-crisis test of the United States legal regime for dealing with major bank failures.

***Media Lens***

A significant factor in the crisis and failure of Silicon Valley Bank was instability in communication. Firstly, Silicon Valley Bank underestimated the sensitivity that depositors would have in reaction to the news, especially in the context of previous financial failures. The critical news was announced at a time when other crypto-centered and FinTech institutions were collapsing. Additionally, this announcement was shared on the same day that Silvergate, the largest crypto bank, declared that it was liquidating (ISOC, 2023). Despite these circumstances already causing panic among depositors, SVB failed to consider how their announcement would be received by others, proceeding without strategy or caution (ABA Banking Journal, 2025).

Additionally, when CEO Greg Becker used a Zoom call, intending to further elaborate and give reassurance, the language used only further generated concerns and failed to reassure their audience. Additionally, Silicon Valley Bank chose to make a single announcement, deciding not to interact or enable any coordinated communication around their announcement by lacking interactive communication or background briefings (ISOC, 2023). SVB's statement was described as “among the worst ever written”, as it not only failed to provide both context and reassurance, but its language was also only targeted towards financial underwriters, unable to be properly understood by the bank’s customers and the media (ISOC, 2023). Because Silicon Valley Bank was unable to present a clear narrative, damaging and twisting narratives were released to the public and the media, becoming what SVB’s customers saw as the truth.

**Enron**

According to Troy Segal (2026), Enron, founded in 1985, was once one of the largest energy companies in the United States. Originally, Enron was formed following a merger between Houston Natural Gas and InterNorth but later expanded into energy trading and other financial markets. Enron used an untraditional, mark-to-market (MTM) accounting method, where MTM measures the fair value of accounts that can change over time, like assets and liabilities. MTM’s purpose is to provide an accurate appraisal of a company’s current financial conditions; however, this method can be manipulated. Investors, analysts, and the media praised the company for its rapid growth and financial performance, causing its stock price to rise rapidly and also enticing trading partners. After partnering with Blockbuster to enter the video-on-demand market, Enron began logging expected earnings based on the estimated growth of the market, inflating the numbers. By mid-2000, Enron Online was executing $350 billion in trades. By building telecom networks when the dotcom bubble began to burst, Enron had gained access to the most volatile parts of the market, causing investors to face decreasing market capitalization (Segal, 2026).

Additionally, through creating Special Purpose Entities, Enron was able to hide billions of dollars in debt and inflate reported profits. This manipulated method of Mark-to-market accounting allowed Enron to record future projected profits that had not been realized, allowing CEO Jeffrey Skilling and Chairman Kenneth Lay to continue promoting strong financial health despite knowing that these statements were false and misleading. In 2001, investigative journalists and financial analysts began to see these inconsistencies in Enron’s financial reports. The company’s stock price collapsed from more than $90 per share to less than $1. Ultimately, Enron filed for bankruptcy on December 2, 2001, resulting in thousands of jobs being lost (Segal, 2026).

***Behavioral Economics***

The Enron scandal demonstrates how behavioral biases and corruption can lead to corporate fraud. The Enron scandal strongly illustrates how overconfidence bias can distort decision-making.

Tang et al. (2024) introduce the concept of the "Enron Effect", which examines how an individual's desire for money can interact with their corporate ethical values, ultimately reducing their moral and ethical standards and exhibiting high dishonesty. This study used 6,382 managers across 31 geopolitical entities and found a positive relationship between the love of money and dishonest behavior (Tang et al., 2024).

This framework explains how Enron's incentive structure could have encouraged unethical behavior. Because monetary rewards became the main measure of success, employees may have valued personal compensation and profitability over ethical and moral standards. The "Enron Effect" shows that people motivated by wealth, power, and financial status are more susceptible to unethical and irrational behavior (Tang et al., 2018).

Additionally, the Enron scandal illustrates how corporate ethical values shape individual behavior. Tang et al. argue that the ethical values of a corporation, often established by senior executives, strongly influence how managers behave. Their research found that managers working in environments of weak ethical values and corruption exhibited significantly more dishonest intentions. This demonstrates how Enron's misconduct is caused not only by the individual motivations of executives, but the organizational environment could have reinforced irrational and unethical decision-making (Tang et al., 2018).

***Law and Economics***

A systemic cause of the Enron scandal was the legal and regulatory system that allowed for this to happen. A major legal failure involved Arthur Andersen, Enron's independent auditing firm. Instead of acting as a monitor, Arthur Andersen earned consulting fees from Enron, weakening the firm's incentive to actually challenge questionable accounting practices (Velasquez, 2002). Current laws and SEC regulations allowed Arthur Andersen to provide consulting services to Enron and provide the audited report about the financial results of the consulting activities (Velasquez, 2002). Arthur Andersen approved and allowed Enron to use special purpose entities to hide billions in debt and inflate profits, ignored warning signs, and additionally shredded audit files once SEC investigations began (Velasquez, 2002). Additionally, most large companies have the ability to manage their own employee pension funds (Whiteside, 2025). This creates a conflict of interest within the legal system, as the companies can use these funds that benefit the company even if they disadvantage employees. This scandal ultimately called for significant legal reform through the Sarbanes-Oxley Act of 2002 (SOX). SOX increased executive accountability by requiring CEOs and CFOs to personally verify financial statements, making them responsible for their accuracy. This law also aimed to strengthen auditor independence, leading to the creation of the Public Company Accounting Oversight Board, which oversees accounting firms, increased internal control requirements, and imposed harsher punishments for securities fraud (Narayanan, 2025). This especially targets the problem of agency abuse, where those in power tend to pursue their own interests instead of protecting and being honest with their clients.

### *Media Lens*

Media and journalism played a significant role in the Enron scandal. Initially, many financial news outlets portrayed Enron as a successful and innovative company (Steiger, 2002). This positive media coverage reinforced investor confidence and supported the company's reputation, making it seem even more trustworthy. However, as Enron's financial statements became more complex, journalists began challenging and investigating the legitimacy behind their statements. Specifically, two Wall Street journalists, Rebecca Smith and John Emshwiller, uncovered the scandal; however, this reporting was largely ignored by many of the nation's new media organizations (Steiger, 2002).

This scandal also demonstrates the significant role of whistleblowers in combination with investigative journalism. Former vice president of corporate development Sherron Watkins warned CEO Kenneth Lay about the company's fraudulent accounting practices in August of 2001 (Steiger, 2002). Although her warnings were initially ignored, media coverage further highlighted the importance of these warnings after Enron's bankruptcy. This collaboration between whistleblowers and journalists further increased the public's awareness and the pressure on regulators to investigate (Steiger, 2002).

## Lehman Brothers

Lehman Brothers, founded in 1850, was one of the older and largest investment banks in the United States (Lioudis, 2026). The firm was involved in corporate mergers, the sale and acquisition of companies, and investment portfolios for individuals, foundations, endowments, and pension funds (Lioudis, 2026). The Lehman Brothers' ability to find investment opportunities greatly contributed to the economic growth of the country, job opportunities, and the standard of living (Lioudis, 2026). In the early 2000s, the firm became increasingly involved in mortgage lending, mortgage-backed securities, and complex financial derivatives (McDonald and Robinson, 2010). Because housing prices were rising, financial institutions assumed that mortgage-related investments carried relatively low risk. However, due to the adoption of an aggressive growth strategy, excessive borrowing, and a risk-taking business model supported by limited equity, Lehman Brothers was highly leveraged, meaning relatively small declines in asset values could eliminate its capital (Bakkar, 2023). Therefore, when the U.S. housing market began declining in 2007, mortgage defaults increased rapidly, causing the value of mortgage-backed securities to decrease significantly, leading to a loss of billions of dollars and failure to secure additional financing (Bakkar, 2023). When no buyer agreed to acquire Lehman without government financial support, Lehman Brothers filed for bankruptcy in September of 2008, making it the largest bankruptcy in the global financial markets (Bakkar, 2023). This failure caused widespread panic, shown in frozen credit markets, stock markets rapidly declined, and financial institutions experiencing liquidity shortages.

### *Behavioral Economics Lens*

The collapse of Lehman Brothers exemplifies how behavioral biases among executives and investors contributed to a significant financial crisis in 2008. A significant bias central to the collapse was overconfidence bias, where CEO Richard Fuld and top executives chose to ignore warning signs of the collapsing housing market, believing that the U.S. housing market would continue to appreciate, trusting inaccurate financial risk models that assumed real estate growth (Bakkar, 2023). Due to this overconfidence, they felt encouraged to increase their investments in mortgage-backed securities, ignoring the growing evidence that the housing market was weakening (Bakkar, 2023).

Another behavioral factor driving this collapse was the sunk cost fallacy (Ahmed, 2011). Executives did not offload declining MBS assets, as they feared reputational and financial loss. They assumed that they would be able to manage downturns and recover losses, showing an unrealistic expectation in their ability to control future outcomes.

Additionally, herd behavior further drove the collapse. Throughout the early 2000s, investment banks, commercial banks, and investors all chose to participate in the rapidly expanding mortgage lending and mortgage-backed securities (Bakkar, 2023). Before almost all major financial institutions were succeeding, financial institutions felt pressure to copy each other's high-risk strategies during the boom, fearing that they would risk falling behind competitors (Ahmed, 2011). This pushed Lehman Brothers to increase its leverage, additionally because of competing firms. Some firms did not independently evaluate the risks, instead following industry trends. This herding behavior was seen again as the housing bubble led to panic within institutions, causing them to sell once market confidence broke.

***Law and Economics***

Legal rules and incentives played a large role in shaping economic behavior during and prior to the Lehman Brothers' collapse. Within the period of 2000 to 2008, incentive structures allowed Lehman Brothers' top executives to receive bonuses based on short-term profits rather than long-term stability. It is estimated that Lehman Brothers derived cash flows of about $1 billion from cash bonuses and equity sales (Bebchuk et al., 2010). This compensation promoted excessive leverage and risk-taking, as they would invest in risky mortgage-backed securities. This incentive structure is flawed, as it was not aligned with shareholders' and society's long-term expectations and needs (Bebchuk et al., 2010).

Additionally, Lehman Brothers' financial conditions were unclear to investors and regulators, who lacked accurate information. The lack of visibility and accuracy was achieved through the use of Repo 105 transactions, enabling Lehman to remove up to $50 billion from its balance sheet for short periods of time and to pay unrelated debts, making leverage appear lower than it actually was. In its Forms 10-K and 10-Q reports, Lehman stated that it sometimes used repos for short-term borrowings that it accounted for as financings. However, Lehman never disclosed its use of repos 105s or that it accounted for any of its repo transactions as "sales". This ultimately caused reduced market efficiency because decisions were made with inaccurate information.

### *Media Lens*

The media played a complex role during the Lehman Brothers' collapse. Before their collapse, the media focused on rising home prices, expanding credit markets, and corporate profits. Because the media did not investigate or further challenge this housing market success and strong financial-sector earnings, this only encouraged optimism and herd behavior (Gold, 2018). Investors who repeatedly heard positive news were more likely to believe that housing prices would continue increasing.

However, as evidence of the Lehman Brothers' financial instability began to emerge, it prompted investigative journalists to report on rising mortgage defaults, declining home prices, and the vulnerability that major financial institutions would face. After the collapse, extensive media coverage followed, educating the public about complex financial instruments, excessive leverage, and weaknesses in financial regulation. However, this onslaught of media coverage, along with the collapse, caused people to question why the media was unable to protect investors and warn them about the housing market collapse. Optimistic news coverage had contributed to unrealistic market expectations that influenced investor decisions. These discrepancies caused journalists to realize their investigative responsibilities (Gold, 2018).

## Madoff Ponzi

Bernard L. Madoff was a highly respected figure on Wall Street as the chair of the Nasdaq exchange for several years. Madoff attracted many investors by claiming to generate stable returns through a technique called split-strike conversion, a trading approach that gave Madoff credibility to his operation. However, regulators later found that client funds were not invested as Madoff presented, and account activity was misreported to maintain investor confidence (Nunez, 2025). Madoff was operating a Ponzi scheme, where Madoff did not invest clients' money as promised, instead using funds from new investors to pay returns to previous investors. This Ponzi scheme relied on a constant flow of new investments to pay earlier investors and maintain the appearance of legitimacy.(Nunez, 2025). Despite Harry Markopolos' repeated warnings and complaints to the SEC, arguing that his returns were mathematically impossible, Madoff's scheme remained undiscovered (Nunez, 2025). During the financial crisis in 2008, Madoff could no longer fulfill his client redemption requests. After confessing to his sons about his scheme, he was arrested in December 2008.

Investigators estimate that investors lost approximately $65 billion in reported account values and around $17-20 billion in principal investments. Madoff pleaded guilty in 2009 and was sentenced to 150 years in prison (Gross, 2024).

### *Behavioral Lens*

Psychological biases were highly influential towards both investors and Madoff himself during the Madoff scandal. Authority bias was central to Madoff's initial success. Madoff was highly respected as a Wall Street figure and a former NASDAQ chairman, leading investors to assume that someone with his experience and reputation must be trustworthy (Roncal, 2008).

Madoff's status ultimately led investors not to do independent research or verify his investment strategies, and instead solely relied on his status. His status ultimately created a self-sustaining cycle. This behavior displayed how investors relied on reputation rather than doing independent verification.

Additionally, Madoff gained further success and attention from other investors due to herd behavior, where investors joined because friends, celebrities, and wealthy individuals also invested (Roncal, 2008). Due to herd behavior, Madoff attracted a large number of investors, giving him more credibility. The discrepancies and warning signs of Madoff's investment strategy were further ignored due to confirmation bias and overconfidence; investors consistently ignored how Madoff claimed positive returns every year regardless of market conditions and how he concealed his strategies (Nunez, 2025). This ignorance was additionally due to investors believing that they were capable of identifying an accurate and safe investment strategy, reducing skepticism and independent research. These experienced investors and financial institutions trusted reputation and status instead of their own knowledge and experience, therefore failing to acknowledge these warning signs.

Ultimately, the Madoff scandal demonstrates how financial fraud develops due to predictable human behavior that distorts proper decision-making and evaluation.

***Law and Economics***

Regulatory inconsistencies and insufficient funding contributed heavily to the Madoff scandal, enabling Madoff to pursue the Ponzi scheme without consequences until the collapse (Langevoort, 2009). During the Ponzi scheme, the SEC received multiple complaints, including several from Harry Markopolos, which consisted of repeated warnings that Madoff's returns were mathematically impossible (Nunez, 2025). However, when his warnings did not lead to any investigation, this demonstrated the huge issue of a lack of incentive for investigations and not listening to whistleblowers. SEC's lack of agency and enforcement of critical rules, due to operating under low budget and staffing limitations, allowed the financial fraud to continue (Langevoort, 2009).

Additionally, Madoff's agency allowed for information asymmetry, where he exploited the distance between regulators and opaque back-office operations, ultimately acting for his own benefit. In doing so, fake account statements were created to cover the Ponzi scheme (Nunez, 2025). Therefore, investors lacked access to accurate financial information and transparency, reducing market efficiency.

The lack of regulatory incentives to investigate suspicious returns and investment strategies in the SEC led to the Dodd-Frank Act (SEC, 2019). The Dodd-Frank Act strengthened financial investigation by creating new oversight bodies, restrictions on risky bank investments, and protecting everyday borrowers (SEC, 2019).

***Media Lens***

According to Gross (2024), in this scandal, the media initially emphasized and promoted Madoff's reputation, illustrating him as successful, experienced, and a respected Wall Street investor. Because the majority of the public relied on and trusted the media for financial news and information, this only further solidified Madoff's reputation. Instead of investigating and challenging his investment behavior and unusually consistent returns, the media did not question it, allowing the Ponzi scheme to continue. However, after the collapse, the media played a significant role in exposing and covering the fraud, bringing awareness and attention to the regulatory failures and major losses (Gross, 2024).

While the media failed to warn the public, it nonetheless played a major role after the collapse in bringing attention to the scandal.

**GameStop Trading Surge**

The GameStop stock surge was a significant stock market event, occurring in 2021. GameStop was a struggling video game retailer facing declining sales due to the increasing prevalence of digital downloads, becoming one of the heavily shorted stocks on Wall Street (Dziak, 2021). Before the stock surge, many investors had already been short selling GameStop shares, expecting the stock to decline. However, members of the Reddit community "r/wallstreetbets" began purchasing shares and also encouraged thousands of retail investors to buy GameStop shares as well. This caused the stock price to increase rapidly, forcing short sellers to buy shares in order to minimize their losses. This pressure to buy created a short squeeze, pushing prices even higher. Over two weeks, GameStop's stock increased dramatically, causing some to profit immensely, while others, like Melvin Capital, lost billions of dollars (Dziak, 2021). This event shows the power of individual investors and social media, showing how their behavior and influence can cause dramatic changes in the financial market.

***Behavioral Economics Lens***

The GameStop stock surge exemplifies how herd behavior can cause a financial collapse within days. Encouraged by discussion on Reddit, millions of retail investors bought GameStop shares. Consequently, those motivated by fear of missing out joined in, buying the shares as the stock price rapidly increased, fearing to miss out on huge profits. Additionally, overconfidence pushed people to take more risks, believing they could "beat" the market without having sophisticated education about financial markets and not being traders for a living (Quinlan, 2021). Additionally, online platforms and tools like Robinhood democratize access to stock trading; however, these tools presume that people who utilize them are able to make educated decisions. While it makes public access to trading and the market more open and easier, it also allows those who trust unreliable sources like social media to access the market. In some situations, easy access to trading and rapid spread of information on social media contribute to herding behavior and increased market volatility.

***Law and Economics***

While journalists and financial media have had a federal regulatory framework created since the 2008-2009 financial crisis, controlling how they educate viewers, social media lacks this regulation. The federal regulations on journalism platforms allow the public to view their reports as analysis, and not as a recommendation. However, there is no regulatory framework to emphasize this in social media. Challenges of regulation and enforcement distorts how viewers are able to differentiate between trustworthy and untrustworthy information, causing them to start chains of herding behavior. (Diwya and Gandhi, 2026).

### *Media Lens*

While the media was able to report how the GameStop stock surge unfolded, it was unable to actively investigate it and warn the public before the collapse happened due to temporal constraints. This reveals the limitations of the capacity that journalists have when reporting events that unfold in a matter of days. Despite only having a role in this case after the fact, the media was able to identify responsible factors as they reported what happened, serving as a lesson for the future (Niccum, 2024).

## FTX-Alameda SBF

FTX was founded in 2019 by Sam Bankman-Fried, quickly becoming a large and successful cryptocurrency exchange. FTX's success created a $2 billion venture fund that funded other crypto firms. He also founded Alameda Research, a quantitative cryptocurrency trading firm. Bankman-Fried's wealth grew to over $16 million in March (Sigalos, 2022). When crypto prices dramatically decreased, he marketed his enterprise as immune to financial collapse. However, FTX customer funds were also being used by Alameda Research, for its trading activities, investments, and other obligations. The internal controls were weak, financial reporting was inaccurate and lacked transparency, and there was a lack of corporate governance (Sigalos, 2022).

In November 2022, a balance sheet exposed the financial conditions of Alameda. Because of this, Binance announced it would sell its FTT holdings, a cryptocurrency token created by FTX. This caused widespread panic throughout the cryptocurrency market, and customers rapidly withdrew billions of dollars from FTX. Because there was a lack of liquid assets to meet these withdrawals, FTX filed for bankruptcy on November 11, 2022 (Sigalos, 2022).

### *Behavioral Economics Lens*

Behavioral economics, specifically in Sam Bankman-Fried, shaped the rise and collapse of FTX and Alameda. Sam Bankman-Fried's behavior may have exemplified overconfidence, shown in his decision not to want a board of directors for FTX (Shefrin, 2023). Bankman-Fried was overconfident about his ability to serve as CFO, putting both his company and his customers in jeopardy (Shefrin, 2023). Additionally, high ambition pushed for risky behavior, as Bankman-Fried felt the need to succeed, always aspiring for more.

Similar to the Madoff scandal case, people were lured to Sam Bankman-Fried's cryptocurrency exchange trading firm due to his celebrity status (Sigalos, 2022). Again, this status was valued by the public, which may have encouraged people to not research his financial behavior. This promoted increasingly risky behavior and ultimately allowed the public to face the losses of FTX's bankruptcy. Herd behavior further encouraged people to participate, leading to widespread participation.

### *Law and Economics Lens*

According to Shefrin (2023), during and preceding the collapse of FTX and Alameda, the emerging market of cryptocurrency had no comprehensive federal regulation, especially regarding offshore exchanges and digital asset trading firms like Alameda. Because crypto developed at a quicker pace than the legal regulatory framework, FTX was able to operate through offshore entities, making regulation and enforcement difficult. This lack of regulation is also caused by a clash of cultures, where existing financial laws couldn't easily and appropriately fit in with new types of digital assets and offshore exchanges. This lack of regulation allowed the FTX and Alameda executives to operate and control the market without supervision. Additionally, users could trade freely, but lacked both legal protection and insurance (Shefrin, 2023). Additionally, there was a conflict of interest, as FTX was closely connected to Alameda, leading it to favor Alameda over its customers. Similar to the Madoff case, there was a significant amount of information asymmetry, where executives knew significantly more about FTX's finances and connection with Alameda than customers and investors did (Shefrin, 2023).

### *Media Lens*

According to McCreesh (2022), similar to FTX and Alameda's customers, journalists and the media also fell for Sam Bankman-Fried's reputation. This automatic trust towards SBF led to a lack of investigation during FTX and Alameda's success, especially when crypto prices dramatically decreased. The media played a role in strengthening SBF's reputation, validating his claims that his company was immune to financial collapse. This further gave him additional credibility and legitimacy, leading investors to be less skeptical of FTX. While several journalists did investigate FTX and Alameda, stories about complex financial risks received less attention than reports of success and rapid wealth (McCreesh, 2022).

Like the Madoff case, some media coverage ended up playing a major role in exposing FTX and Alameda, raising questions about Alameda's balance sheet and how they manipulated their connection to FTX by relying on FTX's FTT token. As the media exposed FTX, this caused widespread panic, where customers rushed to withdraw their funds.

## Conclusion

This paper compares six different financial risk episodes to understand whether they had common behavioral, regulatory, and media related patterns. The paper did not find one single

cause that explained all six cases. Despite cases happening for different reasons, some similar patterns appeared across several of them.

One important finding was the role of human behavior. Overconfidence and herd behavior appeared in several cases, and sometimes it encouraged people to participate in risk-taking behavior and follow the decisions of others. However, some behavior factors did not appear in every case. Ultimately, these behavioral explanations cannot be used for every financial problem. Additionally, a behavioral commonality found across these financial episodes was ambition and the desire for profit. Although greed did not ultimately cause each crisis, risky behavior that prioritized profit was influenced by financial incentives. An example of this is how Bernie Madoff and Sam Bankman-Fried's Alameda acted in this way with the intention to deceive that they were profiting until they were finally exposed. In the Enron scandal, executive compensation and incentives to report successful and stronger earnings contributed to the motives behind the accounting scandal. For Lehman Brothers, the company was motivated by the pursuit of returns through leveraged investments, ultimately increasing their exposure to financial risk. The GameStop stock bubble demonstrates how some investors were motivated by profit gain from rising share prices, while other investors wanted to oppose hedge funds. The Silicon Valley Bank episode involved prioritizing short-term profits, while depositors later withdrew funds to protect their money.

Regulation was also important. The problem was not always lack of regulation. In some cases, problems with enforcement, oversight, incentives, conflicts of interest, and access to information were also important. Information asymmetry was especially significant because investors and the public did not always have the same information as company executives or insiders. New regulations and laws enforced by the government should be designed through the law and economics framework; they should consider what kinds of incentives these laws will create and ensure that they will not create bad incentives.

The media also had a different role across all cases. In some cases, journalists had uncovered important information, while in others, media attention became much stronger after the financial problem was serious. Cases such as Enron, Silicon Valley Bank, and Madoff also show that warning signs can exist before the financial problem becomes visible, but these warnings may not always receive enough attention.

Overall, financial problems are too complicated to explain through only one factor. Human behavior, regulation, incentives, and information can affect each other. For example, stronger enforcement, better oversight, clearer information, and more attention to warning signs may reduce financial risk. However, there is no single solution to every financial problem.

This study still has some limitations. This paper focuses on six financial episodes in the United States, and is not representative of all existing financial episodes. Additionally, these results can vary when studying episodes in different countries other than the United States, as they vary due to different government systems, federal regulatory frameworks, and investor behavior.

Future studies can further expand this research by examining a larger number of financial crises across different countries through different economic lenses. Researchers can also compare the specific effects of different behavioral biases in order to identify which factor influences the overall trajectory of a financial crisis. Additionally, scholars could examine the effectiveness of implementing different regulatory systems to better understand which regulatory approaches will be most successful in preventing future financial instability.

**Works Cited**


ABA Banking Journal Staff. “New York Fed Study Examines Media Role in 2023 Bank Failures.” ABA Banking Journal, 3 Oct. 2025, bankingjournal.aba.com/2025/10/new-york-fed-study-examines-media-role-in-2023-bank-failures/. Accessed 15 Oct. 2026.

Aljifri, Ruqayya. "Herding during different types of crises: The COVID-19 health crisis and Russia–Ukraine political crisis." Heliyon 10.10 (2024).

Ahmed, Arif. “Fall of Lehman Brothers - reasons why the failure could not be stopped” Working Paper, August, 2011. https://www.scribd.com/document/303695848/Lehman-Brothers-Analysis-of-Failure.

Barberis, Nicholas C. "Thirty years of prospect theory in economics: A review and assessment." Journal of economic perspectives 27.1 (2013): 173-196.

Bakkar, Yassine. “Why Did Lehman Brothers Fail?” Economics Observatory, 28 Sept. 2023, www.economicsobservatory.com/why-did-lehman-brothers-fail. Accessed 15 Oct. 2026.

Bebchuk, Lucian A., Alma Cohen, and Holger Spamann. "The wages of failure: Executive compensation at Bear Stearns and Lehman 2000-2008." Yale J. on Reg. 27 (2010): 257.

Brunnermeier, Markus K., and Martin Oehmke. "Bubbles, Financial Crises, and Systemic Risk: A Literature Review." (2012).

Campbell, T. Colin, et al. "CEO optimism and forced turnover." Journal of Financial Economics 101.3 (2011): 695-712.

Casey, Anthony. "Silicon Valley Bank: A Case Study in Post-Crisis Bank Failures." Banking and Finance Law Review, 41.2 (2025).

Chang, Chih-Hsiung. "Examining the Role of Government Intervention in Market Failure and Government Failure." Journal of Applied Business & Economics 24.5 (2022).

Chaudary, Samra, and Daisy Salvador-Adebayo. "Why Regulate Financial Markets The Underlying Rationale for Financial Regulation in the Wake of the Current Crisis." Journal of European Studies 30.1 (2014).

Choi, Dong Beom, Paul Goldsmith-Pinkham, and Tanju Yorulmazer. Contagion effects of the Silicon Valley Bank run. No. w31772. National Bureau of Economic Research, 2023.

Coase, Ronald Harry. "The problem of social cost." The journal of Law and Economics 56.4 (2013): 837-877.

Cooter, Robert, and Thomas Ulen. "Law and economics." (1988).

Daniel, Kent D., David Hirshleifer, and Avanidhar Subrahmanyam. "A theory of overconfidence, self-attribution, and security market under-and over-reactions." Self-Attribution, and Security Market Under-and Over-Reactions (February 19, 1997) (1997).

Dizikes, Peter. "How the debt crisis of 2008–09 fueled populist politics." (2022).

Diwya, Ketut Gede Sri, and I. Kadek Mahesa Parwata Gandhi. "Do Financial Influencers Help or Harm Young Investors? A Dual-Process Model of Financial Literacy, Herd Behaviour, and Investment Intention." Journal of Risk and Financial Management 19.10 (2026): 770.

Doyle, Gillian. “What are Financial Journalists For?”, “The media and the crisis: an information theoretic approach”, Edited by Joseph Stiglitz, Routledge, 2014. 168-180.

Dziak, Mark. “GameStop Short Squeeze | Business and Management | Research Starters | EBSCOhost.” Ebsco, 2021, www.ebsco.com/research-starters/business-and-management/gamestop-short-squeeze/. Accessed 15 Oct. 2026.

FSB. “The Financial Stability Implications of Artificial Intelligence – Financial Stability Board.” Fsb.Org, 2023, www.fsb.org/2024/11/the-financial-stability-implications-of-artificial-intelligence/. Accessed 8 Oct. 2026.

Gold, Howard. “Ten Years After the Financial Crisis, Business Journalism Awaits Its Reckoning.” Columbia Journalism Review, 17 Sept. 2018, www.cjr.org/business_of_news/ten-years-financial-crisis-business-journalism.php. Accessed 15 Oct. 2026.

Gross, Terry. “‘Madoff’ Takes Account of the Biggest Financial Ponzi Scheme in History.” NPR, 10 July 2024, www.npr.org/2024/07/10/g-s1-9114/madoff-takes-account-of-the-biggest-financial-ponzi-scheme-in-history. Accessed 15 Oct. 2026.

Ho, Po-Hsin, CW Huang, CY Lin, JF Yen. "CEO overconfidence and financial crisis: Evidence from bank lending and leverage." Journal of Financial Economics 120.1 (2016): 194-209.

ISOC. “Case Study: SVB -- How Bad Communication Fuelled a Major Bank Failure.” Www.Isoc.Com, 2023, www.isoc.com/blog/SVB-crisis-case-study. Accessed 15 Oct. 2026.

Kahneman, Daniel, and Amos Tversky. "Prospect theory: An analysis of decision under risk." Handbook of the fundamentals of financial decision making: Part I. 2013. 99-127.

Kenton, Will. “Externality: What It Means in Economics, With Positive and Negative Examples.” Investopedia, 18 June 2024, www.investopedia.com/terms/e/externality.asp. Accessed 12 Oct. 2026.

Knauth, Dietrich, and Tom Hals. “Failed Crypto Exchange FTX Has Recovered Over $5 Bln, Attorney Says.” Reuters, 11 Jan. 2023, www.reuters.com/business/finance/ftx-seeks-court-rulings-asset-sales-customer-privacy-2023-01-11/. Accessed 8 Oct. 2026.

Langevoort, Donald. “‘The SEC and the Madoff Scandal: Three Narratives in Search of a Story’ by Donald C. Langevoort.” Georgetown.Edu, 2025, scholarship.law.georgetown.edu/facpub/617/. Accessed 15 Oct. 2026.

Lioudis, Nick. “The Collapse of Lehman Brothers: A Case Study.” Investopedia, 9 Dec. 2024, www.investopedia.com/articles/economics/09/lehman-brothers-collapse.asp. Accessed 15 Oct. 2026.

McDonald, Lawrence G., and Patrick Robinson. A colossal failure of common sense: The inside story of the collapse of Lehman Brothers. Crown Currency, 2010.

Mendez, Katzumi . “UTA Professor Examines How AI Outages Affect Financial Market Behavior.” UTA Professor Examines How AI Outages Affect Financial Market Behavior - College of Business - The University of Texas at Arlington, 2026, www.uta.edu/academics/schools-colleges/business/news-and-events/releases/2026/09/17/rakowski-research. Accessed 8 Oct. 2026.

Narayanan, Devi. “How SOX Compliance Shapes Corporate Executive Responsibilities.” VComply, 18 Feb. 2025, www.v-comply.com/blog/sox-compliance-executive-responsibilities/. Accessed 15 Oct. 2026.

Niccum, Jon. “Social Media Discussions Fueled ‘Meme Stock’ Events and Significant Short Squeezes, Research Finds.” KU News, 2021, news.ku.edu/news/article/social-media-discussions-fueled-meme-stock-events-and-significant-short-squeezes-research-finds. Accessed 15 Oct. 2026.

Nunez, Natalia. “Hiding in Plain Sight: The Madoff Scandal and Regulatory Failure – Michigan Journal of Economics.” Umich.Edu, 4 Apr. 2025, sites.lsa.umich.edu/mje/2025/04/04/hiding-in-plain-sight-the-madoff-scandal-and-regulatory-failure/. Accessed 15 Oct. 2026.

Objectives and Principles of Securities Regulation International Organization of Securities Commissions (IOSCO). 2007.

Ouzan, Samuel. "Loss aversion and market crashes." Economic modelling 92 (2020): 70-86.

Quinlan, Andrew . “How the Internet Almost Took Down Wall Street by: Andrew Quinlan.” Randi L. Sims, Ph.D., 26 Feb. 2021, sites.nova.edu/sims/2021/02/26/how-the-internet-almost-took-down-wall-street/. Accessed 15 Oct. 2026.

Ramasastry, Anita. “The Silicon Valley Bank Collapse Explained.” UW School of Law, University of Washington, 24 Mar. 2023, www.law.uw.edu/news-events/news/2023/svb-collapse. Accessed 12 Oct. 2026.

Roncal, Jose. “Madoff and the Red Flags We Ignored.” Fast Company, 2 Jan. 2009, www.fastcompany.com/1122389/madoff-and-red-flags-we-ignored. Accessed 15 Oct. 2026.

Segal, Troy. “Enron Scandal and Accounting Fraud: What Happened?” Investopedia, 3 Dec. 2024, www.investopedia.com/updates/enron-scandal-summary/. Accessed 12 Oct. 2026.

Schiffrin, Anya. "The press and the financial crisis: A review of the literature." Sociology Compass 9.8 (2015): 639-653.

SEC. “Whistleblower Protections”. 2019. https://www.sec.gov/enforcement-litigation/whistleblower-program/whistleblower-protections.

Shefrin, Hersh. “The Psychology That Brought Down Silicon Valley Bank Is Not Going Away.” Forbes, 19 Mar. 2023,

www.forbes.com/sites/hershshefrin/2023/03/19/the-psychology-that-brought-down-silicon-valley-bank-is-not-going-away/. Accessed 12 Oct. 2026.

Sigalos, MacKenzie. “From $32 Billion to Criminal Investigations: How Sam Bankman-Fried’s Crypto Empire Vanished Overnight.” CNBC, 15 Nov. 2022, www.cnbc.com/2022/11/15/how-sam-bankman-frieds-ftx-alameda-empire-vanished-overnight.html. Accessed 15 Oct. 2026.

Steiger, Paul. “Not Every Journalist ‘Missed’ the Enron Story – Nieman Reports.” Niemanreports.Org, 2002, niemanreports.org/not-every-journalist-missed-the-enron-story/. Accessed 15 Oct. 2026.

Stiglitz, Joseph E. "The media and the crisis: an information theoretic approach." The media and financial crises. Routledge, 2014. 168-180.

Tang, Thomas Li-Ping, et al. "Behavioral economics and Monetary wisdom: the Enron effect—love of money, corporate ethical values, corruption perceptions index (CPI), and dishonesty across 31 geopolitical entities." Monetary Wisdom. Academic Press, 2024. 193-213.

Tang, Thomas Li-Ping, et al. "Monetary intelligence and behavioral economics: The Enron Effect—Love of money, corporate ethical values, Corruption Perceptions Index (CPI), and dishonesty across 31 geopolitical entities." Journal of Business Ethics 148.4 (2018): 919-937.

Tori, Daniele, Eugenio Caverzasi, and Mauro Gallegati. "Financial production and the subprime mortgage crisis." Journal of Evolutionary Economics 33.2 (2023): 573-603.

Velasquez, Manuel. “What Really Went Wrong With Enron? A Culture of Evil?” @SantaClaraUniv, 5 Mar. 2002, www.scu.edu/ethics/focus-areas/business-ethics/resources/what-really-went-wrong-with-enron/. Accessed 15 Oct. 2026.

Whiteside, Eric. “How Do Pension Funds Work?” Investopedia, 2020, www.investopedia.com/articles/investing-strategy/090916/how-do-pension-funds-work.asp. Accessed 15 Oct. 2026.

# A Comparative Analysis of the Influence of Influencer Partnerships and Brand Authenticity on Purchase Intention Among Generation Z and Millennials By Zain Melwani


## Abstract

This review of the literature looks at how working with influencers and being genuine as a brand affect people's desire to buy products, especially among younger generations like Gen Z and Millennials. This review brings together studies from academic sources. It looks at how credible influencers are and how their level of expertise differ. Real and trustworthy influencers can help change how people feel about a product or brand, and make them more likely to buy it. This is true when the influencer and the product or brand match well together. When people feel a brand is genuine, it can make them trust the brand more, connect with it emotionally, and be more likely to buy from it. The results vary depending on the person, the platform, or the type of product. There are differences in how people use social media. Their doubts about advertisements and their level of trust vary. The review points out that current studies have some weaknesses.



## Keywords

Influencer marketing; Brand authenticity; Purchase intention; Generation Z; Millennials; Consumer behaviour


## 1.1 Introduction

Social media has greatly influenced how people find, look at, and buy products and services. Social media platforms like Instagram, TikTok, and YouTube let brands talk directly to customers using engaging and tailored content. As digital marketing has grown, influencer marketing has become a key approach where brands work with people who already have a following and are trusted within certain online groups. Traits like credibility, trust, knowledge, and looks of an influencer can affect how consumers feel about a product and whether they want to buy it (Han and Balabanis, 2023). Influencer partnerships can affect how consumers react by considering things like their knowledge, how much people can connect with them, how trustworthy they are, and how well they match with the brand. These traits can influence how much people interact with content and their desire to buy something, but how much they matter can change based on the social media site, the kind of product, and the type of influencer (Sardar and Vijay, 2025). Besides working with influencers, being genuine has become a key part of how consumers make decisions. Brand authenticity is about how people think a brand is trustworthy, sincere, reliable, and stays true to what it stands for. Credibility, integrity, symbolism, and continuity are key parts of brand authenticity. They also showed that when a brand is authentic, it can affect how consumers feel about the brand and what brand they choose (Morhart et al., 2025). These factors are closely connected to purchase intention, which refers to a consumer’s likelihood or willingness to purchase a product or service. Perceived authenticity and influencer credibility can strengthen consumer trust and contribute to purchase-related responses.

Understanding these relationships is important when comparing Generation Z and Millennials. These generations have developed different relationships with social media, influencers, and digital advertising.

### 1.2 Generational Differences

Generation Z has spent their whole lives growing up with social media and digital content playing a big role in their daily lives. Influencers can help people find new products, affect how others make choices, and help brands connect with their audience. Gen Z cares about how real and trustworthy a person seems. They appreciate when influencers are genuine. They also tend to be wary of promotions that feel too salesy or heavily controlled.

Millennials have a lot of experience with social media and influencer marketing, and the way they use the internet has changed as digital platforms have grown over time. Studies show that people's trust in a brand, the value of the content, how believable the influencer is, and whether the influencer seems genuine can all affect how they react to recommendations from influencers. Millennials might care more about building long-term trust with brands, being transparent, and having strong relationships compared to Gen Z.

### 1.3 Research Problem

Although existing research shows that influencer credibility and brand authenticity can influence consumer trust and purchase intention, their combined effect is not yet clearly established.

Comparing Generation Z and Millennials is useful because differences in social-media behaviour, trust, relatability, and perceptions of authenticity may lead to different responses to influencer partnerships.

### 1.4 Literature Review

- The current research on influencer marketing and how it affects buying decisions looks at different factors that influence how consumers react when influencers team up with brands. Studies vary in what they look at, the groups they study, the platforms they use, and the types of products they examine, so it's important to compare them carefully to understand how influencer marketing, brand authenticity, and consumer behavior are connected overall.
- Influencer credibility is one of the factors that is always looked at closely. Studies show that traits like being expert, trustworthy, attractive, and credible are key factors that can make consumers feel more positively about a product and more likely to buy it. Research shows that factors related to credibility usually help influencers be more effective, but how much they matter can change depending on the situation.
- Influencer–brand fit is also important. Kim and Kim studied in 2021 and discovered that when people feel an influencer matches the product they are promoting, it changes how they understand the sponsored message. A better match can make the partnership look

more real and trustworthy, but a weak match might make people notice the business intention more.

- Relatability is especially important for social media influencers because people often see them as more like themselves compared to regular celebrities. Recent studies show that when people connect with an influencer's way of life or personal experiences, it can make them more engaged and more likely to want to buy something.
- Brand trust helps connect what influencers do with people's desire to buy something. Recent studies on Millennials and Gen Z in Saudi Arabia show that when influencers are seen as experts, trustworthy, and genuine, people tend to trust the brand more, which then makes them more likely to buy the product.
- Purchase intention is often used as the main result being studied, even though many research studies look at intention instead of what people actually do when buying.
- Generation/age group is another important difference. Studies looking at Gen Z and Millennials show that both groups care about honesty and realness, but they might have different ideas about what makes an influencer relationship work. Studies on Gen Z focus more on how relatable a brand is, the impact of friends and peers, and being wary of ads. Meanwhile, research on Millennials tends to stress the importance of trusting a brand and building lasting connections over time.
- Social media platforms and the product category can affect the results. Different social media platforms like Instagram and TikTok have their own ways of sharing content and how influencers interact with their audience. Also, different categories like fashion, beauty, food, and technology can make people feel or respond in various ways.

**2. Methods**

This study looks at how influencer partnerships affect brand authenticity and the desire to buy products among Generation Z and Millennials, using a narrative literature review approach. A story-based method was chosen because it lets us gather and look at results from various kinds of academic studies together, and compare them based on common topics, instead of just focusing on numbers and statistics. This method works for our study because the research includes many connected ideas. For example, influencer credibility, expertise, relatability, how well an influencer matches a brand, disclosing sponsorships, brand authenticity, trust in a brand, engagement on social media, and differences between generations.

The review looked at peer-reviewed academic journal articles. And research studies that discussed influencer marketing, brand authenticity, how consumers behave, their intention to buy, and the perspectives of Generation Z and Millennials. The study looked at academic and scholarly databases and publisher websites. For example, Wiley Online Library, ScienceDirect, Taylor & Francis, Springer, Frontiers. The study focused on research that was based on real data, carefully organized reviews, combined analysis of multiple studies, and well-known theories that directly looked at one or more of the factors being studied.

The literature was grouped into four main topics. The first study looked at how influencer partnerships affect the desire to buy, and it considered factors like credibility, expertise, trust, how relatable the influencer is, how well the influencer matches the brand, the size of the influencer, and whether they clearly mention their sponsored relationship. The second part looked at brand authenticity and purchase intention, covering aspects like authenticity, transparency, consistency, brand trust, and emotional connection. The third part compared Generation Z and Millennials, looking at how they engage with social media, how much they see influencers, what their friends think, how skeptical they are of ads, how much they trust brands, how loyal they are, and how they see brands as genuine. The fourth study combined these variables to look at how working with influencers and being genuine as a brand might connect, and whether these connections are different for the two generations.

## 3. Discussion

### 3.1 Theme 1: Influencer Partnerships and Purchase Intention

Influencers can impact consumers' opinions and behavioral responses through traits including credibility, knowledge, trustworthiness, relatability, and perceived resemblance. Influencer relationships have grown in importance within social media marketing. Through a meta-analysis of 250 effect sizes from 53 research with 25,080 participants, Han and Balabanis (2024) discovered that influencer attributes are crucial in determining how consumers react to influencer marketing. Their results show that customers' perceptions of influencers, content, and promoted brands are significantly influenced by trustworthiness, knowledge, and credibility. Influencer-brand congruence is also identified by the study as a significant factor, indicating that collaborations may be more successful when the influencer and promoted product are seen as compatible. The relationship between influencers and their audiences can be strengthened by perceived homophily or resemblance, and consumers' reactions may be further shaped by parasocial ties. Crucially, the authors discover that influencer effects differ depending on social media platform, product category, influencer type, and geographic context, suggesting that influencer marketing does not provide the same results for every consumer and circumstance. These results offer a solid foundation for investigating how purchase intention may be influenced by influencer credibility, knowledge, trustworthiness, relatability, and influencer–brand fit, as well as if these interactions vary between Generation Z and Millennials (Han and Balabanis, 2023). The effectiveness of influencer marketing is dependent on a number of factors, including the influencer's attributes, the content, and the influencer-customer relationship. Consumer opinions, purchasing intentions, purchase behavior, and sales can all be greatly impacted by influencer marketing. Purchase intention is influenced by the informational and hedonistic value of influencer material, and consumer reactions are also influenced by influencer traits including communication, self-disclosure, and influencer-brand fit. Purchase intention is positively correlated with influencer-brand fit. Customers may react more favorably if they believe the influencer and the promoted brand are compatible. Depending on the kind of social media site

and product category, these parameters' efficacy vary. Because explicitly disclosing economic links might alter how consumers view influencer material, sponsorship disclosure adds another level of complexity. Kim and Kim discovered that through advertising recognition, sponsorship disclosure influences customers' perceptions of an influencer's motivations (Wang et al., 2024). Sponsorship disclosures help people remember a brand better and make them more aware of how advertising works, but they can also make the brand seem less trustworthy. The impact of disclosure depends on factors like how and where it is shared (Liu and Zhao, 2024).

### 3.2 Theme 2: Brand Authenticity and Purchase Intention

Brand authenticity is about how real and true a brand seems to people, and how much it matches the values it says it stands for. Morhart and their team think about how people see a brand's authenticity in four ways: credibility, integrity, symbolism, and continuity.They say that being authentic means a brand is open and truthful, takes responsibility, stays the same over time, and lets customers show parts of who they are (Morhart et al., 2015). Authenticity isn't just about a brand seeming real; it's a complex idea that includes being genuine, staying consistent, and building a real connection with customers (Campagna et al., 2022). Honesty and being open are very important for being genuine because customers look at whether a brand's messages truly show what their products, beliefs, and actions are. Recent studies on how people see brand transparency show that being clear about how value is created, communicating openly, and willingly sharing important information are the main parts of transparency. This transparency affects how customers feel about a brand and whether they want to buy from it (Montecchi et al., 2024). It's also important for a brand's actions to match its values because people's belief in how real and trustworthy the brand is depends, in part, on whether the brand stays true to its stated beliefs over time. Studies show that when a brand is consistent, continuous, and unique, people feel it is more genuine. This sense of authenticity leads to greater trust in the brand and a higher sense of its value. Real authenticity can also build a strong emotional bond with customers. People feel more genuine when they think a brand is authentic, and this makes them emotionally connected to the brand. This connection can affect their decision to choose that brand because it feels like it fits who they are. Later studies show that when certain conditions are met, like how a person feels or the situation they're in, real brands can create even stronger emotional bonds (Hernandez-Fernandez and Lewis, 2019). Brand trust is also important because people tend to trust brands more when they think the brand is truthful, reliable, and really cares about doing the right thing. Studies have shown that when a brand is seen as authentic, people tend to trust it more. This means that being genuine can make customers more confident in the brand and what it says (Moulard et al., 2016). These relationships can influence purchase intention. How authentic a brand seems influences how likely people are to buy from it and what they actually choose to buy, with how dedicated the brand appears and what people expect in terms of quality being key factors in this connection (Rixom and Rixom, 2023). A 2026 study that included three experiments showed that brands seen as real created more positive reactions than those clearly fake. However, these real brands didn't do much better than brands that were just average. This

suggests that people might expect brands to be genuine, and that being authentic might just meet their basic expectations rather than giving them an extra benefit (Kokkoris and Kamleitner, 2026). Brand authenticity can affect how much someone wants to buy from a brand.This happens when the brand is honest and open, stays consistent, builds trust, creates an emotional connection, and seems to match its values with what it does.The strength of these connections can change based on the person and the situation.

### 3.3 Theme 3: Generation Z vs. Millennials

Gen Z and Millennials are different groups, but both of them are very active on social media and care a lot about what influencers share. They might have different ideas about what is real, how trustworthy someone is, and how ads should be presented. Both generations value authenticity, but Gen Z tends to see influencers as friendly experts they can trust for advice, while Millennials are more likely to see influencing as a job that needs to be honest and fair (Singer et al., 2023). Gen Z show are more aware of and sensitive to when brands seem to control things too much or use ads in a heavy way.Pradhan and others found that Gen Z might stop supporting influencers or brands if they feel the brands are too controlling.However, if there's a strong connection between the influencer and the audience, this avoidance behavior can be less likely (Pradhan et al., 2022). For Millennials, connections with influencers can also help build trust in ads and keep them loyal. Kim and Kim discovered that when people share similar interests, feel connected to others, and are seen as attractive, they become more attached to influencers. This attachment makes people more loyal to the influencer and more likely to believe in their ads, whether they're Millennials or Gen Z (Kim and Kim, 2022). Generation Z and Millennials both use social media a lot, but their experiences with digital media have come about in different ways because of the different times they grew up in. Millennials saw the rise of social networking and digital ads, but Gen Z has always had social media, influencers, and digital content around them in their daily lives. This difference might influence how each generation reacts to influencer marketing. Herrando and Martín-De Hoyos looked at 602 Instagram users from Generation Y and Z and discovered that when influencers are seen as credible and when their content is either fun or helpful, it makes people think the ads are more valuable. This, in turn, makes them more likely to have a positive opinion about the ads and want to buy the products. Participants followed influencers because they found them entertaining, wanted lifestyle tips, kept up with trends, and sought product information. This shows that influencers have both social and commercial roles for people of all generations (Herrando and Martín‐De Hoyos, 2022). For Gen Z, social media is closely linked to finding new products, being influenced by friends, and connecting with influencers. Factors like an influencer's credibility, being genuine, and the quality of their content were key things that influenced Gen Z's decision to make a purchase. Just watching or liking an influencer's content didn't always make someone buy something. This shows that following an influencer or engaging with their posts isn't the same as believing in their recommendations. Too much exposure to influencers can harm how consumers view a brand, which shows how important it is to share content that is

relevant and real (Muthulingam and Amirtharaj, 2026). Advertising doubt and realness are key things to think about when looking at what Gen Z buyers care about. Members of Generation Z might steer clear of influencers and brands they endorse if the influencers seem too heavily controlled by the brands or if the promotional content looks too salesy. This means that Gen Z doesn't always dislike influencer ads, but they care more about whether the message feels real and not forced. Gen Z appreciates when influencers communicate in a natural, honest, creative, and consistent way (Venes and Duarte, 2025). For Millennials, influencer marketing is also linked to looking for information, having fun, building trust, and forming connections with brands. When influencers are seen as trustworthy and provide helpful information, it makes people from Generation Y and Z more likely to think positively about ads from those influencers. When influencers show real knowledge and create genuine content, people trust the brand more. This trust then makes them more likely to want to buy the product. This means that for Millennials, influencer suggestions could be very important when they help reinforce how trustworthy and dependable a brand is. Brand loyalty and being genuine are also key factors when looking at Millennials. Influencer partnerships can help strengthen an existing relationship with a brand when the influencer's values and image match the brand's. If an influencer and a brand aren't a good match, people might not trust the brand as much. Singer and their team compared Gen Z and Millennials and discovered that both groups think authenticity is important, but they have different ways of understanding what it means. Gen Z participants usually thought that being real meant being someone you can connect with, like friends who know a lot or are reliable. Millennials cared more about people being professional, acting ethically, and being open and honest (Inkesar et al., 2026). Both generations appreciate honest and reliable communication. Each group judges what's genuine differently.

### 3.4 Theme 4: Comparative Analysis: Influencer Partnerships × Brand Authenticity × Generation

The research shows that influencer collaborations and brand honesty are not entirely separate things because both can help build trust, keep consumers interested, and influence their buying decisions. How trustworthy an influencer is, how much they know about their topic, how friendly they seem, and how well they match with a brand all affect how people see a promotional message. When a brand feels real and genuine, people think it's more honest, consistent, and in line with its values. Recent studies that looked specifically at Millennials and Generation Z showed that when influencers have expertise, are trustworthy, and create genuine content, it builds brand trust and makes people more likely to buy products. Brand trust plays a key role in connecting how influencers are seen with how consumers respond. The relationship isn't exactly the same for every generation. Studies on Gen Z show that trust and how genuine influencers seem can strongly affect whether people want to buy something. However, just being involved with content doesn’t always mean they will actually make a purchase.
(Inkesar et al., 2026). Gen Z may experience more intense parasocial effects compared to Millennials, which suggests that the way influencers connect with their audience might be

different for these two generations (K, 2026). At the same time, research comparing different generations doesn't always show that one group is more affected overall; some studies find no big difference in how Gen Z and Millennials spend their money. So, the research shows that looking at influencer partnerships, brand authenticity, trust, and the intention to buy should be seen as connected factors. Also, instead of thinking that Gen Z and Millennials react the same way, we should consider generation as something that might influence how these factors work together (Dwivedi et al., 2026).

**4. Challenges and Limitations**

One of the key issues is that many studies do not use clear and consistent ways to define and measure brand authenticity. There isn't a commonly agreed-upon definition or way to measure brand authenticity. Different studies look at authenticity in different ways, which makes it hard to compare them directly. Their suggested scale tries to bring together these various views into a more complete measurement (Campagna et al., 2022). Another issue is how fast influencer marketing and social media platforms are always changing. Influencers promote themselves, the websites they use, the rules that decide what people see, the types of content they create, and what people want. This can make old studies less useful in the long run. Several problems in influencer marketing research and practice. These include fake followers and comments, not properly telling when someone is paid to promote a product, choosing the wrong influencers, brands having too much control, using the wrong platforms, too much advertising, and the bad impact on a brand's reputation. These problems make it hard to figure out which traits of influencers always lead to good reactions from consumers ( Alipour et al., 2024). Another problem is that people may intend to buy something but don't actually end up making the purchase. Many studies use surveys where people tell if they plan to buy a product after seeing content from an influencer. Having a plan to buy something doesn't always mean a real purchase happens. This makes it hard to tell how long the effects of working with influencers will last and whether the brand stays genuine. Generational comparison also presents challenges. Generation Z and Millennials are big groups that include people of various ages, different earning levels, diverse cultures, different ways of living, and varying habits when it comes to using social media. Treating each generation as one big group can make it too easy to understand how consumers act. Södergren's review of 73 studies on brand authenticity also points out how cultural and social background plays a big role, and says that people don't all have the same understanding of what authenticity means (Södergren, 2021). The connection between working with influencers, being genuine as a brand, and how much people want to buy something doesn't always go in a straight line. Research that looks at just one factor might not completely explain how consumers make decisions. Future studies can look at these factors over a long period and compare different cultures.

**5. Future directions**

Future studies need to look more closely at influencer marketing, how real a brand seems, and how much people want to buy from them, using methods that are better connected and more suited to specific situations. Existing studies usually look at influencer traits and brand authenticity on their own, which means there is a need for models that explore how things like influencer credibility, authenticity, brand trust, how well the influencer matches the brand, and the intention to buy are connected. Sardar and Vijay looked at 76 studies and found that research on influencer marketing is scattered and not well connected. They suggest that more studies should focus on the different ways and situations that influence how people decide to buy something. Future studies should look into how people really decide if something is authentic, instead of thinking authenticity is just one unchanging trait. Smith and their team suggest that people might judge how real something is by using both logical, fact-focused thinking and feelings or gut reactions. So, research should look into how these processes are different across generations and whether being genuine has varied impacts based on who the consumer is, who the influencer is, which brand it is, or what product it relates to (Sardar and Vijay, 2025). There is also a need for more research that follows people over time and compares different cultures. A lot of current research depends on surveys and brief experiments, which makes it hard to tell if the connections between influencers and how people see their authenticity stay the same over time. Future research might look at how Gen Z and Millennials differ in various countries, cultures, platforms, and types of products to see if their differences stay the same across these areas. Research should also look into newer types of influencer marketing, like virtual and AI-created influencers, because they bring up new issues related to being genuine, clear, trustworthy, and how consumers react. Recent studies about virtual influencers show that authenticity and credibility are key factors that need more research (Barari et al., 2026).

## 6. Conclusion

The research in this study shows that working with influencers and being genuine as a brand are key factors in figuring out why consumers decide to buy products, especially among younger generations like Gen Z and Millennials. Influencer traits like being believable, knowledgeable, trustworthy, easy to connect with, and a good match for the brand can affect how consumers react to promotional messages. How real a brand is can affect how people see it as honest, reliable, trustworthy, and emotionally connected. These factors are connected; being genuine and trustworthy can build brand trust, which then affects how likely someone is to make a purchase. The comparison between Generation Z and Millennials shows that both groups care about honest and trustworthy communication, but they might have different views on how influencers and brands should connect. Studies on Gen Z focus a lot on things like being relatable, influence from friends, feeling connected to influencers, being wary of ads, and liking real, genuine content. Meanwhile, research on Millennials emphasizes trust in brands, their professionalism, and building long-lasting relationships with brands over time. However, the current evidence doesn't show that either group should be treated as a single, uniform group. The literature shows that the desire to buy comes from how different factors work together, like the

traits of the influencer, how the brand is seen, how much the consumer trusts, and the situation around them, not just from seeing the influencer or how authentic they seem alone. The field is limited because there are unclear definitions and ever-changing social media platforms. People report what they plan to buy instead of what they actually do.

**Works Cited**


Alipour, Seyyed Mohammadhossein, et al. “Influencer Marketing Research: A Systematic Literature Review to Identify Influencer Marketing Threats.” Management Review Quarterly, Mar. 2024, https://doi.org/10.1007/s11301-024-00412-5.

Barari, Mojtaba Moji, et al. “Virtual Influencer Marketing: A Meta-Analytic Review.” Journal of the Academy of Marketing Science, July 2026, https://doi.org/10.1007/s11747-026-01186-w.

Campagna, Christopher L., et al. “Brand Authenticity: Literature Review, Comprehensive Definition, and an Amalgamated Scale.” Journal of Marketing Theory and Practice, vol. 31, no. 2, Jan. 2022, pp. 1–17, https://doi.org/https://doi.org/10.1080/10696679.2021.2018937.

Dwivedi, Shalin, et al. “A Study Impact on the Influencer Marketing on Brand Trust and Purchase Intention: A Comparative Study of Millennials and Gen Z in India.” Ijarcce, vol. 15, no. 3, Mar. 2026, https://doi.org/10.17148/ijarcce.2026.153146.

Han, Jiseon, and George Balabanis. “Meta‐Analysis of Social Media Influencer Impact: Key Antecedents and Theoretical Foundations.” Psychology & Marketing, vol. 41, no. 2, Oct. 2023, https://doi.org/10.1002/mar.21927.

Hernandez-Fernandez, Asuncion, and Mathieu Collin Lewis. “Brand Authenticity Leads to Perceived Value and Brand Trust.” European Journal of Management and Business Economics, vol. 28, no. 3, 2019, https://doi.org/10.1108/ejmbe-10-2017-0027.

Herrando, Carolina, and María José Martín‐De Hoyos. “Influencer Endorsement Posts and Their Effects on Advertising Attitudes and Purchase Intentions.” International Journal of Consumer Studies, vol. 46, no. 6, Feb. 2022.

Inkesar, Asra, et al. “Influencer Marketing and Purchase Intention Among Millennials and Generation Z in Saudi Arabia: The Mediating Role of Brand Trust.” Frontiers in Sustainability, vol. 7, July 2026, https://doi.org/10.3389/frsus.2026.1838170.

K, Sowmia Rajan. “Influencer Credibility, Brand Trust, and Impulsive Purchase Intention in Social Commerce: A Cross-Generational Investigation Across Emerging Markets.” International Journal of Commerce and Management Research Studies (IJCMRS), vol. 3, no. 1, Mar. 2026, p. 27, https://doi.org/10.63090/ijcmrs/3049.1908.0032.

Kim, Do Yuon, and Hye-Young Kim. “Social Media Influencers as Human Brands: An Interactive Marketing Perspective.” Journal of Research in Interactive Marketing, vol. 17, no. 1, Feb. 2022, https://doi.org/10.1108/jrim-08-2021-0200.

Kokkoris, Michail D., and Bernadette Kamleitner. “Is Authenticity Rewarded or Inauthenticity Penalized? Reconsidering the Value of Brand Authenticity.” Journal of Retailing and Consumer Services, vol. 88, Jan. 2026, p. 104469, https://doi.org/10.1016/j.jretconser.2025.104469.

Liu, Xia, and Hong Zhao. “A Meta-Analysis of the Effects of Sponsorship Disclosure in Influencer Marketing.” Marketing Letters, vol. 36, no. 3, Nov. 2024, pp. 519–32, https://doi.org/10.1007/s11002-024-09757-z.

Montecchi, Matteo, et al. “Perceived Brand Transparency: A Conceptualization and Measurement Scale.” Psychology & Marketing, vol. 41, no. 10, June 2024, https://doi.org/10.1002/mar.22048.

Morhart, Felicitas, et al. “Brand Authenticity: An Integrative Framework and Measurement Scale.” Journal of Consumer Psychology, vol. 25, no. 2, 2015, pp. 200–18, https://doi.org/https://doi.org/10.1016/j.jcps.2014.11.006.

Moulard, Julie Guidry, et al. “Brand Authenticity: Testing the Antecedents and Outcomes of Brand Management’s Passion for Its Products.” Psychology & Marketing, vol. 33, no. 6, May 2016, pp. 421–36, https://doi.org/10.1002/mar.20888.

Muthulingam, Kodiarasi, and Nixon Amirtharaj. “Influencer Marketing and Gen Z’s Purchase Intentions: A Structural Equation Modeling Approach.” Frontiers in Communication, vol. 11, Apr. 2026, https://doi.org/10.3389/fcomm.2026.1830223.

Pradhan, Debasis, et al. “Influencer Marketing: When and Why Gen Z Consumers Avoid Influencers and Endorsed Brands.” Psychology & Marketing, vol. 40, no. 1, Oct. 2022, Wiley, pp. 27–47, https://doi.org/10.1002/mar.21749.

Rixom, Jessica M., and Brett A. Rixom. “Perceived Brand Authenticity: Examining the Effects of Inferred Dedication and Anticipated Quality on Consumers’ Purchase Intentions and Actual Choices.” Journal of Consumer Behaviour, vol. 22, no. 3, Mar. 2023, pp. 721–37, https://doi.org/10.1002/cb.2156.

Sardar, Sainaz, and T. Sai Vijay. “Social Media Influencers and Purchase Intention: A Review and Research Agenda.” International Journal of Consumer Studies, vol. 49, no. 3, Apr. 2025, https://doi.org/10.1111/ijcs.70046.

Singer, Mervyn, et al. “Differences in Perceived Influencer Authenticity: A Comparison of Gen Z and Millennials’ Definitions of Influencer Authenticity During the De-Influencer Movement.” Online Media and Global Communication, vol. 2, no. 3, Sept. 2023, pp. 351–78, https://doi.org/https://doi.org/10.1515/omgc-2023-0038.

Södergren, Jonatan. “Brand Authenticity: 25 Years of Research.” International Journal of Consumer Studies, vol. 45, no. 4, Feb. 2021, pp. 645–63, https://onlinelibrary.wiley.com/doi/full/10.1111/ijcs.12651.

Venes, Ana, and Alexandre Duarte. “Gen Z’s Code.” Digital Citizenship and Building a Responsible Online Presence, Mar. 2025, pp. 359–74, https://doi.org/10.4018/979-8-3693-6675-2.ch014.

Wang, Katja Spörl, et al. “Predictors of Social Media Influencer Marketing Effectiveness: A Comprehensive Literature Review and Meta-Analysis.” Journal of Business Research, vol. 186, Oct. 2024, pp. 114991–91, https://doi.org/10.1016/j.jbusres.2024.114991.

**Beyond Oil and Ore? Non-Resource Exports from Kazakhstan to China in the Trade Records of Both Countries, 2015-2025 By Inara Beibit**


## Abstract

China is the second-largest market for Kazakhstan's exports after Italy, and most of what Kazakhstan sells there is ores, metals, oil, gas, uranium and gold. This study asks how much of these exports is not minerals, fuels or metals, whether that part has grown, and which products account for the change. It uses China's reported imports from Kazakhstan in the UN Comtrade database for 2015-2025, at the level of four-digit Harmonized System (HS) headings, and checks them against Kazakhstan's own export records, which in 2020-2022 show no exports of copper and several other metals to China. In China's records, non-resource goods made up 2.4-4.8% of imports from Kazakhstan in every year from 2015 to 2022 and 7.1%, 6.9% and 7.6% in 2023-2025; in value they grew about ninefold, from US$157 million to US$1.44 billion. In Kazakhstan's records the share rose from 2.8% in 2015 to 10.5% in 2025. Most of the growth came from food and agri-food goods, above all bran and other milling residues (US$701 million in 2025), sunflower oil, linseed and rapeseed oil. Without bran, the share of processed goods stayed at about 2% from 2019 to 2025. Kazakhstan supplied 62% of China's imports of bran in 2025 but only 1.3-3.1% of its imports of barley, rapeseed oil, raw cotton and polypropylene. Resources still made up more than nine-tenths of the trade.


## Introduction

In 2024 and 2025 China was the second-largest market for Kazakhstan's exports, after Italy. In Kazakhstan's own statistics in UN Comtrade, the share of its exports that went to China rose from 11.9% in 2015 to 19.2% in 2025, close to Italy's 19.7% (*UN Comtrade Database*). For a country whose economy depends heavily on oil (Azretbergenova and Syzdykova), the question is not only how much it sells to China, but what it sells.

This study looks at the whole export basket to China rather than one sector. It asks three questions. First, how large is the non-resource part of what China buys from Kazakhstan, and has it grown between 2015 and 2025? Second, which products are behind the change? Third, how large is Kazakhstan's share of China's total imports of those products, which shows where its sales are already large and where they are still small?

## Literature Review

Research on export structure shows why the composition of exports matters. Across countries, diversification of production and exports rises with income up to a high level of income (Imbs and Wacziarg; Cadot et al.), and the kinds of goods a country exports are related to how fast it grows afterwards (Hausmann et al.). Kazakhstan's agricultural exports to China have already been studied as an opportunity: a study of agri-food trade between the two countries for 2012-2022 found that China's market for food is large and growing, but Kazakhstan holds only a modest share of it (Ibyzhanova et al.).

Studies of trade between two countries also have to choose whose records to use, because the exporter's and the importer's figures for the same flow rarely agree. The BACI database, built from UN Comtrade, combines the two figures using estimates of how reliable each country's reporting is and of transport and insurance costs (Gaulier and Zignago). This study does not reconcile the figures in that way; it uses the importer's records for the reason explained in the Methodology section.

## Methodology

### *Data*

All trade data come from the UN Comtrade database (*UN Comtrade Database*), which publishes the trade statistics that countries report to the United Nations. The data were downloaded through the public UN Comtrade interface for 2015-2025 at two levels of detail, two-digit HS chapters and four-digit HS headings, and for subheading 840130 in 2024 and 2025. Values are in current US dollars. Countries report in the HS edition in force in each year (HS 2012, 2017 and 2022 in this period); the headings discussed in this paper did not change between these editions. Net weights are used to check whether changes in value reflect changes in quantity.

Every trade flow between two countries is reported twice: as an export by the selling country and as an import by the buying country. The two figures rarely match (see the Literature Review). For Kazakhstan's exports to China, China's figures were 7-14% higher than Kazakhstan's in most years, which is in line with China valuing imports including transport and insurance, but 28% higher in 2018, 42-65% higher in 2020-2022 and 25% higher in 2025 (Table 1). In 2020-2022 Kazakhstan's records show no exports to China in HS chapters 74 (copper), 76 (aluminium), 78 (lead), 79 (zinc) and 81 (other base metals), while China recorded imports of copper from Kazakhstan of US$1.83-2.45 billion a year (*UN Comtrade Database*). In the same years Kazakhstan's reported exports of copper to all countries were also close to zero (US$0.05-0.11 billion, against US$2.74 billion in 2019), so the gap is not specific to China; UN Comtrade gives no explanation, and confidential treatment of these data is one possibility. In 2025 most of the gap is gold: China recorded US$1.90 billion of chapter 71 goods from Kazakhstan, Kazakhstan US$2 million. Because whole product groups are missing from one side in some years, this study uses China's reported imports from Kazakhstan, sometimes called mirror data, as the main source, and repeats the main calculation with Kazakhstan's records for the years in which they are complete.

Table 1. Trade from Kazakhstan to China in the two countries' records, US$ billion

| Year | Exports to China, Kazakhstan's records | Imports from Kazakhstan, China's records | Gap, % | Copper, Kazakhstan's records | Copper, China's records |
|---|---|---|---|---|---|
| 2015 | 5.48 | 5.85 | 7 | 0.78 | 0.95 |
| 2016 | 4.21 | 4.81 | 14 | 0.87 | 1.01 |

| | | | | | |
|---|---|---|---|---|---|
| 2017 | 5.80 | 6.38 | 10 | 1.11 | 1.23 |
| 2018 | 6.31 | 8.09 | 28 | 1.36 | 1.47 |
| 2019 | 8.00 | 8.86 | 11 | 1.57 | 1.74 |
| 2020 | 6.84 | 10.37 | 52 | 0.00 | 1.83 |
| 2021 | 6.85 | 11.29 | 65 | 0.00 | 2.45 |
| 2022 | 10.46 | 14.82 | 42 | 0.00 | 2.39 |
| 2023 | 14.76 | 16.36 | 11 | 1.91 | 2.16 |
| 2024 | 14.90 | 15.91 | 7 | 2.69 | 2.74 |
| 2025 | 15.20 | 18.93 | 25 | 2.07 | 2.18 |

Note: Copper is HS chapter 74. The gap is China's figure divided by Kazakhstan's, minus one. China values imports including transport and insurance (CIF), so its figures are expected to be somewhat higher even when both records are complete. Source: UN Comtrade (*UN Comtrade Database*).

***Resource and non-resource goods***

Each four-digit heading was assigned to a group by its HS chapter. Resource goods are mineral fuels (chapter 27), ores, salt, sulphur and stone (chapters 25-26), inorganic chemicals (chapter 28, which includes uranium compounds), precious metals and stones (chapter 71, mainly gold) and base metals (chapters 72-81). Heading 8401 was also counted as a resource good. In China's imports from Kazakhstan in 2024 and 2025 it consisted entirely of nuclear fuel elements, a fabricated form of uranium fuel, which Kazakhstan began to supply after the Ulba-FA plant, a joint venture of Kazatomprom's Ulba Metallurgical Plant and China's CGNPC-URC, began operating in November 2021; the first batch had been delivered to a Chinese nuclear power plant by December 2022 ("First Fuel"). Counting fuel elements as non-resource goods would make the diversification of the basket look larger than it is. Non-resource goods are food and agri-food goods (chapters 1-24) and all other goods, called here other non-resource goods; this second group includes raw cotton, wool and hides as well as manufactured products such as polymers. Chapter 99 (goods not specified by kind) is kept in the total but in neither group; it was below US$15 million a year until 2022 and US$55 million, US$114 million and US$155 million in 2023-2025.

This grouping counts raw crops, such as grain and oilseeds, as non-resource and counts copper cathodes and steel, which are processed, as resource goods. Because both choices affect the result, two other groupings are also reported in the Results section. The first counts only processed goods as non-resource: chapter 11 and chapters 15-24 except unmanufactured tobacco (milling products, vegetable oils and food-industry products), and other non-resource goods except raw cotton, wool, silk, bast fibres and hides. The second also counts base metals as non-resource.

***Measures***

For each year the study calculates the value and share of non-resource goods, the concentration of the export basket and the number of headings with significant sales. Concentration is measured with the Herfindahl-Hirschman index (HHI), the sum of the squared shares of all headings, which is 1 when one heading makes up all exports and falls as sales spread across more headings; chapter 99 is excluded from it. The share of the five largest headings and the number of headings worth at least US$1 million and US$10 million are also reported; these thresholds are in current dollars. For the growing non-resource products, Kazakhstan's share of China's total imports of each heading in 2025 was calculated from China's imports from all countries (*UN Comtrade Database*). Every trade figure in the paper is calculated by a Python script from the downloaded files.

## Results

### *What China buys from Kazakhstan*

China's imports from Kazakhstan grew from US$5.85 billion in 2015 to US$18.93 billion in 2025 (*UN Comtrade Database*). Resources made up most of this in every year. In 2025, ores (mainly copper ores) accounted for 25.7% of the total, base metals for 22.0%, mineral fuels for 17.9%, inorganic chemicals for 12.9%, gold and other precious metals for 10.0% and nuclear fuel elements for 3.0%. The single largest headings were copper ores and concentrates (US$3.24 billion), radioactive elements, which include uranium compounds (US$2.42 billion), refined copper (US$2.16 billion), crude oil (US$2.09 billion), gold (US$1.89 billion) and ferro-alloys (US$1.33 billion).

### *The non-resource share*

In China's records, non-resource goods made up between 2.4% and 4.8% of imports from Kazakhstan in every year from 2015 to 2022, with an average of 3.6%. In 2023-2025 their share was 7.1%, 6.9% and 7.6% (Figure 1). In value they grew about ninefold, from US$157 million in 2015 to US$1.16 billion in 2023, US$1.10 billion in 2024 and US$1.44 billion in 2025, which is 2.5 times their highest value up to 2022. Food and agri-food goods were most of it, US$1.29 billion in 2025, against US$151 million for other non-resource goods.

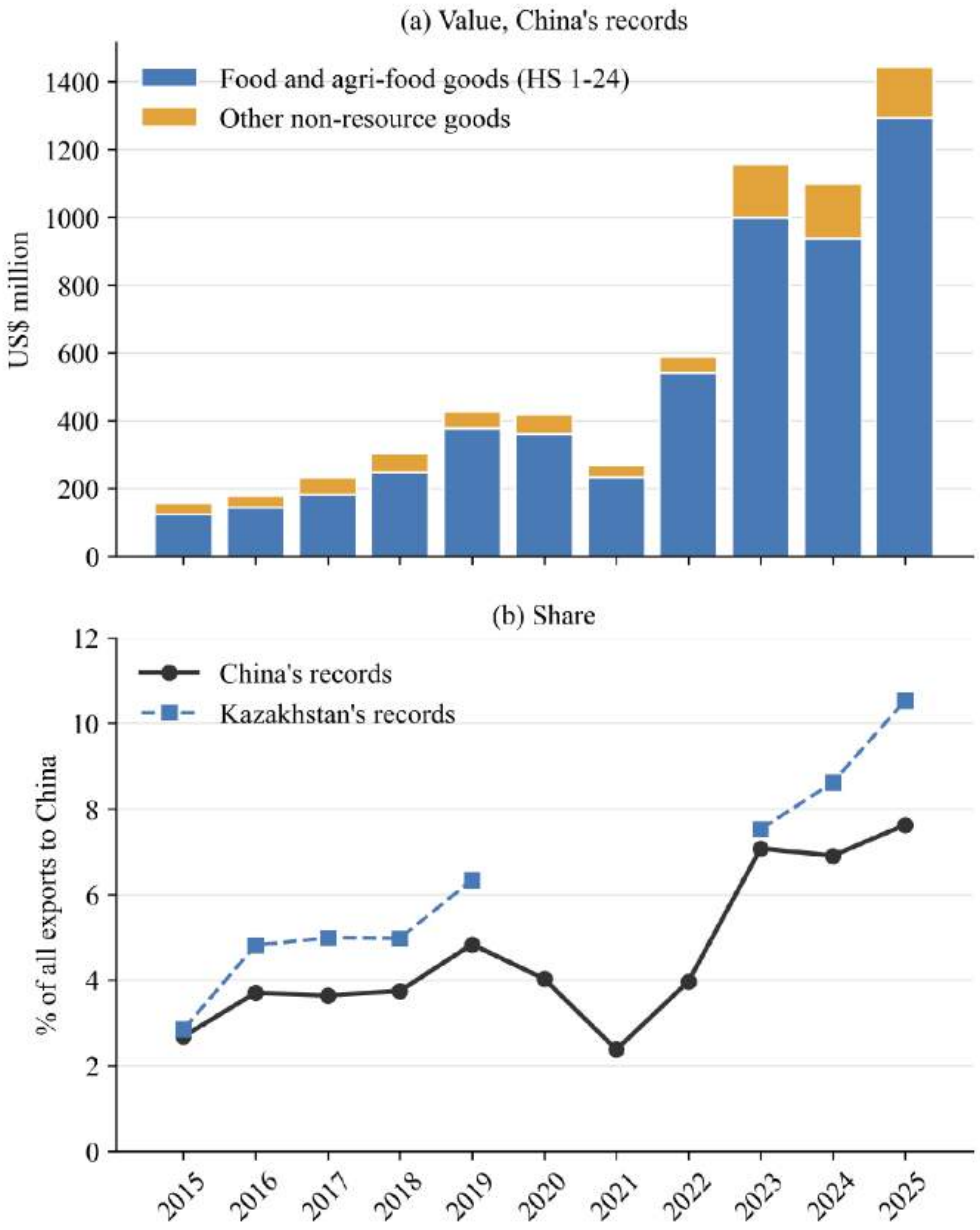



Figure 1. Non-resource goods in Kazakhstan's exports to China, 2015-2025: (a) value in China's records, US$ million; (b) share of all exports to China in China's records and in Kazakhstan's records. Kazakhstan's records are not shown for 2020-2022, when several metal chapters are missing from them. Source: UN Comtrade (*UN Comtrade Database*); grouping as described in the Methodology section.

In Kazakhstan's own records, in the years in which they are complete, the non-resource share was higher than in China's: 2.8% in 2015, 6.3% in 2019, 7.5% in 2023, 8.6% in 2024 and 10.5% in 2025. In both countries' records the share rose about three- to fourfold between 2015 and 2025; in China's records most of the rise came in 2023-2025, while in Kazakhstan's records the share had already reached 6.3% in 2019. Because the non-resource values in China's records are lower than in Kazakhstan's in most years, CIF valuation does not explain the difference for

this group; part of it comes from the two countries classifying the same goods under different headings (see Which Products Grew in the Results).

Changes on the resource side also move the share. China's imports of Kazakh crude oil fell from US$4.18 billion (6.4 million tonnes) in 2023 to US$1.25 billion (2.0 million tonnes) in 2024, mainly because of a lower volume; had crude oil stayed at its 2023 value, the 2024 non-resource share would have been 5.8% instead of 6.9%. In 2025, gold worth US$1.89 billion appeared in China's records and made the total larger.

The result depends on the grouping. If only processed goods are counted as non-resource, their share in China's records rises from 0.7% in 2015 and 1.9% in 2019 to 3.6% in 2024 and 5.7% in 2025, but bran and other milling residues are most of this rise: without them the processed share was 1.7% in 2019, 1.9% in 2023, 1.8% in 2024 and 2.0% in 2025. If base metals are also counted as non-resource, the share moves between 27% and 44% with no upward trend (31.9% in 2015, 42.7% in 2019 and 29.6% in 2025), because base metals then dominate it.

***Which products grew***

Table 2 lists the non-resource headings with China's imports from Kazakhstan of at least US$20 million in 2025, together with nuclear fuel elements. Most of the growth came from a small number of food and agri-food goods, and the largest single change was in bran and other milling residues: from US$0.6 million and 6 thousand tonnes in 2015 to US$701 million and 3,129 thousand tonnes in 2025. Its price per tonne in China's records also rose, from US$103 in 2015 to US$224 in 2025. Kazakhstan records most of these shipments under a different heading, 2309 (preparations used in animal feeding), which China records as almost nil. Taken together, headings 2302 and 2309 in Kazakhstan's records come to US$656 million and 3,196 thousand tonnes in 2025, so the quantity is confirmed, but the product may include feed preparations such as pellets, which could also explain the higher price per tonne. Sunflower oil, linseed and rapeseed oil also grew many times over between 2015 and 2025.

Table 2. Main non-resource products and nuclear fuel elements in China's imports from Kazakhstan

| HS | Product | 2015, US$ million | 2025, US$ million | 2015, kt | 2025, kt | Share of China's imports, 2025, % |
|---|---|---|---|---|---|---|
| 2302 | Bran and other milling residues | 0.6 | 700.8 | 6 | 3,129 | 61.7 |
| 1512 | Sunflower and safflower oil | 6.3 | 148.9 | 7 | 147 | 24.2 |
| 1204 | Linseed | 0 | 132.4 | 0 | 255 | 23.8 |
| 1514 | Rapeseed and mustard oil | 4.4 | 70.3 | 5 | 66 | 3.1 |
| 1003 | Barley | 0 | 64.0 | 0 | 350 | 2.4 |
| 3902 | Polypropylene and other olefin polymers | 4.6 | 61.2 | 5 | 85 | 1.3 |
| 5201 | Cotton, not carded | 0 | 48.1 | 0 | 31 | 2.5 |
| 1206 | Sunflower seeds | 27.5 | 38.3 | 68 | 77 | 53.3 |
| 2306 | Oil-cake | 0 | 31.6 | 0 | 159 | 1.9 |
| 840130 | Nuclear fuel elements (resource group) | 0 | 577.4 | 0 | 0.3 | 48.4 |

Note: Non-resource headings with China's imports from Kazakhstan of at least US$20 million in 2025. Values below US$0.05 million and weights below 0.5 thousand tonnes are shown as 0, except fuel elements, whose weight is shown to one decimal. Weights are net weights as reported by China; kt, thousand tonnes. The last column is China's imports from Kazakhstan divided by China's imports from all countries in 2025; for fuel elements it uses subheading 840130, because heading 8401 also includes reactors and their parts. Source: UN Comtrade (*UN Comtrade Database*).

The comparison of 2015 and 2025 hides large changes from year to year. Barley reached US$190 million in 2023 and US$158 million in 2024 and fell to US$64 million in 2025; wheat, US$151 million in 2023 and US$134 million in 2024, fell to US$15 million in 2025, a year in which China's own imports of wheat from all countries fell from US$3.48 billion to US$1.14 billion. Linseed fell from US$155 million in 2023 to US$61 million in 2024 before rising to US$132 million in 2025. The weights show that the increases since 2015 were mainly increases in quantity rather than in price.

Outside the food chapters, the two largest growing products are polypropylene and other olefin polymers (US$61 million in 2025) and raw cotton (US$48 million). Polypropylene is the clearest new manufactured product. China recorded 4.6 thousand tonnes from Kazakhstan in 2015, almost none in 2016-2021, 3.2 thousand tonnes in 2022, and then 106, 89 and 85 thousand tonnes in 2023-2025. This matches the start of production at the Kazakhstan Petrochemical Industries plant in Atyrau Region, launched in 2022, whose sales table gives 93,109 tonnes to China in 2023 and 89,001 tonnes in 2024 ("Petrochemicals"); the 2024 figure is almost the same as China's 88.9 thousand tonnes. The plant's table also gives 28,859 tonnes to China in 2022, much more than China recorded under heading 3902 for that year, which this study cannot explain. Kazakhstan also reports exports of cyclic hydrocarbons (heading 2902) to China, US$114 million in 2024 and US$56 million in 2025, that China records almost entirely under resource heading 2710 or not at all, which is one reason why the non-resource share is higher in Kazakhstan's records. Nuclear fuel elements grew from nothing before 2022 to US$577 million in 2025, more than three times all other non-resource goods outside the food chapters together; they were counted as a resource good.

***Concentration of the basket***

The basket became less concentrated over the period, although not steadily (Table 3). The HHI fell from 0.206 in 2015 to 0.096 in 2025, the share of the five largest headings from 87% to 63%, and the number of headings worth at least US$10 million rose from 19 to 37. Most of this change, about nine-tenths, came from shifts among resource headings, above all the fall in crude oil's share from 32% of the total in 2015 to 8-11% in 2024-2025, while copper ores, uranium, ferro-alloys and gold grew: if non-resource goods had kept their 2015 share, the 2025 HHI would have been 0.105 instead of 0.096. Most of the fall had already happened by 2019, and concentration rose again in 2021-2022, when crude oil sales to China were high, so the result depends on the years compared.

Table 3. Concentration of China's imports from Kazakhstan by four-digit HS heading

| Year | HHI | Five largest headings, % | Headings of US$1 million or more | Headings of US$10 million or more |
|---|---|---|---|---|
| 2015 | 0.206 | 87 | 40 | 19 |
| 2016 | 0.160 | 84 | 42 | 19 |
| 2017 | 0.142 | 82 | 49 | 22 |
| 2018 | 0.121 | 71 | 53 | 23 |
| 2019 | 0.114 | 68 | 61 | 30 |
| 2020 | 0.117 | 70 | 61 | 33 |
| 2021 | 0.140 | 77 | 52 | 32 |
| 2022 | 0.175 | 80 | 59 | 36 |
| 2023 | 0.138 | 74 | 59 | 35 |
| 2024 | 0.112 | 68 | 62 | 37 |
| 2025 | 0.096 | 63 | 70 | 37 |

Note: HHI is the sum of the squared shares of all headings, excluding chapter 99. Source: UN Comtrade (*UN Comtrade Database*).

***Kazakhstan's share of China's imports***

The last column of Table 2 shows two kinds of products. Kazakhstan was already a large supplier to China in 2025: it supplied 61.7% of China's imports of bran and milling residues, 53.3% of its sunflower seeds and about a quarter of its imports of sunflower oil and of linseed. In others, China is a large buyer and Kazakhstan's share was small: in 2025 China imported US$2.69 billion of barley, US$2.30 billion of rapeseed oil, US$1.93 billion of raw cotton, US$1.66 billion of oil-cake and US$4.62 billion of polypropylene and similar polymers, and Kazakhstan supplied 2.4%, 3.1%, 2.5%, 1.9% and 1.3% of these amounts. For comparison, among resource goods Kazakhstan supplied 48.7% of China's imports of radioactive elements and 48.4% of its imports of nuclear fuel elements, but only 0.7% of its crude oil imports.

**Discussion**

The answer to the first question is that the non-resource part of Kazakhstan's sales to China did grow, but from a very small base. In China's records a share of 2.4-4.8% became a share of about 7-8% in 2023-2025, and in value non-resource goods passed US$1 billion a year; in Kazakhstan's records the share rose from about 3% to about 10%. The two records agree on the direction, but not on the timing or the level, and part of the difference comes from how the same goods are classified.

The second answer is that the growth came mainly from food and agri-food goods: bran and other milling residues, vegetable oils, oilseeds and, in some years, grain. Most of these goods have little processing, and the processed share without bran did not rise after 2019. Only a few processed products, sunflower and rapeseed oil and polypropylene, appear among the growing goods. The new goods are therefore mostly farm goods rather than the manufactured goods

usually associated with upgrading; whether they are sophisticated in the sense of Hausmann et al. was not measured here.

The case of nuclear fuel elements shows why the grouping matters. If they were counted with the other non-resource goods outside the food chapters, that group would appear to have grown about twenty-onefold between 2021 and 2025, from US$35 million to US$729 million, instead of about fourfold. This growth came from adding value to uranium within the resource sector, which is a form of upgrading but not of diversification away from resources.

The third answer suggests where further growth is possible, with caution. In bran, sunflower seeds and sunflower oil Kazakhstan already supplies a large part of what China buys, so further growth depends mainly on China's own demand. In barley, rapeseed oil, raw cotton, oil-cake and polypropylene, China bought between US$1.7 billion and US$4.6 billion each in 2025 and Kazakhstan's share was 3.1% or less. These are the same kind of agri-food opportunities identified for 2012-2022, where the challenges named were technical disparities, logistical limitations and institutional constraints (Ibyzhanova et al.). A small share of a large market is not proof that sales can grow: other suppliers may be cheaper or closer, China's own imports can fall sharply, as wheat did in 2025, and for polypropylene the limit is the output of one plant.

Finally, the gap between the two countries' records is a finding in itself. In 2020-2022 Kazakhstan's export statistics in UN Comtrade show no exports of copper and several other base metals to China, and almost no copper exports to any country. Any analysis of Kazakhstan's trade with China that used only Kazakh data for these years would understate both the total and the role of base metals.

**Limitations**

China reports the value of its imports including transport and insurance, but for non-resource goods its values are lower than Kazakhstan's in most years, so the two records differ for other reasons too, including classification under different headings, and the level of the non-resource share depends on whose records are used. China's records may also contain errors, and goods may be attributed to Kazakhstan as their origin when they only passed through it, or the reverse. Values are in current US dollars and are not adjusted for prices; commodity prices change the size of the total and therefore the share, and only the effect of crude oil in 2024 was calculated, although net weights were used to check the main non-resource products. The result depends on how goods are grouped, as the Results section shows. Kazakhstan's share of China's imports was calculated for a limited list of headings and only for 2025. The study describes what was traded; it does not explain why the changes happened or test whether they will continue.

**Conclusion**

From 2015 to 2025 China became a larger buyer of Kazakhstan's goods, and a small but growing part of what it bought was not minerals, fuels or metals. In China's records non-resource goods rose from 2.4-4.8% of imports from Kazakhstan before 2023 to 7-8% in 2023-2025, or

US$1.1-1.4 billion a year, and in Kazakhstan's records the share rose to 10.5% in 2025. The growth was led by bran and other milling residues, vegetable oils and oilseeds, while the share of processed goods other than bran stayed at about 2%. Resources, including a new processed resource, nuclear fuel elements, still made up more than nine-tenths of the trade, and most of the fall in the concentration of the basket came from changes among resources. Products in which China buys a lot and Kazakhstan supplies little, such as rapeseed oil, barley, oil-cake and polypropylene, are where the non-resource part of this trade has the most room to grow, if supply and access allow.

**Acknowledgments**

The author thanks her mentor, Myrzabek Kurakbay, for collecting and analysing the data and drafting the text, and her scientific supervisor, Yerezhepova Sabina Bolatovna, for her support. The author chose the topic of this study, Kazakhstan's exports to China. The data were downloaded and analysed, the tables and the figure were made, and the text was drafted by the author's mentor with the help of a generative AI assistant (Claude, Anthropic). All trade figures were produced by a reproducible script from the sources listed, and the figures, sources and wording were checked in two further rounds of review by separate AI assistants. The author read and approved the final text.

**Works Cited**

Azretbergenova, Gulmira, and Aziza Syzdykova. "The Dependence of the Kazakhstan Economy on the Oil Sector and the Importance of Export Diversification." International Journal of Energy Economics and Policy, vol. 10, no. 6, 2020, pp. 157-63, https://doi.org/10.32479/ijeep.9997.

Cadot, Olivier, et al. "Export Diversification: What's behind the Hump?" The Review of Economics and Statistics, vol. 93, no. 2, 2011, pp. 590-605, https://doi.org/10.1162/REST_a_00078.

"First Fuel from Ulba-FA Fuel Delivered to Customer." World Nuclear News, 8 Dec. 2022, world-nuclear-news.org/articles/first-fuel-from-ulba-fa-fuel-delivered-to-customer.

Gaulier, Guillaume, and Soledad Zignago. BACI: International Trade Database at the Product-Level: The 1994-2007 Version. CEPII Working Paper no. 2010-23, CEPII, Oct. 2010, www.cepii.fr/PDF_PUB/wp/2010/wp2010-23.pdf.

Hausmann, Ricardo, et al. "What You Export Matters." Journal of Economic Growth, vol. 12, no. 1, 2007, pp. 1-25, https://doi.org/10.1007/s10887-006-9009-4.

Ibyzhanova, Aizhan, et al. "Opportunities for Kazakhstan's Agricultural Exports to the Chinese Market." Journal of Eastern European and Central Asian Research, vol. 11, no. 5, 2024, pp. 871-86, https://doi.org/10.15549/jeecar.v11i5.1570.

Imbs, Jean, and Romain Wacziarg. "Stages of Diversification." American Economic Review, vol. 93, no. 1, 2003, pp. 63-86, https://doi.org/10.1257/000282803321455160.

"Petrochemicals." KazMunayGas Annual Report 2024, KazMunayGas, ar2024.kmg.kz/en/strategic-report/operating-review/petrochemicals. Accessed 17 Oct. 2026.

UN Comtrade Database. United Nations Statistics Division, comtradeplus.un.org. Accessed 17 Oct. 2026.

**Beyond the Fine: Measuring Corporate Risk Across Legal and Regulatory Events**
**By Saksham Singh**


## Abstract

Legal and regulatory events can impose costs on publicly traded companies that extend beyond fines and settlements. This study develops and applies a comparative framework to nine public-company cases across telecommunications, cosmetics and personal care, and banking. Each case is evaluated on five equally weighted dimensions: financial exposure, regulatory severity, market reaction, duration of risk, and business impact. The cases include eight government enforcement actions and one private securities class action. Across the sample, direct financial exposure generally received low scores, while market reaction and the persistence of operational or compliance consequences varied considerably. Wells Fargo received the highest overall score (3.6), followed by Blue Ridge Bankshares and Estée Lauder (3.2 each). The Wells Fargo and Blue Ridge scores reflected sustained restrictions and compliance obligations rather than large direct penalties, whereas Estée Lauder showed the most severe market-reaction score despite a comparatively low financial-exposure score, illustrating that investor response and direct legal costs need not move together. The findings do not establish causation and are limited by the small, non-random sample, event-specific confounding factors, and subjective scoring. Nevertheless, the framework suggests that fines alone are an incomplete measure of corporate legal risk and that duration and operational constraints deserve separate evaluation.



**Keywords** corporate risk; regulatory enforcement; financial exposure; market reaction; compliance; public companies


## Introduction

Legal and regulatory events are often summarized by a single headline figure: the amount of a fine, settlement, or alleged loss. That figure is easy to communicate, but it does not necessarily capture the full consequences for a company. Research on financial regulation supports this concern: in one study of enforcement by United Kingdom regulators, reputational losses were nearly nine times the size of the fines (Armour et al. 1429). A relatively small penalty may accompany years of monitoring, restrictions on business activity, and costly changes to internal controls. Conversely, a large settlement may be manageable for a company with substantial liquidity and may not coincide with a measurable negative stock-market reaction. A more complete assessment of corporate risk therefore needs to distinguish the direct financial cost of an event from its regulatory, market, temporal, and operational consequences.

This paper asks: How are significant legal and regulatory events associated with the financial risk and market performance of publicly traded companies? It further examines three related questions: whether greater financial exposure relative to company size corresponds with a more negative market reaction; whether ongoing regulatory restrictions are associated with greater risk

than one-time penalties; and whether company size helps explain the relative burden of legal exposure.

The project began with the hypothesis that companies facing greater legal exposure relative to revenue, cash, or other financial resources would experience greater negative financial and market consequences. The cross-case evidence did not support that hypothesis consistently. Instead, the cases suggest that direct legal costs are only one component of corporate risk. The length and structure of restrictions, the extent of operational changes, and the market's response can be at least as important within the framework used here.

The study contributes a transparent, repeatable scoring approach rather than a causal estimate of the effect of litigation or enforcement on share prices. Its purpose is to compare documented consequences consistently, identify patterns that merit further study, and make clear where the available evidence does not support strong conclusions.

## Research Design and Methodology

### *Case selection and comparative structure*

The Corporate Risk Ledger (the "Ledger"), the author's unpublished comparative research project, examines nine publicly traded companies selected for having a significant, identifiable legal or regulatory event and publicly available financial and market information. The cases are arranged across three fields—telecommunications, cosmetics and personal care, and banking—with three companies representing different relative scales within each field. The final sample comprises UTStarcom, Shentel, and T-Mobile US; Avon Products, Edgewell Personal Care, and The Estée Lauder Companies; and Regions Financial, Blue Ridge Bankshares, and Wells Fargo. Eight cases concern government enforcement actions; Estée Lauder concerns private securities litigation. The sample is purposive, not random, and is not intended to represent all public companies or legal events.

### *Five dimensions of corporate risk*

Each event receives a score from 1 to 5 in five dimensions, with higher scores indicating greater risk. Financial Exposure measures direct legal costs or documented estimates of non-monetary financial consequences relative to the company's financial capacity. Regulatory Severity measures the intensity of the legal or regulatory response, including penalties, consent orders, monitors, reporting requirements, and restrictions. Market Reaction measures the company's stock performance relative to the S&P 500 around the event. Duration of Risk measures how long the material consequences continued, rather than how long the underlying conduct or investigation lasted. Business Impact evaluates the documented effect on operations, internal controls, business procedures, customers, market access, or core commercial activities.

The overall score is the arithmetic mean of the five dimensions: (Financial Exposure + Regulatory Severity + Market Reaction + Duration of Risk + Business Impact) ÷ 5. Equal weighting makes the calculation transparent and avoids assigning an empirically unsupported

importance to one dimension. The overall average is not treated as a complete description by itself; the distribution of component scores is also examined because two cases with the same average may have very different risk profiles.

### *Financial and market measures*

Financial exposure is assessed using reported monetary penalties or settlements compared with revenue, cash and cash equivalents, and the magnitude of net income or net loss. Where available, operating cash flow and other relevant measures provide additional context. For banks, the analysis adapts the measures to the structure of financial statements: total revenue is used, gross profit is omitted, pre-tax income replaces operating profit, total assets are included, and the Tier 1 Capital Ratio is used instead of the current ratio. These adjustments reflect that conventional product-company metrics do not describe bank operations in the same way. Market Reaction is calculated as the company's return minus the S&P 500 return over consistent event windows (see Appendix A). This approach borrows the logic of event studies, which infer an event's effect on a firm's value from security prices around the event (MacKinlay 13), but it is simpler than a formal event study. A positive result indicates outperformance; a negative result indicates underperformance. The analysis considers the disclosure date, an initial post-disclosure period, and a subsequent short-term window where trading data permit, along with an extended check in some cases. It also checks the 30-day, 90-day, and 52-week pre-event trends, a supplementary industry benchmark, and relevant company-specific developments. Any negative relative performance in a valid supporting window establishes a minimum score of 2/5; a single negative supporting window does not by itself justify a score above 2/5, and stronger scores require evidence of greater magnitude or persistence. The S&P 500 remains the primary benchmark across all cases.

### *Sources, attribution, and limitations of measurement*

The underlying case studies draw on company annual reports and SEC filings, regulatory orders and announcements, court records, and historical stock-price data. Financial measures are compared over a three-year window—typically the year before the event, the event year, and the first full year afterward. For Shentel, whose consent decree was entered in April 2023, the available financial data cover 2021 through 2023, so the event year is the final year of the window and no full post-event year is included (Shenandoah Telecommunications, *Form 10-K 2023*). This period provides context but does not isolate the legal event as the cause of financial changes.The analysis distinguishes documented facts from interpretation and does not infer causation from temporal coincidence. Scores necessarily involve judgment, particularly for regulatory severity and business impact; the rubric is intended to make those judgments explicit and comparable rather than eliminate them.

## Results

### *Overview of the nine cases*

The nine overall scores ranged from 2.0 to 3.6 on the five-point scale. Wells Fargo received the highest overall score (3.6), followed by Estée Lauder and Blue Ridge Bankshares (3.2 each), Avon Products (3.0), UTStarcom (2.8), and T-Mobile US (2.6). Shentel and Edgewell Personal Care each received 2.2, while Regions Financial received 2.0 (see Appendix B). These averages should be read alongside the individual dimensions: the same overall score can reflect either a concentrated, high-severity consequence or a combination of moderate scores.

| Company | Event type (short description) | Overall score |
|---|---|---|
| UTStarcom | FCPA enforcement; DOJ and SEC; December 2009 | 2.8/5 |
| Avon Products | FCPA enforcement; DOJ and SEC; December 2014 | 3.0/5 |
| Shentel | FCC consent decree (911 outage); April 2023 | 2.2/5 |
| T-Mobile US | FCC consent decree (customer data breaches); September 2024 | 2.6/5 |
| Edgewell Personal Care | FTC action against acquisition of Harry's; February 2020 | 2.2/5 |
| Estée Lauder | Securities class action; November and December 2023 | 3.2/5 |
| Wells Fargo | Federal Reserve asset-growth cap; February 2018 | 3.6/5 |
| Blue Ridge Bankshares | OCC consent order and fintech/capital restrictions; January 2024 | 3.2/5 |
| Regions Financial | Multi-agency enforcement and consent order; June 2014 | 2.0/5 |

***Financial exposure did not reliably predict market reaction***

The clearest cross-case result is the lack of a consistent relationship between direct financial exposure and market reaction. Financial Exposure scores clustered narrowly between 1/5 and 2/5, while Market Reaction scores spanned the full range from 1/5 to 5/5 (see Appendix A). Estée Lauder had the most severe Market Reaction score (5/5) despite a low Financial Exposure score; Wells Fargo and Blue Ridge Bankshares also recorded elevated market-reaction scores while their direct monetary exposure was limited or their principal consequences were non-monetary. By contrast, UTStarcom and T-Mobile received Market Reaction scores of 1/5, and Edgewell's score was similarly low, despite each facing a significant legal or regulatory event: UTStarcom's December 2009 resolution of SEC and DOJ Foreign Corrupt Practices Act investigations (UTStarcom Holdings, *Form 10-K 2010*), T-Mobile's September 2024 FCC consent decree over customer data breaches (Federal Communications Commission, *In the Matter of T-Mobile*), and the Federal Trade Commission's February 2020 challenge to Edgewell's proposed acquisition of Harry's (Federal Trade Commission, "Statement of Daniel Francis").

The Estée Lauder case is especially instructive because its $210 million settlement of a private securities class action, which the court approved as fair, reasonable, and adequate in October 2026 (*In re Estee Lauder*, Opinion and Order), was substantial in absolute terms but still represented about 1.32 percent of fiscal 2023 net sales in the financial comparison used by the Ledger (Estée Lauder Companies, *Form 10-K 2024*). The parties reached an agreement in principle on April 2, 2026, and the company reported that insurance carriers funded part of the amount, so $210 million is a gross figure and not necessarily Estée Lauder's out-of-pocket cost (Estée Lauder Companies, *Form 10-K 2026*). The stock's measured relative reaction was more severe than the financial-exposure ratio alone would suggest. That measurement centered on the company's November 1, 2023 disclosure of weaker sales and a lowered outlook (Estée Lauder Companies, "Estée Lauder Companies Reports"), which preceded the first complaint on December 7, 2023 (Estée Lauder Companies, *Form 10-K 2026*), so it should not be read as the market's response to the litigation itself. The case also involved pre-existing business pressures and other possible influences on the share price (Estée Lauder Companies, *Form 10-K 2024*). It does, however, demonstrate why a penalty-to-revenue ratio cannot serve as a substitute for measuring market reaction separately.

Avon provides a complementary example. Its approximately $135 million FCPA resolution represented about 1.77 percent of 2014 revenue, 14.06 percent of cash and equivalents, and 39.19 percent of the magnitude of its loss from continuing operations (U.S. Securities and Exchange Commission, "SEC Charges Avon"; Avon Products, *Form 10-K 2015*). These ratios show that the settlement was more meaningful relative to liquidity and earnings than to revenue. Yet Avon's broader financial condition was already deteriorating before the 2014 resolution: revenue fell from about $8.50 billion in 2013 to $7.65 billion in 2014, cash from $1.11 billion to $960.5 million, and operating cash flow from $470.5 million to $288.9 million (Avon Products, *Form 10-K 2015*). It would therefore be inaccurate to attribute the company's overall deterioration to the enforcement action alone.

***Duration and operating restrictions mattered***

The highest overall score (Wells Fargo, 3.6/5) and one of the two tied second-highest scores (Blue Ridge Bankshares, 3.2/5) were associated with prolonged regulatory restrictions rather than a direct monetary penalty as the central consequence. Wells Fargo's Federal Reserve asset-growth cap remained in effect for more than seven years (Board of Governors, "Responding to Widespread"; Board of Governors, "Federal Reserve Announces"). The February 2018 order barred the company from growing total assets beyond their year-end 2017 level of about $1.95 trillion until it improved its governance and risk management (Board of Governors, "Responding to Widespread"), and Wells Fargo disclosed that reductions in commercial deposits reflected actions taken in response to the cap (Wells Fargo & Company, *2018 Annual Report*). Its direct disclosed financial impact was small relative to its resources: the estimated $300–$400 million reduction in 2018 after-tax net income equaled about 0.34 to 0.45

percent of 2017 revenue (Wells Fargo & Company, *2018 Annual Report*). The risk profile instead reflected duration, market performance, and constraints on important business activity.

Blue Ridge Bankshares' January 2024 Office of the Comptroller of the Currency consent order continued and expanded earlier requirements, including restrictions on fintech activities and minimum capital requirements (Blue Ridge Bankshares, *Form 8-K*). The OCC terminated the order effective November 13, 2025, after about 22 months (Blue Ridge Bankshares, "Blue Ridge Bankshares, Inc. Announces"), so the restrictions were multi-year but not permanent. The bank's revenue and total assets declined during the three-year period examined, although net income improved from a loss in 2023 to positive earnings in 2025 (Blue Ridge Bankshares, *Form 10-K 2025*). These trends are consistent with a company experiencing significant financial and operational pressures, but the available data cannot establish that the consent order alone caused them. The case received a high overall score because the restrictions and compliance obligations persisted across multiple years and were structurally important, even though they ended before the end of the period studied.

The comparison with Edgewell prevents an overly broad conclusion that all restrictions are more damaging than fines. On February 3, 2020, the Federal Trade Commission voted to challenge Edgewell's proposed $1.37 billion acquisition of Harry's (Federal Trade Commission, "Statement of Daniel Francis"; Edgewell Personal Care Company, "Edgewell Personal Care and Harry's"). Edgewell abandoned the acquisition on February 10, 2020 (Federal Trade Commission, "Statement of Daniel Francis"), and the documented consequences were more limited in duration than the long-running restrictions in the Wells Fargo and Blue Ridge cases (Edgewell Personal Care Company, *Form 10-K 2021*). Its overall score was 2.2/5 and its Duration of Risk score was 2/5.

***Regulatory severity and business impact are distinct***

Regions Financial illustrates that a significant enforcement response does not necessarily translate into severe market or operational disruption. The June 2014 actions, brought by the Securities and Exchange Commission, the Federal Reserve, and the Alabama State Banking Department, involved a total of approximately $51 million in penalties, as well as requirements to improve policies and procedures (Regions Financial Corporation, "Regions Financial Reaches"). Yet the event-window analysis did not show a meaningful sustained negative relative stock reaction (see Appendix A), and the documented requirements did not substantially restrict core banking activities (Regions Financial Corporation, "Regions Financial Reaches"). Regions received a Regulatory Severity score of 3/5 but an overall score of 2.0/5, the lowest in the sample.

Estée Lauder illustrates a different separation: a severe measured market reaction can coexist with a lower Regulatory Severity score. Because the case involved private securities litigation rather than government enforcement, it did not involve the same agency monitoring, consent-order, or reporting mechanisms evaluated in the other cases. The difference is a limitation of the unified rubric, but also a substantive reminder that private litigation and

government enforcement can generate different risk profiles. Prior research finds significantly negative stock price reactions to shareholder-initiated class actions and shows that investors partially anticipate such suits, capitalizing some of the loss before filing (Gande and Lewis 823). A framework that combines them should not treat their legal mechanisms as interchangeable.

***Company size and financial resilience***

The evidence suggests that company size alone does not determine the relative burden of legal risk. A fixed-dollar cost can weigh more heavily on a smaller company, but profitability and liquidity at the time of the event also matter. Shentel's $227,200 penalty under an April 2023 FCC consent decree resolving a 2022 911 outage (Federal Communications Commission, *In the Matter of Shenandoah*) was minimal relative to its financial capacity, about 0.08 percent of 2023 revenue, yet it equaled about 2.84 percent of 2023 net income because the company's profitability was low (Shenandoah Telecommunications, *Form 10-K 2023*). Blue Ridge Bankshares' after-tax remediation expenses of roughly $3.6 million in 2024 were noticeable against its 2023 net loss but equaled only about 0.12 percent of total assets, so the consent order's main burden came from its operating constraints, not its direct cost (Blue Ridge Bankshares, *Form 10-K 2024*; Blue Ridge Bankshares, *Form 10-K 2025*).

A larger revenue base did not guarantee comparable insulation. Avon's roughly $135 million resolution equaled 1.77 percent of revenue but 14.06 percent of cash and 39.19 percent of its loss from continuing operations (U.S. Securities and Exchange Commission, "SEC Charges Avon"; Avon Products, *Form 10-K 2015*), and Estée Lauder's $210 million settlement equaled 1.32 percent of fiscal 2023 revenue but 12.13 percent of operating cash flow (Estée Lauder Companies, *Form 10-K 2026*; Estée Lauder Companies, *Form 10-K 2024*). Conversely, T-Mobile's 2024 FCC consent decree required a $15.75 million civil penalty and an additional $15.75 million investment in cybersecurity programs (Federal Communications Commission, *In the Matter of T-Mobile*), an amount comparatively absorbable given its substantial financial capacity (T-Mobile US, *Form 10-K 2025*). Wells Fargo likewise had substantial capacity (Wells Fargo & Company, *2018 Annual Report*), even though it faced a long-lasting operational restriction (Board of Governors, "Responding to Widespread"; Board of Governors, "Federal Reserve Announcements"). These cases suggest that the capacity to absorb legal costs cannot be read from revenue or market capitalization alone. Cash, operating cash flow, earnings, capital strength, and the type of consequence all shape the practical burden. The banking adaptations used in this project (total revenue, pre-tax income, and the Tier 1 capital ratio) also caution against comparing every ratio across banks and non-financial companies, because bank balance sheets and capital requirements differ structurally.

**Discussion**

The initial hypothesis—that greater financial exposure relative to company resources would correspond with greater negative market and financial consequences—was not supported consistently. Across this sample, Financial Exposure scores were low and narrowly distributed,

while the other dimensions varied more widely. This does not mean that fines are unimportant or that direct legal costs never threaten a company. It means that, for the selected cases, direct monetary costs often did not capture the main source of risk.

The findings support three cautious conclusions. First, a company's direct legal cost and the market's reaction should be measured separately. Stock prices reflect expectations about future cash flows, uncertainty, management credibility, business conditions, and many other factors; a simple event-window comparison cannot identify which factor drove a movement. Second, duration and operational effects should be evaluated explicitly. An indefinite restriction can shape business decisions for years without appearing as a one-time expense in an income statement. Third, company scale should be considered alongside financial conditions. A large revenue base can coexist with weak liquidity or declining earnings, while a smaller firm may be resilient in some dimensions and vulnerable in others.

The five-part score is best understood as a structured diagnostic tool, not a predictive model or legal judgment. Equal weighting makes cases easier to compare, but it assumes that each dimension contributes equally to overall risk. That is a practical design choice rather than an empirically validated relationship. In addition, the five dimensions are not fully independent: a restriction can affect operations, extend the duration of risk, and influence investor expectations simultaneously. The overall average can obscure these connections, so the component scores and case narratives remain essential.

**Limitations**

Several limitations constrain the conclusions. The sample contains only nine cases and was selected based on data availability and comparability; it is neither random nor representative. Eight cases involve government enforcement and only one involves private securities litigation, preventing a balanced comparison of legal mechanisms. Events also differ in date, underlying conduct, industry conditions, and the amount of information disclosed.

Second, market reaction is measured using short event windows and relative returns rather than a formal econometric event study (see MacKinlay 13–39). Pre-event trends, broad market volatility, overlapping company news, and investor anticipation may affect the observed returns. This is particularly important in cases such as Wells Fargo, Estée Lauder, and Blue Ridge Bankshares. Wells Fargo's post-announcement windows coincided with a broad market selloff in early February 2018, when the S&P 500 fell more than 4 percent in a single session on February 5 (Yahoo Finance, "Historical Data"). Estée Lauder's stock had already declined sharply over the preceding year, and the court noted five alleged corrective disclosure dates (*In re Estee Lauder*, Opinion and Order), so the Ledger's single event window captures only one of several candidate dates. Blue Ridge Bankshares' score depends on an extended window and a severe pre-event decline, making it among the less certain scores. The S&P 500 provides a consistent benchmark, but it cannot fully control for industry-specific or company-specific influences.

Third, the three-year financial window may miss consequences that develop later. Public filings also do not always separately quantify compliance expenses, reputational effects, foregone growth, or management time. For events with non-monetary consequences, company-disclosed estimates may be uncertain or unavailable. Finally, the scoring process includes qualitative judgments. Another researcher applying the same evidence could reasonably assign different scores, especially for regulatory severity and business impact. These limitations mean that the results should be treated as exploratory patterns that generate questions for larger studies, not as proof of causal relationships.

## Conclusion

A fine or settlement is an incomplete measure of the corporate risk created by a legal or regulatory event. In this nine-case study, direct financial exposure rarely corresponded consistently with market reaction or overall risk. Wells Fargo had the highest overall score, followed by Blue Ridge Bankshares and Estée Lauder at 3.2 each. The Wells Fargo and Blue Ridge cases involved multi-year restrictions and compliance obligations as their principal consequences, while Estée Lauder's severe market-reaction score despite low financial-exposure scoring, and Regions Financial's significant enforcement response alongside limited documented operational disruption, further demonstrate that corporate risk is multidimensional.

The project's central finding is therefore not that financial penalties are unimportant, but that they are insufficient on their own. A more complete assessment considers the scale of the cost relative to financial capacity, the severity of the legal response, the persistence of market underperformance, the duration of consequences, and the degree of operational constraint. Future research should expand the number of companies and industries, include more private-litigation cases, apply formal event-study methods, and test whether the framework produces consistent results when independent researchers score the same events. Until then, the Ledger offers a transparent starting point for evaluating corporate legal risk without confusing an observed association with proof of causation.

## Appendix A
## Event-Window Returns and Market Reaction Scores

Table A1

*Cumulative Return Relative to the S&P 500 Following Each Event, in Percentage Points*

| Company (ticker) | Disclosure date | Immediate | Primary | Subsequent | Extended | Market Reaction score |
|---|---|---|---|---|---|---|
| UTStarcom (UTSI) | 12/31/2009 | +1.92 (12/31/09) | +6.33 (1/4/10) | +2.79 (1/5/10) | — | 1/5 |
| Avon Products (AVP) | 12/17/2014 | +0.37 (12/17/14) | +0.09 (12/19/14) | −2.18 (12/22/14) | — | 2/5 |
| Shentel (SHEN) | 4/24/2023 | −2.27 (4/24/23) | −2.76 (4/25/23) | −2.69 (4/26/23) | — | 3/5 |
| T-Mobile US (TMUS) | 9/30/2024 | +0.16 (9/30/24) | +1.91 (10/2/24) | +1.29 (10/4/24) | — | 1/5 |
| Edgewell (EPC) | 2/3/2020 | +12.63 (2/3/20) | +19.43 (2/5/20) | +14.88 (2/7/20) | — | 1/5 |
| Estée Lauder (EL) | 11/1/2023 | −19.96 (11/1/23) | −16.74 (11/3/23) | −13.85 (11/7/23) | −13.17 (11/9/23) | 5/5 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Wells Fargo (WFC) | 2/2/2018 | −5.13 (2/5/18) | −7.28 (2/7/18) | −6.98 (2/8/18) | — | 4/5 |
| Blue Ridge Bankshares (BRBS) | 1/24/2024 | +0.58 (1/24/24) | +0.12 (1/26/24) | −3.23 (1/30/24) | −9.82 (2/1/24) | 4/5 |
| Regions Financial (RF) | 6/25/2014 | +0.46 (6/25/14) | +0.29 (6/27/14) | +0.50 (7/1/14) | +0.92 (7/3/14) | 1/5 |

*Note.* Each figure is the company's cumulative return minus the S&P 500's cumulative return from the closing price on the trading day before the disclosure date through the window end date shown in parentheses. Positive values indicate outperformance and negative values indicate underperformance. For Wells Fargo, the first window ends on the first trading day after the Federal Reserve's February 2, 2018 announcement. An extended check is reported for the three cases in which one was run (Estée Lauder, Blue Ridge Bankshares, and Regions Financial). Scores follow the rubric described in the Research Design and Methodology section, including the supporting-window persistence rule applied to Avon, Shentel, and Blue Ridge Bankshares. Blue Ridge Bankshares' score reflects a severe pre-event decline, so it is among the less certain scores in the sample.

*Source.* Independent calculations from daily stock prices (Yahoo Finance, "Historical Data").

## Appendix B

## Component Scores by Case

Table B1

*Risk Scores by Dimension (1–5)*

| Company | Field | Size | Financial Exposure | Regulatory Severity | Market Reaction | Duration of Risk | Business Impact | Overall |
|---|---|---|---|---|---|---|---|---|
| UTStarcom | Telecommunications | Medium | 2 | 4 | 1 | 4 | 3 | 2.8 |
| Avon Products | Cosmetics | Medium | 2 | 4 | 2 | 4 | 3 | 3.0 |
| Shentel | Telecommunications | Small | 1 | 3 | 3 | 2 | 2 | 2.2 |
| T-Mobile US | Telecommunications | Large | 1 | 4 | 1 | 4 | 3 | 2.6 |
| Edgewell | Cosmetics | Small | 1 | 4 | 1 | 2 | 3 | 2.2 |
| Estée Lauder | Cosmetics | Large | 2 | 3 | 5 | 4 | 2 | 3.2 |
| Wells Fargo | Banking | Large | 1 | 4 | 4 | 5 | 4 | 3.6 |
| Blue Ridge Bankshares | Banking | Small | 2 | 3 | 4 | 4 | 3 | 3.2 |
| Regions Financial | Banking | Medium | 2 | 3 | 1 | 2 | 2 | 2.0 |

*Note.* Overall score is the arithmetic mean of the five dimension scores. Higher scores indicate greater risk. Scores are from the Ledger, the author's unpublished research project.

**Works Cited**


Armour, John, et al. “Regulatory Sanctions and Reputational Damage in Financial Markets.” Journal of Financial and Quantitative Analysis, vol. 52, no. 4, 2017, pp. 1429–48, https://doi.org/10.1017/S0022109017000461.

Avon Products, Inc. Form 10-K for the Fiscal Year Ended December 31, 2015. U.S. Securities and Exchange Commission, 23 Feb. 2016, www.sec.gov/Archives/edgar/data/8868/000000886816000104/avp10k2015.htm.

Blue Ridge Bankshares, Inc. “Blue Ridge Bankshares, Inc. Announce 2025 Fourth Quarter and Full Year Results.” Exhibit 99.1, Form 8-K, U.S. Securities and Exchange Commission, 29 Jan. 2026, www.sec.gov/Archives/edgar/data/842717/000119312526029711/brbs-ex99_1.htm.

———. Form 10-K for the Fiscal Year Ended December 31, 2024. U.S. Securities and Exchange Commission, 10 Mar. 2025, www.sec.gov/Archives/edgar/data/842717/000095017025036301/brbs-20241231.htm.

———. Form 10-K for the Fiscal Year Ended December 31, 2025. U.S. Securities and Exchange Commission, 12 Mar. 2026, www.sec.gov/Archives/edgar/data/842717/000119312526104410/brbs-20251231.htm.

———. Form 8-K. U.S. Securities and Exchange Commission, 25 Jan. 2024, www.sec.gov/Archives/edgar/data/842717/000119312524015560/d870968d8k.htm.

Board of Governors of the Federal Reserve System. “Federal Reserve announces Wells Fargo Is No Longer Subject to the Asset Growth Restriction from the Board’s 2018 Enforcement Action against the Bank.” Press release, 3 June 2025, www.federalreserve.gov/newsevents/pressreleases/enforcement20250603a.htm.

———. “Responding to Widespread Consumer Abuses and Compliance Breakdowns by Wells Fargo, Federal Reserve Restricts Wells’ Growth until Firm Improves Governance and Controls.” Press release, 2 Feb. 2018, www.federalreserve.gov/newsevents/pressreleases/enforcement20180202a.htm.

Edgewell Personal Care Company. “Edgewell Personal Care and Harry’s, Inc. Comment on FTC Decision.” Press release, 3 Feb. 2020, ir.edgewell.com/news-and-events/press-releases/2020/02-03-2020-152557573.

———. Form 10-K for the Fiscal Year Ended September 30, 2021. U.S. Securities and Exchange Commission, 19 Nov. 2021, www.sec.gov/Archives/edgar/data/1096752/000109675221000032/epc-20210930.htm.

The Estée Lauder Companies Inc. “The Estée Lauder Companies Reports Fiscal 2024 First Quarter Results.” Exhibit 99.1, Form 8-K, U.S. Securities and Exchange Commission, 1 Nov. 2023, www.sec.gov/Archives/edgar/data/1001250/000100125023000175/fy2024q1exhibit991.htm.

———. Form 10-K for the Fiscal Year Ended June 30, 2024. U.S. Securities and Exchange Commission, 19 Aug. 2024, www.sec.gov/Archives/edgar/data/1001250/000100125024000116/el-20240630.htm.

———. Form 10-K for the Fiscal Year Ended June 30, 2026. U.S. Securities and Exchange Commission, 19 Aug. 2026, www.sec.gov/Archives/edgar/data/0001001250/000100125026000041/el-20260630.htm.

Federal Communications Commission. In the Matter of Shenandoah Telecommunications Company. Order, DA 23-337, 38 FCC Rcd 3975, 24 Apr. 2023, www.fcc.gov/document/eb-settles-911-outage-investigation-shentel-227200.

———. In the Matter of T-Mobile US, Inc. Order, DA 24-860, 30 Sept. 2024, docs.fcc.gov/public/attachments/DA-24-860A1.pdf.

Federal Trade Commission. "Statement of Daniel Francis, Deputy Director of FTC Bureau of Competition, Regarding Announcement That Edgewell Personal Care Company Has Abandoned Its Proposed Acquisition of Harry's, Inc." Press release, 10 Feb. 2020, www.ftc.gov/news-events/news/press-releases/2020/02/statement-daniel-francis-deputy-director-ftc-bureau-competition-regarding-announcement-edgewell.

Gande, Amar, and Craig M. Lewis. "Shareholder-Initiated Class Action Lawsuits: Shareholder Wealth Effects and Industry Spillovers." Journal of Financial and Quantitative Analysis, vol. 44, no. 4, Aug. 2009, pp. 823–50, https://doi.org/10.1017/S0022109009990202.

In re The Estee Lauder Co., Inc. Securities Litigation. Opinion and Order. No. 1:23-cv-10669-AS, ECF No. 138, U.S. District Court, Southern District of New York, 13 Oct. 2026. Labaton Keller Sucharow, edge.sitecorecloud.io/lksfirmsite98d9-labatonedc2-prodf3da-8e79/media/Project/Labaton/Master/Master-Site/Cases/Case-Materials/Resolved-Cases/In-re-The-Estee-Lauder-Companies/Estee-Lauder---Opinion-and-Order.pdf.

MacKinlay, A. Craig. "Event Studies in Economics and Finance." Journal of Economic Literature, vol. 35, no. 1, Mar. 1997, pp. 13–39.

Regions Financial Corporation. "Regions Financial Reaches Agreements with the SEC, Federal Reserve and Alabama State Banking Department to Resolve 2009 Loan Accounting Issue." Exhibit 99.1, Form 8-K, U.S. Securities and Exchange Commission, 25 June 2014, www.sec.gov/Archives/edgar/data/0001281761/000119312514249140/d748733dex991.htm.

Shenandoah Telecommunications Company. Form 10-K for the Fiscal Year Ended December 31, 2023. U.S. Securities and Exchange Commission, 21 Feb. 2024, www.sec.gov/Archives/edgar/data/354963/000162828024005781/shen-20231231.htm.

T-Mobile US, Inc. Form 10-K for the Fiscal Year Ended December 31, 2025. U.S. Securities and Exchange Commission, 11 Feb. 2026, www.sec.gov/Archives/edgar/data/1283699/000128369926000010/tmus-20251231.htm.

U.S. Securities and Exchange Commission. "SEC Charges Avon with FCPA Violations." Press release 2014-285, 17 Dec. 2014, www.sec.gov/news/press-release/2014-285.

UTStarcom Holdings Corp. Form 10-K for the Fiscal Year Ended December 31, 2010. U.S. Securities and Exchange Commission, 16 Mar. 2011, www.sec.gov/Archives/edgar/data/0001030471/000104746911002190/a2202701z10-k.htm.

Wells Fargo & Company. 2018 Annual Report to Shareholders. U.S. Securities and Exchange Commission, 27 Feb. 2019, www.sec.gov/Archives/edgar/data/72971/000007297119000227/wfc-12312018xex13.htm.

Yahoo Finance. "Historical Data." Yahoo Finance, Yahoo, finance.yahoo.com. Accessed 12 Oct. 2026.