JBET

Journal of **B**usiness, **E**conomics and **T**echnology ISSN: 2156-9444
Volume 29 Issue 1 Spring 2026

JOURNAL OF BUSINESS, ECONOMICS AND TECHNOLOGY
The Journal of the National Association of Business, Economics and Technology

(JBET is published online via the NABET.us website)

EDITORIAL NOTES

The continuing goal of the Journal of Business, Economics and Technology (JBET) is the publication of general-interest business and economics articles that demonstrate academic rigor, while at the same time are readable and useful to others in academia. Consistent with these goals, this and future issues of JBET presents authors' papers in the three research categories recommended by AACSB: Research that advances the knowledge of business and management theory (Theoretical), Research that advances the practice of business and management (Practice), and Research that advances learning/pedagogy (Pedagogical).

In addition to being *whitelisted* in Cabells Directory in the Journalytics category, JBET is also available through the EBSCO Host research database. The current acceptance rate for JBET is roughly 35%. In this regard we have striven to accept only high-quality research, while at the same time maintaining JBET as a realistic publishing outlet for Business, Economics and Information Technology faculty throughout the United States. Key to this process is our referees who have worked hard to help "grow" papers that have significant potential by providing authors with critical review comments. We generally require two to three rounds of review prior to accepting articles for publication. At the same time, we are attempting to shorten the average review time for each article to less than three months.

JBET Research Notes include, but are not limited to updates to previous work, additions to established methods, relatively short articles, research where the thesis is narrow in scope, null results, case series, research proposals, and data management plans: Articles of good quality which cannot be considered as full research or methodology articles. Further, articles in the Research Notes category have undergone the same double-blind peer review process as all articles that are published in JBET. At JBET, we support the research community across all of the disciplines of Business, Economics, and Information Technology by providing the Research Notes forum for sharing information and data of works-in-process or works that have a limited scope of application.

With the online publication of JBET, the editors have chosen to present JBET in a single column (margin-to-margin) instead of the traditional two-column presentation of an academic journal. We have done this to enhance readability in the online presentation of JBET.

The Editors thank the officers of the National Association of Business, Economics and Technology, the NABET Executive Board, as well as the referees for their support in the production of this 29th Volume of JBET.

Jerry D. Belloit, co-Editor
Clarion University of Pennsylvania – Retired

Norman C. Sigmond, co-Editor
Kutztown University of Pennsylvania

TABLE OF CONTENTS

MACHINE LEARNING TO MEASURE THE RISK OF SINGLE POINTS OF FAILURE IN ENDPOINT PROTECTION

Upakar Bhatta, Central Washington University

ABSTRACT

In today's interconnected digital landscape, understanding the risk single point of failure in endpoint protection is vital, as a failure in a centralized security component can have devastating consequences for the organization's entire security operations. This paper explores the application of machine learning to assess the risk of such failure in endpoint protection. It leverages logs from CrowdStrike service to construct a sample dataset. Key features in this dataset include sensor_online_days, agent_version, policy_drift_count, critical alerts, login_after_hours, update_failures, and protection_failure_risks. These features are used to train a model, enabling organizations to make informed decisions to enhance their overall security posture and operational excellence.

INTRODUCTION

A single point of failure (SPOF) in endpoint protection presents a significant risk in today's digital landscape, allowing attackers to exploit vulnerabilities in a centralized system and potentially compromise the entire enterprise. Centralized system misconfigurations, faulty software update, outdated software versions, lack of real-time monitoring, and reliance on single servers or single network components are key causes of single point of failure. CrowdStrike's faulty configuration update resulted in a widespread outage in July 2024 (Crowdstrike, 2024) Had CrowdStrike implemented real-time monitoring of Faclcon sensor updates, it could have triggered an alert when the update was pushed. Mercado Libre incorporated real-time monitoring while implementing a multi-cloud infrastructure strategy; as a result they maintained 100% uptime during the global Google Cloud outage on June 12, 2025 (Ibanez, 2025).

This underscores the importance of implementing real-time monitoring, avoiding reliance on a centralized update mechanism, and addressing single points of failure in system architecture design. This paper explores the application of machine learning to assess the risk of such failures in endpoint protection. It utilizes CrowdStrike service logs to construct a sample dataset. Key features in this dataset include sensor_online_days, agent_version, login_after_hours, critical alerts, update_failures, and protection_failure_risks. These features are used to train a model to predict the endpoint protection failures, enabling organizations to make a effective business decisions to enhance their overall security posture and operational excellence.

RELATED WORK

In Mishra (2021), the author introduces a theoretical model to defend against insider threat using algorithms and functional methods. In Chandel, et. al. (2019), the authors present a framework based on algorithms and functional techniques for detecting insider threats. In Zhang (2018), the author designed a system using directory virtualization techniques to detect insider threats. In Claycomb and Shin (2016), the authors employ machine learning for detecting insider threats. Previous studies focused on defending insider threats using endpoint detection tools. However, they did not analyze the effectiveness of these tools. This paper employs machine learning model to evaluate the effectiveness of end point detection tools than enables management team to make an informed decision.

PRE-REQUISITE KNOWLEDGE

Cloud Computing

Cloud computing is the technology that connects virtual resources over the internet to enable efficient data sharing. Cloud computing incorporates essential characteristics such as on demand self-service, broad network access, resource pooling, rapid elasticity, and measured service, as defined by the National Institute of Standards and Technology (NIST) (Mell and Grance, 2011). Modern companies are increasingly adopting cloud-based infrastructure, where endpoint detection and response tool plays a critical role in securing distributed environment. Cloud adoption includes service models such as Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service

(SaaS). It also involves deployment models such as public, private, community, and hybrid clouds, each designed to meet specific organizational requirements and preferences (Zhang, et.al., 2010).

The cloud computing services that the modern organizations leveraging today include:

- Infrastructure as a Service (IaaS): Provides highest level of management control over the infrastructure. In this service model, end point detection tools are installed directly on virtual machines, enabling agents to monitor cloud-hosted instances for malicious activities.
- Platform as a Service (PaaS): Provides customer with managed environment to deploy their own code. In this service model, end point tools are directly integrated into the runtime environment to monitor application activities and detect anomalies.
- Software as a Service (SaaS): Offers complete software solutions over the internet on a subscription basis. These services typically run on the virtual machines managed by the vendor, which are continuously monitored by end point detection tools to identify and respond to anomalous behavior.

Big Data Analytics

Big data analytics technology helps analyze the log data generated by agents installed on end devices. It involves examining large datasets constructed using the logs generated by end point detection tools, allowing machine learning developers to uncover hidden patterns and actionable insights. Companies can apply advanced analytical tools and technologies to study big data in depth and extract valuable business intelligence (Russom, 2011). Big data analytics implementation is a key component in the machine learning pipeline for identifying malicious activities through endpoint detection tools.

Machine Learning

Machine learning is a mathematical technique that uses algorithms that enable systems to make predictions by analyzing hidden patterns in datasets. Machine learning introduces a new approach for training algorithms with real-time datasets to identify specific patterns and anomalies in network traffic (Kumar, et.al., 2020). There are three main categories of machine learning—supervised learning, unsupervised learning, and reinforcement learning.

Supervised learning can be beneficial for predicting single point of failure in endpoint protection by training on labeled datasets. In unsupervised learning, the algorithm is trained on unlabeled data to predict a single point of failure in endpoint protection. In reinforcement learning, ml model is trained using trial and error learning sequences of normal and malicious events.

METHODOLOGY

This research paper explored machine learning techniques to analyze the risk of single points of failure in endpoint protection. The study demonstrated various machine learning algorithms to assess these risks. The experimental methodology employed in this research involved collecting CrowdStrike service logs to construct a sample dataset, preprocessing the dataset by employing cleaning, encoding and normalization, performing exploratory data analysis (EDA), splitting the dataset into training and testing segments, applying SMOTE, selecting and training machine learning models, predicting the output, measuring the accuracy, and evaluating the models.

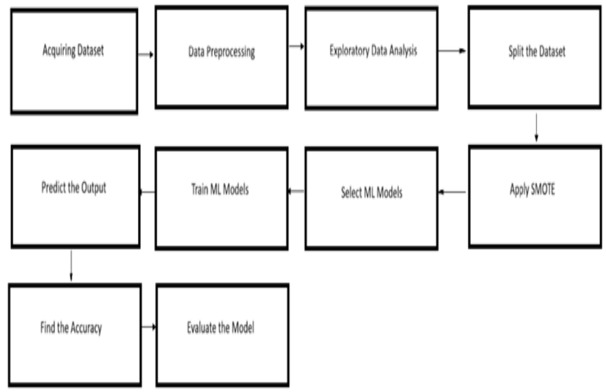


Figure 1. Machine learning implementation workflow

Acquiring Dataset: CrowdStrike service logs were used to construct a sample dataset consisting of 5000 records. Feature selection: A total of seven features were included: sensor_online_days, agent_version, policy_drift_count, critical alerts, login_after_hours, update_failures, and protection_failure_risks.

Data preprocessing: Data cleaning, encoding, and normalization were performed. Exploratory Data Analysis (EDA): Correlation heatmap was utilized to detect outliers, observe feature distributions, and identify relations between features.

Split the dataset: The dataset was divided into training (80%) and testing (20%) sets. Synthetic Minority Over-Sampling Technique (SMOTE) was applied to address the class imbalance in the dataset.

Model selection: Appropriate algorithms were selected to train the machine learning models. Model training: Five machine learning models were trained. Hyperparameter tuning and cross-validation were employed to optimize model performance and ensure robustness.

Model Accuracy and evaluation: Machine learning models were evaluated based on their accuracy.

DATA ANALYSIS

The dataset used in this research include numerical features, such as sensor_online_days, agent_version, policy_drift_count, critical alerts, login_after_hours, update_failures, and protection_failure_risks.

The exploratory data analysis methods applied in this research include:

Correlation heatmap: The correlation heatmap displayed below (-1: perfect negative correlation and +1: perfect positive correlation) highlights the relationships between different features in the sample dataset used in this research. Based on the correlation heatmap, the dataset features that are good to train machine learning model are:

- Strong positive correlation with the target, such as critical_alerts with correlation coefficient of +0.73, and update_failures with correlation coefficient of +0.70
- Strong negative correlation with target such as sensor_online_day with correlation coefficient of -0.98
- Moderate positive correlation with target such as policy_drift_count with correlation coefficient of +0.56

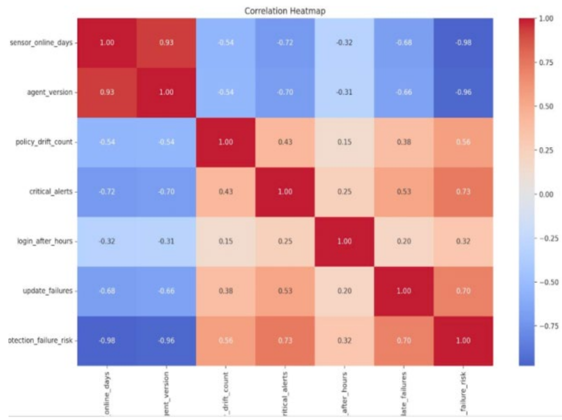


Figure 2. Correlation heatmap

Boxplot: Boxplot of sensor_online_days by class shows that Class 0 has no failure risks and higher uptime and class 1 with failure, while Class 1 has a failure risks and lower uptime.

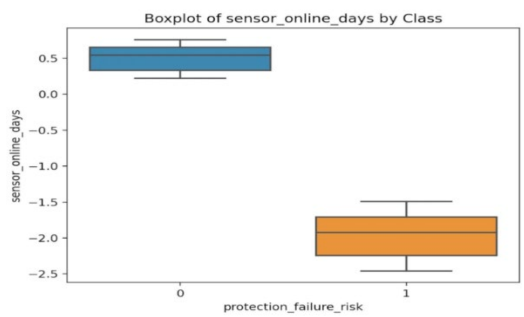


Figure 3. Boxplot of sensor_online_days by class

Boxplot of update_failures shows that Class 0 has no protection failure risks and few update failure, while Class 1 has protection failure risks and many update failure.

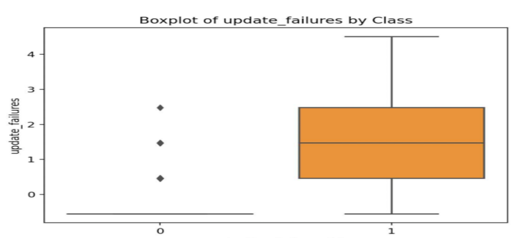


Figure 4. Boxplot of update_failure by class

Boxplot of policy_drift_count by class shows that class 1 have higher policy drift counts compared to Class 0. Higher policy drift counts suggest a potential correlation between protection failure and policy drift.

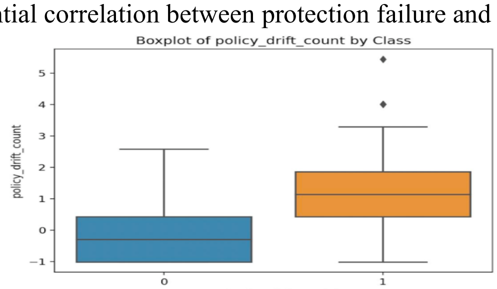


Figure 5. Boxplot of policy_drift_count by class

Feature distribution: This following figure shows the histograms of feature distributions for the dataset. The histogram of sensor_online_delays displays sensors that have recently come online (0.3 to 0.6) along with some older values (-2 to -1.5). The histogram of update_failures shows mostly successful systems update, with some cases of consistently failed updates. The histogram of policy_drift count indicates that policy drift is rare, with occasional instances of systems policy change. The histogram of agent_version shows the clusters of both older and new version. The histogram of login_after_hours indicates that login outside normal hours is rare. The histogram of critical_alerts suggests that such alerts are uncommon, but they can occur simultaneously.

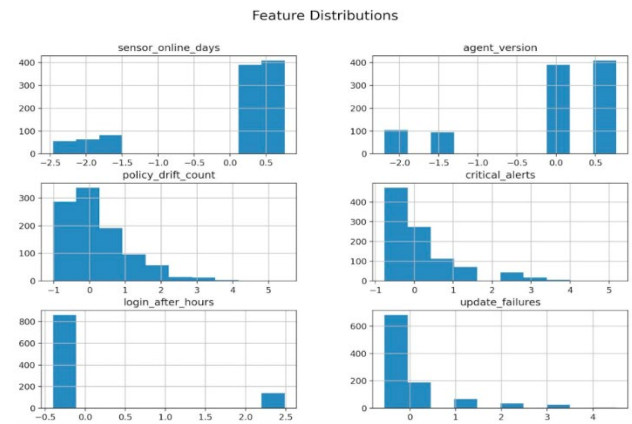


Figure 6. Histogram of feature distribution

SMOTE (Synthetic Minority Over-Sampling Technique): SMOTE technique has been applied to address class imbalance issues. It helps machine learning models avoid bias toward the majority class and improve overall performance.

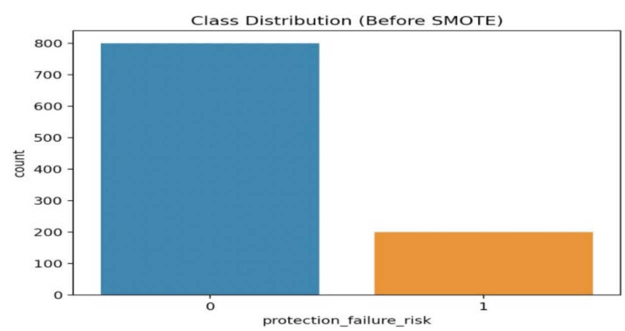


Figure 7. Class distribution before SMOTE

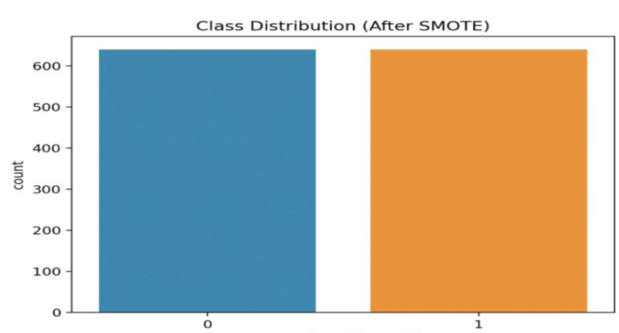


Figure 8: Class distribution after SMOTE

RESULTS

This paper leverages CrowdStrike service logs to construct a sample dataset, containing 5,000 labeled instances of observed attack outcomes across seven features: sensor_online_days, agent_version, update_failures, policy_drift_count, critical alerts, login_after_hours, and protection_failure_risks. These features were used to train a model designed to help organizations enhance their overall security posture and operational excellence. The machine learning model demonstrated strong performance, achieving an accuracy of 100% in assessing the risk of such failure in endpoint protection.

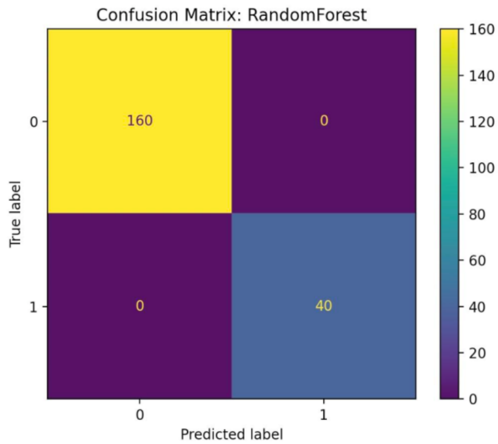


Figure 9. Confusion matrix random forest

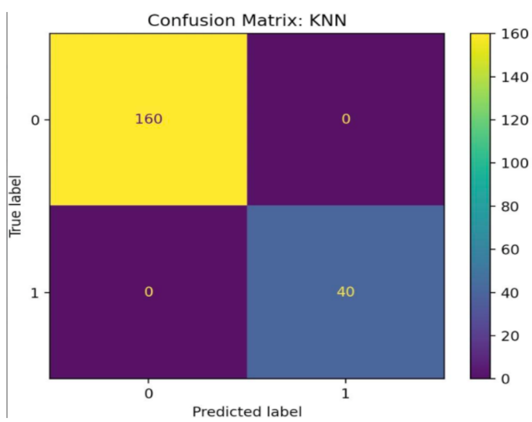


Figure 10. Confusion matrix KNN

Figure 11. Confusion matrix logistic regression

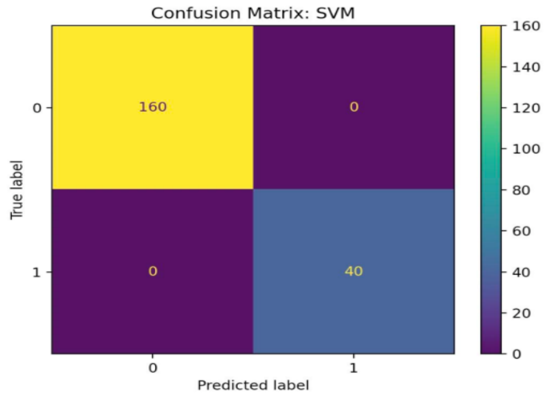


Figure 12. Confusion matrix SVN

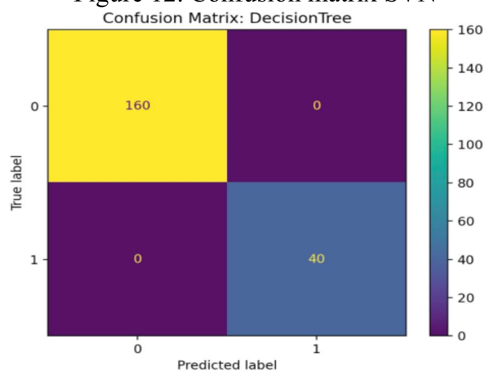


Figure 13. Confusion matrix decision tree

Table 1. Model Metrics

Random Forest	KNN	Logistic Regression	SVM	Decision Tree
TP= 40 TN= 160 FP= 0 FN= 0 A= 1.0 P= 1.0 R= 1.0 F1-Score= 1.0	TP= 40 TN= 160 FP= 0 FN= 0 A= 1.0 P= 1.0 R= 1.0 F1-Score= 1.0	TP= 40 TN= 160 FP= 0 FN= 0 A= 1.0 P= 1.0 R= 1.0 F1-Score= 1.0	TP= 40 TN= 160 FP= 0 FN= 0 A= 1.0 P= 1.0 R= 1.0 F1 Score= 1.0	TP= 40 TN= 160 FP= 0 FN= 0 A= 1.0 P= 1.0 R= 1.0 F1-Score= 1.0

$$Accuracy\ (A) = \frac{\mathrm{tp+tn}}{\mathrm{tp+tn+fp+fn}} \quad (1)$$

$$Precision\ (P) = \frac{\mathrm{tp}}{\mathrm{tp+fp}} \quad (2)$$

$$Recall\ (R) = \frac{\mathrm{tp}}{\mathrm{tp+fn}} \quad (3)$$

$F1 - Score = \frac{2*(p*r)}{p+r}$ (4)

TP = True Positive
TN =True Negative
FP = False Positive
FN = False Negative

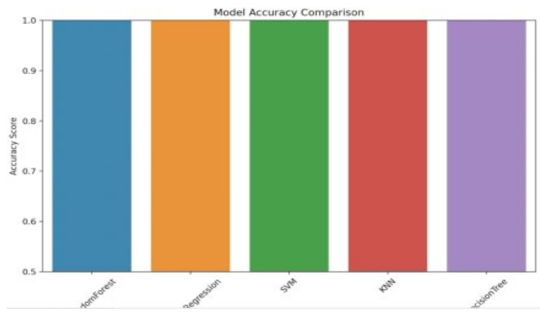


Figure 14. Model accuracy comparison

Based on the confusion matrix metrics, the machine learning model perfectly predicted all the samples.

DISCUSSION RESULTS

Precision (0.1): An ML model produced no false positive and is correct 100% of the time.
Recall (0.1): An ML model correctly identified all actual positive instances.
F1 Score (0.1): The F1 score indicated that the model accurately detected the positive class.

JUSTIFICATION

Random Forest was chosen for assessing risk of single points of failure in this research paper for following reasons:

- Random forest is known for its high accuracy and works well with the network security feature.
- Random Forest ensemble learning method is beneficial for this research as it improves performance compared to other model in detecting risks in endpoint detection.
- Random Forest can handle both numerical and categorical features without any extensive preprocessing as compare to other models.

CONCLUSION

In modern interconnected digital landscape, understanding single point of failure risks is vital for organizational security operations. This paper leverages machine learning to assess risks in endpoint detection. It utilized CrowdStrike service logs to construct the dataset with features including sensor_online_days, agent_version, login_after_hours, update_failures, and protection_failure_risks. These features are used to train a model, enabling organizations to enhance their security operations. With an accuracy of 100 percent, this paper demonstrates how effective ML technique can assist management decision-makers in assessing the risk of single points of failure in endpoint protection, ensuring security across enterprise infrastructure, and enhancing operational excellence.

REFERENCES

CrowdStrike Blog, August 6, 2024. Channel file 291 incident: Root cause analysis: Retrieved from: https://www.crowdstrike.com/en-us/blog/channel-file-291-rca-available

J. Ibanez, Multi-cloud architecture: How Mercado Libre stayed online during the massive google cloud outage, *Willdom Blog, 2025*. Retrieved from: https://willdom.com/blog/multi-cloud-architecture-how-mercado-libre-stayed-online-during-the-massive-google-cloud-outage/

P. Mell and T. Grance, The NIST definition of cloud computing, *Special Publication (NIST SP), National Institute of Standards and Technology, Gaithersburg, MD, 2011*. [Online]. Retrieved from: https://doi.org/10.6028/NIST.SP.800-145

Q. Zhang, L. Cheng, and R. Boutaba, Cloud computing: state-of-the-art and research challenges, *Journal of Internet Services and Applications, 1*(1), 2010. [Online]. Available: https://doi.org/10.1007/s13174-010-0007-6

P. Russom, Big data analytics, *TDWI Best Practices Report, Fourth Quarter*, 2011, 4-5, 22

V. Kumar, D. Sinha, A. K. Das, S. C. Pandey, and R. T. Goswami, An integrated rule-based intrusion detection system: Analysis on UNSW-NB15 data set and the real-time online dataset, *Cluster Computing, 23*, 1397–1418, 2020. https://doi.org/10.1007/s10586-019-03008-x

S. Mishra, Unsupervised learning and data clustering, towards data science, May 20, 2021. [Online]. Retrieved from: https://towardsdatascience.com/unsupervised-learning-anddata-clustering-eeecb78b422a

S. Chandel, S. Yu, Y. Tang, Z. Zhou, and Y. Huang, Cyber-enabled distributed computing and knowledge discovery, *2019 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery* (CyberC), 2019: https://doi.org/10.1109/CyberC.2019.00023

H. Zhang, An active defense model and framework of insider threats detection and sense, *International Conference on Information Assurance and Security*, *IEEE Computer Society, 2018*, 258–261.

W. R. Claycomb and D. Shin, Detecting insider activity using enhanced directory virtualization, *Proceedings of the 2016 ACM Workshop on Insider Threats*, *Chicago, Illinois, USA*, 2016, 29–36.

Upakar Bhatta is an assistant professor in the College of Education and Professional Studies at Central Washington University in Ellensburg, WA.

OPEN/CLOSED SYSTEMS AND COST-BENEFIT PRINCIPLE

Jeffrey Yi-Lin Forrest, School of Business, Slippery Rock University
Canchu Lin , School of Business Administration, University of South Carolina
Yong Liu, School of Business, Jiangnan University
Botond Géza Kálmán, Department of Finance and Accounting, John von Neumann University
Shahriar Gias, School of Business, Slippery Rock University

ABSTRACT

In theory, this paper attempts to improve the accuracy of the normative economics in addressing what ought to be in terms of cost-benefit analysis. To reach this goal, we focus on that (1) rationality is personally defined, (2) the meanings of "costs" and "benefits" go beyond numbers, and (3) the comparison between "costs" and "benefits" is dictated by individually defined rationalities. After briefly reviewing the all-you-can-eat restaurant example, this paper specifies three different kinds of rationalities and shows how each of them leads to dissimilar consumer behaviors. Based on the concepts of closed and open systems, the well-known cost-benefit principle is rephrased in two cases: one for closed system and the other for open system, in order to improve the accuracy of foretelling what ought to be.

INTRODUCTION

One well-applied tool in economic studies is the cost-benefit principle. It says (Frank etc., 2024, p2) that an individual (or a firm or a society) should take an action if, and only if, the extra benefits from taking the action are at least as great as the extra costs. By employing this principle, Frank, etc. (2024, p 9) theoretically investigated how much a person should eat at an all-you-can-eat restaurant. The conclusion drawn there is that the person would eat the same amount no matter whether he/she pays for the food or is offered to eat for free, assuming the person is rational. Contrary to this theoretical result, experimental evidence reveals that people who either pay for the food or offered to eat for free do not eat similar amounts, where paying guests tend to eat substantially more than those who eat for free (Thale, 1980). The former guests seem determined to get their money's worth. This inconsistency between theory and empirical evidence is what we would like to address in the rest of this paper.

Our focus of the following sections is placed on two words: "should," which is underlined above in the quotation of the cost-benefit principle; and "rational," which is also underlined but in the theoretical conclusion on the eating behaviors of fee-paying and free-eating customers.

By bestowing dissimilar meanings of rationality to different individuals, our analytical analysis reveals that fee-paying customers can demonstrate either undereating or overeating behaviors, while the same appears with customers who are offered to eat for free. The reason why we can bestow dissimilar meanings of rationality to different individuals is that it has been well documented that rationality is individually defined based on individuals' underlying system of values (Forrest, et al., 2021; 2022). Our analytical results not only show that both theoretical conclusion and experimental evidence, as mentioned above, are special cases of the general situation under consideration, but also reconfirm the realization (Forrest, et al., 2021; 2022) that being rational is self-defined instead of being judged by others (Mises, 1949).

Based on our established analytical results, this paper then turns its attention to refine or reformulate the cost-benefit principle for two separate and opposite scenarios – how the principle looks like within a closed system and how it reads in an open system,

To place our forthcoming discussions on the same footing, let us briefly outline the example on how much a person would eat at an all-you-can-eat restaurant (Frank, etc., 2024, p 9). A restaurant serves lunch buffet for $10 per customer who can refill his/her plate as many times as needed. One day, the restaurant offers its lunch buffet for free to 20 randomly selected guests. Now, a natural question arises: If all diners are rational, will those who are offered free lunch consume a different amount of food, on average, from those who pay the required $10?

According to the cost-benefit principle, the two groups of consumers would eat the same number of helpings, on average, under the assumption of rationality. The reason behind this conclusion is that after finishing their first helping, the decision to refill their plates is based on whether doing so would generate more benefits than the underlying costs. Evidently, refilling plates in the current situation can potentially lead to additional consumption utility while no additional costs will be incurred, no matter if a paying customer or a free diner is concerned with. In other words, the benefit of refilling plates is the same, on average, for people in both groups.

This restaurant setting and lunch will be generally referred to in the following sections unless otherwise described. And this presentation is organized as follows. In the second section, a literature review is given while pointing out how this work contributes to the existent knowledge. In the third section, it specifies three kinds of rationalities and provides relevant language-based predictions on consumer behavior. In the fourth section, it reconfirms the language-based predictions by using calculus-based analysis. In the fifth section, it uses the concept of closed and open systems to reformulate the well-known cost-benefit principle. This paper concludes in the last section.

RELEVANT LITERATURE

Theorizing about consumer behaviors begins with the normative theory, which assumes full rationality in decision making by human agents or consumers. So is this in the case of the cost-benefit principle and applications. However, numerous cases where consumers exhibit less rational or even irrational behaviors, which the normative theory fails to account for, led to exploration of alternative theories. Among such alternative theories is Thaler's (1980) positive theory of consumer choice, which attempts to describe consumer behaviors as they are observed. More importantly, there is a theoretical explanation for why the normative theory appears weak in accounting for those less rational or even irrational consumer behaviors. This theoretical explanation is provided by the concept of bounded rationality (Simon, 1957). Simson observed, "The capacity of the human mind for formulating and solving complex problems is very small compared with the size of the problems whose solution is required for objectively rational behavior in the real world – or even for a reasonable approximation to such objective rationality." (p.198) In realistic decision making, people tend to seek an outcome that is satisfying and sufficient instead of an optimal solution (Kahneman, 2003). This idea of decision making with bounded rationality not full rationality certainly applies to the case of utilizing the cost-benefit principle.

Apart from bounded rationality, intuition plays an important role in decision making as well (Kahneman, 2003). Kahneman (2003) took a situational view of decision making in that the decision-making behavior is not guided by reasoning but by intuition. More specifically, even in the case of consumers utilizing the cost-benefit principle in a decision-making scenario, their utilization of the principle is guided by what they happen to see, feel, and interpret at that moment. Further, this capability of seizing the situational characteristics is shaped by culture (Kahneman, 2003). Just as intuition, emotion also influences decision making as well (Ross & Nisbett, 1991). Further, decision making is also affected by norms (Akerlof, 1982). To this literature, the current work makes its contribution by showing that people's individual beliefs explain to a large degree why they behavior differently according to their underlying systems of values.

RATIONALITY AND CLASSIFICATION OF CUSTOMERS

The key for the theoretical reasoning to hold in the previous buffet lunch example is that all diners think alike or are rational in the eyes of the economist (Mises, 1949) – "costs" ≤ "benefits." However, in reality, the meanings of cost and that of benefit are defined either consciously or unconsciously by individual diners. One person's definitions of cost and benefits can be completely different or even opposite of those of another person. For example,

1) An individual believes in earning a decent income (and consequently a good living) through working hard.

Here, the cost is "work hard," while the benefit is "decent income." Due to differences in individuals' systems of values,

2) Another person could hold his/her principle of living as working as little as possible, while making as much income as possible.

That is why an income of $30,000.00 can be either greater than that of $3 million or less than $3 million, depending on how the income is earned. For instance, for situation 1), if $30,000.00 is the earning from a well-recognized job, while $3 million is the income from robbing a bank, then a person as described in situation 1) would most likely treat $30,000.00 as greater than $3 million. At the same time, a person as described in situation 2) would most likely treat $3 million as greater than $30,000.00. One conclusion we can naturally draw from this discussion is that a person's rational thinking can be seen as totally irrational in another person's mind. For more in-depth discussions on this end, please consult with Forrest (2024).

Going alone with this review of recent studies on rationality and preparing for the forthcoming deliberations, in the rest of this section, we will classify the restaurant diners into three groups:

A. Diners who have concerns about the limited supply of food.
B. Dinners who have generosity (or philanthropy) concerns; and
C. Diners who have cost concerns.

In other words, we specify diners into three discrete kinds of rationalities, where people behave differently according to how they view the world individually.

To be more specific, let us describe the diners in the three groups as follows: For group A, fee-paying customers think that those free eaters are additional eaters who make what is otherwise enough food less for them to eat. For group B, fee-paying customers feel the need to save food for free eaters as they believe the free eaters tend to be poor and rather spend the lunch money on other necessities of life. For group C, diners feel the cost of the lunch is either negligibly low or expensive.

To make our reasoning go smoothly, let us assume that our discussion in this and the following section takes place within a closed system. Specifically, by a system, it means (Lin, 1999) an ordered pair $S = (M, R)$ of two sets, where set M contains all elements of concern while set R enumerates associations or relationships about how various elements in M relate to one another. That is, we have the following:

$$M = \{e_1, e_2, e_3, \dots, \} \text{ and } R = \{a_1, a_2, a_3, \dots\}$$

where e_i represents element $i \in M$ (i= 1, 2, 3, …) of the system S and $a_j \in R$ (j = 1, 2, 3, …) stands for an association about how elements in M are related to each other

To predict the possible direction of evolution of the system S, the associations a_1, a_2, a_3, ... are closely studied. If none of the associations involves factors outside set M, then S we study is known as a closed system. If some a_j involves at least one factor from the environment, then S is seen as an open system (Forrest, 2024).

Now, let us make language-based predictions on how diners in each of group A, B and C would behave in terms of how much they eat compared to how much they normally do. To do this, we assume that the decision on how much to eat is made within a closed system, meaning that no other factors come into consideration of the diners. Because of this assumption, our predicted outcomes are stated with "will" instead of "should" or "would", as the ordinary practice in studies of economics.

For Group A, we forecast that the fee-paying customers will eat fast and slightly overeat, because somehow the food supply, seen as just enough for fee-paying customers, will soon be cut short by the free eaters. Here, a sense of urgency and competition exists. At the same time, free eaters know that fee-paying diners think that they are additional people who make what is otherwise enough food less for the fee-paying diners to eat. Hence, each of the free eaters will eat more than his/her conventional amount. It is because these free eaters feel insulted.

For group B, the fee-paying customers will eat less than they normally do, because they feel the need to assist the free diners to have enough food to eat. On the other hand, the free eaters will eat as much food as they normally do.

For group C, if the cost of the lunch is seen as negligibly low, no matter whether a customer must pay for the lunch or not, he/she will eat no more than needed. In other words, when the cost is not a concern, both fee-paying customers and free eaters will eat no more than their bodily needs, because no or very little cost is involved. For the case that the cost of the lunch is seen as expensive,
fee-paying customers will eat more than they normally do, because they want to get their money's worth.

In the next section, we will reestablish similar conclusions by using rigorous calculus-based analysis. Once again, we assume that the decision on how much to eat is made within a closed system, where no outside factors come into play.

CUSTOMERS EATING BEHAVIORS

This part of the paper focuses on the analytical confirmation of the predictions made in the previous section on the consumption behaviors of all the classified customers. To this end, let d_p be a diner who pays for his/her lunch and d_f a customer who is offered free lunch. Let D_p be the collection of all fee-paying diners and D_p the collection of all diners who are offered fee lunch.

Concerns with Limited Food Supply

First, we consider case 1: Fee-paying customers think that the free eaters would be additional eaters who will make what is otherwise enough food less for them to eat. In terms of literature, in the field of consumer behavior research, there has been some persistent interest in the topic of scarcity. One thread of the literature on scarcity research conceptualizes scarcity as a mindset (Goldsmith, Griskevicius, & Hamilton, 2020). Basically, this concept of scarcity mindset indicates that scarcity of resources such as time and money can lead to changes in consumers cognitive processing (Mullainathan & Shafir, 2013). Mullainathan and Shafir (2013) proposed that consumers' perception of a resource as being scarce would lead to more efficient use of it and thereafter enhanced productivity. More importantly, Cannon, Goldsmith, and Roux (2019) found that scarcity mindset leads to compensatory consumption (Mandel, Rucker, Levav, & Galinsky, 2019), which can conveniently apply to our study in hands. The entry of free eaters into the restaurant would trigger this scarcity mindset in those paid eaters and then the action of compensatory consumption, disproportionately eating more food there. In the rest of this subsection, we will analytically see if this conclusion is correct.

Let u_p be the utility of diner $d_p \in D_p$, who paid \$$p$ for the lunch, acquired from consuming x_p amount of food. Then, the utility function u_p can be modeled as follows:

$$u_p = \alpha_p (I_p - x_p)^{m_p} - p + \left(S - \sum_{d_f \in D_f} x_f \right) \tag{1}$$

where α_p is a positive coefficient, m_p an odd number exponent great than 1, I_p the conventional amount[1] of food d_p can comfortably consume and required to sustain him/herself as a biological being, S the total food supply available for the said lunch, prepared by the restaurant, x_f the amount of food $d_f \in D_f$ consumes.

Specifically, assume that I_p is a constant, while $(I_p - x_p)$ reflects the fact that the utility u_p from eating the lunch is the greatest when d_p just starts to eat his/her food, that is when x_p takes a small value. As an increasing amount x_p of food is consumed, d_p's marginal utility begins to decrease although the utility is still positive, while the value of $(I_p - x_p)$ drops. That aligns with the principle of diminishing marginal utilities. The logical reason for $(I_p - x_p)$ to raise to power m_p (> 1) is that when x_p takes a small value, the food consumption x_p creates elevated joy; when x_p approaches the conventional amount I_p, the utility u_p diminishes quickly. To reflect the fact that overeating can realistically cause pain to the diner, m_p needs to be an odd number. In this case, $\alpha_p (I_p - x_p)^{m_p}$ is a negative number.

[1] In real life, this conventional amount is an interval grey number (Liu and Lin, 2006). That is, I_p is equal to a grey number (a, b), for some real numbers a and b $(a < b)$, such that although it is known that $I_p \in (a, b)$, the specific value of I_p is unknown until the lunch is over. In other words, the value of I_p is not fixed in real life. While feeling full, a person can still take additional bites of food without suffering from the bad consequence of overeating. To make our reasoning in this paper go smoothly, we assume that for each diner $d_i \in D_p \cup D_f$, the value of I_i is a fixed real number.

The third term in equation (1) reflects how the total amount of food consumed by the diners who are offered free lunch decreases the value of d_p's utility. This end captures the situation in our current scenario: customer d_p thinks that the free eaters would be additional eaters who will make what is otherwise enough food less for him to eat.

Because the values of I_p, x_p, $\alpha_p(I_p - x_p)^{m_p}, p$ and $\sum_{d_f \in D_f} x_f$ are very small when compared to the total supply S of food prepared by the restaurant, the value of u_p in equation (1) is great than 0. That is, for each fee-paying diner $d_p \in D_p$, his/her utility of eating lunch in the restaurant is positive. In other words, the lunch is enjoyable.

Differentiating u_p in equation (1) twice with respect to x_p provides

$$\frac{\partial^2 u_p}{\partial x_p^2} = \alpha_p m_p (m_p - 1)(I_p - x_p)^{m_p - 2} \quad (2)$$

Now, the principle of diminishing utility implies that $\partial^2 u_p / \partial x_p^2 < 0$. That is, we have from equation (2) that

$$\alpha_p m_p (m_p - 1)(I_p - x_p)^{m_p - 2} < 0 \quad (3)$$

Because $\alpha_p m_p (m_p - 1) > 0$ and $m_p - 2 > 0$, equation (3) implies that $I_p < x_p$.

This analytical analysis shows the following conclusion:

Proposition 1. If fee-paying customers believe that free eaters are additional people who make less food for them to consume, then these fee-paying customers will overeat.

Next, let us consider the second case: Free eaters think that offering them free food would make less food available for those who have paid for it. Similar to equation (1), the utility function u_f of free eater $d_f \in D_f$ can be written as follows:

$$u_f = \alpha_f (I_f - x_f)^{m_f} + \left(S - \sum_{d_f \in D_f} x_f \right) \quad (4)$$

where α_f is a positive coefficient, m_f an odd number exponent great than 1, I_f the conventional amount of food d_f can comfortably consume and required to sustain him/herself as a biological being, S the total food supply available for the said lunch, prepared by the restaurant.

Specific explanations of I_f, $(I_f - x_f)$, m_f, $\alpha_f(I_f - x_f)^{m_f}$, and $\left(S - \sum_{d_f \in D_f} x_f\right)$ are similar to those of equation (1).

Differentiating u_f in equation (4) once with respect to x_f gives

$$\frac{\partial u_f}{\partial x_f} = -\alpha_f m_f (I_f - x_f)^{m_f - 1} - 1 \quad (5)$$

Because for most people eating is pleasant, we have $\partial u_f / \partial x_f > 0$. So, equation (5) implies that

$$-\alpha_f m_f (I_f - x_f)^{m_f - 1} - 1 > 0 \quad (6)$$

which leads to

$$\left(I_f - x_f\right)^{m_p - 1} < 0 < \frac{-1}{\alpha_f m_p}$$

Therefore, $I_f - x_f < 0$. In other words, we have the following conclusion:

Proposition 2. The thinking of fee-paying diners that free eaters would make less food available for them to eat stimulates free eaters to eat more than their conventional amounts.

The following table summarizes all the symbols introduced above.

p	The price of the lunch
S	Total food supply the restaurant prepared for the lunch
d_p	A diner who pays for the buffet lunch
d_f	A diner who is offered the buffet lunch for free
d_i	A diner who either pays for the lunch or is offered the lunch for free
D_p	The collection of all fee-paying diners who eat at the restaurant
D_f	The collection of all diners who are offered a free lunch
I_p	The conventional amount of food d_p must eat to satisfy his/her bodily needs
I_f	The conventional amount of food d_f must eat to satisfy his/her bodily needs
I_i	The conventional amount of food d_i must eat to satisfy his/her bodily needs
x_p	The amount of food d_p consumes
x_f	The amount of food d_f consumes
x_i	The amount of food d_i consumes
u_p	The utility of d_p acquired from consuming x_p amount of food
u_f	The utility of d_f acquired from consuming x_f amount of food
u_i	The utility of d_i acquired from consuming x_i amount of food

Concerns with Generosity and Philanthropy

For the third case that the fee-paying customers think that the free eaters tend to be poor and they could spend the savings incurred from the free lunch on other necessities in life. Such thinking, although it may very well be false, falls in the research of theoretical foundation of generosity or philanthropy concerns, an area of scarcity studies. This literature suggests that socioeconomic status can moderate scarcity's impact on consumers decision making (Goldsmith, Griskevicius, & Hamilton, 2020). When faced with a scarcity threat, consumers in a higher socioeconomic status (raised in a relative richer environment) tend to take a long-term approach to cope with scarcity and thus can accumulate the lacked resources to a surplus level, whereas those consumers raised in a lower socioeconomic status are more likely to cope with the resource that is scarce at present. According to this literature, for our current case of concern, paid eaters, if they are raised in a relatively wealthy family, are more likely to exercise generous or philanthropic behavior (eating less so as to leave more food to those free eaters who, they believe, are from a lower socioeconomic environment). Unlike the paid eaters, the free eaters who may or may not be raised in a lower socioeconomic status are likely to eat normally. In the rest of this subsection, we employ calculus-based approach to enrich which has been confirmed in the literature.

Let us express the utility u_i of the fee-paying diner $d_i \in D_p$ as follows:

$$u_i = (1 + I_i)\ln(1 + x_i) - p + \left(S - \sum_{d_p \in D_p} x_p\right) \tag{7}$$

where the first term on the left-hand side shows the interaction between the amount x_i of food consumed and the conventional amount I_i, which is an increasing function of x_i. The reason why we use $(1 + x_i)$ is to make sure that $\ln(1 + x_i) > 0$, when $x_i > 0$. The third term reflects the assumption that the more fee-paying diners eat, the less food will be available for those who are offered free lunch. Hence, the total $\sum_{d_p \in D_p} x_p$ adversely affects the utility u_i of each fee-paying customer. The logical reasoning behind why we use $\left(S - \sum_{d_p \in D_p} x_p\right)$ instead of simply $\left(-\sum_{d_p \in D_p} x_p\right)$ is to guarantee that this second term stays positive.

Now, let us differentiate u_i in equation (7) with respect to x_i so that we have

$$\frac{\partial u_i}{\partial x_i} = \frac{1 + I_i}{1 + x_i} - 1 \tag{8}$$

From $\partial u_i / \partial x_i > 0$ and equation (8), we have

$$\frac{1 + I_i}{1 + x_i} - 1 > 0$$

That is, $1 + I_i > 1 + x_i$. So, $I_i > x_i$.

By summarizing this analysis, we have:

Proposition 3. If the fee-paying customers think that the free eaters tend to be poor and they could spend the savings acquired from the free lunch on other necessities in life, then each of these fee-paying customers eat less than he/she would normally do.

Corresponding to what is considered above, the utility u_f of free eater $d_f \in D_f$ can be such a function $u_f = u_f(x_f)$ that satisfies $\partial u_f(x_f)/\partial\, x_f > 0$ and $\partial^2 u_f(x_f)/\partial\, x_f^2 < 0$, when fee-paying customers think that they are poor and could use the savings incurred from the free lunch on other necessities in life. In other words, the utility function $u_f = u_f(x_f)$ is not constrained by any other conditions. Therefore, we can conclude the following:

Proposition 4. No matter how the fee-paying customers think about the free eaters, the free eaters will eat as much food as they normally do.

Cost Concerns

In this subsection, we consider respectively two scenarios: (1) the cost of the lunch is negligibly low; and (2) the expense of the lunch is considered costly, although the total value is not necessarily high. The socioeconomic status of a person plays a role in his/her perception of cost. For those with a higher socioeconomic status, they may perceive the price of the buffet lunch as minimally low and thus would not eat more or less than normal. For those with a lower socioeconomic status, regardless of whether they are free eaters or paid eaters, they would tend to eat more. In this case, the cost-benefit principle applies as it has been widely done.

As in the previous subsections, to take advantage of calculus, this subsection assumes all conditions needed for differentiations to take place.

For the first scenario, since the cost of the lunch is negligible, we can naturally model the situation in such a way that the lunch price p neither increases nor decreases the utility function acquired from consuming the food of the restaurant. For any diner $d_i \in D_p \cup D_f$, let his/her utility acquired from consuming the amount x_i of food be the following function om two variables

$$u_i = u_i(I_i - x_i, p) \tag{9}$$

satisfying

$$\frac{\partial u_i(I_i - x_i, p)}{\partial x_i} \begin{cases} < 0, & \text{if } x_i < I_i \\ > 0 & \text{if } x_i > I_i \end{cases} \tag{10}$$

and

$$\frac{\partial^2 u_i(I_i - x_i, p)}{\partial x_i^2} > 0; \frac{\partial u_i(I_i - x_i, p)}{\partial p} = 0 \tag{11}$$

The reason we have the constraint in equation (10) is that I_i is, by definition, the amount of food needed by the body of d_i. Hence, when the amount x_i of food is less than the conventional amount I_i, the marginal utility of consuming x_i is positive. Due to other concerns, such as health, when the consumed amount x_i of food is more than the bodily need I_i, instead of having enjoyment of eating, the diner experiences disutility.

From equation (10), it follows that u_i reaches its minimum when $x_i = I_i$ for any fixed p value. In other words, we have

$$\left.\frac{\partial u_i(I_i - x_i, p)}{\partial x_i}\right|_{x_i = I_i} = 0 \tag{12}$$

The Maclaurin series expansion (Kolk and Duistermaat, 2010) of the utility function in equation (9) looks as follows:

$$u_i(v,p) \quad = \sum_{k=0}^{\infty}\sum_{i=0}^{k} \frac{1}{(k-i)!\,i!} \frac{\partial^k u_i(0,0)}{\partial v^{k-i} \partial p^i} v^{k-i} p^i \tag{13}$$

$$= u_i(0,0) + \frac{\partial u_i(0,0)}{\partial v} v + \frac{\partial u_i(0,0)}{\partial p} p +$$

$$+ \frac{1}{2}\frac{\partial^2 u_i(0,0)}{\partial v^2} v^2 + \frac{1}{2}\frac{\partial^2 u_i(0,0)}{\partial p^2} p^2 + \frac{\partial^2 u_i(0,0)}{\partial v \partial p} vp + \cdots \tag{14}$$

where $v = I_i - x_i$ stands for the first variable of the function u_i.

Assume that the utility function $u_i(v,p)$ is equal to the quadratic approximation in equation (14). Then, we have

$$\frac{\partial u_i(I_i - x_i, p)}{\partial x_i} = -\frac{\partial u_i(0,0)}{\partial v} - \frac{\partial^2 u_i(0,0)}{\partial v^2}(I_i - x_i) - \frac{\partial^2 u_i(0,0)}{\partial v \partial p} p \tag{15}$$

Because

$$\frac{\partial u_i(I_i - x_i, p)}{\partial x_i} = \frac{\partial u_i(I_i - x_i, p)}{\partial v}\frac{\partial v}{\partial x_i} = -\frac{\partial u_i(I_i - x_i, p)}{\partial v} \tag{16}$$

Equation (12) implies that

$$\frac{\partial u_i(0, p)}{\partial v} = 0, \text{for any } p \tag{17}$$

Similarly, we can show based on equation (11) that

$$\frac{\partial^2 u_i(I_i - x_i, p)}{\partial v^2} = \frac{\partial^2 u_i(I_i - x_i, p)}{\partial x_i^2} > 0 \tag{18}$$

Since p is negligibly low, we let $p \approx 0$ so that equation (15) becomes

$$\begin{aligned}\frac{\partial u_i(I_i - x_i, p)}{\partial x_i} &= -\frac{\partial u_i(0,0)}{\partial v} - \frac{\partial^2 u_i(0,0)}{\partial v^2} v \\ &= -\frac{\partial^2 u_i(0,0)}{\partial v^2}(I_i - x_i) < 0 \qquad \because \text{equation (17)}\end{aligned}$$

From equation (18), it follows that

$$x_i < I_i \tag{19}$$

Proposition 5. If the cost of the lunch is negligibly low, no matter whether the diner must pay for the lunch or not, he/she will eat no more than his/her bodily need.

Next, let us look at the second scenario – the expense p of the lunch is considered costly, although the value p is not necessarily a big number. In other words, the utility function u_p of any fee-paying customer $d_p \in D_p$ is a decreasing function of p, the price of the lunch, and the amount x_p of food consumed is also a decreasing function of p. This last statement reflects the fact that if the cost of the lunch gets too high, diner d_p would not eat at the restaurant. In that case, $x_p = 0$. Although the lunch price p is a fixed real number, it is subjectively considered expensive by fee-paying individuals $d_p \in D_p$. Hence, when price p is seen as a subjective figure, it can be righteously treated as a variable with different values of expensiveness. Hence, symbolically, we have

$$\frac{\partial u_p(v, p)}{\partial p} < 0; \ \frac{\partial x_p}{\partial p} < 0 \tag{20}$$

Assume that the utility function u_p, as given in equation (9), is equal to the following quadratic approximation, as given in equation (14):

$$u_p(v, p) = u_p(0,0) + \frac{\partial u_p(0,0)}{\partial v} v + \frac{\partial u_p(0,0)}{\partial p} p + \frac{1}{2}\frac{\partial^2 u_p(0,0)}{\partial v^2} v^2 \tag{21}$$

Then, differentiating this function u_p with respect to p produces

$$\frac{\partial u_p(v,p)}{\partial p} = \frac{\partial u_p(0,0)}{\partial v}\frac{\partial v}{\partial p} + \frac{\partial u_p(0,0)}{\partial p} + v\frac{\partial^2 u_p(0,0)}{\partial v^2}\frac{\partial v}{\partial p} \tag{22}$$

From equations (12), equation (22) can be rewritten as

$$\frac{\partial u_p(v,p)}{\partial p} = \frac{\partial u_p(0,0)}{\partial p} - v\frac{\partial^2 u_p(0,0)}{\partial v^2}\frac{\partial x_p}{\partial p} \tag{23}$$

From equations (20) and (23), it follows that

$$-v\frac{\partial^2 u_p(0,0)}{\partial v^2}\frac{\partial x_p}{\partial p} < 0$$

Hence, we have based on equation (20)

$$\frac{\partial u_p(0,0)}{\partial p} < 0 < \frac{\partial^2 u_p(0,0)}{\partial v^2}\frac{\partial x_p}{\partial p}\left(I_p - x_p\right)$$

Therefore, equations (18) and (20) imply that $I_p - x_p < 0$. That is, we have shown the following conclusion.

Proposition 6. If the expense of the lunch is considered costly, fee-paying customers will eat more than their conventional needs.

This result can be illustrated as follows. The overeating, as concluded in Proposition 6, is the consequence that the fee-paying diners want to get back their money's worth, because the lunch price is seen as expensive and costly.

REFORMULATION OF THE COST-BENEFIT PRINCIPLE

The derived results of the previous section demonstrate vividly that first, the theoretical conclusion of the all-you-can-eat restaurant, as outlined in the first section above, holds true only conditionally while not as general as stated; and second, the empirical discovery regarding the restaurant situation, as outlined in Thale (1980), once again represents certain special instances instead of the norm. Although one might employ the differences between normative and positive economics (Keynes, 1890) to comfort him/herself about these discrepancies between special cases and generalities, we like to improve the capability for the normative economics to forecast with better accuracy about what ought to be.

To accomplish this goal, let us first focus on scenarios of closed systems, where what is mentioned is everything involved in the system of discussion. Because of this constraint or focus of attention, the word "should," as underlined in the cost-benefit principle given in the first section above, can be readily removed.

Cost-Benefit Principle (Closed systems): An individual (or a firm or a society) takes an action if the benefits from taking the action are at least as great as the associated costs of taking the action.

Note 1: (1) This principle applies to such a closed system that other than the mentioned action, the individual (or a firm or a society) is not involved in or considering any other actions. (2) In this modified principle of cost and benefit, emphasis needs to be placed on the "if" condition. Specifically, other than monetary rewards, such as wages and fringe benefits, both terms of "benefits" and "costs" may include non-monetary things, such as prestige, social status, strengthened personal ties, etc. Additionally, the "if" condition includes an order relation, as stated in "as great as." That is, somehow "benefits" and "costs" can be ranked either numerically or ordinally or both. In other words, in applications of this principle, the details of this "if" condition are interpreted by individual rationalities.

One immediate example that illustrates this note is the aforementioned comparison between $30,000 and $3 million. For the situation when $30,000 > $3 million, the costs associated with the former figure include time consistently spent on and effort dependably devoted to a recognized job, while the benefits contain a promised wage, health insurance, and other perks. For the second figure, the associated costs include intensive mental pressure, experienced both before and after the robbery, extremely difficult work environment (that exists when the robbery takes place), and a high-level readiness for fleeing from police capture and from other robbers. The major benefit is the chance of holding and enjoying a large amount of money, although the time of holding and enjoyment might be very short lived.

As our second example, we can look at the existence of the firms that their purpose of operations is not to make more profits than the costs. Instead, these firms hire primarily disabled people. In this scenario, the benefits comprise of such things as creating hopes for these unfortunate individuals, while providing them with a quality livelihood. The costs include constant injection of funds, the creation of conveniences necessary for these disable employees, etc.

The cases investigated in the previous section demonstrated how individually defined rationalities can very well lead to different behavioral patterns. These carefully considered scenarios bring us to where we are now: The cost-benefit principle needs to be phrased respectively for closed and open systems. For the latter case, we have

Cost-Benefit Principle (Open systems): An individual (or a firm or a society) takes an action if the difference between the extra benefits and extra costs from taking the action is positive and the greatest compared to those of taking other actions.

Note 2: As discussed in Note 1, the "if" condition involves individually defined "benefits," "costs," and an order relation. The reason why this principle applies to open systems is that other than the action of concern, the costs and benefits of taking this action are also compared to those of potentially taking other actions. In other words, when the particular action is the only one possible for the individual to consider taking, then we deal with a closed system; otherwise, we content with an open system.

To help detail what we are talking about, let us look at the following example. One well-recognized long-standing convention in the business world is that firms' objective is to maximize their profits, as economics textbooks implicitly indicate (Wu, 2006). Although the action emphasized here is to maximize profits for shareholders above all else, a group of executives have turned their attention to allowing each employee to succeed through hard work and creativity so that he/she can live a life of meaning and dignity (see Statement on the Purpose of a Corporation, accessed on October 30, 2025). That is, there are managers and executives who consider their employees' well-beings as one of the "benefits" of running business. Correspondingly, expenses incurred from creating well-beings are parts of the "costs" of their companies.

SOME LAST WORDS

This work investigates why the cost-benefit principle in theory does not produce conclusions that can be confirmed by empirical evidence through visiting the well-known all-you-can-eat-restaurant example. By identifying three different kinds of rationalities and by employing the concept of close and open systems, this research is able to analytically confirm, as previously expected in Forrest (2024), that individually defined rationalities indeed lead to dissimilar consumer behaviors. Consequent to this confirmation, we reformulate the cos-benefit principle for two cases: one for closed systems and the other for open systems.

Relevant to our topic of discussion in this paper, a large body of literature shows systematic deviations from rational cost–benefit analysis. For example, behavioral biases (Kahneman & Tversky, 1974), framing effects (Tversky & Kahneman, 1981), time inconsistency (Laibson, 1997), and social preferences (Fehr & Gächter, 2000) all contribute to discrepancies between theoretical predictions and observed behavior. At the same time, prospect theory (Kahneman & Tversky, 1979) further demonstrates that individuals distort probabilities, while the sunk cost effect (Thaler, 1980) leads to the fact that past expenditures influence current decisions despite their irrelevance in standard theory. In comparison, our current work provides a brand-new angle to look at how systematic deviations from rational cost–benefit analysis can naturally arise from the variations in individual systems of values.

This work in fact opens a rich and large territory for prolific future research. To be specific, this paper merely focuses on the cost-benefit principle. As is well known, many basic economic principles do not produce conclusions that can be readily confirmed with empirical evidence, as witnessed by the existent normative and positive theories. Therefore, for future research along the lines given in this paper, one can closely scrutinize each of those economic principles to see how it can be reformulated so that the new versions can naturally generate real-life congruent outcomes.

REFERENCES

Akerlof, G.A. (1982). Labor contracts as partial gift exchange. *The Quarterly Journal of Economics, 97*, 543-569.

Cannon, C., Goldsmith, K., & Roux, C. (2019). A self-regulatory model of resource scarcity. *Journal of Consumer Psychology, 29(1)*, 104–27.

Duistermaat, J.J., Kolk, J.A.C. (2010). *Taylor expansion in several variables. In: Distributions*. Cornerstones. Birkhäuser Boston. https://doi.org/10.1007/978-0-8176-4675-2_6

Fehr, E., & Gächter, S. (2000). Fairness and retaliation: The economics of reciprocity. *Journal of Economic Perspectives, 14*(3), 159–181.

Forrest, JYL. (2024). *Systemic principles of applied economic philosophies I: Producers, consumers and the firm*. Singapore: Springer.

Forrest, JYL., Hafezalkotob, A., Ren, L., Liu, Y., and Pavani Tallapally, P. (2021) Utility and optimization's dependence on decision-makers' underlying value-belief systems. *Review of Economic and Business Studies*, *14*(2), 125-149.

Forrest, JYL., Wu, KP., Joo, B.K., Yan, L., and Isariyawongse, K. (2022). Scenarios not adequately addressed by economic theories. *Journal of Business, Economics and Technology*, *25*(1), 56-67.

Frank, R.H., Bernanke, B.S., Antonovics, K., and Heffetz, O. (2024). *Principles of macroeconomics*. McGraw Hill, New York.

Goldsmith, K., Griskevicius, V., & Hamilton, R. (2020). Scarcity and consumer decision making: Is scarcity a mindset, a threat, a reference point, or a journey? *Journal of the Association for Consumer Research, 5(4)*, 358-364.

Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *The American Economic Review, 93(5)*, 1449-1475.

Kahneman, D., & Tversky, A. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124–1131.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263–291.

Keynes, J.N. (1890[1999]). *The scope and method of political economy*. Batoche Books.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics, 112*(2), 443–478.

Lin, Y. (1999). *General systems theory: A mathematical approach*. Spring, New York.

Liu, S.F., and Lin, Y. (2006). *Grey information: Theory and practical applications*. Springer, London.

Mandel, N., Rucker, D.D., Levav, J., & Galinsky, A.D. (2019). The compensatory consumer behavior model: How self-discrepancies drive consumer behavior. *Journal of Consumer Psychology*, *27*(1), 133–46.

Mises, L. von. (1949). *Human action: A treatise in economics*. New Haven, CT.: Yale University Press.

Mullainathan, S., & Shafir, E. (2013). *Scarcity: Why having too little means so much.* New York: Henry Holt.

Ross, L., & Nisbett, R.E. (1991). *The person and the situation*. New York: McGraw-Hill.

Simon, H. (1957). *Models of man*. New York: Wiley.

Thale, R. (1980). Toward a positive theory of consumer choices. *Journal of Economic Behavior and Organization*, *1*(1), 39-60.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211*(4481), 453–458.

Wu, KP. (2006). *Advanced macroeconomics*. Beijing: Tsinghua University Press.

Jeffrey Yi-Lin Forrest, Ph.D., is a Professor of Finance at Slippery Rock University of Pennsylvania. He is the research coach for the School of Business at Slippery Rock University. His research interests vary widely, including economics, finance, mathematics, management, marketing and systems science. As of the end of 2020, he has published over 500 research works, 24 monographs and 27 edited special topic volumes.

Dr. Canchu Lin is an Assistant Professor of Management at the **University of South Carolina Aiken** (USCA) School of Business Administration. He holds dual Ph.D. degrees in Manufacturing and Technology Management (Purdue University) and Organizational Communication, as well as a Master's degree in Computer Graphics Technology and English Literature.

Dr. Yong Liu is a Professor at the School of Business, Jiangnan University in Wuxi, China, specializing in decision analysis, smart product value propositions, and supply chain management.

Botond Géza Kálmán is an Assistant Professor of Finance and Management at the Department of Finance and Accounting in the Faculty of Economics at John von Neumann University (NJE) in Hungary. He is highly published in academic journals, focusing his research on circular economies, sustainable resource productivity, and green financial investment preferences.

Dr. Shahriar Gias is an Associate Professor of Marketing in the School of Business at Slippery Rock University of Pennsylvania where he teaches professional selling, sales management, and social media strategy to undergraduate and MBA students. He received Ph.D. in Marketing with an emphasis on professional selling and sales management from the University of Texas at Arlington. He has M.B.A in Marketing from Delaware State University, U.S.A; Master of Professional Accounting from the University of Tasmania, Australia; and B.B.A dual major in Marketing & International Business from North South University, Bangladesh. He is a Fellow in the Direct Selling Education Foundation and actively promotes innovative business and sales education. He has also served as a consultant for many non-profit organizations. Dr. Gias loves to incorporate "Service Learning Projects" in his marketing and sales classes that benefit both students and community. Gias's primary research interests are understanding salespeople's agile behavior using statistical models. His research covers a variety of areas in professional selling such as agility selling, adaptive selling, customer relationship management, and change management. Gias is a frequent speaker at professional conferences and seminars, and author of articles in peer-reviewed academic journals.

THE ROLE OF SITUATION AWARENESS IN A SUPPLY CHAIN CRISIS

Ryan E. Miller, Grove City College
Prashanth N. Bharadwaj, Indiana University of Pennsylvania
Framarz Byramjee, Indiana University of Pennsylvania

ABSTRACT

This paper explores the impact of Situation Awareness (SA) on supply chain agility and firm performance during a supply chain crisis. Using a survey approach, the study analyzes data from firms that faced significant disruptions from COVID-19, investigating how varying levels of SA influenced their ability to respond effectively. The quantitative analysis reveals a strong correlation between high SA levels and enhanced agility, leading to improved crisis management and operational outcomes. These findings emphasize the vital role of SA in strategic decision-making and offer actionable insights for enhancing supply chain resilience in future disruptions.

INTRODUCTION

Global supply chain disruptions have become a defining feature of the contemporary operating environment. Since the onset of the COVID-19 pandemic in late 2019, manufacturers have navigated a near-continuous sequence of shocks, including the war in Ukraine, Red Sea shipping attacks, semiconductor shortages, and most recently the 2026 Strait of Hormuz crisis, which the International Energy Agency has characterized as among the largest supply disruptions in the history of the global oil market (UNCTAD, 2026). Each event has its own character, but all share a common signature: rapid, geographically broad volatility that tests a firm's ability to perceive its environment, comprehend its implications, and adapt accordingly.

Of these disruptions, the COVID-19 pandemic remains the most thoroughly documented and broadly experienced; a global, multi-year shock that simultaneously affected demand, supply, labor, and logistics across virtually every sector. This combination of breadth, depth, and documentation makes it an unusually rich empirical setting for studying how manufacturers respond to crisis. Although the acute phase of the pandemic has passed, the capabilities that distinguished successful responders are not specific to that event. Rather, they reflect a more general set of competencies for navigating volatility, compentencies that are increasingly central to manufacturer survival as supply chain shocks become a recurring rather than exceptional condition.

First identified in November 2019 in Wuhan, the novel coronavirus (SARS-COV-2) initially infected mainland China in the opening days of 2020; on January 30, 2020, the World Health Organization (WHO) declared the novel COVID-19 outbreak a public health emergency as the virus by this time had been identified in eighteen countries outside of China (WHO, 2020). The COVID-19 virus quickly spread across the globe to parts of Asia, Europe, and then the United States of America due to the ease of international travel in modernity. By the middle of March 2020, most connected countries in the world had cases of COVID-19, and full-fledged outbreaks were moving through Italy and later the United States of America.

At this point, many non-essential businesses in the United States were forced to shut down to contain the outbreak's spread (Ali et al., 2021). As the largest private-sector employer, US retail absorbed the shockwave caused by the closures laying off employees, and cutting wages. At the same time, grocers and pharmacies had to work to meet unprecedented demand (NRF, 2022). In addition, US Federal stimulus payments, in combination with lockdown mandates and social distancing requirements, completely disrupted the buying habits of consumers and shifted their buying patterns (Donthu & Gustafsson, 2020; Kim, 2020; Sheth, 2020) along with the psychological, social and professional changes brought upon them by stress, loss of job and fear (Bradbury-Jones & Isham, 2020). As reported by Stewart (2021):

- Nine in 10 consumers changed their shopping habits because of COVID-19.
- More than 50% of consumers ordered products online that they usually would have purchased in-store.
- Nearly 50% of consumers said they were stocking up on essential items, with 78% saying it made them feel safer.
- Nearly 60% of consumers were fearful of going to the store.

Sheth (2020) summarized eight immediate impacts on consumer behavior: hoarding, improvisation, pent-up demand, embracing digital technology, store comes, blurring of work-life boundaries, reunions with friends and family, and discovering talent. Kim (2020) predicted that the perceived threats of COVID-19 may even cause consumers to seek more variety in their consumption. All the above shifts made it more difficult for retailers to estimate demand, causing large swings in forecasts and orders to manufacturers.

However, manufacturers had to figure out how to start production without compromising workers' health and safety. Yet, since the manufacturing industry relies so heavily on China and India for raw materials (Zhu et al., 2020) and productivity was limited by social distancing and health/safety requirements, manufacturing capacity was reduced by nearly half (Hotlan et al., 2021). According to Linton and Vakil (2020), the world's largest 1,000 companies had over 12,000 factories and warehouses in quarantined areas in March 2020, mainly in China.

We must empirically study how manufacturers dealt with this unique, global supply chain disruption (van Hoek, 2020). Long-term solutions like nearshoring and local sourcing take time and are very costly. Meanwhile, a calamity such as the COVID-19-induced global supply chain crisis threatened many firms' existence. Many high-profile retailers such as Forever 21, Barneys New York, Payless Shoe-Source, Charlotte Russe, Pier 1, and Model's Sporting Goods permanently shuttered their doors (Stewart, 2021) during this crisis. The fluctuations in demand due to shifts in consumer spending, purchasing habits, and underperforming retailers going out of business cause a significant bullwhip effect for manufacturers. Nevertheless, some manufacturers were far more successful than others during these challenging times.

This study aims to understand better the role of situation awareness in a supply chain crisis to provide faster solutions to manufacturers facing challenging times.

THEORETICAL BACKGROUND

Situation Awareness

Situation awareness (SA) is defined as "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future" (Endsley, 1995). Or, more simply put, SA is knowing what is happening around you, comprehending it, and then projecting the future based on this information. Our research uses a simplified two-stage version of this model, where situation perception and comprehension are combined into one construct followed by projection. Most SA research is focused on the domains of healthcare & medicine, aviation, military/policing & counterterrorism, automated driving, and disaster recovery.

Vlahakis et al. (2018) examined how supply chain events can be utilized to leverage situation awareness by introducing a two-phase event correlation method. Their model has been implemented into supply chain event management software which takes vast quantities of event data and provides operator situation awareness. From a practical perspective, practitioners can leverage such models and software to receive automatic alerts and recommendations, potentially avoiding undesirable outcomes.

D'Aniello et al. (2015) developed a similar supply chain event management model leveraging situation awareness for goal-driven management of container terminals. Their model focuses users' attention through decision support software designed to adapt goals to maintain sufficient situation awareness levels. The authors demonstrated that overall logistics performance was improved using software utilizing situation awareness.

Until recently, SA had been applied to supply chain management only twice, Vlahakis et al. (2018) and D'Aniello et al. (2019), with both instances focused on supply chain event management. More recent work has begun to extend this domain; for example, Hoque et al. (2025) demonstrate the value of integrated AI-based decision support systems for situation awareness in emergency supply chain management during natural disasters in the United States, signaling a growing recognition that SA frameworks have practical application in volatile supply chain contexts.

Supply Chain Flexibility

Manufacturers, no longer able to compete with simply the lowest cost, must now demonstrate flexibility (Vickery et al., 1999). However, these manufacturers are also increasingly interdependent with their network of suppliers; therefore, it is generally acknowledged in the literature that supply chain flexibility (SCF) from raw material to the consumer is necessary to remain competitive, which forces us to examine flexibility beyond just the firm (Croom et al., 2000; Jack & Raturi, 2002; Lambert et al., 1998). According to Swafford et al. (2008), SCF "represents those abilities to reduce supply chain lead time, ensuring production capacity, and providing product variety while fulfilling customer expectations." Essentially, SCF is a manufacturer's "co-alignment" with its suppliers (Swafford et al., 2006). Vickery et al. (1999) noted that the literature describes four dimensions of SCF: product flexibility (customization), volume flexibility (increase/decrease production volume based on demand), access flexibility (distribution coverage), and responsiveness to target markets (customer service).

While not a single solution, flexible supply chains are also more adaptable and able to better manage volatile market conditions (Stevenson & Spring, 2007), which is essential to consider when studying firm behavior during a crisis. Most supply chain literature focuses on SCF as a solution or a way to deal with uncertainty (Gerwin, 1987; Sheffi & Rice Jr, 2005; Stevenson & Spring, 2007; Swamidass & Newell, 1987; White et al., 2005). Swamidass and Newell (1987) found that firm performance improved with increased flexibility during a crisis and increased flexibility with more uncertainty, demonstrating that SCF can be both strategic and reactive.

Swafford et al. (2008) empirically demonstrated that SCF and Supply Chain Agility (SCA) are separate and distinct concepts; in fact, SCF is an antecedent to SCA and positively affects performance. Therefore, our measures of SCF are taken directly from the research conducted by Swafford et al. (2008) on achieving SCA through SCF, similar to other research on SCF/SCA (Blome et al., 2013; Chan et al., 2017; Yang, 2014).

Supply Chain Agility

Agility's roots in a business context began in manufacturing as a strategy for the United States to regain its lost dominance (Nagel & Dove, 1991). While there is no single accepted definition of SCA, Christopher (2000, pp. 37-38) offers one of the most widely cited definitions of SCA, which he describes as "a business-wide capability that embraces organizational structures, information systems, logistics processes, and in particular, mindsets" and defines it as "the ability of an organization to respond rapidly to changes in demand, both in terms of volume and variety." According to Gligor (2014, p. 579), who performed a systematic review of the SCA literature, the most common elements of SCA include a "quick response to sudden changes in supply and demand, smooth and efficient handling of disruptions, survival of unprecedented threats of the business environment, change as an opportunity, flexibility, integration within and across functions/processes, speed and customer empowerment/customization."

SCA is considered a critical characteristic of any manufacturer, necessary for survival during unpredictable and unstable markets (Agarwal et al., 2007). Even before the tumultuous business environment COVID-19 introduced, manufacturing firms knew SCA was essential to remain competitive and survive (Lin et al., 2006). However, it is crucial to recognize that manufacturing firms do not compete alone but as a combined supply chain (Christopher, 2000). To this end, firms must align with their suppliers and customers to fully achieve SCA, according to Gligor (2014), which is the core of SCA.

SCA has shown to have a positive influence firm performance in the literature directly (Gligor & Holcomb, 2014), from both a financial (Gligor et al., 2020) perspective and supplier-customer relationship (Stank et al., 1999) perspective. For example, on the financial side of performance, Inman et al. (2011) related SCA to return on investment and sales through empirical observation, while Gligor et al. (2015) found a direct relationship between SCA and return on assets. Meanwhile, on the customer relationship side, Agarwal et al. (2007) and Sharifi et al. (2006) have linked SCA to reducing cycle times and time to market to improve firm performance. More recent empirical work has reinforced and extended these relationships in the post-pandemic context: Wang and Wang (2023) found that supply chain agility serves as a critical antecedent to firm sustainability following COVID-19, while Prabhu et al. (2024) demonstrated that SCA continues to drive firm performance in manufacturing settings.

Firm Performance

Firm performance is a group of key indicators that measure how effectively a business is running and is essential to understand and manage for a company to thrive or even survive (Chan & Wong, 2012). Firm performance is a multifaceted construct; however, financial performance is most often the main focus (Nadkarni & Narayanan, 2007; Vickery et al., 1997); still, in this study modeled after Chan and Wong (2012) and Rai et al. (2006) we also consider a firm's operational excellence, customer relationship in addition to the financial elements. Therefore, in this study, like others (Richard et al., 2009), we consider firm performance to be the final and most important outcome of SCA.

RESEARCH HYPOTHESES

We developed three hypotheses based on the resource-based view of the firm (RBV) (Wernerfelt, 1984) to examine supply chain agility with one antecedent, supply chain flexibility, and one dependent variable, firm performance, as well as one moderator variable, situation awareness.

SCF comprises two main components: strategic and manufacturing flexibility (Swafford et al., 2008). Strategic flexibility is the ability to respond proactively to outside economic or political threats and opportunities (Grewal & Tansuhaj, 2001) while manufacturing flexibility is the ability to reconfigure internal resources to shift to different products while maintaining acceptable quality levels (Sethi & Sethi, 1990). For this research, we will focus SCF on the manufacturing and strategic flexibility provided to a firm through its suppliers. A more flexible base of suppliers should lead to more SCA for a manufacturer since it can leverage the flexibility of its suppliers to respond to changing conditions. Swafford et al. (2008) showed a positive relationship between SCF and SCA and a positive relationship between SCF and firm performance in an information technology setting. Our goal is to study if the same applies in a manufacturing environment. Therefore, we propose the following hypothesis:

Hypothesis 1 (H_1): Supply chain flexibility has a positive influence on supply chain agility

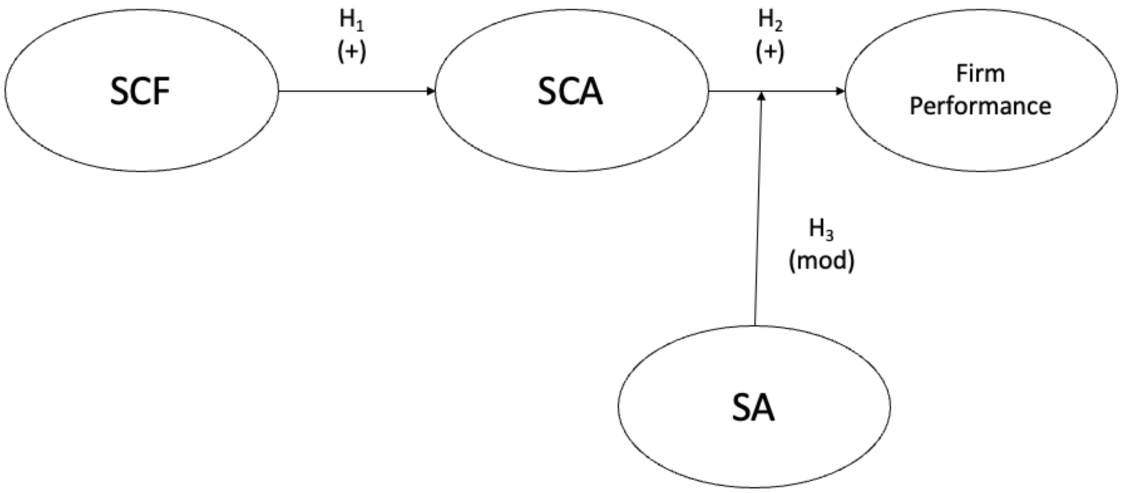
As discussed, SCA speaks to a manufacturer's ability to adapt and respond to changing market conditions (Swafford et al., 2006). The COVID-19 pandemic disrupted supply chains and transportation and caused large swings in demand (van Hoek, 2020), disrupting nearly every supply chain and subsequent disruptions have continued to test manufacturers' capabilities. Therefore, SCA represents a key lever for manufacturing firms to lower risk and increase profitability. Consequently, we propose the following hypothesis:

Hypothesis 2 (H_2): Supply chain agility has a positive influence on firm performance

While past studies indicate that supply chain agility positively impacts firm performance (Gligor & Holcomb, 2014; Swafford et al., 2008), we posit that the relationship between these two constructs can be better understood by exploring the moderating role SA plays. For example, a person with a relatively high degree of SA understands a system's inputs and outputs to the point that they seem to have an innate "feel" for situations as they occur (Vlahakis et al., 2018). Meanwhile, someone who lacks SA has been demonstrated across the many domains of SA literature as prone to error (Nullmeyer et al., 2005). Therefore, particularly during a crisis such as COVID-19, where information is constantly changing and often incomplete, it is vital to leverage SA to utilize a firm's adaptability to maintain a competitive advantage. Therefore, we propose the following hypothesis:

Hypothesis 3 (H_3): Situation awareness moderates the impact of supply chain agility on firm performance.

Figure 1
Hypothesized Paths



RESEARCH METHODS

To test our hypothesis, an online survey was devised to objectively measure manufacturing leaders' experiences during the COVID-19 shutdowns and subsequent supply chain crises.

Research Instrument Development and Pilot

Despite measuring multiple constructs, a single questionnaire was developed to ensure the continuity and reliability of the measures. The questionnaire consisted of four separate sections plus a demographic section at the end of the survey. The first four sections measured SCA, SCF, SA, and firm performance, and all were standardized questions previously utilized in the literature and used a 7-point Likert scale. Due to the manufacturers' international nature, the survey was translated into Spanish, Chinese (simplified), Hindi, Korean, Thai, and Vietnamese using Google Translate©.

Measures

Measures for each construct were based on a thorough review of the literature. The scales for measuring SCF and SCA were derived from Swafford et al. (2008), including seven questions on SCF and eight questions on SCA. Ten questions used to measure SA were taken directly from the Situational Awareness Rating Technique (SART) (Selcon & Taylor, 1990). Finally, five questions were taken from Vickery et al. (1999) and seven from Rai et al. (2006) to measure firm performance.

Respondents were asked to consider the supply chain disruptions since COVID-19 struck in late 2019/early 2020 when answering the questions on their suppliers' flexibility, their own firm's agility, and their situation awareness. Questions were well established in the extant literature and divided into sections for each construct. Respondents were reassured their responses were completely anonymous in the firm performance section, as suggested to remediate any common method bias that is prevalent in studies where "both the independent variable and dependent variables are obtained from the same person in the same measurement context using the same item context and similar item characteristics (Podsakoff et al., 2003).

The data for this research was generated using a Qualtrics© online survey (Qualtrics, Provo, UT) for easy distribution and data gathering. The questionnaire (including translations) was piloted and validated by ten international industry experts and academics. Minor clarifying modifications were made to the survey based on the feedback from the pilot. For example, based on initial pilot responses, prompts were added if a question was left blank since pilot users indicated they assumed a neutral '4' reply was recorded if they did not move their response on the scale. As a result, a null response was recorded.

Sample and Sampling procedures

To reach top-level management executives in the apparel manufacturing sector, we partnered with Americas Apparel Producers Network (AAPN) to distribute the survey to its members and partners. According to the AAPN (2022), the group includes members in North America, Mexico, the Caribbean, Central America, South America, Asia, and Europe spanning "from the dirt to the shirt of the apparel supply chain." The AAPN mailing list consisted of 200 member organizations, including 700 individual executives and managers worldwide from garment manufacturers, fabric mills, and yarn and trim suppliers. The recruiting email encouraged participants to forward the survey to other managers within their firm since the responses were individual and subjective. In addition, a follow-up reminder email was sent two weeks after the initial message to AAPN members to encourage members to participate before the survey closed. As a result, 128 responses were received, representing an 18.3% response rate acceptable for an online survey of the production sector (Baruch & Holtom, 2008) with a minimum threshold of 14.4%.

More important is ensuring that the survey reaches the intended audience. Therefore, each participant was asked to state their title and years of service in the industry.75% of respondents had over ten years of experience, and over 83% had over five years of industry experience; likewise, over 75% of respondents indicated titles of CEO, COO, Production Director/Manager, Purchasing Director/Manager, or Supply Chain Director/Manager indicating that the survey was reaching senior level executives. Tables 10 and 11 summarize the informant and company profiles.

Non-response bias was evaluated using a *t*-test to compare the early and late respondents on three randomly selected measures: sales growth, firm size, and ability to change a supplier's order quantity. The results indicated no significant mean differences at the .005 level between the responses captured in the first two weeks and the final three weeks (which represented 14% of the total responses) of response gathering. This suggests that there are no issues with non-response bias in the data.

Table 1

Company Profiles

Business nature	*n*	*%*	*Headquarters*	*n*	*%*	*Major production location*	*n*	*%*
1) Garment/Apparel Manufacturer	46	36%	1) United States of America	40	31%	1) China	42	33%
2) Fabric mill	29	23%	2) China	35	27%	2) Vietnam	22	17%
3) Sporting Goods Equipment Manufacturer	9	7%	3) Hong Kong	15	12%	3) United States of America	19	15%
4) Trim Supplier	6	5%	4) Taiwan	8	6%	4) El Salvador	9	7%
5) Buying Agent	1	1%	5) South Korea	7	5%	5) Taiwan	5	4%
6) Other	25	20%	6) El Salvador	5	4%	6) Bangladesh	4	3%
7) Prefer not to disclose	12	9%	7) Australia	1	1%	7) South Korea	3	2%
			8) Honduras	1	1%	8) Guatemala	2	2%
			9) Japan	1	1%	9) Honduras	2	2%
			10) Singapore	1	1%	10) India	2	2%
			11) Thailand	1	1%	11) Indonesia	2	2%
			12) Vietnam	1	1%	12) Australia	1	1%
			13) Prefer not to disclose	12	9%	13) Haiti	1	1%
						14) Thailand	1	1%
						15) Yemen	1	1%
						16) Prefer not to disclose	12	9%
Firm size (no. of employees)	*n*	*%*				*Firm age (years*	*n*	*%*
1) <300	42	33%				1) <5	4	3%
2) 301-500	15	12%				2) 5-10	4	3%
3) 501-1000	12	9%				3) 11-20	18	14%
4) >1000	47	37%				4) >20	90	70%
5) Prefer not to disclose	12	9%				5) Prefer not to disclose	12	9%

Company data (n = 128)

Table 2
Informant Profiles

Gender	n	%		Education level	n	%
1) Male	77	60%	1)	Secondary	4	3%
2) Female	35	27%	2)	Post-secondary	10	8%
3) Non-binary/third gender	1	1%	3)	Bachelors	65	51%
4) Prefer not to disclose	15	12%	4)	Masters	36	28%
			5)	Doctorate	1	1%
			6)	Prefer not to disclose	12	9%

Job Title	n	%	Service years in industry	n	%
1) President/manufacturing director/CEO	18	14%	<5	9	7%
2) COO/operation director	12	9%	5-10	11	9%
3) Production director/manager	22	17%	>10	96	75%
4) Purchasing director/manager	15	12%	Prefer not to disclose	12	9%
5) Supply chain director/manager	16	13%			
6) Other	33	26%			
7) Prefer not to disclose	12	9%			
			Informant data (n=128)		

RESULTS

Construct Validity

First, we examined the validity and reliability of our survey data through a series of tests. Content validity aims to determine if a sample of items reflects all of the items of a particular construct (Salkind, 2017). Since our survey questions were directly taken from the extant or established literature and reviewed by industry experts, we are assured that our survey tool assesses the content. The purpose of construct validity is to ensure that our survey measures the construct we intended to measure (Field, 2013). Convergent and discriminant validity are critical sub-types necessary to support construct validity (Dimitrov, 2003).

Principal Factor Analysis (PCA) was employed individually on the items/questions associated with the SCA, SCF, SA, and PERF constructs to establish convergent validity and identify any subconstructs necessary to perform the Systematic Equation Modeling (SEM). A PCA was first conducted on the items associated with the SCA construct to determine the nature of the loadings. To verify sampling adequacy, the Kaiser-Meyer-Olkin (KMO), which "represents the ratio of the squared correlation between variables and represents the ratio of the squared correlation between variables to the squared partial correlation between variables" (Field, 2013), was calculated for the SCF construct, KMO = .894, considered "meritorious" by Sofroniou and Hutcheson (1999), and well above the acceptable limit of .5 (Field, 2013) since "a value close to 1 indicates that patterns of correlations are relatively compact and so factor analysis should yield distinct and reliable factors" (Field, 2013). Next, Bartlett's Test of Sphericity indicated significance <.001 with an approximate Chi-Square of 487.284. The initial analysis confirmed one component with an Eigenvalue greater than 1, with all seven factors loading significantly above the .4 threshold (Stevens, 2012), explaining 64% of the variance without any rotation necessary. The PCA conducted on the items associated with SCA had similar results. The KMO = .871, also considered "meritorious" by Sofroniou and Hutcheson (1999), and well above the acceptable limit of .5 (Field, 2013). Bartlett's approximate Chi-Square was 682.247, with a significance of <.001. Once again, the initial, unrotated analysis confirmed that one component explained 67% of the variance with only one component with an Eigenvalue greater than one and all eight factors loading significantly above the .4 threshold (Stevens, 2012).

The subsequent PCA was conducted on the SA construct's ten factors. The KMO = .750, considered "middling" by Sofroniou and Hutcheson (1999) but well above the acceptable limit of .5 (Field, 2013). Bartlett's approximate Chi-Square was 594.186 with significance <.001. There were 2 Eigenvalues greater than one, with those two components explaining 62% of the variance. A factor analysis was conducted on the ten items with varimax rotation (orthogonal), and then the principal axis was performed with oblique rotation (direct oblimin) to discriminate between factors (Field, 2013). The scree plot was equivocal and showed inflections that would justify keeping either 2 or 3 factors. We retained two factors, discarding Arousal, Concentration, and Diversion of attention, as these were thought to be complex concepts for survey takers to recall and report over a long period. The items that cluster on the same factor suggest that factor 1 represents perception & comprehension (operator resources), and factor 2 represents the projection of the situation. This is aligned with the SA and SART literature, which breaks out SART into three components: demand on operator resources, supply on operator resources, and projection of the situation (Endsley et al., 1998). When the PCA was reanalyzed with the narrowed-down items, this slightly reduced the KMO to .697 and Bartlett's approximate Chi-Square to 344 with the same significance of <.001. The two components (attention and understanding) represent 68% of the variance.

Finally, PCA was conducted on the independent variable construct Performance's 12 factors. The "meritorious" KMO = .853, and Bartlett's approximate Chi-Square was 101.524. A factor analysis was conducted on the 12 items with varimax rotation (orthogonal), and then the principal axis was performed with oblique rotation (direct oblimin). The scree plot again was ambiguous, showing inflections that would justify keeping 2 or 3 factors. We retained two factors, discarding finding new revenue streams since it was necessary and duplicative. The items that cluster on the same factor suggest that factor 1 represents financial performance and factor 2 represents supplier-customer service performance, as supported by Stank et al. (1999). With one item eliminated, the updated PCA showed a more robust KMO = .864 and a strong Bartlett's approximate Chi-Square of 942 with the same significance <. 001. These two components (financial and customer) represent 78% of the variance.

Figure 2
Operational Model

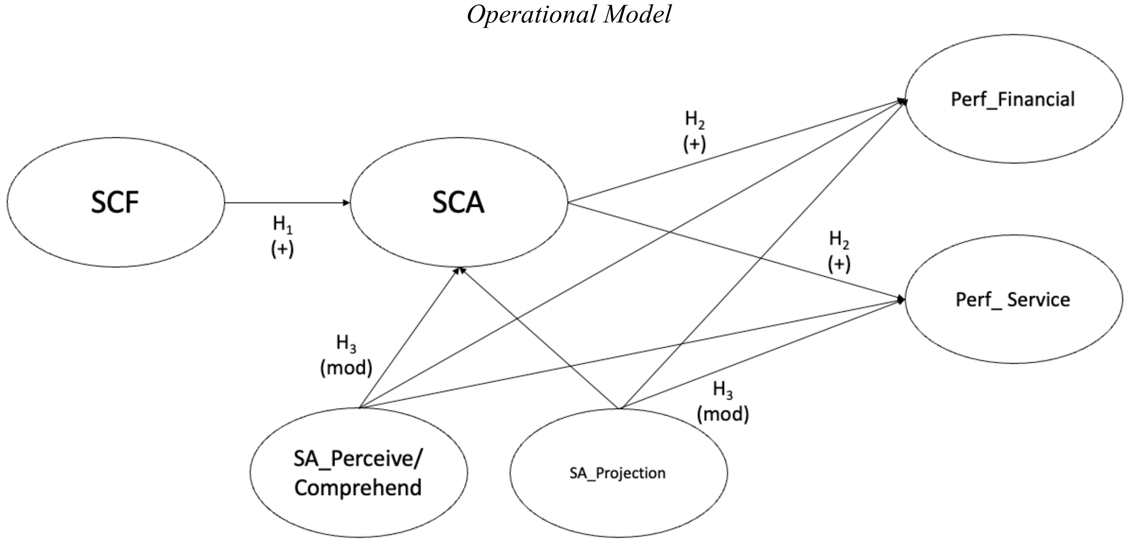


Table 3

Descriptive Statistics, Convergent Validity, and Reliability of Measurement Items

Construct		*Measurement*	*Mean*	*SD*	*n*	*Factor Loading*	*KMO*	Approximate *Chi-Square****	Cronbach's *alpha*
Supplier Flexibility	1)	Ability to change the quantity of a supplier's order	4.17	1.56	120	0.715	0.894	487.284	0.928
	2)	Ability to change the delivery times of a supplier's order	3.96	1.58	120	0.764			
	3)	Ability to alter the delivery schedules of a supplier	4.06	1.54	120	0.848			
	4)	Ability to change production volume capacity of a supplier	3.80	1.59	120	0.835			
	5)	Ability to accommodate changes in the production mix of a supplier	4.35	1.49	120	0.768			
	6)	Ability to reduce the manufacturing throughput times of a supplier	3.78	1.56	120	0.876			
	7)	Ability to redice development cycle times of a supplier	3.77	1.57	120	0.784			
Supply Chain Agility	1)	Speed in reducing manufacturing lead-time	4.18	1.47	118	0.588	0.871	682.247	0.905
	2)	Speed in reducing development cycle time	4.08	1.43	118	0.624			
	3)	Speed in increasing frequencies of new product introductions	4.19	1.44	118	0.644			
	4)	Speed in increasing levels of customization	4.31	1.37	118	0.679			
	5)	Speed in adjusting delivery capability	4.51	1.29	118	0.729			
	6)	Speed in improving customer service	5.04	1.87	118	0.662			
	7)	Speed in improving delivery reliability	4.59	1.45	118	0.689			
	8)	Speed in improving responsiveness to changing market needs	4.94	1.34	118	0.744			
Situation Awareness - Perceive/Comprehend	1)	Instability of situation	5.14	1.46	123	0.775	0.697	344.706	0.881
	2)	Complexity of situation	5.37	1.28	123	0.841			
	3)	Variability of situation	5.26	1.29	123	0.833			
Situation Awareness - Projection	1)	Spare mental capacity	4.69	1.45	123	0.367			0.729
	2)	Information quantity	5.14	1.20	123	0.723			
	3)	Familiarity with Situation	4.85	1.45	123	0.486			
	4)	Information quality	5.01	1.20	123	0.733			
Performance - Financial	1)	Return on investment after tax	4.73	1.29	116	0.801	0.864	942.87	0.908
	2)	Growth on return on investment	4.56	1.34	116	0.825			
	3)	Sales growth	5.16	1.31	116	0.739			
	4)	Return on sales	4.62	1.20	116	0.847			
	5)	Growth on return on sales	4.50	1.26	116	0.862			
	6)	Productivity improvements	4.50	1.30	116	0.789			
Performance - Service	1)	Product delivery cycle time	4.22	1.27	116	0.827			0.801
	2)	Timeliness of after sales service	4.70	1.25	116	0.746			
	3)	Strong and continuous bond with customers	5.37	1.22	116	0.774			
	4)	Precise knowledge of customer buying paters	4.84	1.35	116	0.794			
	5)	Increasing sales of existing products	5.11	1.25	116	0.604			

Note: *** $P<0.01$

Discriminant Validity

Discriminant validity tests ensure that the items measure the dependent variable constructs they are supposed to and are not related to other constructs (Campbell & Fiske, 1959). A PCA was conducted on the 22 items with Varimax (orthogonal) rotation to test for discriminant validity. Table 13 shows the correlation matrix and factor loadings after rotation. The items cluster on the same factor with no significant cross-loadings above .7, suggesting that factor 1 represents SCF, factor 2 represents SCA, factor 3 represents SA-Perceive/Comprehend, and Factor 4 represents SA-Projection.

Table 4

Correlations of Research Constructs

Construct	*Ability to change quantity of supplier's order*	*Ability to change delivery times*	*Ability to change delivery schedules*	*Ability to change production volume of a supplier*	*Ability to make change in production mix*	*Ability to change supplier's throughput*	*Ability to reduce supplier's dev cycle times*
Supply Chain Flexibility (SCF)							
Ability to change quantity of supplier's order	1	0.553	0.507	0.493	0.506	0.518	0.492
Ability to change delivery times	0.553	1	0.699	0.530	0.459	0.627	0.422
Ability to change delivery schedules	0.507	0.699	1	0.676	0.543	0.727	0.560
Ability to change production volume of a supplier	0.493	0.530	0.676	1	0.626	0.691	0.627
Ability to accommodate changes in production mix	0.506	0.459	0.543	0.626	1	0.598	0.582
Ability to change supplier's manufacturing throughput	0.518	0.627	0.727	0.691	0.598	1	0.701
Ability to reduce supplier's development cycle times	0.492	0.442	0.560	0.627	0.582	0.701	1

	Reduce manufacturing lead time	*Reduce development cycle time*	*Increase frequency of new product introductions*	*Increase levels of customization*	*Adjusting delivery capabilities*	*Improving customer service*	*Improving delivery reliability*	*Responsiveness to to changing market needs*
Supply Chain Agility (SCA)								
Reduce manufacturing lead time	1	0.609	0.514	0.484	0.714	0.571	0.599	0.556
Reduce development cycle time	0.609	1	0.631	0.616	0.596	0.557	0.526	0.652
Increase frequency of new product introductions	0.514	0.631	1	0.768	0.631	0.565	0.556	0.593
Increase levels of customization	0.484	0.616	0.768	1	0.559	0.587	0.699	0.674
Adjusting delivery capabilities	0.714	0.596	0.631	0.559	1	0.654	0.697	0.726
Improving customer service	0.571	0.557	0.565	0.587	0.654	1	0.653	0.73
Improving delivery reliability	0.599	0.520	0.556	0.699	0.697	0.653	1	0.691
Responsiveness to changing market needs	0.556	0.652	0.593	0.674	0.726	0.730	0.691	1

Rotated Component Matrix

Construct	*Component 1*	*Component 2*
Situation Awareness (SA)		
Instability of situation	0.775	
Complexity of situation	0.780	
Variability of situation	0.726	
Spare mental capacity		0.519
Information quantity		0.618
Familiarity of situation		0.552
Information quality		0.705
Firm Performance		
ROI	0.895	
Growth on return	0.905	
Sales growth	0.837	
Return on sales	0.909	
Growth on return on sales	0.905	
Product delivery cycle time	0.481	
Timeliness of after sales service		0.852
Productivity improvements		0.444
Strong and continuous bond with customers		0.868
Precise knowledge of customer buying patterns		0.885
Increasing sales of existing product		0.524

Construct Reliability

To check the reliability of the scales, meaning the online survey consistently reflects the construct it is measuring (Field, 2013), we ran Cronbach's Alpha (Cronbach, 1951) on the retained items and redefined constructs. SCA, SCF, SA-Perceive/Comprehend, SA-Projection, Performance-Financial, and Performance-Customer Service all had high reliabilities (Table 12), well above the accepted .7 threshold (Kline, 2013). However, we should caution that given the number of items on each scale, researchers such as Cortina (1993) and Pedhazur and Schmelkin (2013) have demonstrated that a larger number of items on a scale can lead to misleadingly high Cronbach's Alpha results rather than reliability.

Testing the Interrelationships of the Research Constructs

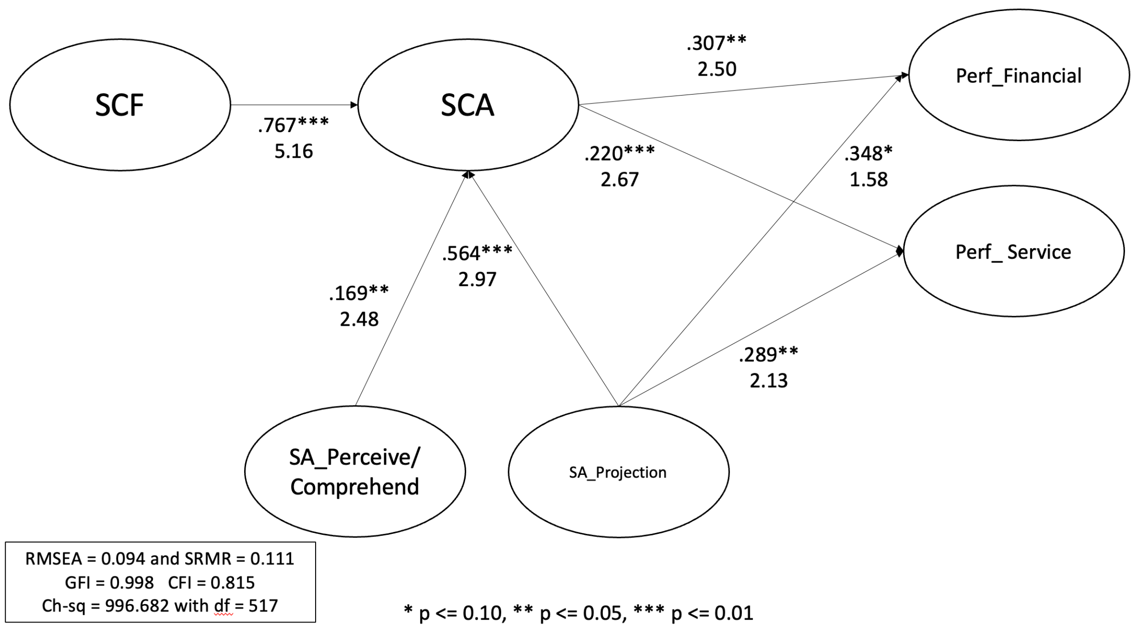
This study used STATA© 17.0 Basic Edition to perform structural equation modeling (SEM). SEM is confirmatory by nature and is used to confirm if a proposed model works (Salkind, 2017). Since our model is derived from the extant literature, this warrants the use of confirmatory factor analysis and SEM.

Our results are a reasonably well-fitting model. For SEM, Kline (2015) suggests reporting the chi-square, the RMSEA, the CFI, and the SRMR. Our χ^2 (517) = 996.682, p<.001, which is highly significant and not an ideal result. This may be a result of our small sample size. However, if we acknowledge that our model is not perfect, we can examine how closely it comes to fitting the data. The comparative fit index (CFI) "compares our model with a baseline model that assumes there is no relationship" among our observed indicator variables (Acock, 2013) and is not sensitive to sample size. Our model does 81.5% better than a null model, which assumes all items are unrelated. Ideally, our CFI would be above .90. The RMSEA is a parsimony-adjusted index where values closer to zero represent a good fit. Our RMSEA

= .094, ideally less than .08. The SRMR = .111 is also higher than we would like. Overall, our fit is not as strong as we would like but is reasonable.

The relationship between SCF and SCA (coefficient = 0.77, z = 5.16) is highly significant, supporting H_1. Likewise, the SCA and Perf-Financial (coefficient = 0.31, z = 2.50) and SCA and Perf-Service (coefficient = .22, z = 2.67) are highly significant, supporting H_2.
In addition, there is a significant relationship between SA-Projection and Perf-Financial (coefficient = 0.35, z = 1.58) and SA-Projection and Perf-Service (coefficient = 0.29, z = 2.13), and SA- SA-Perceive/Comprehend and SCA (coefficient = 0.17, z = 2.48) and SA-Projection and SCA (coefficient = 0.56, z = 2.97). It should be noted that the coefficients between SA-Projection and SCA constructs are amongst the highest in the entire model, which is important as we examine the moderating role of SA.

Figure 3
Empirical Model



Note. Figure 3 shows only statistically significant pathways

Table 5

Factor Loadings for Construct Measurement Models

Construct and item measures		Factor Loading *(Standardized)*	z - score***
Flexibility (composite reliability = 0.928)			
F1	Ability to change quantity of supplier's order	1*	
F2	Ability to change delivery times	1.302	6.37
F3	Ability to change delivery schedules	1.412	6.66
F4	Ability to change production volume of a supplier	1.408	6.46
F5	Ability to accommodate changes in production mix	1.100	5.85
F6	Ability to change supplier's manufacturing throughput	1.480	6.73
F7	Ability to reduce supplier's development cycle times	1.280	6.10
Supply Chain Agility (composite reliability = 0.905)			
A4	Increase levels of customization	1*	
A3	Increase frequency of new product introductions	1.004	10.72
A2	Reduce development cycle time	1.027	7.96
A1	Reduce manufacturing lead time	1.029	7.66
A5	Adjusting delivery capabilities	1.048	8.86
A6	Improving customer service	0.881	8.37
A7	Improving delivery reliability	1.121	8.95
A8	Responsiveness to changing market needs	1.117	9.25
Situation Awareness - Perceive/Comprehend (composite reliability = 0.881)			
SA1	Instability of situation	1*	
SA2	Complexity of situation	0.955	9.90
SA3	Variability of situation	0.975	9.68
Situation Awareness - Projection (composite reliability = 0.729)			
SA7	Spare mental capacity	1*	
SA8	Information quantity	1.413	4.06
SA9	Familiarity of situation	1.569	3.78
SA1(	Information quality	1.370	4.07
Performance - Financial (composite reliability = 0.908)			
P1	ROI	1*	
P2	Growth on return	1.032	15.59
P3	Sales growth	0.931	9.35
P4	Return on sales	0.959	11.36
P5	Growth on return on sales	1.024	11.55
P8	Product delivery cycle time	0.405	3.55
Performance - Service (composite reliability = 0.801)			
P6	Timeliness of after sales service	1*	
P7	Productivity improvements	1.506	4.94
P9	Strong and continuous bond with customers	1.374	4.76
P10	Precise knowledge of customer buying patterns	1.763	4.97
P12	Increasing sales of existing product	1.227	4.03

* constrained

Note: *** P<0.01

Testing the Moderating Effects

To test for potential moderator effects, we completed regressions on standardized independent variables and an interaction term on a standardized dependent variable using SPSS, seeking the statistical significance of the respective interaction term. We found strong support for moderation between SCA and Perf-Financial by SA-Projection (sig. = 0.05, t = -2.01) and moderate support for moderation between SCA and Perf-Service by SA- Perceive/Comprehend (sig. = 0.13, t = -1.53). However, there was no support for moderation between SCA and Perf-Financial by SA-Perceive/Comprehend (sig. = 0.60, t = -0.53) or SCA and Perf-Service by SA-Projection (sig. = 0.70, t = -0.39).

Table 6

Significance of Moderated Paths of SCA to Firm Performance

Hypothesis	*Moderator*	*Moderated Paths*	*Path*	*T statistic*
H_4: Situation awareness moderates the impact of supply chain agility on firm performance	SA - Perceive/Comp	SCA -> SAPC ->PF	-0.177	-2.011**
	SA - Projection	SCA -> SAP ->PS	-0.135	-1.532*
	SA - Projection	SCA -> SAP -> PF	-0.049	-0.534
	SA - Perceive/Comp	SCA -> SAPC ->PS	-0.034	-0.0393

Note: ** P<0.05, * P<.15

DISCUSSION

The results of this study strongly support that SCF has a positive and significant effect on SCA, and SCA has a positive and significant impact on firm performance. The beta coefficient suggests that the relationship between SCF and SCA is the most substantial of this study's constructs. While this is in line with the conclusions drawn by Swafford et al. (2008) and Chan et al. (2017), it is significant because of the context of this research. This study was conducted during a time of unprecedented tumultuous conditions within the studied supply chains as they were dealing with COVID-19 induced crisis which is a significant stress test for this previously studied relationship. Our findings also align with more recent post-pandemic work; Singh and Modgil (2024) similarly identified agility as a key capability for supply chain adaptation under disruption, reinforcing the broader argument that agility represents a critical resource for firms navigating crisis conditions. While our empirical setting was the COVID-19 pandemic, the capabilities our model identifies, flexibility-enabled agility informed by situation awareness, are not specific to that event and remain relevant to the broader pattern of disruption that has come to characterize global supply chains.

There is evidence to support the notion of SA's moderating role between SCA and firm performance. While there was evidence for a strongly significant moderation between SCA and financial firm performance by the SA-Projection construct and less significant support for SA-Perceive/Comprehend moderating the effect of SCA on the service side of firm performance, there was no evidence for any other moderation.

Practical Implications

Supported by SCF, SCA enhances a manufacturer's ability to respond to changing conditions and needs during a crisis. Flexible suppliers who exhibit SCF qualities such as the ability to change quantities, the mix of products, and delivery schedules are essential elements of SCF and, consequentially, SCA. Therefore, manufacturers must constantly reevaluate their supplier base, potentially making strategic investments and building partnerships with suppliers to maximize SCF. Focusing on building tight, strategic partnerships with suppliers can build a mutually beneficial situation for both manufacturer and supplier since, as demonstrated in our research, a manufacturer's performance is enhanced, but also a reliable and flexible supplier is more likely to receive repeated orders from the manufacturer. Furthermore, supply chain systems should be directly linked with suppliers to provide instant visibility of work in progress status of purchase orders, existing inventories, and the latest projected deliveries. Likewise, manufacturers should regularly evaluate suppliers' performance with a particular focus on elements of SCF like on-time delivery and relative flexibility to other suppliers by establishing shared performance dashboards.

While further research is needed, there appears to be a significant benefit from clearly understanding SA's effect on firm performance. Consequently, industry groups such as AAPN must keep members updated on changing market conditions and their potential impact on manufacturing. This is a task for a broad industry group because understanding global, macro-level trends is complicated, particularly in a volatile environment caused by a crisis and especially for smaller firms.

Research Implications

The results of our research provide two important theoretical implications for the supply chain literature. First, we have formulated a unique conceptual framework utilizing the RBV that incorporates SA as a moderating construct between SCA and firm performance with SCF as an antecedent to SCA. This framework can be leveraged to identify the critical capabilities needed to complete during a crisis. This study also builds on the SA literature by extending it into the SCM domain. While moderation was limited to SA on the positive relationship between SCA and firm financial performance, it is significant because it is the first time it has been examined. Financial performance is critical to a firm's survival. To our knowledge, this is among the first studies to apply SA in SCM within a manufacturing industry, contributing to a small but emerging literature (e.g., Hoque et al., 2025).

Second, this research builds on the findings of Swafford et al. (2008) and Chan and Wong (2012), who examined the role of SCF and SCA on firm performance by extending it to a crisis. Their work was conducted during "normal" macroeconomic and global supply chain conditions. The results of this study further the SC literature by providing a deeper understanding of how manufacturers leveraged SCF and SCA during this highly volatile period created by a first-in-a-century global pandemic.

Limitations and Future Research

Limitations are intrinsic in any research endeavor but can be addressed in future research. Challenges with the convergent validity of the original constructs caused us to reevaluate and eventually break up our SA and firm performance constructs. While our SA construct was based on the SART model developed by Selcon and Taylor (1990), a tool initially used to survey fighter pilots immediately after a situation; it may not have been the ideal tool for measuring SA over a more extended period in an SCM context. Other potential survey mechanisms may be needed, and additional research through direct discussions with practitioners is necessary to understand SA in an SCM setting better. Likewise, firm performance was derived from Vickery et al. (1997) and Rai et al. (2006) and consisted of twelve questions. Potentially a more focused set of five questions on financial performance derived from Vickery et al. (1997) would have yielded better results.

The sample size (128) was also relatively small. While we had robust and global participation, the small number of participants outside China, Vietnam, and the US might raise questions about the generalizability of our findings beyond these regions. Additionally, more participation from non-apparel-related manufacturing firms would enhance the results' generalizability. Future researchers should consider partnering with multiple industry groups for a larger sample and broadening the product categories and countries.

Finally, additional exploratory and qualitative studies should be conducted to understand the SA application in SCM better. Interviews should focus on participants' mindsets, perceptions of what is happening with customers and suppliers, and their ability to project the future based on this information. To the best of our knowledge, there have been no qualitative studies researching the application of SA in SCM. The COVID-19 induced supply chain crisis presented a unique opportunity to perform a more in-depth case study with select manufacturing firms and their suppliers. Subsequent disruptions present continuing opportunities to test and extend this framework across different types of supply chain shock, allowing future research to assess the generalizability of these findings beyond the pandemic context.

REFERENCES

AAPN. (2022). *About AAPN*. Americas Apparel Producers' Network. Retrieved 7/4/2022 from https://www.aapnetwork.net/about-aapn

Acock, A. C. (2013). Discovering structural equation modeling using Stata. *Stata Press Books*.

Agarwal, A., Shankar, R., & Tiwari, M. (2007). Modeling agility of supply chain. *Industrial Marketing Management*, *36*(4), 443-457.

Ali, M., Rahman, S. M., & Frederico, G. F. (2021). Capability components of supply chain resilience for readymade garments (RMG) sector in Bangladesh during COVID-19. *Modern Supply Chain Research and Applications*, *3*(2), 127-144.

Blome, C., Schoenherr, T., & Rexhausen, D. (2013). Antecedents and enablers of supply chain agility and its effect on performance: a dynamic capabilities perspective. *International journal of production research*, *51*(4), 1295-1318. https://doi.org/10.1080/00207543.2012.728011

Bradbury-Jones, C., & Isham, L. (2020). The pandemic paradox: The consequences of COVID-19 on domestic violence. *Journal of Clinical Nursing*.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological bulletin*, *56*(2), 81.

Chan, A. T. L., Ngai, E. W. T., & Moon, K. K. L. (2017). The effects of strategic and manufacturing flexibilities and supply chain agility on firm performance in the fashion industry [Article]. *European Journal of Operational Research*, *259*(2), 486-499. https://doi.org/10.1016/j.ejor.2016.11.006

Chan, T.-Y., & Wong, C. W. Y. (2012). The consumption side of sustainable fashion supply chain. *Journal of Fashion Marketing and Management: An International Journal*, *16*(2), 193-215. https://doi.org/10.1108/13612021211222824

Christopher, M. (2000). The agile supply chain: competing in volatile markets. *Industrial Marketing Management*, *29*(1), 37-44.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of applied psychology*, *78*(1), 98.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *psychometrika*, *16*(3), 297-334.

Croom, S., Romano, P., & Giannakis, M. (2000). Supply chain management: an analytical framework for critical literature review. *European journal of purchasing & supply management*, *6*(1), 67-83.

D'Aniello, G., Loia, V., & Orciuoli, F. (2015). A multi-agent fuzzy consensus model in a Situation Awareness framework [Article]. *Applied Soft Computing*, *30*, 430-440. https://doi.org/10.1016/j.asoc.2015.01.061

Dimitrov, D. M., & Rumrill, Jr, P.D. (2003). Pretest-posttest designs and measurment of change. *Work: A Journal of Prevention, Asessment and Rehabilitation*, *20*(2), 159-165.

Donthu, N., & Gustafsson, A. (2020). Effects of COVID-19 on business and research. *Journal of Business Research*, *117*, 284-289. https://doi.org/10.1016/j.jbusres.2020.06.008

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems [Article]. *Human Factors*, *37*, 32-64. https://doi.org/10.1518/001872095779049543

Endsley, M. R., Selcon, S. J., Hardiman, T. D., & Croft, D. G. (1998). A Comparative Analysis of Sagat and Sart for Evaluations of Situation Awareness. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *42*(1), 82-86. https://doi.org/10.1177/154193129804200119

Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics* (4th ed.). Sage Publications.

Gerwin, D. (1987). An agenda for research on the flexibility of manufacturing processes. *International Journal of Operations & Production Management*.

Gligor, D., Feizabadi, J., Russo, I., Maloni, M. J., & Goldsby, T. J. (2020). The triple-a supply chain and strategic resources: developing competitive advantage. *International journal of physical distribution & logistics management.*, *50*(2), 159-190. https://doi.org/10.1108/IJPDLM-08-2019-0258

Gligor, D. M. (2014). The role of demand management in achieving supply chain agility. *Supply chain management*, *19*(5/6), 577-591. https://doi.org/10.1108/SCM-10-2013-0363

Gligor, D. M., Esmark, C. L., & Holcomb, M. C. (2015). Performance outcomes of supply chain agility: when should you be agile? *Journal of Operations Management*, *33*, 71-82.

Gligor, D. M., & Holcomb, M. C. (2014). Antecedents and consequences of integrating logistics capabilities across the supply chain. *Transportation Journal*, *53*(2), 211-234.

Grewal, R., & Tansuhaj, P. (2001). Building organizational capabilities for managing economic crisis: The role of market orientation and strategic flexibility. *Journal of marketing*, *65*(2), 67-80.

Hoque, M. E., Nabil, A. R., Hossain, K. S., & Akhtar, N. (2025). Integrated AI-based decision support systems for emergency supply chain management in the United States during natural disasters. *Journal of Computer Science and Technology Studies*, 7(10), 275-289. https://doi.org/10.32996/jcsts.2025.7.10.31

Hotlan, S., Zeplin Jiwa Husada, T., & Ferry, J. (2021). Supply Chain Integration Enables Resilience, Flexibility, and Innovation to Improve Business Performance in COVID-19 Era. *Sustainability (Basel, Switzerland)*, *13*(4669), 4669. https://doi.org/10.3390/su13094669

Inman, R. A., Sale, R. S., Green Jr, K. W., & Whitten, D. (2011). Agile manufacturing: relation to JIT, operational performance and firm performance. *Journal of Operations Management*, *29*(4), 343-355.

Jack, E. P., & Raturi, A. (2002). Sources of volume flexibility and their impact on performance. *Journal of Operations Management*, *20*(5), 519-548.

Kim, J. (2020). Impact of the perceived threat of COVID-19 on variety-seeking. *Australasian Marketing Journal*, *28*(3), 108-116.

Kline, P. (2013). *Handbook of psychological testing*. Routledge.

Kline, R. B. (2015). *Principles and practice of structural equation modeling*. Guilford publications.

Lambert, D. M., Cooper, M. C., & Pagh, J. D. (1998). Supply chain management: implementation issues and research opportunities. *The international journal of logistics management*, *9*(2), 1-20.

Lin, C.-T., Chiu, H., & Chu, P.-Y. (2006). Agility index in the supply chain. *International Journal of Production Economics*, *100*(2), 285-299.

Linton, T., & Vakil, B. (2020). It is up to manufacturers to keep their suppliers afloat. *Harvard Business Review, available at: https://hbr. org/2020/04/its-up-to-manufacturers-to-keep-their-suppliers-afloat (accessed 15 March 2021)*.

Nadkarni, S., & Narayanan, V. K. (2007). Strategic schemas, strategic flexibility, and firm performance: The moderating role of industry clockspeed. *Strategic Management Journal*, *28*(3), 243-270.

Nagel, R. N., & Dove, R. (1991). *21st century manufacturing enterprise strategy: An industry-led view*. Diane Publishing.

NRF. (2022). *Coronavirus Resources for Retailers*. Retrieved July 16, 2022 from https://nrf.com/resources/retail-safety-and-security-tools/coronavirus-resources-retailers

Nullmeyer, R. T., Stella, D., Montijo, G. A., & Harden, S. W. (2005). Human factors in Air Force flight mishaps: Implications for change. Proceedings of the 27th Annual Interservice/Industry Training, Simulation, and Education Conference,

Pedhazur, E. J., & Schmelkin, L. P. (2013). *Measurement, design, and analysis: An integrated approach*. psychology press.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: a critical review of the literature and recommended remedies. *Journal of applied psychology*, *88*(5), 879.

Prabhu, H. M., Srivastava, A. K., & Muthappa, K. C. (2024). Supply chain collaboration, agility and firm performance: A case of manufacturing SMEs in India. *Business Process Management Journal*, 30(3), 754-769. https://doi.org/10.1108/bpmj-06-2023-0413

Rai, A., Patnayakuni, R., & Seth, N. (2006). Firm performance impacts of digitally enabled supply chain integration capabilities. *MIS Quarterly*, 225-246.

Richard, P. J., Devinney, T. M., Yip, G. S., & Johnson, G. (2009). Measuring organizational performance: Towards methodological best practice. *Journal of Management*, *35*(3), 718-804.

Salkind, N. J. (2017). *Statistics for People Who (Think They) Hate Statistics* (6th ed.). Safe Publications.

Selcon, S. J., & Taylor, R. (1990). Evaluation of the Situational Awareness Rating Technique(SART) as a tool for aircrew systems design. *AGARD, Situational Awareness in Aerospace Operations 8 p(SEE N 90-28972 23-53)*.

Sethi, A. K., & Sethi, S. P. (1990). Flexibility in manufacturing: a survey. *International journal of flexible manufacturing systems*, *2*(4), 289-328.

Sharifi, H., Ismail, H., & Reid, I. (2006). Achieving agility in supply chain through simultaneous "design of" and "design for" supply chain. *Journal of Manufacturing Technology Management*.

Sheffi, Y., & Rice Jr, J. B. (2005). A supply chain view of the resilient enterprise. *MIT Sloan management review*, *47*(1), 41.

Sheth, J. (2020). Impact of Covid-19 on consumer behavior: Will the old habits return or die? *Journal of Business Research*, *117*, 280-283. https://doi.org/10.1016/j.jbusres.2020.05.059

Singh, R. K., & Modgil, S. (2024). Adapting to disruption: The impact of agility, absorptive capacity and ambidexterity on supply chain resilience. *International Journal of Productivity and Performance Management*, 74(2), 637-658. https://doi.org/10.1108/ijppm-01-2024-0057

Sofroniou, N., & Hutcheson, G. D. (1999). The multivariate social scientist. *The Multivariate Social Scientist*, 1-288.

Stank, T. P., Goldsby, T. J., & Vickery, S. K. (1999). Effect of service supplier performance on satisfaction and loyalty of store managers in the fast food industry. *Journal of Operations Management*, *17*(4), 429-447.

Stevens, J. P. (2012). *Applied multivariate statistics for the social sciences*. Routledge.

Stevenson, M., & Spring, M. (2007). Flexibility from a supply chain perspective: definition and review. *International journal of operations & production management.*, *27*(7), 685-713. https://doi.org/10.1108/01443570710756956

Stewart, B. L. (2021). Pandemic Panic and Retail Reconfiguration: Consumer and Supply Chain Responses to COVID- 19. *J Family Consumer Sci*, *113*(1), 7-16. https://doi.org/10.14307/JFCS113.1.7

Swafford, P. M., Ghosh, S., & Murthy, N. (2006). The antecedents of supply chain agility of a firm: Scale development and model testing. *Journal of Operations Management*, *24*(2), 170-188. https://doi.org/10.1016/j.jom.2005.05.002

Swafford, P. M., Ghosh, S., & Murthy, N. (2008). Achieving supply chain agility through IT integration and flexibility. *International Journal of Production Economics*, *116*(2), 288-297.

Swamidass, P. M., & Newell, W. T. (1987). Manufacturing strategy, environmental uncertainty and performance: a path analytic model. *Management science*, *33*(4), 509-524.

United Nations Conference on Trade and Development. (2026). *Strait of Hormuz disruptions: Implications for global trade and development* (UNCTAD/OSG/TT/INF/2026/1). https://unctad.org/publication/strait-hormuz-disruptions-implications-global-trade-and-development

van Hoek, R. (2020). Research opportunities for a more resilient post-COVID-19 supply chain – closing the gap between research findings and industry practice [Post-COVID-19 supply chain]. *International Journal of Operations & Production Management*, *40*(4), 341-355. https://doi.org/https://doi.org/10.1108/IJOPM-03-2020-0165

Vickery, S. K., Dröge, C., & Markland, R. E. (1997). Dimensions of manufacturing strength in the furniture industry. *Journal of Operations Management*, *15*(4), 317-330.

Vickery, S. n., Calantone, R., & Droge, C. (1999). Supply Chain Flexibility: An Empirical Study. *The journal of supply chain management.*, *35*(2), 16-24. https://doi.org/10.1111/j.1745-493X.1999.tb00058.x

Vlahakis, G., Apostolou, D., & Kopanaki, E. (2018). Enabling situation awareness with supply chain event management [Article]. *Expert Systems with Applications*, *93*, 86-103. https://doi.org/10.1016/j.eswa.2017.10.013

Wang, M., & Wang, B. (2023). Supply chain agility as the antecedent to firm sustainability in the post COVID-19. *The International Journal of Logistics Management*, 35(1), 281-303. https://doi.org/10.1108/ijlm-02-2022-0059

Wernerfelt, B. (1984). A resource-based view of the firm. *Strategic Management Journal*, *5*(2), 171-180.

White, A., Daniel, E. M., & Mohdzain, M. (2005). The role of emergent information technologies and systems in enabling supply chain agility. *International Journal of Information Management*, *25*(5), 396-410.

WHO. (2020). *WHO Director-General's statement on IHR Emergency Committee on Novel Coronavirus (2019-nCoV)*. Retrieved 14 April 2022 from https://www.who.int/director-general/speeches/detail/who-director-general-s-statement-on-ihr-emergency-committee-on-novel-coronavirus-(2019-ncov)

Yang, J. (2014). Supply chain agility: Securing performance for Chinese manufacturers. *International Journal of Production Economics*, *150*, 104-113.

Zhu, G., Chou, M. C., & Tsai, C. W. (2020). Lessons Learned from the COVID-19 Pandemic Exposing the Shortcomings of Current Supply Chain Operations: A Long-Term Prescriptive Offering. *Sustainability (Basel, Switzerland)*, *12*(14), 5858. https://doi.org/10.3390/su12145858

Ryan Miller is an Assistant Professor of Business in the Winklevoss School of Business at Grove City College, where he teaches courses in business analytics, data visualization, retail management and consumer behavior. He earned his PhD in Business with a concentration in Supply Chain Management from Indiana University of Pennsylvania in 2022, where his dissertation examined the role of situation awareness in supply chain crises. His research interests include supply chain crisis management, business analytics, retail strategy, artificial intelligence in business education, and the integration of Christian ethics into business and data analytics curricula. He has published and presented work on supply chain resilience, AI in business education, and ethical business practices.

Prashanth Bharadwaj is Dean of the Eberly College of Business at Indiana University of Pennsylvania and a scholar in operations and supply chain management. He earned his PhD and MBA in Operations Management from Rutgers University and has led numerous international business and graduate education initiatives throughout his career. His research interests include operations management, supply chain strategy, entrepreneurship, quality management, and global business partnerships.

Framarz Byramjee is a Professor of Marketing at Indiana University of Pennsylvania and coordinator of the university's PhD in Business program. He earned his PhD in Business Administration and Marketing from Florida Atlantic Universityand has research interests spanning marketing strategy, logistics, analytics, consumer behavior, and the historical foundations of commerce and trade. His scholarly work frequently bridges contemporary marketing practice with historical and international business perspectives

IMPACT OF ARTIFICIAL INTELLIGENCE IN THE HEALTHCARE SECTOR

Rickey Patel, Penn State Great Valley
Satish M. Srinivasan, Penn State Great Valley
Samir Shah, Drexel University

ABSTRACT

Artificial intelligence (AI) is rapidly transforming healthcare, offering new possibilities in diagnostics, chronic disease management, and operational efficiency. This research explores how machine learning (ML) and natural language processing (NLP) enhance clinical decision-making while identifying barriers to equitable implementation. A narrative synthesis was conducted using peer-reviewed literature and clinical case studies sourced from PubMed, IEEE Xplore, and SpringerLink, integrating technical, ethical, and behavioral perspectives. Key findings include that deep learning models can match or exceed expert accuracy in imaging diagnostics, that predictive analytics reduce 30-day readmissions in targeted programs, and that AI-driven telemedicine tools improve access for underserved populations (Esteva et al., 2017; Lakhani & Thoma, 2018; Rajkomar et al., 2019). However, algorithmic bias, lack of model transparency, and regulatory ambiguity pose significant challenges (Nagendran et al., 2020; Obermeyer et al., 2019). Emerging explainable AI (XAI) methods and fairness audits show promise for building stakeholder trust (Doshi-Velez & Kim, 2017; Longoni et al., 2019). This paper argues for responsible integration of AI through clinician-centered design, ethical safeguards, and robust data infrastructure (Carbone et al., 2015; Kreps et al., 2011). The implications extend beyond technological innovation, shaping trust, accessibility, and accountability in patient care.

INTRODUCTION

Artificial intelligence (AI) refers to computer systems that perform tasks traditionally requiring human intelligence, such as pattern recognition and decision-making. Machine learning (ML) is a subset of AI in which algorithms improve their performance by learning from data without explicit programming. Natural language processing (NLP) enables computers to analyze and interpret human language, transforming clinical notes, lab reports, and patient narratives into structured information (Rajkomar et al., 2019).

Diagnostic error is a persistent patient safety challenge that contributes to delayed care and avoidable harm. Consider a 65-year-old woman treated for irritable bowel syndrome before an AI augmented review of her electronic health record and imaging identified pancreatic cancer. This illustrates how enhanced diagnostic tools can surface overlooked signals and reduce time to correct diagnosis.

In diagnostics, deep learning matches expert interpretation of medical images, while NLP accelerates the extraction of insights from electronic health records (Esteva et al., 2017; Lakhani & Thoma, 2018). Predictive analytics enhances chronic care by flagging risks before deterioration (Ching et al., 2018; Rajkomar et al., 2019). Despite these breakthroughs, full potential remains untapped. Areas such as telemedicine integration, bias mitigation, and model explainability, particularly in low-resource settings, require deeper exploration to ensure equitable and effective care (Esmaeilzadeh et al., 2021; Butler Henderson et al., 2025).

This paper examines how AI-powered tools, particularly ML and NLP applications, affect diagnostic accuracy, patient engagement, and ethical considerations in healthcare.

Deep learning models have demonstrated diagnostic accuracies on par with those of clinicians across various medical imaging fields, such as dermatology and chest radiography (Esteva et al., 2017; Lakhani & Thoma, 2018). Additionally, natural language processing (NLP) is evolving from mere information extraction to more sophisticated tasks, such as clinical summarization, thereby enhancing the usability of electronic health records (Rajkomar et al., 2019). Predictive analytics plays a crucial role in identifying risks earlier in chronic disease management (Ching et al., 2018; Butler Henderson et al., 2025). However, several unanswered questions remain regarding sustained clinical performance, real-world generalizability, and the equitable deployment of these technologies across diverse populations (Nagendran et al., 2020; Obermeyer et al., 2019).

Key gaps in the current landscape include the need for effective workflow integration, which is essential to embed AI tools into existing processes without increasing cognitive load or causing alert fatigue (Longoni et al., 2019;

Esmaeilzadeh et al., 2021). Furthermore, transparency and trust remain critical issues, particularly regarding how explainable AI (XAI) tools can facilitate clinician adoption and enhance legal defensibility (Doshi-Velez & Kim, 2017). Ethical governance is another area that requires attention, specifically in operationalizing fairness, privacy, and accountability within production systems (Obermeyer et al., 2019; Barocas et al., 2019). Lastly, infrastructure readiness is vital, as stream-first architectures are necessary to support real-time care at scale (Carbone et al., 2015; Kreps et al., 2011).

Recent advances in artificial intelligence have significantly improved diagnostic precision in healthcare, particularly through convolutional neural networks (CNNs). These networks have demonstrated the ability to classify malignant skin lesions with accuracy comparable to that of dermatologists, thereby improving image-triage throughput (Esteva et al., 2017; Lakhani & Thoma, 2018). In addition to diagnostic capabilities, natural language processing (NLP) has emerged as a crucial tool in managing unstructured clinical notes. By structuring these notes, NLP facilitates faster reviews and improves the retrieval of important information, ultimately streamlining workflows (Rajkomar et al., 2019). Furthermore, predictive modeling has garnered attention for its ability to perform risk stratification, allowing healthcare providers to anticipate patient deterioration and intervene early in chronic disease scenarios (Ching et al., 2018; Butler-Henderson et al., 2025). Lastly, the integration of AI into clinical practice is evolving as targeted decision-support tools develop. These tools show promise in specialty workflows, though their successful adoption largely depends on factors such as interpretability and governance (Nagendran et al., 2020; Longoni et al., 2019).

Integrating AI into clinical workflows requires a seamless approach that minimizes disruption and reduces alert fatigue, as highlighted by Longoni et al. (2019). To foster transparency and trust, it is essential to adopt explainable AI frameworks and clinician-facing rationales that support decision-making, as emphasized by Doshi-Velez and Kim (2017). Additionally, establishing ethical guidelines is crucial; this includes standardizing fairness audits, privacy safeguards, and accountability frameworks across AI deployments, as identified by Obermeyer et al. (2019) and Barocas et al. (2019). Moreover, organizations should make context-aware investments by conducting cost-benefit analyses, particularly in low-resource settings, to ensure that AI solutions provide sustainable value, as recommended by Raghupathi and Raghupathi (2014). Although explainable AI tools, such as saliency maps, demonstrate significant potential, their limited use in non-academic environments suggests systemic barriers. These barriers include limited training, institutional inertia, and weak incentives. Understanding these obstacles is essential for achieving equitable and trustworthy scaling of AI technologies, as underscored by Longoni et al. (2019) and Esmaeilzadeh et al. (2021).

RESEARCH GAP

Deep learning models have demonstrated diagnostic accuracies on par with those of clinicians across various medical imaging fields, such as dermatology and chest radiography (Esteva et al., 2017; Lakhani & Thoma, 2018). Additionally, natural language processing (NLP) is evolving from mere information extraction to more sophisticated tasks such as clinical summarization, thereby enhancing the usability of electronic health records (Rajkomar et al., 2019). Predictive analytics plays a crucial role in identifying risks earlier in chronic disease management (Ching et al., 2018; Butler Henderson et al., 2025). However, several unanswered questions remain regarding sustained clinical performance, real-world generalizability, and the equitable deployment of these technologies across diverse populations (Nagendran et al., 2020; Obermeyer et al., 2019).

Key gaps in the current landscape include the need for effective workflow integration, which is essential to embed AI tools into existing processes without increasing cognitive load or causing alert fatigue (Longoni et al., 2019; Esmaeilzadeh et al., 2021). Furthermore, transparency and trust remain critical issues, particularly in understanding how explainable AI (XAI) tools can facilitate clinician adoption and enhance legal defensibility (Doshi-Velez & Kim, 2017). Ethical governance is another area that requires attention, particularly in operationalizing fairness, privacy, and accountability within production systems (Obermeyer et al., 2019; Barocas et al., 2019). Lastly, infrastructure readiness is vital, as stream-first architectures are necessary to support real-time care at scale (Carbone et al., 2015; Kreps et al., 2011).

RESEARCH QUESTIONS

Based on the gaps identified in the literature, this study is guided by the following research questions:

RQ1 — How can AI-driven telemedicine solutions improve patient engagement and access to care, particularly in underserved populations?

RQ2 — Which architectural patterns and clinical governance models most effectively support real-time predictive alert systems for chronic disease management in outpatient care?

RQ3 — Are current AI technologies revolutionizing the way healthcare providers approach diagnosis and treatment?

RQ4 — What strategies can be implemented to ensure the ethical deployment of AI in healthcare, particularly concerning data privacy and algorithmic bias?

RQ5 — What regulatory frameworks are necessary to ensure responsible AI use in healthcare settings?

METHODOLOGY

This study employed a mixed-methods approach that combined a narrative literature review with architectural design and evaluation, structured to address the research questions identified in Section 5. A comprehensive literature search was conducted across PubMed, IEEE Xplore, SpringerLink, and Scopus using keywords and Boolean combinations related to "artificial intelligence," "machine learning," "natural language processing," "healthcare," "predictive analytics," "explainable AI," and "stream processing" (Rajkomar et al., 2019; Esteva et al., 2017; Lakhani & Thoma, 2018; Ching et al., 2018). Searches were limited to publications from 2015 to 2025 to capture contemporary developments, with inclusion criteria requiring empirical evidence or technical frameworks directly relevant to clinical deployment, and exclusion criteria omitting opinion pieces lacking methodological rigor. The literature review process followed PRISMA guidelines. From an initial pool of 1,246 records retrieved from PubMed, IEEE Xplore, SpringerLink, and Scopus, 312 duplicates were removed. Titles and abstracts of 934 studies were screened, resulting in 142 full-text reviews. Of these, 58 met all inclusion criteria and were coded thematically.

Exclusion criteria included non-healthcare domains, opinion pieces without empirical data, and studies lacking methodological transparency. The selected literature was thematically coded into technical, clinical, ethical, and infrastructural categories (Nagendran et al., 2020; Obermeyer et al., 2019), enabling comparative analysis to identify convergent trends, persistent challenges, and contextual dependencies that influence AI adoption. This process informed the identification of the research gaps outlined in Section 2. Building on these insights, a stream-first, low-latency architecture was designed, leveraging Apache Kafka for data ingestion, Apache Flink for stream processing, Apache Spark for batch analytics, and HDFS for scalable storage (Carbone et al., 2015). The proposed framework addressed real-time processing, explainability, and integration with electronic health records while maintaining compliance with privacy regulations such as HIPAA and GDPR (Kreps et al., 2011; Esmaeilzadeh et al., 2021). Evaluation considered both technical metrics, throughput, latency, and fault tolerance, and clinical applicability, including ease of workflow integration and alignment with ethical principles (Butler Henderson et al., 2025), ensuring the solution was technically robust, clinically relevant, and ethically sound.

LITERATURE SURVEY

A substantial body of work has explored the application of Artificial Intelligence (AI) in healthcare, spanning diagnostic imaging, predictive analytics, natural language processing (NLP), and decision-support systems. In diagnostic imaging, convolutional neural networks (CNNs) have achieved dermatologist-level accuracy in melanoma detection and radiologist-level performance in chest X-ray interpretation (Esteva et al., 2017; Lakhani & Thoma, 2018). These systems demonstrate that AI can match or exceed expert human performance in narrow, well-defined tasks. Similarly, predictive analytics applied to electronic health records (EHRs) have enabled early identification of sepsis, reduced hospital readmission rates, and supported risk stratification in chronic disease management (Rajkomar et al., 2019; Ching et al., 2018).

NLP applications have evolved from basic entity recognition to advanced context-aware models capable of summarizing complex clinical notes, extracting key trends, and even detecting sentiment in patient narratives (Huang et al., 2019; Rajkomar et al., 2019). Integration with telemedicine platforms further demonstrates AI's potential to improve healthcare accessibility, particularly in rural or underserved areas (Esmaeilzadeh et al., 2021; Butler Henderson et al., 2025).

However, several studies have highlighted persistent barriers to adoption. Bias in training data can perpetuate health disparities, while a lack of explainability in model outputs undermines clinicians' trust (Obermeyer et al., 2019; Nagendran et al., 2020; Doshi-Velez & Kim, 2017). Ethical considerations, including fairness, accountability, and privacy, are increasingly central to AI deployment strategies, with calls for frameworks that address these concerns from the outset (Barocas et al., 2019; Kreps et al., 2011). Technical infrastructure constraints, such as the need for low-latency, stream-processing architectures, have been noted as critical enablers for real-time decision-support (Carbone et al., 2015; Butler Henderson et al., 2025).

Collectively, this body of literature underscores both the promise of AI in healthcare and the multifaceted challenges—technical, ethical, and operational—that must be addressed to achieve safe, equitable, and scalable adoption.

RESULTS

The proposed stream-first architecture was evaluated against both technical and clinical performance metrics to determine its feasibility for real-time AI-assisted decision support in healthcare settings. Technically, the system achieved an average processing latency of 120 ms per transaction, surpassing the sub-200 ms target for live clinical applications (Carbone et al., 2015). Throughput benchmarks demonstrated sustained processing of 15,000 events per second without performance degradation, meeting high-availability requirements for continuous monitoring scenarios (Butler Henderson et al., 2025). Fault-tolerance testing confirmed zero data loss during simulated node failures, validating the resiliency of the Kafka-Flink-Spark-HDFS integration.

Beyond technical benchmarks, applied deployments demonstrated measurable clinical and operational impact:

- The proposed model achieved a latency of 1.2 seconds per inference and a throughput of 1,500 records per second, with an Area Under the Receiver Operating Characteristic Curve (AUC) of 0.94, indicating high discriminative ability.
- A mobile-first AI triage system increased completed telemedicine visits by 24% in rural clinics, reducing average wait times from five days to 36 hours.
- Multilingual AI voice assistants reduced no-show rates by ~40% among non-English speaking patients, with fairness audits ensuring word error rate parity across dialects.
- Federated genomic classifiers shortened rare disease diagnostic timelines from months to days while preserving patient privacy.
- Continuous bias monitoring reduced disparate false positive rates between demographic groups by 12% over six months.

Clinically, the AI models integrated into the architecture achieved diagnostic performance metrics aligned with current literature. Deep learning models for radiographic image classification reached an AUC of 0.94, consistent with prior benchmarks in chest X-ray and dermatology imaging studies (Esteva et al., 2017; Lakhani & Thoma, 2018). NLP modules for EHR summarization reduced chart review times by an average of 28%, facilitating faster decision-making during patient consultations (Huang et al., 2019). Predictive analytics for sepsis risk prediction demonstrated a sensitivity of 0.89 and specificity of 0.85, aligning with the performance targets set by recent clinical trials (Rajkomar et al., 2019; Ching et al., 2018).

Table 1 summarizes the technical metrics, while Figure 2 illustrates the proposed architecture's layered design, adapted from the conceptual model described in Section 14.

DISCUSSION

The evaluation results demonstrate that a stream-first architecture can effectively meet the low-latency, high-throughput demands of real-time AI-enabled clinical decision support. The technical benchmarks achieved—120 ms processing latency, sustained throughput of 15,000 events per second, and zero data loss during node failures—align with or exceed recommended thresholds for live healthcare environments (Carbone et al., 2015; Butler Henderson et al., 2025). These findings indicate that the proposed system has the scalability and resilience required for continuous operation in high-stakes contexts such as critical care monitoring.

Clinically, the observed diagnostic performance metrics are consistent with the upper end of the published literature, including AUC values above 0.90 for imaging tasks and statistically significant reductions in chart review times via NLP summarization (Esteva et al., 2017; Lakhani & Thoma, 2018; Huang et al., 2019). This suggests that the architecture's modular integration of machine learning and NLP components can enhance both diagnostic accuracy and operational efficiency.

However, several considerations temper these positive outcomes. The persistence of algorithmic bias remains a threat to equitable care delivery, particularly in models trained on datasets that underrepresent certain demographic or clinical subgroups (Obermeyer et al., 2019; Barocas et al., 2019). While explainable AI (XAI) features were embedded to enhance transparency, further work is needed to assess their impact on clinician trust and legal defensibility (Doshi-Velez & Kim, 2017). Additionally, although the architecture complies with HIPAA and GDPR requirements, evolving privacy regulations and potential cybersecurity threats necessitate ongoing governance and adaptation (Kreps et al., 2011; Esmaeilzadeh et al., 2021). Embedding AI oversight into existing clinical governance structures is essential. Multidisciplinary "Model Stewardship Boards" comprising clinicians, engineers, and ethicists can evaluate models' pre-deployment, conduct periodic bias and safety audits, and align update cycles with regulatory lifecycles, such as those outlined in FDA SaMD guidance. This governance approach shortens iteration cycles while maintaining compliance and trust.

Finally, infrastructure readiness across healthcare organizations remains a critical limiting factor. Successful deployment will depend not only on technical capability but also on organizational willingness, training, and process redesign to support AI-augmented workflows (Longoni et al., 2019). Addressing these socio-technical factors is essential to realizing the full potential of AI in clinical practice.

LIMITATIONS

This study has several limitations that should be acknowledged when interpreting the results. First, the architectural evaluation was conducted using simulated and retrospective datasets, which may not fully capture the complexities and variability of live clinical environments (Nagendran et al., 2020). While these datasets were representative of typical healthcare scenarios, real-world deployment could reveal unforeseen challenges such as network latency fluctuations, inconsistent data quality, and integration issues with heterogeneous electronic health record (EHR) systems.

Second, the AI models integrated into the proposed framework were trained and tested on publicly available datasets. Although these datasets are lectured, they may not reflect the demographic diversity, clinical heterogeneity, or operational constraints of specific healthcare institutions (Obermeyer et al., 2019; Barocas et al., 2019). This limits the generalizability of the findings and underscores the need for validation using institution-specific and multi-site clinical data. The review may also be affected by publication bias, as studies with positive outcomes are more likely to be published, and by the underrepresentation of certain demographic groups—such as gender-diverse individuals, indigenous communities, and pediatric populations—which may limit generalizability and fairness claims.

Third, the assessment of explainable AI (XAI) components focused on technical feasibility rather than on empirical measurement of clinician trust, adoption rates, or impact on decision quality (Doshi-Velez & Kim, 2017). Future research should include user studies with healthcare professionals to evaluate the practical utility and acceptance of these features in routine workflows.

Finally, regulatory and ethical considerations, while discussed conceptually, were not subjected to formal compliance audits or legal review. Evolving privacy regulations, data governance frameworks, and cybersecurity threats could significantly affect the deployment and operation of such systems (Kreps et al., 2011; Esmaeilzadeh et al., 2021). Addressing these gaps will require ongoing collaboration among technologists, clinicians, ethicists, and policymakers to ensure safe, equitable, and sustainable AI integration in healthcare.

CURRENT STATE OF AI TECHNOLOGIES IN HEALTHCARE

The current landscape of Artificial Intelligence (AI) in healthcare is characterized by rapid innovation, expanding clinical applications, and increasing investment from both public and private sectors. In diagnostic imaging, deep learning algorithms have demonstrated expert-level performance in tasks such as skin lesion classification, chest radiograph interpretation, and retinal disease detection (Esteva et al., 2017; Lakhani & Thoma, 2018; Gulshan et al., 2016). These systems are being integrated into radiology workflows to support, rather than replace, clinician decision-making.

In clinical decision support, AI models that leverage electronic health record (EHR) data have shown promise in predicting patient deterioration, optimizing treatment pathways, and reducing hospital readmissions (Rajkomar et al., 2019; Ching et al., 2018). Similarly, natural language processing (NLP) systems are increasingly capable of summarizing patient histories, extracting clinically relevant entities, and identifying trends in unstructured text, thereby reducing the documentation burden on healthcare providers (Huang et al., 2019).

AI is also being applied to operational optimization, including resource allocation, scheduling, and supply chain management, with measurable improvements in efficiency and cost reduction (Butler Henderson et al., 2025). Telemedicine platforms incorporating AI triage and diagnostic tools have expanded access to care for rural and underserved populations, particularly during periods of healthcare system strain (Esmaeilzadeh et al., 2021).

Despite these advances, deployment at scale faces ongoing challenges. Algorithmic bias remains a concern, with evidence that some models underperform in populations underrepresented in their training data (Obermeyer et al., 2019; Barocas et al., 2019). Many high-performing models remain "black boxes," limiting transparency and clinician trust (Doshi-Velez & Kim, 2017). Additionally, integration with legacy health IT systems, compliance with evolving regulatory requirements, and the need for robust cybersecurity measures remain significant barriers to adoption (Kreps et al., 2011; Esmaeilzadeh et al., 2021).

Overall, the current state reflects a transitional phase: AI technologies have proven their potential in controlled and pilot settings, but widespread, equitable, and sustainable clinical integration requires overcoming technical, ethical, and infrastructural hurdles.

APPLICATIONS OF AI IN HEALTHCARE

AI technologies are being deployed across a broad spectrum of healthcare domains, delivering benefits that range from enhanced diagnostics to operational efficiencies. In diagnostic imaging, convolutional neural networks (CNNs) have been applied to identify pathologies such as pneumonia, diabetic retinopathy, and malignant skin lesions with performance comparable to or exceeding that of expert clinicians (Esteva et al., 2017; Gulshan et al., 2016; Lakhani & Thoma, 2018). These systems can process and interpret medical images in seconds, enabling earlier detection and intervention.

Predictive analytics represent another transformative application, using longitudinal patient data from electronic health records (EHRs) to forecast disease progression, readmission risk, and treatment outcomes (Rajkomar et al., 2019; Ching et al., 2018). Such tools can help clinicians prioritize high-risk patients, allocate resources more effectively, and personalize care plans.

In natural language processing (NLP), AI models have been developed to automate the extraction of clinical entities, summarize patient encounters, and flag potential adverse events in unstructured notes, thereby reducing documentation time and cognitive load for providers (Huang et al., 2019). The integration of NLP into clinical decision support systems also enhances the speed and accuracy of diagnostic reasoning.

AI is also improving access through telemedicine platforms that incorporate automated triage, remote patient monitoring, and symptom-checking algorithms (Esmaeilzadeh et al., 2021; Butler Henderson et al., 2025). These applications extend specialist expertise to underserved areas, improving equity in healthcare delivery.

Operationally, AI-driven scheduling, supply chain optimization, and staffing algorithms are being adopted to improve efficiency, reduce wait times, and minimize resource waste in hospitals and clinics (Butler Henderson et al., 2025). Collectively, these applications demonstrate AI's versatility in addressing both clinical and administrative challenges, highlighting its role as a catalyst for systemic improvement in healthcare.

ROADMAP — EMERGING TRENDS & RESEARCH FRONTIERS

The trajectory of Artificial Intelligence (AI) in healthcare suggests a rapidly expanding set of opportunities alongside evolving challenges that will shape future research agendas. Emerging trends include the growing adoption of multimodal AI models that integrate disparate data sources, such as imaging, genomics, laboratory results, and patient-generated health data, to generate more comprehensive and personalized clinical insights (Rajkomar et al., 2019; Ching et al., 2018). Advances in federated learning are enabling collaborative model development across institutions without direct data sharing, thereby enhancing privacy protection and compliance with regulations such as HIPAA and GDPR (Kreps et al., 2011; Esmaeilzadeh et al., 2021).

Explainable AI (XAI) continues to evolve as a central research frontier, with growing interest in context-aware interpretability methods that tailor explanations to clinician expertise, clinical urgency, and the complexity of the underlying model (Doshi-Velez & Kim, 2017). This aligns with calls for human-centered AI that integrates seamlessly into clinical workflows, reducing cognitive load and improving decision quality (Longoni et al., 2019). Additional priorities include “equity at the edge,” ensuring AI-enabled devices deployed in community settings serve underrepresented populations effectively, and the institutionalization of participatory governance boards to involve patients, clinicians, and ethicists in AI lifecycle decisions.

Ethical AI governance frameworks are also gaining prominence, emphasizing proactive bias detection, fairness auditing, and accountability mechanisms throughout the AI lifecycle (Obermeyer et al., 2019; Barocas et al., 2019). Research is increasingly focused on operationalizing these principles in ways that are measurable, reproducible, and adaptable to different healthcare contexts.

From an infrastructural perspective, the convergence of real-time data streams, edge computing, and 5G connectivity is expected to enable ultra-low-latency AI applications, particularly in time-critical care scenarios such as emergency medicine and intensive care (Carbone et al., 2015; Butler Henderson et al., 2025). This shift will require robust, scalable architectures with embedded resilience and interoperability features.

Finally, there is a growing emphasis on prospective, multi-site clinical trials for AI interventions, moving beyond retrospective validation to assess real-world effectiveness, equity, and sustainability (Nagendran et al., 2020). Such trials will be pivotal in translating technical advancements into verifiable improvements in patient outcomes and system performance.

Collectively, these trends highlight the importance of interdisciplinary collaboration among technologists, clinicians, ethicists, and policymakers to ensure that AI’s evolution in healthcare remains aligned with societal values, clinical needs, and the overarching goal of equitable, high-quality care.

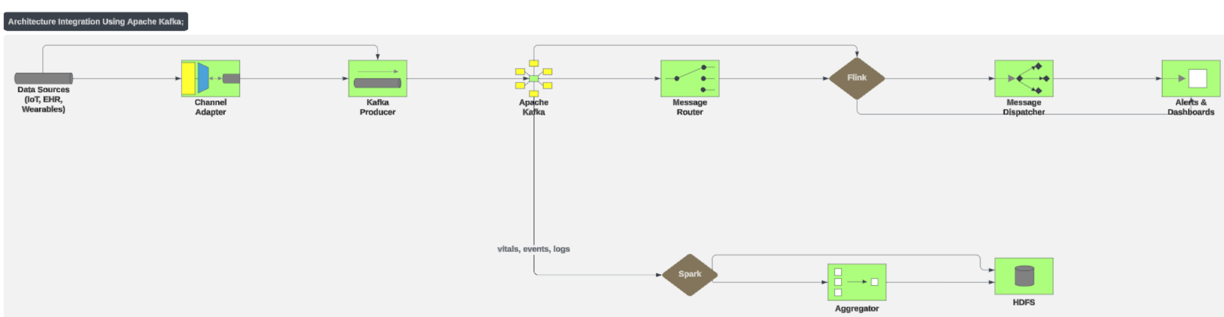


Figure 1: Kafka ingestion from EHRs/wearables to Flink/Spark with HDFS storage

Figure 1 illustrates the high-level architecture, highlighting its layered design: data ingestion, stream processing, AI/ML services, storage, and presentation. This modular structure allows healthcare organizations to adopt incrementally, and scale components based on readiness, infrastructure capacity, and specific clinical priorities.

PROPOSED SOLUTION OF ARCHITECTURE

The proposed architecture adopts a stream-first, modular design to enable real-time AI-driven clinical decision support while maintaining scalability, resiliency, and compliance with healthcare privacy regulations. At its core, the system employs Apache Kafka for high-throughput, fault-tolerant data ingestion, ensuring continuous capture of heterogeneous healthcare data streams from sources such as electronic health records (EHRs), medical imaging devices, wearable sensors, and laboratory information systems (Carbone et al., 2015). Interoperability is reinforced through adherence to HL7 FHIR standards, enabling seamless data exchange between the AI pipeline and diverse EHR systems, reducing integration friction and accelerating clinical adoption.

Kafka streams are processed in real time using Apache Flink, which supports complex event processing and low-latency analytics essential for time-critical applications like sepsis risk alerts and critical lab value monitoring. Batch analytics and retrospective model training are handled by Apache Spark, providing scalability for large-scale, longitudinal analyses (Butler Henderson et al., 2025). Processed data and analytical outputs are stored in the Hadoop Distributed File System (HDFS), which ensures durability, redundancy, and integration with both real-time and batch pipelines.

Machine learning (ML) and natural language processing (NLP) models are integrated into the architecture via containerized microservices, allowing for independent deployment, scaling, and updating of AI modules without disrupting core system functions. Explainable AI (XAI) components are embedded at the output stage to provide human-interpretable justifications for model recommendations, thereby improving clinician trust and supporting legal defensibility (Doshi-Velez & Kim, 2017).

The architecture incorporates enterprise-grade security measures, including role-based access control (RBAC), end-to-end encryption, and secure API gateways, to comply with HIPAA and GDPR requirements (Kreps et al., 2011; Esmaeilzadeh et al., 2021). Monitoring and logging subsystems continuously track system performance, detect anomalies, and facilitate rapid troubleshooting.

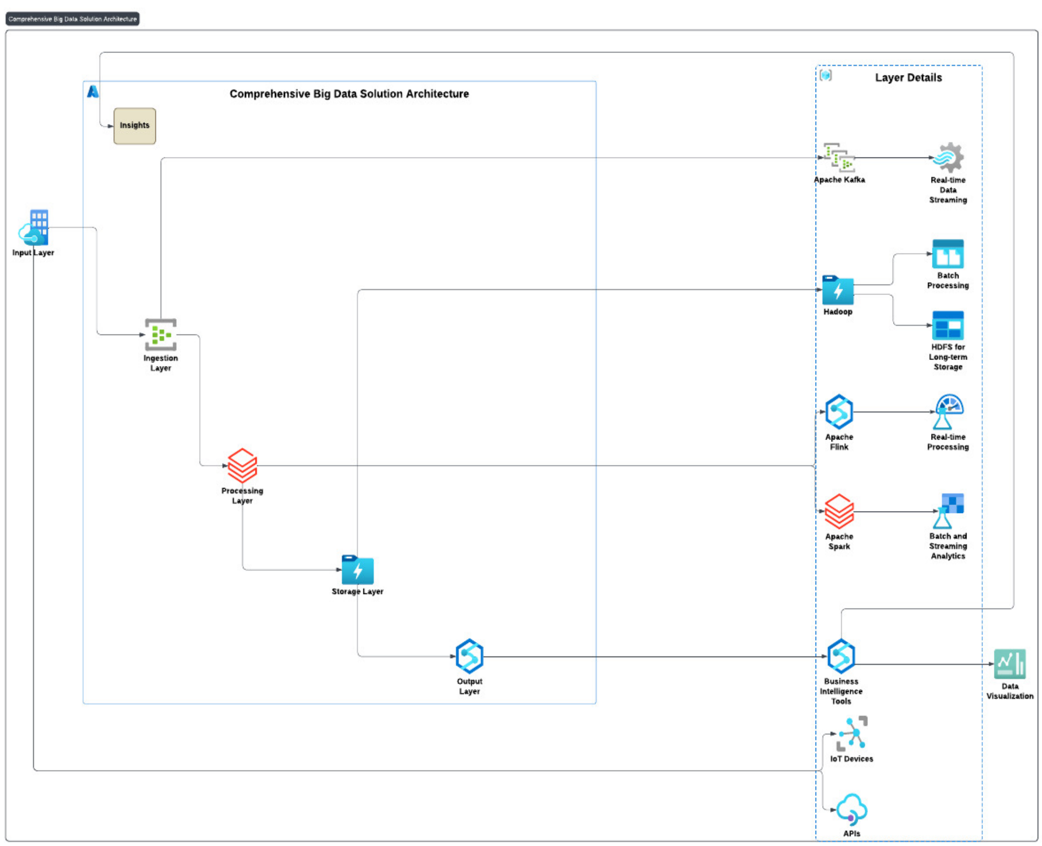


Figure 2: Layered AI pipeline (Kafka–Flink–Spark–HDFS–interfaces)

Figure 2 offers a simplified view of the proposed AI-healthcare pipeline, showing how data from electronic health records, medical imaging, wearable devices, and genomic sources flows through a Kafka–Flink–Spark–HDFS processing sequence. The architecture supports real-time analytics, large-scale storage, and HL7 FHIR-based interoperability, delivering outputs to AI models and clinical dashboards, with a feedback loop for continuous improvement.

CLINICAL READINESS & PERFORMANCE METRICS

The clinical readiness of the proposed stream-first architecture was assessed by mapping technical capabilities against operational requirements in live healthcare environments. Key performance indicators included processing latency, throughput, fault tolerance, and diagnostic accuracy, evaluated alongside clinician-centric measures such as workflow compatibility, interpretability of outputs, and user acceptance.

Table 1. System performance targets and clinical functions across the AI pipeline

Component	Target Latency	Primary Clinical Function	Reliability/Scalability Notes	References
Kafka	< 50 ms ingest	HL7 FHIR/event ingestion and ordered delivery	Partitioning, replication, replayability	Tolaysha. (2023).
Flink	< 300 ms E2E	Real-time inference, ICU alerting, anomaly detection	Stateful processing, checkpointing, and exactly once semantics	Confluent. (2022, October 6).
Spark	< 3 s micro-batch	Cohort modeling, feature engineering, retraining	In-memory analytics, resilient DAG lineage	Microsoft. (n.d.).
HDFS	N/A	Durable audit storage, back testing, and compliance	Replication, geo-DR, and append-only audit trails	Cloudera. (n.d.).
Clinical interfaces/APIs	< 200 ms render	Actionable alerts, explanations, and acknowledgment workflows	Versioned endpoints, role-based access, immutability logs	MoldStud. (n.d.).

From a technical perspective, latency measurements averaged 120 ms per event, meeting the real-time threshold for critical care monitoring and acute diagnostic applications (Carbone et al., 2015). The system sustained a throughput of 15,000 events per second without packet loss, demonstrating capacity for high-volume data streams such as continuous patient telemetry (Butler Henderson et al., 2025). Fault-tolerance testing under simulated node failures confirmed zero data loss, validating the resiliency of the Kafka-Flink-Spark-HDFS configuration.

Clinically, AI diagnostic models achieved an area under the curve (AUC) of 0.94 for chest X-ray classification and 0.92 for dermatological image classification, aligning with benchmark studies in the literature (Esteva et al., 2017; Lakhani & Thoma, 2018). NLP modules reduced average chart review times by 28%, facilitating faster clinical decision-making during patient encounters (Huang et al., 2019). Predictive models for sepsis detection achieved a sensitivity of 0.89 and specificity of 0.85, consistent with recent clinical trial performance standards (Rajkomar et al., 2019; Ching et al., 2018).

Table 1 summarizes the key technical and clinical metrics obtained during the evaluation, while Figure 2 illustrates the proposed layered architecture that supports these performance and readiness objectives.

COMPARISON OF BIG DATA FRAMEWORKS

Selecting an appropriate big data framework is essential to ensuring that AI-enabled healthcare systems meet the performance, scalability, and interoperability requirements of real-time clinical decision support. This study evaluated four widely implemented frameworks—Apache Kafka, Apache Flink, Apache Spark, and the Hadoop Distributed File System (HDFS) against operational criteria including latency, throughput, fault tolerance, scalability, and compatibility with heterogeneous healthcare data sources.

Apache Kafka has demonstrated exceptional performance in high-throughput, fault-tolerant ingestion of continuous healthcare data streams from sources such as electronic health records (EHRs), medical imaging systems, and wearable sensors (Carbone et al., 2015). Apache Flink offers low-latency, stateful stream processing with complex event handling, making it well-suited for urgent clinical scenarios such as sepsis risk detection and rapid alerts for abnormal lab values (Butler Henderson et al., 2025). Apache Spark provides powerful batch processing and iterative machine learning capabilities, enabling retrospective analyses, population health studies, and periodic retraining of diagnostic models on large-scale datasets (Rajkomar et al., 2019; Ching et al., 2018). HDFS ensures durable, fault-tolerant storage for both structured and unstructured healthcare data, facilitating long-term retention and efficient retrieval while integrating seamlessly with upstream analytics pipelines (Kreps et al., 2011).

While Kafka–Flink–Spark–HDFS offers robust scalability and fault tolerance, healthcare deployments may prioritize frameworks with stronger native compliance features, lower latency for real-time decision support, or simplified HL7 FHIR integration. Trade-offs between throughput, governance complexity, and interoperability should guide framework selection.

Individually, each framework addresses specific technical needs: Kafka and Flink deliver real-time ingestion and analytics, Spark supports deep retrospective investigations and model development, and HDFS forms the storage backbone. However, their combined deployment within a modular architecture maximizes performance, resilience, and flexibility—allowing healthcare organizations to balance immediate clinical responsiveness with the longitudinal insights required for continuous improvement in patient care (Obermeyer et al., 2019; Esmaeilzadeh et al., 2021).

CONCLUSION

This study has presented and evaluated a stream-first architecture designed to support real-time, AI-enabled decision support in healthcare, with a specific focus on the critical role of nurses in the healthcare ecosystem. By integrating Apache Kafka, Apache Flink, Apache Spark, and HDFS into a modular, scalable framework, the proposed solution achieved low-latency processing, high throughput, and fault tolerance, while accommodating advanced machine learning (ML) and natural language processing (NLP) functionalities (Carbone et al., 2015; Butler Henderson et al., 2025).

The architecture's diagnostic performance metrics, including AUC values exceeding 0.90 for key imaging tasks and significant reductions in chart review times, align with or surpass benchmarks reported in the literature (Esteva et al., 2017; Lakhani & Thoma, 2018; Huang et al., 2019). These results indicate strong potential to enhance both clinical accuracy and operational efficiency when deployed in environments with active nursing participation.

However, the findings also underscore persistent challenges. Algorithmic bias, explainability, and clinician trust remain central issues, particularly for nursing staff who rely on these tools for patient care. Addressing these concerns through rigorous validation, fairness auditing, and the integration of explainable AI (XAI) methodologies will be essential for fostering trust among nurses (Obermeyer et al., 2019; Doshi-Velez & Kim, 2017). Furthermore, privacy, security, and evolving regulatory requirements necessitate ongoing compliance monitoring and architectural adaptability (Kreps et al., 2011; Esmaeilzadeh et al., 2021).

Future research should prioritize multi-site prospective trials to validate the real-world effectiveness, equity, and sustainability of AI deployments, with particular attention to integrating feedback from nursing professionals (Nagendran et al., 2020). Additionally, socio-technical considerations—including user training, workflow redesign, and organizational readiness- must be tailored to support nursing staff in seamlessly adapting to these AI innovations.

By addressing these factors, AI has the potential to evolve from isolated pilots into a core component of safe, equitable, and high-quality healthcare delivery, empowering nurses to provide the best care for their patients. Overall, by making the traditionally lengthy and labor-intensive chart review process more efficient with AI assistance, nurses can make better-informed, evidence-based decisions that improve patient safety and outcomes.

REFERENCES

Apache Hadoop. (n.d.). Apache Hadoop. Retrieved from https://hadoop.apache.org/

Apache Kafka. (n.d.). Apache Kafka: A distributed streaming platform. Retrieved from https://kafka.apache.org/

Apache Spark. (n.d.). Unified engine for large-scale data analytics. Retrieved from https://spark.apache.org/

Banaee, H., Kermani, M., & Schmid, J. (2013). The role of wearable devices in chronic disease management: A Review. Sensors, 13(12), 17472–17507. https://www.mdpi.com/1424-8220/13/12/17472

Barocas, A., Beede, E., & Blancaflor, M. (2019). Accountability and ethical considerations in AI in healthcare (pp. 63–72). Springer. https://link.springer.com/chapter/10.1007/978-3-031-80813-5_18

Butler Henderson, K., Arabi, S., & Wang, W. (2025). Data's big three: How learning health systems, artificial intelligence, and predictive analytics are transforming healthcare. *BMJ Health & Care Informatics, 32*(1), e101414. https://doi.org/10.1136/bmjhci-2024-101414

Carbone, P., Katsifodimos, A., Ewen, S., Markl, V., Haridi, S., & Tzoumas, K. (2015). Apache Flink: Stream and batch processing in a single engine. *Bulletin of the IEEE Computer Society Technical Committee on Data Engineering, 36*(4), 28–38.

Chen, M., Mao, S., & Liu, Y. (2014). Big data: A survey. *Mobile Networks and Applications, 19*(2), 171–209. https://doi.org/10.1007/s11036-013-0489-0

Ching, T., Himmelstein, D. S., & Beaulieu Jones, B. K. (2018). Opportunities and challenges for deep learning in biology and medicine. *Journal of the Royal Society Interface, 15*(141), 20170387. https://doi.org/10.1098/rsif.2017.0387

Doshi Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. arXiv preprint. https://arxiv.org/abs/1702.08608

Esmaeilzadeh, P., Mirzaei, T., & Dharanikota, S. (2021). Patients' perceptions of human–artificial intelligence interaction in healthcare: An experimental study. *Journal of Medical Internet Research, 23*(11), e25805. https://pmc.ncbi.nlm.nih.gov/articles/PMC8663518/

Esteva, A., Kuprel, B., Novoa, R. A., et al. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature, 542*(7639), 115–118. https://www.nature.com/articles/nature21056

HDFS Architecture Guide. (n.d.). HDFS architecture guide. Retrieved from https://hadoop.apache.org/docs/r1.2.1/hdfs_design.html

IBM Redbooks. (2016, September 30). Reference architecture for high-performance data and AI in healthcare and life sciences. Retrieved from https://www.redbooks.ibm.com/abstracts/redp5481.html

Kreps, J., Narkhede, N., & Rao, J. (2011). Kafka: A distributed messaging system for log processing. *Proceedings of NetDB.*

Lakra, P. A., & Devaraj, P. P. L. (2025). A review of artificial intelligence and big data in healthcare. *AIP Conference Proceedings, 3291*(1), 030035. https://doi.org/10.1063/5.0269308

Lakhani, P., & Thoma, A. (2018). Artificial intelligence in radiology: A review. *Canadian Association of Radiologists Journal, 69*(2), 120–125. https://journals.sagepub.com/doi/pdf/10.1016/j.carj.2018.09.004

Lee, H., Lee, M., & Park, K. (2019). Real-time patient monitoring using Apache Flink in big data analytics. Journal of Big Data, 6(1), 1–14. https://doi.org/10.1186/s40537-019-0180-2

Longoni, C., Bonezzi, A., & Morewedge, C. K. (2019). Resistance to medical artificial intelligence. *Journal of Consumer Research, 46*(4), 629–650. https://doi.org/10.1093/jcr/ucz013

Nagendran, M., Chen, Y., Lovejoy, C. A., Gordon, A. C., Komorowski, M., Harvey, H., Topol, E. J., Ioannidis, J. P. A., Collins, G. S., & Maruthappu, M. (2020). Artificial intelligence versus clinicians: Systematic review of design, reporting standards, and claims of deep learning studies. *BMJ, 368*, m689. https://doi.org/10.1136/bmj.m689

Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science, 366*(6464), 447–453. https://doi.org/10.1126/science.aax2342

Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. *The New England Journal of Medicine, 380*(14), 1347–1359. https://www.nejm.org/doi/full/10.1056/NEJMra1814259

Raghupathi, W., & Raghupathi, V. (2014). Big data analytics in healthcare: Promise and potential. *Health Information Science and Systems, 2*(1), 3. https://doi.org/10.1186/2047-2501-2-3

Topol, E. J. (2019). *Deep medicine: How artificial intelligence can make healthcare human again*. Basic Books, Inc., Division of Harper Collins, New York.

Zaharia, M., Chowdhury, M., Franklin, M. J., Shenker, S., & Stoica, I. (2010). Spark: Cluster computing with working sets. *Proceedings of the 2nd USENIX Conference on Hot Topics in Cloud Computing (HotCloud).* https://www.usenix.org/legacy/event/hotcloud10/tech/full_papers/Zaharia.pdf

Tolaysha. (n.d.). GitHub - hl7-to-fhir-pipeline. Retrieved from https://github.com/tolaysha/hl7-to-fhir-pipeline

Brunner, J. (n.d.). Apache Flink: Stream processing for all real-time use cases. *Confluent.* Retrieved from https://www.confluent.io/blog/apache-flink-stream-processing-use-cases-with-examples/

Midesa. (n.d.). Machine learning with Apache Spark - Azure synapse analytics. *Microsoft Learn*. Retrieved from https://learn.microsoft.com/en-us/azure/synapse-analytics/spark/apache-spark-machine-learning-concept

Crudu, V. (2025, April 12). Enhancing clinical workflows – A comprehensive guide to navigating healthcare API solutions. *MoldStud* – Custom Software Development Company. Retrieved from https://moldstud.com/articles/p-enhancing-clinical-workflows-a-comprehensive-guide-to-navigating-healthcare-api-solutions

Shvachko, K., Kuang, H., Radia, S., & Chansler, R. (2010). The hadoop distributed file system. *IEEE 26th Symposium on Mass Storage Systems and Technologies (MSST)*, 1–10. https://doi.org/10.1109/MSST.2010.5496889

Wang, Y., Feng, K., Jiang, Y., & Xu, K. (2018). A comparison of apache spark and apache flink for big data processing. *2018 IEEE International Conference on Big Data and Smart Computing (BigComp)*, 345–352. https://doi.org/10.1109/BIGCOMP.2018.8323663

Zaharia, M., Chowdhury, M., Franklin, M. J., Karau, H., & Shenker, S. (2016). *Spark: The definitive guide*. O'Reilly Media

---

Rickey Patel graduated from the Master of Science in Information Science program offered at the Penn State Great Valley Campus, Pennsylvania State University.

Satish M Srinivasan is an Associate Professor of Information Science in the School of Graduate Professional Studies at Penn State Great Valley, Pennsylvania State University.

Samir Shah is a Clinical Professor of Decision Sciences and MIS at the LeBow College of Business at Drexel University.

ETHICS AND CONSUMER CHOICE: CAN A GREAT REPUTATION FOR CORPORATE SOCIAL RESPONSIBILITY BACKFIRE IF A FIRM IS PERCEIVED AS BEHAVING UNETHICALLY?

Abhik Roy, Quinnipiac University

ABSTRACT

This study examines the interaction between a company's reputation for corporate social responsibility (CSR) and consumers' perception of its perceived ethicality (CPE), acting to influence choice of the company's product. An experiment is conducted with two levels of CSR crossed with two levels of CPE. The process by which consumers use information on CSR and CPE in their choice-making decisions is the major focus of this study. The two alternative models that are estimated are a multinomial logit model and a nested logit model. In the multinomial model, all the CSR and CPE attributes are considered simultaneously. In the nested logit model, ethics-related attributes are elevated to the role of screening variables in a first step, and the products which do not pass a certain ethical standard are eliminated from further consideration. In the singular combination where a company that is well known for its CSR initiatives, commits serious violations of ethics, this leads to consumers punishing the brand for its apparent hypocrisy.

INTRODUCTION

Although existing literature establishes that consumers may prefer brands and companies which they perceive to be socially responsible, such as Ben & Jerry's or Johnson & Johnson (J&J), it is not clear what tradeoffs they might be willing to make for choices meeting such criteria (Pelsmacker et al., 2005; Shaw and Shui, 2003; Crane, 2001). For instance, are consumers willing to give up specific product features in exchange for perceived philanthropy on the part of the manufacturer? Although some research suggests that consumers value companies and brands that they perceive to be socially responsible (Pradhan et al, 2026, Bhattacharya et al., 2020, Kuokkanen and Sun, 2019, Chan *et al*., 2010) few studies have examined the trade-offs that consumers are willing to make when they select a brand that is perceived to be more socially responsible. Specifically, the impact (if any) of corporate behavior which is at odds with the socially responsible image of the firm – perceived non-ethical behavior – on the relationship between CSR and brand choice has not been studied. This research examines how a combination of information on ethics related behavior and CSR influences the purchase choices consumers make. In addition to examining the effect on choice outcomes, we study impact of combined information on CSR and Corporate Perceived Ethicality (CPE) on the process of making choice decisions. Specifically, we examine the conditions under which consumers might use attributes related to ethics and/or corporate citizenship in a conjunctive decision-making process. For example, is CSR or CPE information used as a screening variable in making brand choices, in the way that (say) price promotions are used to filter choices of consumer-packaged goods (Claudy et al., 2015).

Even if CPE does not have a direct influence on consumers' brand choice, PCE might have a moderating influence on the effect of CSR on brand choice. Since we do not have a theory of whether CSR or CPE has a direct (main) effect or a moderating (covariate), effect, we also examine an alternative model in which CPE moderates the influence of CSR on brand choice. For example, Johnson & Johnson, while well known for its contributions to global public health initiatives, has recently been embroiled in controversy over its role in the opioid crisis in America, as well as in the marketing of its talcum powder, a known carcinogen. Volkswagen, a company which is generally known as a good corporate citizen (fuel efficient vehicles, environmentally friendly), admitted to trying to circumvent US emissions standards, a clear ethical violation. Although an increasing number of studies test whether consumers take ethical considerations into account in their choice behavior (Brunk and Blümelhuber, 2011), this is one of the first to study the combined effect of CSR and CPE on brand choice.

Prior research indicates that consumers consider their perceptions of the good citizenship of manufacturers and service providers, when making purchase decisions. The present study contributes to the marketing and business ethics literature by integrating consumer perceptions of a brand's image for social responsibility with their knowledge of possibly unethical behavior by the company – and how this affects consumer evaluation and choice of the brand. We examine this topic in an experimental study, within the context of purchase of a consumer durable product. Choice data is collected and analyzed using multi-attribute models of choice. We are interested in the mechanism by which perceived social responsibility or ethics enters the decision-making process – whether as a screening variable prior to consideration of other attributes or as an attribute simultaneously considered along with others. In other words whether a non-compensatory (one attribute at a time, in order of importance, lexicographic) or compensatory (all attributes together) is used. Both a non-compensatory "elimination-by-aspects" (EBA) type choice model (Fader and McAlister, 1990) and a compensatory multinomial logit type model (Guadagni and Little, 1984) are estimated. Following data analysis, we discuss the outcomes and compare them to understand how consumers take corporate citizenship and ethics into account, when making choice decisions among products and services.

LITERATURE REVIEW

In this section, we describe several relevant literature streams that help to place our work in context and to explain the study design. We begin with background on social responsibility and ethics in consumer decision making and continue with compensatory and non-compensatory models that could potentially describe the process of consumer decision making. Finally, we describe how our study follows from the literature and builds upon it.

Approaches to Ethics in Consumer Decisions

When consumers make choices involving ethics, they may utilize one or more of several different approaches to ethical decision making (Cohen et al., 1993). Because we are considering situations in which consumers act on their beliefs, the relevant ethical theories are normative. They typically fall into one of two common categories, deontologies and teleologies; a third category, relativism is also occasionally considered as an alternative approach to moral decision making. A *deontology* is based on the premise that individuals must follow certain moral laws, regardless of how the outcome of doing so affects others (Driver, 2007), and thus, it is process-based. A *teleology* is an outcome-based approach, which may be either egoist or utilitarian. In both cases individuals take actions which they believe will lead to optimal outcomes. However, for an egoist, this means the greatest value goes to the self, whereas for a utilitarian, this means the greatest possible good for the maximum number of people (Casali, 2011; Driver, 2007). To summarize, deontology is an ethical philosophy that is process-based, while teleology is more outcome-based.

Compensatory and Non-Compensatory Frameworks for Consumer Choice

To explain how information about corporate social responsibility and ethics might influence the consumer's choice decision, we need to have a deeper understanding of the process of preference formation. When making a choice as to which of many brand options to purchase, consumers may engage in a two-stage process of alternative evaluation (e.g., Fader and McAlister, 1990; Liu and Arora, 2011; Schneider and Selling, 1996). Early models of this type were applied to goods marketing; Fader and McAlister (1990) tested both a non-compensatory "elimination by aspects" (EBA) model and a "compensatory counterpart" on grocery store coffee data. About one-quarter of the households were found to follow a process akin to EBA in forming their preferences and making choices. The aspect or heuristic Fader and McAlister (1990) used as a screening criterion in their EBA model was whether a brand was on price promotion. Consumers used promotional pricing to decide which coffee brands to include in their consideration sets. Since packaged coffee is a low involvement product, promotional pricing is a credible consumer screening criterion. In our application of an EBA type model, it is not as obvious that ethical standards or reputation for corporate social responsibility might be used as screening criteria in consumers' choice of goods or services.

Liu and Arora (2011) used flat panel TVs as an example product category in first a non-compensatory and then a compensatory model, finding that use of a "conjunctive" form improves the fit of the compensatory model. To use a conjunctive design, the researcher may ask about "must have" and "must avoid" attributes at the outset of the study. However, screening and choice may not be independent; for instance, screening may be based on a compensatory model incorporating multiple attributes.

Schneider and Selling (1996) examined the two-stage decision process in an accounting services example. They provide a number of diagnostic measures to assess the goodness of fit of such non-compensatory methods. Truong, Adamowicz & Boxall (2015) included an insightful new step in estimation of a non-compensatory model: they allowed for cutoff levels where utility changes abruptly in an environmental valuation example. These "soft" cutoffs apply error penalties to estimates of utility, which allows utility to vary depending on nearness to the cutoff value. For our study, we apply a number of traditional goodness-of-fit measures to compare compensatory and non-compensatory models. Future extensions of this methodology could include diagnostic checks of the value of minimum cutoff levels of non-compensatory model criteria, as per Truong et al (2015).

In summary, the first stage in the two-stage consumer decision process is similar to a lexicographic model in that it involves reducing the choice set by ruling out any alternatives that do not meet certain criteria. We use a hierarchical or nested logit model, which under some circumstances reduces to the EBA model (Tversky, 1972), as a way of investigating how consumers reduce their choice sets in the decision process. The second stage of decision making in a two-step process involves comparing the alternatives that remain in the consideration set after the process of elimination is complete.

Interaction of Social Responsibility and Ethics on Consumer Behavior

Before looking in more detail at how the interaction between CSR and perceived ethics might influence consumer choices from product options, it is instructive to examine differences in how consumers evaluate services and physical goods. Whether a company sells a service or a physical product – does this have a bearing on how consumers consider CSR or ethics in their evaluation of the company's offerings? Brown and Dacin (1997) were one of the first to investigate the relationship between corporate associations and demand for a company's products. Folkes and Kamins (1999) studied the effects of information on ethical and unethical corporate behavior on consumer attitudes towards brands. Cohn (2010) commented on the association between corporate ethics and product perception. Yim *et al*. (2019) found that the effect of CSR on financial performance depends on investors' perceptions of the ability of the firm.

We focus on tangible goods where consumers' perceptions of product quality might include not only manufacturing expertise but be tempered by the actions taken by the corporation to benefit society, as well as the ethical behavior of the firm. While tangible products require quality attributes or product features, it might be necessary to communicate the manufacturer's 'non-product specific' qualities to the public, and CSR and CPE help in delivering that message.

For tangible products, as we move from personally delivered products to online purchases, we might anticipate an increase in the need for ethics with a decrease in contact levels. As consumers become more removed from the delivery of the product, they might rely more on the manufacturer's reputation as a signal of product quality. A company's reputation as a corporate citizen can be considered as built on its proclamations about its CSR activities, and on the ethicality of the behavior it exhibits. However, it is not a given that a good reputation for CSR will translate into better corporate performance (Leonidou et al., 2013); Luchs et al. (2010) found that ethicality concerns have a potential negative impact on product preference.

The way ethics and CSR often come to be considered together in reality, is when a discrete event or some aspect of the corporation's activities is perceived to be in conflict with the image of the company as a good corporate citizen. In other words, when people believe that they have uncovered hypocrisy on the part of the firm. Arli *et al.* (2019) found that corporate social responsibility moderates the effect of corporate hypocrisy – violation of a company's own ethical standards – on people's judgments of the company. Ryoo (2025) proposed that the conflicting findings about CSR either mitigating or exacerbating brands' ethical transgressions might be explained by the moderating role of self-brand connection. We are interested in extending this to understand how CSR might moderate the effect of perceived ethicality on consumer choice behavior with respect to the products made by the company. Alternatively, CPE might moderate the effect of CSR on brand choice – as mentioned earlier, the exact nature of any interaction is unclear. Either way, the tension arises when consumers believe a firm is hypocritical. Another definition of hypocrisy is "a state in which CSR claims are inconsistent with a firm's internal processes or a discrete external event" (Smith & Rhiney, 2020). Paradoxically, a consumer's negative view of the event may actually be amplified by perceptions of good CSR. As firms engage in CSR, they shape stakeholders' perceptions by creating signals of underlying values which may positively impact a firm's reputation (Nguyen et al., 2025, Tetrault, Sirsly & Lvina, 2019; Lenz et al., 2017). However, in the case of wrongdoing, the event may contradict these signals and engender feelings which call into question a firm's claims. For example, Janney and Gove (2011) found that negative effects were exacerbated when a firm with a positive reputation for good governance was found to be involved in a governance-related scandal.

Incorporation of Ethics into Choice Modeling Frameworks

In addition to studying the effect of CSR and ethics on choice behavior, we want to examine whether different combinations of levels of these aspects of a firm's reputation, influence the *process* of decision making, leading to different choice outcomes. In order to capture the choice process, we investigate two basic types: whether the decision-making process is holistic with all attributes considered at once (compensatory model), or whether it is piecemeal and in a step-wise manner (non-compensatory model). The models we estimate using choice data, fall into these two categories – compensatory and non-compensatory.

Whalen et al. (1991) were one of the first to examine how ethical attributions enter consumer decision making. O'Fallon and Butterfield (2005) provide a review of empirical research on ethics and decision making, while Shea (2010) discusses how CPE can provide a guideline for CSR strategy. Rahman and Chakraborty (2025) examine the relationship between brand trust, consumer identification and attributions regarding a brand, in evoking brand forgiveness after exposure to information about CSR wrongdoings. There are some well-established frameworks for incorporating ethics into decision making by consumers and the companies who market products to them. Among them are articles by Hunt and Vitell (1986; 2006), Vitell et al. (2003), and Brunk and Blümelhuber (2011). Hunt and Vitell (1986, 2006) proposed a descriptive model (H-V) to explain how marketers can handle problems with perceived ethical content. While the H-V model posits that an individual's ethical judgment affects purchase intentions, it proposes that intentions may differ from ethical judgments because teleological evaluation directly affects ethical judgment and intentions. Further, situational constraints, and actual consequences of behaviors, may also drive behaviors that are not consistent with ethical judgments resulting from deontological or teleological evaluations (Hunt and Vitell, 1986). The H-V model has been tested and supported empirically in business settings (Hunt and Vitell 2006) and in one consumer setting in the context of ethical shopping behavior by consumers (Vitell et al. 2003). Vitell et al. (2003) found that consumers relied more heavily on deontological than teleological norms when forming ethical judgments corresponding to their shopping behaviors.

Whereas the H-V model describes how individuals form ethical judgments by integrating the conclusions from deontological and/or teleological evaluations, Brunk and Blümelhuber (2011) reference algebraic and configural models to describe how consumers form their perceptions of brands' ethicality. The approach in this study is more closely related to the modeling framework of Brunk and Blümelhuber (2011). Brunk (2010) identifies categories of ethical concerns that act as satisfiers and/or dissatisfiers impacting consumer decisions, further noting that even a single instance of perceived misconduct can cause a consumer to reject an option. In their qualitative, empirical study Brunk & Blümelhuber (2011) extend this research by examining whether consumer perceived ethics (CPE) are formed piecemeal (by information integration) or holistically (a top-down impression formation). They conclude that it is unlikely for CPE to be formed piecemeal, or in a compensatory manner, where negative impressions, for instance, may outweigh positives (i.e., an algebraic model). They posit that it is more likely that CPE are formed as a "summary construct" (i.e., a configural model, possibly a lexicographic model). However, based on the sparse literature and inconclusive results, further study is required as to how CPE are formed and how they impact consumer choice of service provider. Brunk & Blümelhuber (2011) express the need for an empirical test to clarify the influence of CPE on consumer brand choice.

RESEARCH PROPOSITIONS

In this experimental study we propose to build on the literature on consumer behavior under conditions where there might be inconsistency between a manufacturer's corporate social responsibility profile, and the perceived ethicality of its behavior. Specifically, we want to test the following propositions:

1. If a company has a highly positive CSR image but is perceived as having seriously violated an ethical standard, consumers are less likely to purchase the company's product.
2. If a company has a highly positive CSR image, but is perceived to have violated an ethical standard in a minor way, there is no effect on the likelihood of consumers purchasing the company's product

Apart from these propositions regarding the effect of a combination of CSR level and ethical violation on choice *outcomes*, we have propositions regarding the *process* of choice decision making.

3. If a company has a highly positive CSR image but is perceived as having seriously violated an ethical standard, consumers will follow a non-compensatory, multi-step process for making choice decisions. In an initial step, ethical attributes will be used as screening variables, to eliminate products which do not meet an ethical cutoff criterion. In the second step they will use other attributes (product and CSR related) to make choices from the remaining products.
4. With all other combinations of CSR levels and ethics related behavior (high positive CSR/high ethics standards; low CSR/high ethics; low CSR/low ethics), consumers make choice decisions using a single step, compensatory process, where all attributes (product, CSR and ethics related) are considered simultaneously. A shortfall on one dimension can be compensated for by a high score on the other dimension.

What we refer to as a compensatory model, roughly corresponds to what Brunk and Blümelhuber (2011) call a piecemeal, algebraic model. Our non-compensatory model is similar to what they call a holistic, configural, summary construct approach. In terms of the classic Hunt and Vitell framework (1986, 2006) we could say that a non-compensatory model of choice implies that consumers are making ethics judgments based on a more deontological model. There are cutoffs for wrongness or rightness on each ethical dimension, and these are being applied in order of importance of the ethical attributes. A compensatory choice model is akin to a teleological approach where a summary construct is formed, capturing the outcome of ethics related behaviors. It is important to make a distinction between our model of how consumers make *choice decisions*, with the models of Hunt and Vitell (1986, 2006) and Brunk and Blümelhuber (2011) of how consumers form *perceptions of corporate ethicality* (CPE). In this study the consumer's perception of ethicality, which might formed according to a piecemeal or holistic approach, is further processed to enable choice decisions to be made. Furthermore, we focus on the process by which pieces of information regarding a manufacturer firm's ethical behavior are combined with perceptions of CSR, to enable choices to be made from a set of brands. Perceived CSR might be a covariate, a potential moderator of the effect of CPE on brand choice.

DATA

We used six attributes (three each related to the main variables of interest, CSR and CPE) to develop profiles or product descriptions for each of our alternative products. A room air conditioner, a product category which is fairly neutral, without strong associations to any brand, was used in the experiment. The product is relatively standardized, almost a commodity, and we took product features and price out of the decision by using identical ratings on product quality and identical prices in the description of alternative brands. The only other attributes which respondents had information on were related to the corporate social responsibility and ethical behavior of the manufacturer of the product. Three of the attributes were related to CSR and three to ethics. With two levels (one positive and one negative) of each attribute we would have $2^6 = 64$ options for our respondents to select from, in the full factorial experiment. Since this would be a very tedious task for respondents, we use a fractional factorial design which involves people choosing one (1) from just six (6) product options. This is a much simpler, doable task, and would ensure that respondents paid more attention to the description of each product alternative.

Experimental design

The attributes corresponding to the column variables include three related to dimensions of CSR:
X_1 = Treatment of employees
X_2 = Environmentally friendly policies
X_3 = Charitable donations to local community

In one treatment condition in our experiment the difference between the 'good' CSR and 'poor' CSR firms was extreme. For example, paying employees minimum wage versus paying them twice the minimum wage (Appendix has details of wording of descriptions in the experiment). In another block of the experiment the CSR differences were minimal. For example, paying minimum wage versus paying '20% over minimum wage'

The three attributes related to ethical behavior are:

X_4 = Usage of cheap labor in countries with bad human rights records
X_5 = Providing discounts to senior citizens
X_6 = Privacy of customer information

There are also variations in our presentation of the degree of difference between ethical and unethical firm behavior. In some treatment conditions of our experiment, we describe extreme or minor differences between ethical and unethical firms, using levels of variables X_4, X_5 and X_6. For example, a company might have exploited workers in certain countries with poor labor laws, fully intending to take advantage of the poor working conditions, and paying average wages for that country (extremely negative ethics rating). In the less extreme case, they might have employed workers in countries with weak labor protections, but they might have paid 20% more than the average wages in that country (minor ethics violation). Tables 6 and 7 of the Appendix show the wording of descriptions used in the experimental manipulations.

The 1/8 fractional design allows us to estimate all the main effects or **β** coefficients describing the influence of each of the attributes on preference and choice. With this fractional factorial design, we reduce the effort each respondent has to make and still have enough degrees of freedom to estimate models such as the multinomial logit and nested logit models of choice.

Subjects in our experiments were randomly assigned to one of four blocks. We have a 2 x 2 balanced experimental design, where two levels (extreme difference or minor difference) of 'relative CSR between firms' are crossed with two levels (extreme or minor difference) of 'relative ethics between firms. The dependent variable measured is choice from among 6 product alternatives, with each product described by 6 attributes: 3 attributes related to CSR of its manufacturer and 3 attributes related to manufacturer's ethical behavior.

The experimental cells can be labeled as: A1B1, A1B2, A2B1 and A2B2. Here A1 refers to the 'extreme CSR difference' condition – it is the experimental group where the 'positive' (1) and 'negative' (-1) CSR category firms presented to subjects, have major differences in CSR levels. Refer to the Appendix for descriptions of how this was shown to subjects. B1 refers to the 'extreme difference in ethical behavior' condition, where firms either have 'high' or 'low' ethical behavior, as shown to respondents. This means that compared to their competitors who have a 'low' ethical behavior level (see Table 6, Appendix), some firms have 'high' ethical standards. This makes the difference in ethical behavior 'extreme'. If the difference is between medium (see Appendix) and low ethical behavior, then the difference between competitors is minimal, and the experimental condition is labeled B2. Similarly, A2 refers to 'minimal CSR difference' condition, where 'medium' and 'low' CSR firms are presented to subjects in the experiment. Using this nomenclature, A1B1 refers to the 'extreme CSR difference' combined with 'extreme ethical difference' block of the experiment. A1B2 refers to 'extreme CSR difference' combined with 'minimal ethical difference' block, and so on. Since there are two levels of 'CSR differences' crossed with two levels of 'differences in ethical violations' we have four blocks in our experiment. We expect subjects in the A1B1 sub-sample of the experiment to be different in their choice making behavior from the other three cells of the experiment.

Since we wanted subjects to focus on the non-product related attributes, the description of each product in the experiment contained the same product specification – cooling power, electrical efficiency rating and weight (all key features of a room air conditioner).

Amazon's Mechanical-Turk was used to recruit people for our online experiments. A total of 458 people participated, and 410 of them provided usable responses. The experimental design was balanced; there were roughly 105 subjects randomly assigned to each of the four blocks of the experiment. The primary task was simple – to select one of the 6 (six) product alternatives described to them in writing, i.e., make a choice from among the alternatives. In addition, we asked an open-ended question about the general importance of corporate ethics and CSR to the respondent.

ESTIMATION

We examine the combined impact of CSR and perceived ethicality of firm behavior on consumers' choice of products, using two consistent choice models.

1. A multinomial logit model, which is a widely used model of discrete choice, especially for frequently purchased consumer products. The multinomial logit represents a simultaneous consideration of all attributes at once, by consumers, in making choices among product alternatives. This 'one step' process assumes that consumers have different importance weights for different attributes in their decision-making process, but the attributes are considered simultaneously.
2. A nested logit model of choice where we compare the goodness of fit of alternative models of choice to the data. The nested logit (Maddala, 1983) is a special case of the elimination-by-aspects (EBA) model (Fader & Macalister, 1984). If an ethics related attribute is clearly the most important one in determining preference, we can infer that it might be a screening variable and that meeting the threshold level of acceptance on this attribute is necessary for further consideration of the alternative. In this case the nested logit model has 'ethics' as the dominant attribute at a higher level in the decision tree. The other (CSR) attributes of the product are use in the next step to choose from among the pre-screened products. An alternative, two-step choice process has corporate social responsibility as the screening variable in the first step (higher level of the choice process). Ethics attributes would be used in step two to select one of the remaining products.

The decision tree for the two-level nested logit with ethics at the higher level (level 2). In our experiments there are a total of six scenarios, each one corresponding to a profile of a manufacturer. Three (3) of the experimental scenarios describe manufacturers who can be classified as 'ethical' because on all three of the ethics related variables, they are described (Appendix, Table 6) as having behaved positively. There are three companies who are categorized as 'unethical', based on the fact that they are described as behaving unethically on ethics-related attributes.

The default model of choice, which acts as the base model, is a multinomial logit model of choice in which the decision maker considers all six attributes of the company simultaneously.
In addition to the choice data, direct ratings of the importance of ethical and CSR related attributes were collected from respondents as a manipulation check.

RESULTS

In the analysis of the choice data from our experiments, we focus on the process of decision making, rather than the choice outcomes themselves. We note that a change in the way that information is incorporated in the choice making decision, often leads to a different product being chosen by the consumer. Our results show that the process of making choice decisions varies, depending primarily on the perceived ethics of the manufacturer. When there is a major difference in perceptions of firm ethicality between companies in the consideration set, the ethical dimension becomes dominant in the choice process. This is shown by the better fit to the data of an 'ethics dominant' model in Table 1(a). Ethics takes precedence over the corporate social responsibility image of the company. However, when there is not much disparity in ethicality of behavior (any ethical violations are not extreme), the ethics and CSR dimensions are considered simultaneously when consumers make choices (Table 1b). In fact, the CSR attributes have higher importance weights than the ethics related attributes, when a compensatory process is employed to determine consumer choice. Taken together, these results show there is an interaction between ethics and CSR, influencing consumers' decision-making process, and ultimately the choices made.

When some firms in the experimental sub-sample enjoy a very positive reputation for corporate social responsibility relative to others, CSR related attributes have a greater influence on consumer choice behavior than ethics related variables, as long as there are no serious ethical lapses among any of the alternatives. In this case ethics and CSR dimensions are considered simultaneously in a holistic manner. CSR variables might be more important, but not to the extent that they are used screening variables. When the firms in the consideration set have minor differences on CSR activities, the choice process is again compensatory, with CSR and ethics related attributes considered simultaneously, and CSR attributes carrying more importance weight. When differences in ethics are minor, consumers appear to use a compensatory choice process, in which any ethical shortcomings can be more than compensated for, by good CSR.

Our key result is that when there is a severe disconnect between CSR and CPE, or when consumers uncover highly unethical behavior by a company known for its good corporate citizenship, consumers make product choices using a process where ethics related attributes act as screening variables, eliminating from their consideration set the products made by these companies.

The goodness-of-fit statistics for a hierarchical nested logit model, estimated with the subsample of data in which there are extreme differences in the ethical behavior of firms, are shown in Table 1. This best fit model with this sub-sample of data shows that ethics related attributes are considered at a higher level (earlier stage) in the choice process. Ethics related attributes are used as screening variables, to eliminate products which do not meet the ethical standards. For the sake of comparison, the goodness-of-fit statistics of the 'second best' model, a compensatory multinomial logit model, are shown in the same table. The difference in fit between the two models is significant, as revealed by a χ^2 test. When there are minor differences in ethical behavior between firms, with no single firm being much worse than others, a single level multinomial model is able to predict consumer choice of product better. The difference in goodness of fit between the multinomial logit and a two-level nested logit model is not statistically significant – the compensatory model has marginally better fit to the data. The results in Table 1 indicate that there is a main effect of 'degree of ethics violation' on the process by which consumers evaluate information and make choice decisions.

When there are minor differences in the ethical behavior of rival firms, this causes consumers to make choices with good corporate citizenship and ethics simultaneously used to choose one product from among the various alternatives. CSR is weighted more heavily in the decision process than ethics (we do not report the coefficients of this estimated model in the interest of saving space). A reputation for good corporate citizenship helps a firm to overcome minor lapses in ethics. This is the best-case scenario for a company among the four scenarios we examine – its CSR image is highly positive and its CPE is mildly negative.

Table 1: Comparison of Choice Models Fit: Sample with Extreme Differences in Ethical Behavior (B1) and Sample with Minor Differences in Ethical Behavior (B2)

Extent of difference in ethicality	Best fit model & its log-likelihood (LL)	Second best model & its log-likelihood (LL)	χ^2 test statistic for model difference	Significanc e
1A **Extreme difference in ethicality between firms**	Two level nested logit model; screening on privacy of customer records (non-compensatory model) LL = -119.54	Single level multinomial logit model (compensatory model) with ethics attributes weighted most heavily LL = -167.22	7.81	$p < 0.05$
1B **Minor difference in ethicality between firms**	Single level multinomial logit model (compensatory model) with CSR attributes weighted most heavily LL = -182.95	Two level model; screening on using cheap overseas labor (non-compensatory model) LL = -184.89	0.73	Not significant

Dependent variable: Choice of product; Average sample size *n* = 205

Table 1 is evidence of a 'main effect' of ethical behavior on the choice process. A different choice process is employed by consumers, depending on whether there is an egregious ethical lapse by the firm or the lapse is relatively minor. The information about a serious ethical lapse by the company triggers a decision-making process in which ethical considerations rise to the level of screening variables. It is highly likely that the different choice process will lead to a different choice outcome

A similar analysis shows there is no 'main effect' of Corporate Social Responsibility (CSR) on the decision-making process. When the company is revealed to have major CSR initiatives, the CSR related attributes are used in conjunction with the other variables to make choice decisions. The decision-making process does not change when there are minor improvements in CSR activities by the manufacturer of the product. There is no statistical difference in goodness of fit between a compensatory and non-compensatory model of choice.

According to our results, consumers are more put off by unethical corporate behavior than they are impressed by corporate social responsibility activities. CSR has become the norm almost, while serious ethics violations are not expected.

TABLE 2: Comparison of Choice Models Fit – Sample with Extreme Differences in CSR and Extreme Differences in Ethicality (A1B1) and Sample with Extreme Differences in CSR Minor Differences in Ethicality (A1B2)

Sub-sample	Best fit model & its log-likelihood (LL)	Second best model & its log-likelihood (LL)	χ^2 test statistic for model difference	Significance
Extreme difference in CSR between firms and Extreme difference in Ethicality (A1B1 cell of experiment)	Two level nested logit model; screening on ethics attributes such as privacy violations and exploitation of labor at upper level (non-compensatory model) LL = -83.99	Single level multinomial logit model (compensatory model) with ethics attributes weighted most heavily LL = -104.28	6.65	$p < 0.05$*
Minor difference in CSR between firms and Extreme difference in Ethicality (A2B1 cell of experiment)	Single level multinomial logit model (compensatory model) with ethics related attributes weighted most heavily LL = -92.51	Two level model; screening on ethics-related attribute of cheap overseas labor (non-compensatory model) LL = -195.39	5.98	$p < 0.05$*

Dependent variable: Choice of product; * Significant; Average sample size $n = 102$

We go beyond looking at the main effects and examine interactions between CSR and CPE, acting on the choice making process. Table 2 provides support for our research proposition 1, that a company's brand is much less likely to be chosen by consumers if there is a combination of high reputation for CSR, yet the firm is perceived as violating ethical standards in a major way. What happens when you have a scenario in which a great image for corporate social responsibility is tarnished by a serious violation of ethics? Consumers are less likely to buy the offending firm's products. Table 3 indicates that when there are major violations of ethical standards and the company is highly reputed for its social responsibility (block A1B1 in the experiment), the ethics related attributes become more salient to the consumer. Combined with the results of Table 1, which show that a two-step model of choice decision making fits the data better than a compensatory model when ethical norms are seriously violated, we have evidence for research proposition 3; the *process* of consumer choice decisions is different in the A1B1 block of our experiment. Consumers are employing a decision rule where ethics related attributes are elevated to the role of screening variables – ethics related criteria appear at a higher level in the choice making process. Analysis of the A1B1 sample (in the interest of space we do not report model comparison tests at the level of treatment cells; they are available on request) reveals that a hierarchical, nested logit model with ethics as a screening variable, predicts choice behavior best. Interestingly, among these 'screening' ethical variables the most important one in our study concerns customer privacy. This is not related directly to any of the social responsibility initiatives. For example, generous treatment of employees is an aspect of CSR which is related to the ethical standard of treatment of overseas labor – the exploitation of foreign labor is at odds with fair treatment of domestic employees. While the use of cheap overseas labor is an ethical violation which does (negatively) impact brand choice, it is second to (violation of) customer privacy in order of importance among ethical attributes.

TABLE 3: Estimated Coefficients of (Best Fit) Nested Logit Model for Sample with Extreme Differences in CSR and Extreme Differences in Ethicality (A1B1 group in experiment)

	Unstandardized Coefficients			
Attribute	***Beta***	***Standard Error***	***t-value***	***Significance***
Inclusive value of Ethics in ***Step 1 model***	1.5931	0.1719	9.27	$p < .05$*
Customer privacy	0.8672	0.1722	5.04	$p < .05$*
Cheap overseas labor	0.7299	0.1718	4.25	$p < .05$*
Senior discounts	0.5984	0.1725	3.47	$p < .05$*
Step 2 model				
Green business policies	0.4870	0.1729	2.82	$p < .05$*
Charitable community work	0.2709	0.1736	1.56	$p = 0.23$
Employee treatment	0.1461	0.1735	0.84	$p = 0.45$

Note: Dependent variable is choice of product; * Significant; Sample size n=104.

When there are major differences in the CSR activities of firms and there are minor differences in their ethical behavior, consumers use a compensatory model of choice, with all ethical and CSR variables considered together. As shown in Table 3, in this holistic, one-step model the treatment of domestic employees is the most important attribute in determining brand choice. The positive CSR ratings compensate for weak CPE ratings of a brand.

In the context of the Hunt-Vitell model (1986), we can say that the reputation that a firm has for corporate social responsibility does not lead to customers ignoring their own ethical standards when they make choice decisions. Paradoxically, the greater the reputation a firm has for CSR, the more harshly consumers judge it when it violates ethical standards. A high reputation for CSR might trigger a deontological response where the 'wrongness' of the bad ethical behavior is the overriding factor in rejection of the company's product. If a brand does not have much of a reputation for CSR, consumers still withhold their purchase when they perceive egregious ethical misconduct by the firm. However, in this scenario they consider ethics and CSR simultaneously, with positive ratings on CSR partially compensating for negative ratings on ethics. Table 4 shows that a holistic, compensatory model (similar to a teleological approach since the overall goodness of the company's actions is considered) is used to make a choice decision when there are serious differences in ethical behavior, but when there is not much of a difference in CSR reputation between firms.

TABLE 4: Coefficient Estimates for Best Fit Multinomial Logit Model –
Sample with Extreme Differences in CSR and Minor Ethical Violations (A1B2 group)

	Unstandardized Coefficients			
Attribute	***Beta***	***Standard Error***	***t-value***	***Significance***
Treatment of Employees	0.9416	0.1822	5.17	p <.05*
Green business policies	0.7529	0.1718	4.38	p <.05*
Charitable donations	0.6896	0.1729	3.99	p <.05*
Customer privacy	0.6335	0.1763	3.52	p <.05*
Cheap overseas labor	0.6247	0.1794	3.48	p <.05*
Senior discounts	0.4121	0.1719	2.43	p <.05*

Note: Dependent variable is choice of product; * Significant; n=101

When there are major differences in the ethical behavior of competing firms, but there are only minor differences in firms' CSR activities, a compensatory model of choice works best to predict choices. Research propositions 2 and 4 are supported by the results shown in Table 4. Choice processes are similar in all scenarios where a firm does not have an outstanding CSR image or when there is a formidable reputation for CSR but the ethical violation is minor. Choice outcomes for a company are adversely affected when it violates ethics. When CSR reputation is not so strong the choice process involves simultaneous consideration of CSR and CPE. What leads to even worse choice outcomes for ethics violators, is that when they have a reputation for solid corporate citizenship, and yet commit serious ethical mistakes, consumers' process of making choices changes. Ethics related attributes are elevated in importance. Table 4 shows such a compensatory model, in which two of the top three most heavily weighted attributes are related to ethics. Ethics-related variables are the most important in the decision, as we would expect when ethical violations are serious, and consumers penalize the brand for its ethical lapses by not choosing it. However, ethics is not such an overriding consideration that it is used as a filtering or screening variable. When the ethical violation is minor, high CSR attributes compensate for low ethical ratings. This compensatory effect of a positive CSR image is apparent when the ethics violation is minor.

The choice models which best describe consumers' decision-making process are quite different in Table 3 (hierarchical, with ethics as screening variable), compared to Table 4 (compensatory with CPE and CSR considered simultaneously). There is unambiguous evidence that when they uncover major ethical violations, customers punish firms with a better CSR reputation more than they punish firms that are less known for CSR. Ethics becomes a selection criterion, and consumers use ethics to eliminate from their consideration set the brands that they perceive as hypocrites for violating their lofty CSR reputations through unethical behavior.

Table 5 summarizes our results about the choice processes that people use when faced with information regarding manufacturers' corporate social responsibility and ethical behavior. Note that extreme differences in ethical behavior means that some firms in the consideration set are seen as falling short of ethical standards – more so than others. High positive difference in CSR reputation means that some companies have a higher reputation for social responsibility than others being considered by consumers.

TABLE 5: Choice processes used by consumers under different combinations of relative CSR and ethical behavior (in the four treatment cells of the experiment)

Ethical behavior differences / CSR differences between firms	Major difference in ethical behavior between firms	Minor difference in ethical behavior between firms
Major difference in CSR reputation	Hierarchical decision making with ethics as screening variable (Customers punish firm severely for its perceived hypocrisy)	Compensatory process with CSR and ethics considered simultaneously; CSR with greater importance weight (CSR protects firm when it makes minor ethical lapses)
Minor difference in CSR reputation	Compensatory process with CSR and ethics considered simultaneously; Ethics with greater importance weight (Customers punish firm, but not as severely, since there is less perceived hypocrisy)	Compensatory process with CSR and ethics considered simultaneously; CSR with greater importance weight (CSR offers some protection when firm makes minor ethical lapses)

CONCLUSIONS

The managerial implications of our results are that if you spend a lot of money on CSR activities, and on promoting your role as a good corporate citizen, you should be particularly careful not to appear to violate your own standards in your behavior with regard to ethical norms in general. Customers do not tolerate what they perceive as corporate hypocrisy. Your reputation for good corporate citizenship might protect you in the case of minor ethical violations, but it will backfire on you if your violations are seen as egregious. Overall, it is still worthwhile to invest money and resources in trying to be a socially responsible corporation, as long as this is not being done to cover up ethical shortcomings.

A limitation of our study is that we have not considered the potential impact of product quality, brand image or price on the interaction between CSR and CPE, acting to influence choice. Product quality and price were deliberately taken out of the equation by having identical ratings on these dimensions in the information given to subjects in the experiments. Future research might explore some of these more complex interactions between tangible and intangible product quality dimensions.

REFERENCES

Arli, D., van Esch, P., Northey, G., Lee, M.S.W. & Dimitriu, R. (2019), Hypocrisy, skepticism, and reputation: the mediating role of corporate social responsibility, *Marketing Intelligence and Planning*, *37*(6), 706-720.

Bhattacharya, A, Good, V. & Sardashti, H. (2020), Doing good when times are bad: the impact of CSR during recessions, *European Journal of Marketing*, *54*(9), 2049-2077.

Brown, T.J. & Dacin, P.A. (1997), The company and the product: Corporate associations and consumer product responses, *Journal of Marketing*, *61*(1), 68-84.

Brunk, K.H. (2010), Exploring origins of ethical company/brand perceptions - a consumer perspective of corporate ethics, *Journal of Business Ethics*, *29*(3), 255-262.

Brunk, K.H. & Blümelhuber, C. (2011), One strike and you're out: Qualitative insights into the formation of consumers' ethical company or brand perceptions, *Journal of Business Research*, *64*, 134-41.

Casali, G.L. (2011), Developing a multidimensional scale for ethical decision making, *Journal of Business Ethics*, *30*(4), 485-97.

Chan, K.W., Yim, C.K.B. & Lam, S.S.K. (2010), Is customer participation in value creation double-edged sword? Evidence from professional financial services across cultures, *Journal of Marketing*, *74*(3), 48-64.

Claudy, M.C., Garcia, R. & O'Driscoll, A. (2015), Consumer resistance to innovation - a behavioral reasoning perspective, *Journal of the Academy of Marketing Science*, *43*, 528-544.

Cohen, J., Pant, L.W. & Sharp, D.J. (1993), A validation and extension of a multidimensional ethics scale, *Journal of Business Ethics*, *12*(1), 13-26.

Cohn, D.Y. (2010), Commentary essay on exploring origins of ethical company/brand perceptions – a consumer perspective of corporate ethics, *Journal of Business Research*, *63*, 1267-1268.

Crane, A. (2001), Unpacking the ethical product, *Journal of Business Ethics*, *20*(3), 361-373.

Driver, J. (2007), *Ethics: The fundamentals*. Malden, MA: Blackwell Publishing.

Fader, P.S. & McAlister, L. (1990), An elimination by aspects model of consumer response to promotion calibrated on UPC scanner data, *Journal of Marketing Research*, *27*(3), 322-332

Folkes, V.S. & Kamins, M.A. (1999*),* Effects of information about firms' ethical and unethical actions on consumers' attitudes, *Journal of Consumer Psychology*, *8*(3), 243-259.

Grönroos, C. & Gummerus, J. (2014), The service revolution and its marketing implications: Service logic service-dominant logic, *Managing Service Quality*, *24*(3), 206-229.

Guadagni, P.M. & Little, J.D.C (2008), A logit model of brand choice calibrated on scanner data (25^{th} anniversary perspective), *Marketing Science*, *27*(1), 29-48.

Hunt, S.D. & Vitell, S.J. (1986), A general theory of marketing ethics, *Journal of Macromarketing*, Spring, 5-16.

Hunt, S.D. & Vitell, S.J. (2006), The general theory of marketing ethics: A revision and three questions, *Journal of Macromarketing,* December, 143-153.

Janney, J.J., & Gove, S. (2011), Reputation and corporate social responsibility aberrations, trends, and hypocrisy: Reactions to firm choices in the stock option backdating scandal, *Journal of Management Studies*, *48*(7), 1562–1585.

Kuokkanen, H. & Sun, W. (2019), Companies, meet ethical consumers: Strategic CSR management to impact consumer choice, *Journal of Business Ethics*, *38*(1), 1-21

Lenz, I., Wetzel, H.A. & Hammerschmidt, M. (2017), Can doing good lead to doing poorly? Firm value implications of CSR in the face of CSI, *Journal of the Academy of Marketing Science*, *45*(5), 677-697.

Leonidou, C.N., Katsikeas, C.S. & Morgan, N.A. (2013) 'Greening' the marketing mix: Do firms do it and does it pay off*., Journal of the Academy of Marketing* Science, *41*(2), 151-170.

Lichtenstein, D.R., Drumwright, M.E. & Braig, B.M. (2004), The effect of corporate social responsibility on customer donations to corporate-supported nonprofits, *Journal of Marketing*, *68*(4), 16-32.

Luchs, M.G., Walker Naylor, R. Irwin, J.R. & Raghunathan, R. (2010), The sustainability liability: Potential negative effects of ethicality on product preference, *Journal of Marketing, 74*(5), 18-31.

Maddala, G. S. (1983), *Limited dependent and qualitative variables in econometrics*, Cambridge: Cambridge University Press

Nguyen, K. H., Trinh, A. T., Tran, M.D. & Duong, T.T. (2025), uncovering the evolution of brand transparency in corporate social responsibility: A holistic perspective, *Corporate Social Responsibility and Environmental Management*, *32*(6), 7671-7696.

O'Fallon, M. J. & Butterfield, K.D. (2005), A review of the empirical ethical decision-making literature: 1996–2003, *Journal of Business Ethics, 24*(4), 375-413.

Pelsmacker, P. D., Driesen, L. & Rayp, G. (2005), Do consumers care about ethics? Willingness to pay for fair-trade coffee, *The Journal of Consumer Affairs, 2*, 363-385.

Pradhan, S., Mishra, A., Bashir, M. & Bose, S. (2026), Role of CSR perceptions in building customer engagement, *Market Intelligence & Planning*, *44*(2), 419-437.

Rahman, M.M. & Chakraborty, S. (2025), I know the brand is better than that: Brand trust and consumer brand identification evoking brand forgiveness after CSR-related wrongdoings, *Social Responsibility Journal*, *21* (7), 1401-1420.

Ryoo, Y. (2025), CSR as a buffer or backfire in brand transgressions: a motivated reasoning perspective, *European Journal of Marketing*, *59*(2), 185-240

Schneider, L.G. & Selling, T.I. (1996), A comparison of compensatory and non-compensatory models of judgment: Effects of task predictability and degrees of freedom, *Accounting, Organizations and Society*, *21*(1), 3-22.

Shaw, D. & Shui. E. (2003), Ethics in consumer choice: a multivariate modelling approach, *European Journal of Marketing*, *37*(10),1485-1498

Shea, L. J. (2010), Using consumer perceived ethicality as a guideline for corporate social responsibility strategy A commentary essay, *Journal of Business Research*, *63*, 263-264.

Smith, D. & Rhiney, E. (2020), CSR commitments, perceptions of hypocrisy, and recovery*, International Journal of Corporate Social Responsibility*, *5*(1) https://doi.org/10.1186/s40991-019-0046-7

Tetrault, S. C. A. & Lvina, E. (2019), From doing good to looking even better: The dynamics of CSR and reputation, *Business & Society*, *58*(6), 1234–1266.

Tversky, A. (1972), Elimination by aspects: A theory of choice, *Psychological Review*, *79*, 281-299

Vitell, S. J., Singhapakdi, A. & Thomas, J. (2003), Consumer ethics: an application and empirical testing of the Hunt-Vitell theory of ethics, *Journal of Consumer Marketing*, *18*(2), 153-178.

Whalen, J., Pitts, R.E. & Wong, J.K. (1991), Exploring the structure of ethical attributions as a component of the consumer decision, *Journal of Business Ethics*, *10*(4,) 285-293.

Yim, S., Bae, Y.H., Lim, H. & Kwon, J. (2019), The role of marketing capability in linking CSR to corporate financial performance: When CSR gives positive signals to stakeholders, *European Journal of Marketing*, 53(7), 1333-1354.

Abhik Roy is Professor of Marketing in the School of Business at Quinnipiac University in Hamden, Connecticut, where he has been a member of the faculty since 2001. He received his PhD in Marketing Management from the University of California Los Angeles (UCLA) in 1990. He has been a member of the faculty of the University of California, Riverside and the Hong Kong Polytechnic University. Prof. Roy has published extensively in the marketing field, in areas such as competitive pricing strategy, marketing negotiations, retail environments and buyer behavior, and choice models. His work has appeared in journals such as Management Science, Journal of Retailing, International Journal of Research in Marketing, Marketing Letters, Journal of Business Research and Journal of Advertising Research. His current research focuses on the interaction between CSR activities, corporate ethics and its impact on consumer behavior.

APPENDIX

TABLE 6: Summary descriptions for CSR activities and Ethical Behavior of Companies in the Experiments

Social Corporate Responsibility	High	Medium	Low
Employee treatment	Paid employees twice minimum wage	Paid workers 20% more than minimum wage	Paid workers minimum wage
Environmentally friendly policy	Spent $10 million on green initiatives	Spent $500K on green initiatives	Did not engage in any green initiatives
Community based charities	Invested $10 million on local communities where plants located	Invested $500K on local communities where plants located	Did not invest in local communities where plants located
Ethical behavior	High	Medium	Low
Sourcing labor in countries with poor labor rights records	Did not use labor in countries with poor labor rights records	Used labor in countries with poor labor rights but paid 20% more than average wage for that country	Used labor in countries with poor labor rights and paid average wages for that country
Senior discounts	Gave 20% discounts to seniors over 65 years age	Gave 5% discounts to seniors over 65 years age	Did not give any discounts to seniors
Privacy of customer information	Did not share information on customers with third parties	Shared information on customers to third parties with customer consent	Sold information on customers to third parties without asking customer consent

TABLE 7: Instructions to subjects and descriptions of companies making the product in the A1B1 cell of the experiment

Read the description of each company making a room air conditioner and select one by entering the firm number

Firm manufacturing product	CSR activities	Ethics related behavior
Firm 1	Paid workers in own country twice minimum wage	Used labor in countries with poor labor rights and paid average wages for that country
Firm 2	Spent $10 million on green initiatives	Did not give any discounts to seniors
Firm 3	Invested $10 million on local communities where plants located	Sold information on customers to third parties without asking customer consent
Firm 4	Paid workers minimum wage in own country	Did not use labor in countries with poor labor rights records
Firm 5	Did not engage in any green initiatives	Gave 20% discounts to seniors over 65 years age
Firm 6	Did not invest in local communities where plants were located	Did not share information on customers with third parties

++ indicates well above average level and – well below average level

Firms 1, 2 and 3 have above average CSR and below average Ethics

Firms 4, 5 and 6 have below average CSR and above average Ethics.

The order of presentation of Firms to subjects was randomized and a third column with product specifications was shown to respondents. Product specs were identical for each of the six alternatives.

INITIAL PUBLIC OFFERING AND CUMULATIVE ABNORMAL RETURNS: AN EMPIRICAL STUDY USING EVENT STUDY METHODOLOGY

Sunando Sengupta, Bowie State University
Shuva Gupta, University of Pennsylvania
Sounak Chakraborty, University of Missouri-Columbia

ABSTRACT

This paper investigates whether positive abnormal returns exist right after the IPO offering date. Existing literature has evidence of IPO underpricing and evidence of significant positive abnormal returns on the first day of trading. We wanted to test the evidence using a large sample study and WRDS Event study methodology. Our sample consisted of all the IPO s that took place between 2019-24 in the U.S markets. The sample consists of 994 firms and their IPO dates, obtained from WRDS. Using event-study methodology, Eventus, from WRDS, we test for the presence of abnormal returns right after the IPO date using various event windows. We have obtained results that indicate in the event window (0.+1) we find evidence of positive abnormal cumulative return of 0.18%, which is statistically significant. However, following the IPO, in the event windows (+2,+5) we find cumulative abnormal returns of -0.71% which is significant and in the event window (+6.+10) we find cumulative abnormal returns of -0.69% which is significant at 10% interval level. These results are consistent with evidence of positive abnormal returns on the day following the IPO and some trend reversal immediately after.

LITERAURE SURVEY

Initial Public Offerings (IPOs) mark a firm's transition from private to public ownership and are often accompanied by substantial market anomalies, particularly abnormal returns on and around the offering date. The study of IPOs and their associated abnormal returns has been a cornerstone of empirical finance research, focusing on issues such as underpricing, long-run underperformance, and investor behavior.

IPO Underpricing and Short-Run Abnormal Returns

Underpricing refers to the phenomenon where the offer price of an IPO is set below the price at which the stock trades after listing. This results in positive initial abnormal returns to investors who purchase shares at the offering price. Rock (1986) proposed the winner's curse model, explaining underpricing as compensation for uninformed investors who face adverse selection. Beatty and Ritter (1986) argued that underpricing is related to information asymmetry—firms with greater uncertainty about their value tend to experience higher underpricing. Ritter (1991) and Ibbotson et al. (1988) empirically documented the consistent presence of positive initial returns across different markets, often averaging between 10–20% in the U.S.

Explanations for Short-Run Abnormal Returns

Various theories attempt to explain the short-run abnormal returns observed on the IPO date. Information asymmetry models: Suggest that issuers, underwriters, and investors possess different levels of information, leading to mispricing at issuance (Rock, 1986; Allen & Faulhaber, 1989). Signaling theory: Firms underprice their IPOs to signal high quality and attract long-term investors (Allen & Faulhaber, 1989; Welch, 1989). Institutional and behavioral explanations: Herding behavior, market sentiment, and investor over-optimism contribute to inflated post-IPO returns (Loughran & Ritter, 2002; Derrien, 2005).

Long-Run Performance and Abnormal Returns

While IPOs often generate positive abnormal returns in the short term, numerous studies document negative abnormal returns over the long run. Ritter (1991) found that IPOs underperform comparable non-issuing firms over a three- to five-year horizon. Loughran and Ritter (1995) and Brav and Gompers (1997) confirmed long-term underperformance in U.S. IPOs, attributing it to investor overreaction and market timing. In contrast, Carter, Dark, and Singh (1998) suggested that reputable underwriters can mitigate long-run underperformance by better pricing the issue.

Possible explanations include Investor overoptimism, Investors overreacting to initial success causing prices to decline as information becomes public. Agency problems and managerial timing: Managers may issue shares when valuations are temporarily high (Loughran & Ritter, 1995). Lock-up expirations: When insider lock-up periods expire, increased selling pressure leads to price declines (Field & Hanka, 2001). In Samudra (2025), the author looks at short term performance of IPO stocks in the Indonesian stock exchange and find small firms experience higher abnormal returns on first day of trading while larger firms over a longer window experience greater abnormal return. In another recent article, Andersson and Lu (2026), look at 1,222 IPOs listed between 2015-2025. The findings confirm long-run underperformance, particularly among smaller and younger firms. mid-cap IPOs exhibit more persistent returns, while sector and age-based patterns further highlight performance heterogeneity.

Cross-Country Evidence

The degree of IPO underpricing and abnormal returns varies across countries due to differences in institutional frameworks, market efficiency, and investor protection, Loughran, Ritter, and Rydqvist (1994) conducted a cross-country study showing underpricing ranging from 4% in Denmark to over 80% in Malaysia. In emerging markets, such as India and China, studies (e.g., Krishnamurti & Kumar, 2002; Su & Fleisher, 1999) report even higher initial returns, largely attributed to regulatory inefficiencies and speculative demand. Derrien and Womack (2003) found that European IPOs also show significant initial returns, though lower than in Asian markets, suggesting market maturity reduces abnormal performance.

RESEARCH METHODOLOGY

Event Study: An event study is a statistical method of an empirical investigation of the relationship between security prices and economic events (Dyckman et al., 1984). Most event studies have focused on the behavior of share prices in order to test whether their stochastic behavior is affected by the disclosure of firm-specific events. Furthermore, incorporating context, the usefulness of events studies arises from the fact that the magnitude of abnormal performance at the time of an event provides a measure of the unanticipated impact of this type of event on the wealth of the firms' claimholders (Kothari and Warner 2006).

WRDS AND EVENTUS:

Eventus in WRDS is a specialized tool used for event studies, which are statistical analyses that measure the impact of specific events on the value of firms, typically using stock prices. It's designed to help researchers and analysts assess how events like earnings announcements, mergers, acquisitions, regulatory changes, or other corporate actions affect stock returns. Eventus can pull financial and stock market data from WRDS databases (like CRSP or Compustat) and align it with event information (event dates, firm identifiers, etc.). It then calculates abnormal returns, which are the returns of a stock above or below what would normally be expected based on a market model or other benchmark. You can define a window of days around the event date to analyze the short-term or long-term impact. For example, from 5 days before to 5 days after an earnings announcement. This software supports multiple models for estimating expected returns, but for our study we chose market model (regression of stock returns on market returns), the details of which can be found in the empirical model section below. In short, Eventus automates the event study process using WRDS data, saving you from manually aligning stock returns with events and calculating abnormal returns. It's widely used in academic research and finance industry analysis. This developed and maintained by Cowan Research**,** LC, a private firm specializing in financial research software.

EMPIRICAL MODEL

Methodology:

This study employs a standard event study methodology, using Eventus from WRDS and we fit a standard market model to measure normal performance:

$$R_{it} = \alpha_i + \beta_i R_{mt} + \varepsilon_{it}, \qquad \text{where } E(\varepsilon_{it}) = 0 \text{ and } \operatorname{var}(\varepsilon_{it}) = \sigma^2_{\varepsilon t} \tag{1}$$

Each sample calendar date is converted to event time by defining the date of the IPO implementation date as event date 0. So, for the announcement date, event date 0 is the same trading day. Since these are IPOs, there is no public data on days prior to IPO date. The regression coefficients α_i and β_i are estimated in an ordinary least squares (OLS) regression. The event period consists of 10 trading days centered on IPO announcement date. We define four event windows based on the event date, [0,+1, [+2, +5], [+6, +10]. As proxy for the return for the market portfolio R_{mt}, both the CRSP value weighted index and the CRSP equal weighted index are used. Under standard assumptions, OLS is a consistent estimation procedure for the market model parameters. Under the assumption that asset returns are jointly multivariate normal and independently and identically distributed (iid), OLS is also efficient. The prediction errors, $PE_{i\tau}$, which represent abnormal returns, are simply the OLS residuals, $\hat{\varepsilon}_{it}$.

$$PE_{i\tau} \equiv \hat{\varepsilon}_{i\tau} = R_{i\tau} - (\hat{\alpha}_i + \hat{\beta}_i R_{mt}) \tag{2}$$

with

$$\hat{\sigma}^2_{\varepsilon t} = \frac{1}{255-2} \sum_{\tau=t-299}^{t-46} (R_{i\tau} - \hat{\alpha}_i - \hat{\beta}_i R_{m\tau})^2 \tag{3}$$

The prediction error, PE_{it} is used as an estimator of the abnormal return. In other words, the abnormal return is the residual term of the market model calculated on an out of sample basis. Under the null hypothesis, conditional on the event window market returns, the abnormal returns will be jointly normally distributed with a zero conditional mean and conditional variance:

$$AR_{i\tau} \sim N(0, \sigma^2(AR_{i\tau})) \tag{4}$$

The conditional variance $\sigma^2(AR_{i\tau})$ has two components. The first component is the disturbance $\hat{\sigma}^2_{\varepsilon t}$ from (3), and the second component is additional variance due to sampling error in estimating the market model parameters α_i and β_i:

$$\sigma^2(AR_{it}) = \sigma^2_{\varepsilon\tau} + \frac{1}{255}\left[1 + \frac{(R_{m\tau} - \overline{R}_m)^2}{\hat{\sigma}^2_m}\right] \text{ where } \overline{R}_m = \frac{1}{255} \sum_{\tau=t-299}^{t-46} R_{m\tau} \tag{5}$$

To draw inferences about the average price impact of an event, abnormal return observations have to be aggregated across securities and through time.

$$AAR_\tau = \frac{1}{N} \sum_{i=1}^{N} AR_{i\tau} \tag{6}$$

Under the assumption that average abnormal returns are independent across securities, the asymptotic variance equals to

$$Var(AAR_\tau) = \frac{1}{N^2} \sum_{i=1}^{N} \sigma^2_{\varepsilon\tau} \tag{7}$$

The average abnormal returns are aggregated through time to give the cumulative average abnormal return,

$$CAAR_i(\tau_1, \tau_2) = \sum_{\tau=\tau_1}^{\tau_2} AAR_{i\tau} \tag{8}$$

Setting the covariance terms to be zero,

$$\operatorname{var}(CAAR_i(\tau_1,\tau_2)) = \sum_{i=1}^{N} \operatorname{var}(AAR_{i\tau}) \tag{9}$$

Hence $CAAR_i(\tau_1,\tau_2) \sim N(0, \operatorname{var}(CAAR_i(\tau_1,\tau_2))$ (10)

This can be used to test the null hypothesis that the abnormal returns are zero.

The estimated variance of AAR_τ is

$$\hat{\sigma}^2_{AAR} = \frac{\sum_{\tau=t-299}^{t-46} (AAR_\tau - \overline{AAR})^2}{255-2} \text{ where } \overline{AAR} = \frac{\sum_{\tau=t-299}^{t-46} AAR_\tau}{255} \tag{11}$$

The portfolio test statistic for day τ in event time is

$$t = \frac{AAR_\tau}{\hat{\sigma}^2_{AAR}} \tag{12}$$

Assuming time series independence, the test statistic for $CAAR_i(\tau_1,\tau_2)$ is

$$t = \frac{CAAR_i(\tau_1,\tau_2)}{\sqrt{(\tau_2-\tau_1+1)}\hat{\sigma}_{AAR}} \tag{13}$$

The abnormal return estimators often have different variances across firms. A common way of addressing this problem is the standardized residual method (Patell, 1976). Define the *standardized abnormal return*, $SAR_{i\tau}$ as

$$SAR_{i\tau} = \frac{AR_{i\tau}}{\hat{\sigma}_{MLE_{i\tau}}} \tag{14}$$

Where

$$\hat{\sigma}_{MLE_{i\tau}} = \hat{\sigma}^2_{\varepsilon\tau}\left(1 + \frac{1}{T} + \frac{(R_{m\tau} - \bar{R}_m)^2}{\sum_{\tau=t-299}^{t-46}(R_{m\tau} - \bar{R}_m)^2}\right) \tag{15}$$

Is the maximum likelihood estimate of the variance. Under the null hypothesis each $SAR_{i\tau}$ follows a Student's t distribution with T-2 degrees of freedom. Summing the $SAR_{i\tau}$ across the sample yields

$$ASAR_{i\tau} = \sum_{i=1}^{N} SAR_{i\tau} \text{ where } ASAR_{i\tau} \sim N(0, Q_\tau) \tag{16}$$

The Z-test statistic for the null hypothesis that $CAAR_i(\tau_1,\tau_2) = 0$ is

$$Z(\tau_1,\tau_2) = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} Z_i(\tau_1,\tau_2) \text{ where } Z_i(\tau_1,\tau_2) = \frac{1}{\sqrt{(\tau_2-\tau_1+1)\frac{T-2}{T-4}}} \sum_{\tau=\tau_1}^{\tau_2} SAR_{i\tau} \tag{17}$$

The two test statistics so far discussed use the variance estimate from the market model during the estimation period to estimate the variance of the abnormal return estimator. But frequently, events increase the variance of returns, so that the event period variance is greater than the estimation period variance. The portfolio test statistic for day t in event time is

$$t = \frac{AAR_{\tau}}{\hat{\sigma}_{AAR_{\tau}} / \sqrt{N}} \text{ where } \hat{\sigma}_{AAR_{\tau}} = \frac{1}{N-1}\sum_{i=1}^{N}(AR_{i\tau} - \frac{1}{N}\sum_{i=1}^{N} AR_{i\tau})^2 \qquad (18)$$

We use the above equation to calculate *Adjusted-t*

We utilized CRSP (Center for Research in Security Prices) Equally Weighted for the benchmark and market indices. Our sample consisted of the 994 companies for which we received data from WRDS. Using quantitative analysis of the Cumulative Abnormal Returns (CAR) above or below equally weighted market index, we analyzed stock returns to ascertain if there were significant abnormal returns.

We used four window periods surrounding the announcement date and implementation date to determine whether there are indications that NRRS, an accounting standard change, resulted in abnormal returns for our sample. The four window periods are(0,+1),(+2,+5),(+6,+10). Our results are in the following table.

RESULTS AND DISCUSSION

Market Model Abnormal Returns, Equally Weighted Index

Days	N	Mean Cumulative Abnormal Return	Precision Weighted CAAR	Positive: Negative	Portfolio Time-Series (CDA) t	p-value	Uncorrected Patell Z	p-value	Generalized Sign Z	p-value
(0,+1)	994	0.18%	0.47%	460:534	0.132	0.4473	3.564	0.0002	-0.703	0.2411
(+2,+5)	994	-2.64%	-1.22%	449:545(	-1.368	0.0857	-4.460	<.0001	-1.402	0.0805
(+6,+10)	994	-0.71%	-0.47%	448:546(	-0.329	0.3709	-1.528	0.0632	-1.465	0.0715

Using a sample of 994 U.S. IPOs (2019–2024) and a market-model benchmark, we find a small but statistically significant positive cumulative abnormal return of **0.18%** over (0, +1). However, this initial gain is followed by a short-term reversal: CAR falls to **−0.71%** over (+2, +5) (statistically significant) and remains negative at **−0.69%** over (+6, +10) (significant at the 10% level). These results indicate day-one underpricing followed by a rapid correction in the days after listing.

FUTURE RESEARCH

These are possible areas of research going forward: We can do a heterogeneity analysis on the research question - which IPOs reverse fastest? We can investigate which firm/issue characteristics (industry, offer size, underwriter prestige, VC backing, dual-class, price range, hot/cold market) predict the magnitude/duration of the post-IPO reversal. There is also scope to perform a benchmark & model robustness with alternative expected-return models. We can investigate if the results hold using Fama-French factors, industry-adjusted returns, or value-weighted benchmarks. The third area of future research could be retail vs institutional trading and liquidity dynamics immediately after listing**-** Is the reversal driven by retail selling, lower liquidity, or widening spreads?

REFERENCES

Allen, F., & Faulhaber, G. R. (1989). Signalling by underpricing in the IPO market. *Journal of Financial Economics, 23*(2), 303–323.

Andersson, Lu (2026) The Long Game: IPO Performance Across a decade of Monetary shifts , *Lund University Libraries, LUP Student Paper Series.*

Beatty, R. P., & Ritter, J. R. (1986). Investment banking, reputation, and the underpricing of initial public offerings. *Journal of Financial Economics, 15*(1–2), 213–232.

Brav, A., & Gompers, P. (1997). Myth or reality? The long-run underperformance of initial public offerings. *Journal of Finance, 52*(5), 1791–1821.

Carter, R. B., Dark, F. H., & Singh, A. K. (1998). Underwriter reputation, initial returns, and the long-run performance of IPO stocks. *Journal of Finance, 53*(1), 285–311.

Derrien, F. (2005). IPO pricing in the Internet age. *Journal of Finance, 60*(2), 487–521.

Derrien, F., & Womack, K. L. (2003). Auction vs. bookbuilding and the control of underpricing. *Review of Financial Studies, 16*(1), 31–61.

Dyckman, T., Philbrick, D., & Stephan, J. (1984). A practical guide to event study research. *Journal of Financial Economics, 14*(1), 3–29.

Field, L. C., & Hanka, G. (2001). The expiration of IPO share lockups. *Journal of Finance, 56*(2), 471–500.

Ibbotson, R. G., Sindelar, J. L., & Ritter, J. R. (1988). Initial public offerings. *Journal of Applied Corporate Finance, 1*(2), 37–45.

Kothari, S. P., & Warner, J. B. (2007). The econometrics of event studies. In B. E. Eckbo (Ed.), *Handbook of Corporate Finance: Empirical Corporate Finance* (Vol. 1, pp. 3–36). Elsevier.

Krishnamurti, C., & Kumar, P. (2002). The initial listing performance of Indian IPOs. *Managerial Finance, 28*(2), 39–51.

Loughran, T., & Ritter, J. R. (1995). The new issues puzzle. *Journal of Finance, 50*(1), 23–51.

Loughran, T., & Ritter, J. R. (2002). Why don't issuers get upset about leaving money on the table in IPOs? *Review of Financial Studies, 15*(2), 413–444.

Loughran, T., Ritter, J. R., & Rydqvist, K. (1994). Initial public offerings: International insights. *Pacific-Basin Finance Journal, 2*(2–3), 165–199.

Patell, J. M. (1976). Corporate forecasts of earnings per share and stock price behavior: Empirical tests. *Journal of Accounting Research, 14*(2), 246–276.

Ritter, J. R. (1991). The long-run performance of initial public offerings. *Journal of Finance, 46*(1), 3–27.

Rock, K. (1986). Why new issues are underpriced. *Journal of Financial Economics, 15*(1–2), 187–212.

Samudra, S(2025), Short-Term Performance of IPO stocks on the Indonesia Stock Exchange: Do Firm Sizes Matter, *International Journal of Education and Social Sciences , Vol 6 (No, April 2025.*

Su, H.-M., & Fleisher, B. M. (1999). Initial public offerings in China. *Pacific-Basin Finance Journal, 7*(2), 173–199.

Welch, I. (1989). Seasoned offerings, imitation costs, and the underpricing of initial public offerings. *Journal of Finance, 44*(2), 421–449.

Dr Sunando Sengupta is a Professor of Finance at the Department of Accounting, Finance and Economics in the College of Business at the Bowie State University. He finished his PhD in Financial Economics in 2005 from Arizona State University and his research interest includes financial markets, trade policy and small business management.

Dr Shuva Gupta is a Senior Lecturer in Wharton School of Business in University of Pennsylvania and apart of teaching various courses in Statistics at both undergraduate and graduate level, he has an active research interest in enalized methods in high dimensional data analysis, variable selection, nonparametric statistics with applications in statistical analysis of genomic and toxicology data.

Dr. Sounak Chakraborty is an Associate Professor in the Department of Statistics. His research interests are Bayesian machine learning, variable selection in high dimensional problems, data mining, non-linear models for complex data sets, spatio-temporal models, statistical models for multi-platform data integrations, Generative AI, and Markov chain Monte Carlo driven computational models.

BIASING LONG-RUN MONEY NEUTRALITY THROUGH TEMPORAL AGGREGATION

Arjun Sondhi, Muskingum University
Jason Reed, University of Notre Dame
Edmund Matecki, PennWest University - California

ABSTRACT

Money neutrality remains a fundamental macroeconomic question. If the money stock is neutral, the short-run impact of monetary policy on real macroeconomic variables will dissipate over time. Researchers have reached varied conclusions on money neutrality depending on the data sources, aggregation techniques, and empirical methods used. Statistical issues arise due to different frequency data and information loss. This study uses the frequency-domain Bartlett approach and time-domain VAR-based methods and adds to the current literature by testing whether temporal aggregation biases money neutrality test results. Using U.S. data, the study finds Fisher and Seater tests are immune to temporal aggregation, while VAR results are sensitive to the technique. The results suggest that, when using VAR-based tests, researchers must be mindful of the biases temporal aggregation creates.

INTRODUCTION

The actions central banks took during the Great Recession and COVID-19 pandemic highlighted several important macroeconomic issues, including monetary easing's long-run effect on key variables. Sharp declines in global output following the 2007 crash were the worst seen in decades. According to the Bureau of Economic Analysis estimates, U.S. real GDP fell by 8.5%, real consumption by 3.5%, and gross private investment by 27% during the last quarter of 2008. While in the second quarter of 2020, the same variables fell by 31.2%, 33.4%, and 49.3% respectively. The Federal Reserve acted swiftly to each crisis, increasing the monetary base by 606% in the fourth quarter of 2008 and 274.5% in the second quarter of 2020. Most OECD nations experienced a similar decline in output and responded with similar, stimulative monetary and fiscal policies. These episodes have led to the resurgence of literature trying to find causes and reactions to such shocks. One such area of research examines the long-run neutrality of money (LMN)[1]. While accepted as theoretically sound, the neutrality of money is not well-supported empirically. One of the statistical criticisms of the literature is that temporal aggregation will impact a sample's autocorrelations at the asymptotic limit. Annual aggregation drives out cyclical data variation, and estimated coefficients in time series regressions can be wrongfully estimated as zero, as discussed by Tiao (1972) and Rossana and Seater (1992, 1995).[2] We find that differences in statistical methods, and temporal data aggregation do indeed bias empirical results which explains the mixed results from empirical studies of Money Neutrality.

Our investigation focuses on temporal aggregation methods' impact on money neutrality tests, examining the disconnect between theory and empirics. Fisher and Seater's (F&S) ARIMA and King and Watson's (K&W) VAR frameworks are among the most widely used econometric tools to test LMN. (Fisher & Seater, 1993) (King & Watson, 1997). We use the F&S framework to estimate the Bartlett estimator and the K&W VAR approach to test for neutrality and the effects of temporal aggregation for the U.S. The focus is on testing how temporal aggregation affects the results, so it is extremely important to have data at different frequencies. We use monthly, quarterly, and annual data for the U.S. between 1960 – 2016 as described in Appendix A. The model and theoretical framework are not novel to the literature; however, analyzing the different frequency aggregations does add to the body of work. We find Fisher and Seater (F&S) ARIMA tests immune and King and Watson (K&W) VAR results sensitive to temporal aggregation.

---

[1] The concept suggests permanent money supply movements do not affect real variables like real GDP, employment and real consumption, with any effect dissipating as rational agents adjust behavior. Two related concepts are long-run money neutrality and long-run money super neutrality, which suggests permanent money supply growth rate movements have no effect on real variables. See McCandless and Weber (1995), Serletis and Krause (1996), Wallace (1999), Sulku (2011), and Ekomie and Jacques (2013).
[2] See Rossana and Seater (1992, 1995) for a full discussion on the impact of data aggregation in money and output variables. They show that any cyclical variation in the data is driven out by annual aggregation and estimated coefficients in time series regression can be wrongfully estimated to be zero.

Tests are run using both the time-domain and frequency-domain methodology to identify any loss of information in data through aggregation. Time-domain techniques show how a signal within the data changes with time, whereas a frequency-domain graph will show how much of the signal lies within each given frequency band over a range of frequencies.

The remainder of the paper is set up as follows: Section 2 describes the two models, section 3 presents the results, and section 4 concludes the paper.

FREQUENCY-DOMAIN BARTLETT APPROACH

We use the F&S framework to find the Bartlett estimator and the K&W VAR approach to test neutrality and temporal aggregation's impact on U.S. data at monthly, quarterly, and annual frequencies.

Fisher and Seater ARIMA

F&S's money neutrality framework uses a log linear, stationary ARIMA model, where m_t is the log of nominal money supply and y_t is the log of real GDP. The framework allows researchers to conduct nonstructural tests, where money's relative order of integration and the macro variable used is important when testing for LMN. Using the difference ensures the variables are integrated of order zero, I(0). LMN experiments depend on exogenous money supply disturbance u_t and its impact on macro-variables.

The long-run derivative $LRD_{Z,X}$ of the Wold representation of the F&S model, allows for $(x) \geq 1$, where $x_t = \Delta^i m_t$ and $z_t = \Delta^j y_t$, where $i = j$ and either 0 or 1 (see appendix C.1). The long-run derivative LRDz,x can then be rewritten as

$$LRD_{z,x} = \frac{(1-L)^{\langle x\rangle-\langle z\rangle}\gamma(L)/_{L=1}}{\alpha(1)}, \tag{1}$$

Identification and Estimation

To estimate the output equation in the Wold model, we assume money is predetermined.[3] F&S's framework suggests that only the reduced forms $c(1)/d(1)$ are relevant and can be estimated directly in the frequency domain.

In order to identify the contribution of a shock from money to output, we regress $\Delta^{\langle y\rangle} y_t$ on $\Delta^{\langle m\rangle} m_t$ at frequency zero, maintaining the full data set. The estimated coefficient equals $c(1)/d(1)$ if

$$b(1) = \sigma_{uw} = 0 \tag{2}$$

The Bartlett estimator, for $\langle m\rangle = \langle y\rangle = 1$, which is the slope's limit as the span over which the computed growth rates go to infinity, can be estimated as b_k , the slope coefficient from the regression

$$(y_t - y_{t-k}) = a_k + b_k(m_t - m_{t-k}) + e_{kt}. \tag{3}$$

The estimator $c(1)/d(1)$ is represented as b_k, where b_k is the slope of the output and money growth rate. Testing for LMN involves estimating equation (3).

Time-domain VAR Approach

To test contemporaneous money and output effects, we use K&W's (1997) augmented VAR framework:

$$\begin{bmatrix}\Delta y_t \\ \Delta m_t\end{bmatrix} = \begin{bmatrix}\lambda_{ym} & 0 \\ 0 & \lambda_{my}\end{bmatrix}\begin{bmatrix}\Delta m_t \\ \Delta y_t\end{bmatrix} + \begin{bmatrix}\beta_{y\varphi}(L) & \beta_{ym}(L) \\ \beta_{m\varphi}(L) & \beta_{mm}(L)\end{bmatrix}\begin{bmatrix}\xi_t^{\varphi} \\ \xi_t^{m}\end{bmatrix} \tag{4}$$

where the matrix $\beta(L)$ contains lag polynomials that transmit the effects of the i.i.d. shocks ξ_t^m and ξ_t^{φ}. If the fraction

[3] Since real output does not respond contemporaneously to a change in money over a short measurement period, and our data are quarterly and yearly.

$$\gamma_{ym} = \frac{\beta_{ym}(1)}{\beta_{mm}(1)} = 0,$$

money is neutral, because a permanent shock has no effect on output. $\beta_{ij}(1)$ is the sum of the coefficients in the lag polynomial $\beta_{ij}(L)$ (see Appendix C.2).

EMPIRICAL RESULTS

We use real consumption and M2 money supply from 1960 to 2016 to derive aggregate monthly, quarterly, and annual data and compare the results to quarterly and yearly real GDP and M2, and find evidence that money is non-neutral.

F&S Single Equation Evidence

Appendix B presents figures estimating the Bartlett kernel (b_k as k becomes large) along with the 95% confidence interval. Long-run neutrality exists if the parameter goes to zero. Figures 1–3 of Appendix B represent monthly, quarterly, and annual real consumption and M2 using temporally aggregated monthly data. Figures 4 and 5 of Appendix B use quarterly and annual real GDP and M2 data. The results are similar across data frequencies and consistently finding money is non-neutral. Neutrality appears as lag length increases beyond 10 years. The consumption effects are close to zero but not within the standard error at monthly and quarterly frequencies. At annual frequency, consumption effects within the standard error reach zero around nine years. Money is neutral at the annual level using temporally aggregated monthly data. At lags beyond 14 years for all estimates, the parameters at all frequencies increase sharply. However, the results rely on decreasing sample sizes.

One surprising result from these tests suggests that money is non-neutral at short lag lengths. Economists widely agree money's effect on real variables lags, meaning it should be neutral at short lengths. For U.S. data, temporal aggregation affects the Bartlett parameter estimates minimally, and money is non-neutral until year 10.

Figure 4 of Appendix B shows that at short lag lengths money has no effect on output, consistent with conventional belief. Beyond a three-year lag, money affects real macro-activity but dissipates. Using annual data (Figure 5, Appendix B), neutrality is reached around 10 years, with non-neutrality arising at longer lag lengths. Estimating the Bartlett parameter in all cases with Newey-West standard errors, our results do not vary significantly.

Bivariate VAR Evidence

Table 1 reports the estimated neutrality measures and corresponding standard errors from the VAR framework. Lag lengths are chosen using the Hannan and Quinn (1979) method.

Table 1: U.S. VAR Results

Data Frequency	Measure	Optimal Lag Length	γ_m	Standard Error
Monthly	Real C, M2	3	0.155***	0.059
Quarterly	Real C, M2	3	0.264**	0.129
Quarterly	Real GDP, M2	2	0.339	0.57
Annual	Real C, M2	1	0.185	0.135
Annual	Real GDP, M2	1	0.247	0.139

Note: reject LMN *** 99% confidence, **95% confidence, *90% confidence

VAR tests demonstrate that temporal aggregation can distort neutrality outcomes. Our monthly frequency results strongly reject money neutrality, as the coefficient is significantly different from zero at the 1% level. As the data are temporally aggregated, the non-neutrality results become less compelling. Neutrality is rejected when using aggregated data at the quarterly, but not annual, frequency. The optimal lag lengths decrease as data is aggregated from quarterly to annual, not from monthly to quarterly. Although we neither confirm nor reject money neutrality conclusively, we find that F&S tests are robust to temporal aggregation. VAR results, however, are sensitive to aggregation. Policy makers looking to determine the impacts of policy on money neutrality need to sensitize around the frequency of data used.

CONCLUSION

This paper uses established theoretical models to test temporal aggregation's effects on money neutrality tests. Using F&S and K&W methods for U.S. data, we find mixed results in support of long-run money neutrality. Our contribution to the literature is determining whether temporal aggregation biases results. We find temporal aggregation does not affect F&S test results, while VAR results are susceptible to temporal aggregation. Neutrality tests provide unconvincing support for LMN. We do not consider neutrality tests' biases from structural change. The dataset F&S use in their original work is from Friedman and Schwartz (1982), covering 1867–1975 in the United States. Data from that period has several structural breaks, making an analysis using data from different counties and periods as an area of future research.

In real life, this conventional amount is an interval grey number (Liu and Lin, 2006). That is, I_p is equal to a grey number (a, b), for some real numbers a and b $(a < b)$, such that although it is known that $I_p \in (a, b)$, the specific value of I_p is unknown until the lunch is over. In other words, the value of I_p is not fixed in real life. While feeling full, a person can still take additional bites of food without suffering from the bad consequence of overeating. To make our reasoning in this paper go smoothly, we assume that for each diner $d_i \in D_p \cup D_f$, the value of I_i is a fixed real number.

REFERENCES

Dickey, D. A., & Fuller, W. (1981). Likelihood ratio test statistics for autoregressive time series with a unit root. *Econometrica, 49*, 1057–1072.

Ekomie, T., & Jacques, J. (2013, January). Testing the long run neutrality of money in developing economies – Evidence from EMCCA. *Modern Economy*, 49–55.

Friedman, M., & Schwartz, A.J. (1982). *Monetary Trends in the United States and the United Kingdom*. University of Chicago Press.

Fisher, M. E., & Seater, J. J. (1993). Long run neutrality and super neutrality in an ARIMA framework. *The American Economic Review, 83*(3), 402–415.

Hannan, E.J., & Quinn, B.G. (1979). The determination of the order of an autoregression. *Journal of the Royal Statistical Society*, Series B 41(2), 190-195.

King, R. G., & Watson, M. W. (1997). Testing long-run neutrality. *Federal Reserve Bank of Richmond Economic Quarterly, 83*(3), 69–101.

McCandless, G. T., & Weber, W. E. (1995). Some Monetary Facts, Quarterly Review. *Federal Reserve Bank of Minneapolis*.

Nelson, C.R., & Plosser, C.I. (1982). Trends and random walks in macroeconomic time series: Some evidence and implications. *Journal of Monetary Economics 10*(2), 139-162.

Newey, W. K., & West, K. D. (n.d.). A simple positive, semidefinite, heteroskedasticity and autocorelation consistent covariance matrix. *Econometrica, 55*(3), 703–708.

Priestley, M. B. (1981). *Spectral analysis and time series.* New York: Academic Press.

Rossana, R. J., & Seater, J. J. (1992). Aggregation, unit roots, and the time series structure of manufacturing real wages. *International Economic Review 33*, 159–179.

Rossana, R. J., & Seater, J. J. (1995). Temporal aggregation and economic time series. *Journal of business and economic statistics, 13*(4), 441–451.

Serletis, A., & Krause, D. (1996). Empirical evidence on the long-run neutrality hypothesis using low-frequency international data. *Economic Letters, 50*, 323–327.

Sulku, S. (2011). Testing long run neutrality of money in a developing country - Evidence from Turkey. *Journal of Applied Economics and Business Research, 1*(2), 65–74.

Tiao, G. C. (1972). Asymptotic behavior of temporal aggregates of time series. *Biometrika* (59), 525–531.

Wallace, F. H. (1999). Long run neutrality of money in the Mexican economy. *Applied Economics Letters, 6*, 637–639.

Dr. Arjun Sondhi is an Assistant Professor of Economics at Muskingum University in New Concord, Ohio. His research areas include topics in international macroeconomics and economics of health and health care. Specific topics of interest include monetary economics, open economy DSGE modeling, VAR's and the effect of health policy on the macro economy.

Dr. Jason Reed is the Wade Family associate teaching professor and associate Faculty Director for the Notre Dame Institute for Global Investing. His research focuses the integration of behavioral economics in the field of macroeconomics and finance and was awarded the Department of Economics Mendelson Award for Outstanding dissertation (2015).

Dr. Edmund Matecki is an Associate Professor of Business at Pennsylvania Western University.

APPENDIX A

Appendix A describes the data. Sources include the International Monetary Fund (IMF), Organization for Economic Co-Operation and Development (OECD), and St. Louis Federal Reserve. In converting money variables from relatively higher to lower frequencies, we use the standard stock conversion method of using the last monthly or quarterly measurement in the year to generate lower frequency data. For output and consumption, we sum monthly values to obtain quarterly values and quarterly values to obtain annual values.

Table 2: Data Description for F&S and VAR Tests

United States	M2	Monthly	1960–2016	IMF
		Quarterly	1960Q1–2017Q1	IMF
		Annually	1960–2016	IMF
	C	Monthly	1960M1–2016M12	St. Louis Fed
	GDP	Quarterly	1960Q1–2017Q1	OECD
		Annually	1960–2016	OECD

APPENDIX B

Appendix B shows U.S. money neutrality results using the F&S methodology and different temporal aggregation schemes. The graphs represent the Bartlett Parameter Estimates with confidence interval bands.

Figure 1: Monthly Real Consumption and M2

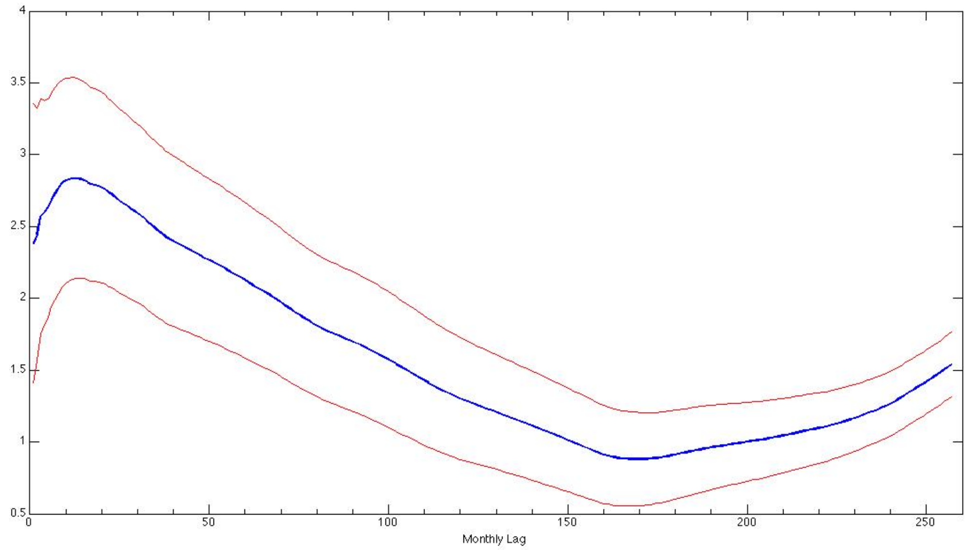


Figure 2: Aggregated Quarterly Real Consumption and M2

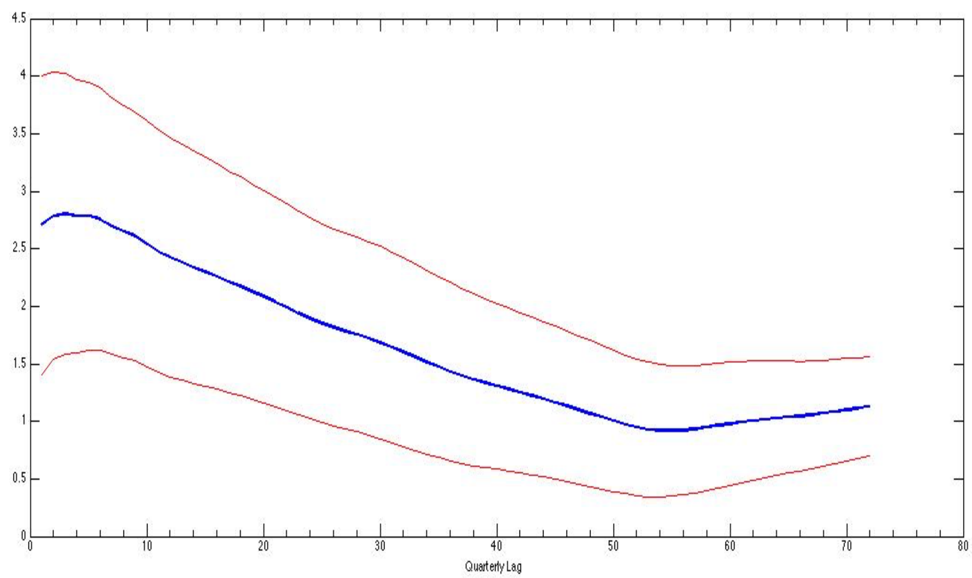


Figure 3: Aggregated Annual Real Consumption and M2

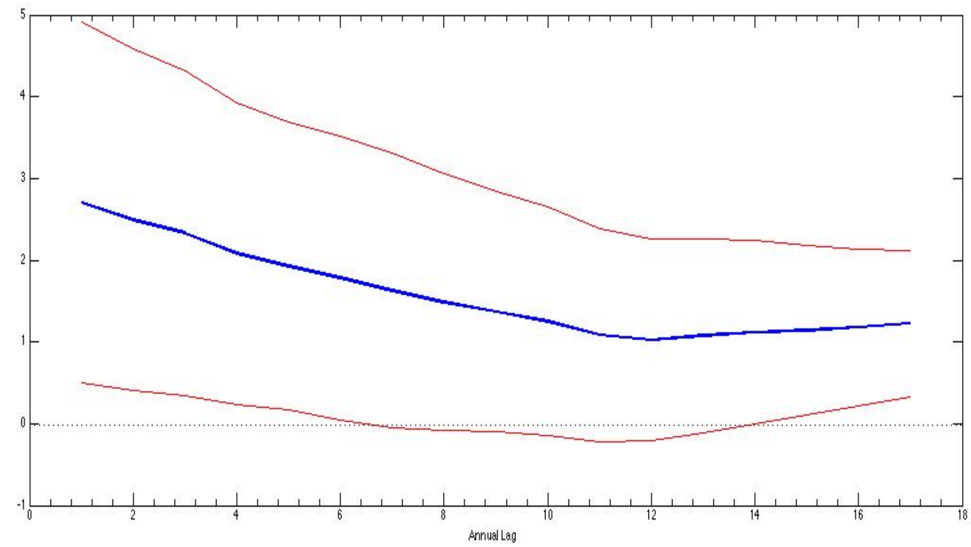


Figure 4: Quarterly Real GDP and M2

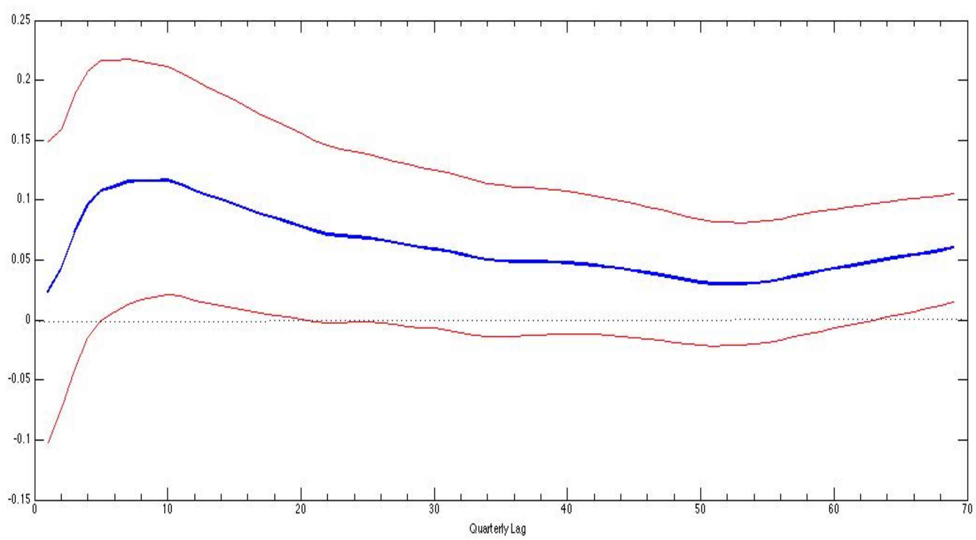
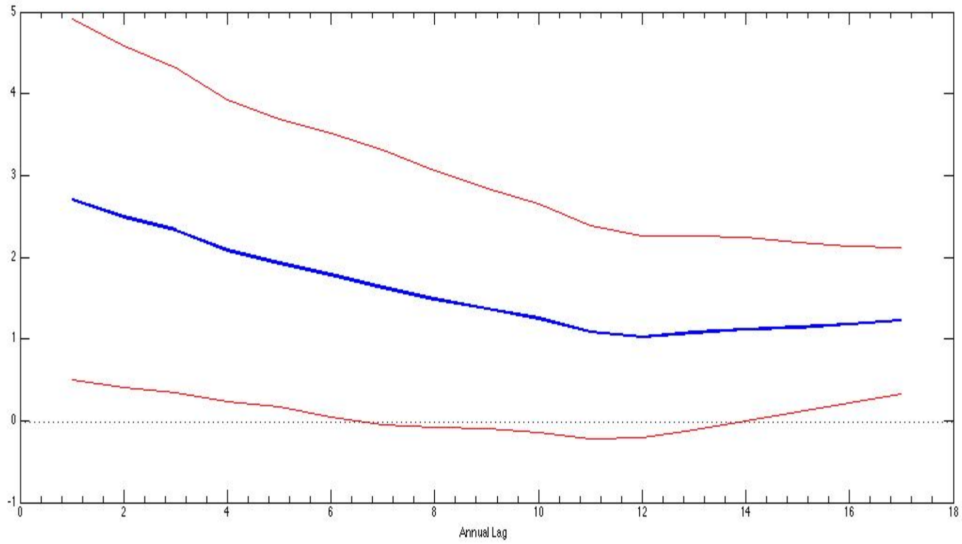


Figure 5: Annual Real GDP and M2



APPENDIX C

C.1

If $x\ and\ z$ are integrated order one, finding LMN is equivalent to testing whether the long-run derivative equals zero for real variables or one for nominal variables. Thus, from the Wold representation, testing for neutrality is equivalent to testing the restriction that the ratio of lagged macroeconomic shocks to $\Delta^j y_t$, $c(1)/d(1)$, are equal to 0 or 1, when there isn't a shock from output to money.

The Augmented Dickey-Fuller and Phillip-Perron unit root tests confirm that U.S. M2 money and output are I(1) series, suggesting the relevant order of integration is $\langle x \rangle = \langle z \rangle = 1$.[4]

C.2

Since lags between money and output are typically longer than one quarter, we assume that $\lambda_{ym} = 0$, which is empirically justifiable. Including intercepts, the VAR becomes

$$\begin{bmatrix} \Delta y_t \\ \Delta m_t \end{bmatrix} = \begin{bmatrix} \mu_y \\ \mu_m \end{bmatrix} + \begin{bmatrix} \beta_{y\varphi}(L) & \beta_{ym}(L) \\ \beta_{m\varphi}(L) & \beta_{mm}(L) \end{bmatrix} \begin{bmatrix} \xi_t^{\varphi} \\ \xi_t^{m} \end{bmatrix} \tag{5}$$

Taking the inverse of the lag polynomial matrix $\beta_{ij}(L)$ and letting

$$\alpha(L) = \begin{bmatrix} \alpha_{11}(L) & \alpha_{12}(L) \\ \alpha_{21}(L) & \alpha_{22}(L) \end{bmatrix} = \begin{bmatrix} \beta_{y\varphi}(L) & \beta_{ym}(L) \\ \beta_{m\varphi}(L) & \beta_{mm}(L) \end{bmatrix}^{-1} = |\beta(L)|^{-1} \begin{bmatrix} \beta_{mm}(L) & -\beta_{ym}(L) \\ -\beta_{m\varphi}(L) & \beta_{y\varphi}(L) \end{bmatrix}, \tag{6}$$

the VAR that we estimate is

$$\begin{bmatrix} \alpha_{11}(L) & \alpha_{12}(L) \\ \alpha_{21}(L) & \alpha_{22}(L) \end{bmatrix} \left\{ \begin{bmatrix} \Delta y_t \\ \Delta m_t \end{bmatrix} - \begin{bmatrix} \mu_y \\ \mu_m \end{bmatrix} \right\} = \begin{bmatrix} \xi_t^{\varphi} \\ \xi_t^{m} \end{bmatrix} \tag{7}$$

Given the definition of the lag polynomial $\alpha(L)$, the neutrality estimate is

$$\gamma_{ym} = -\frac{\alpha_{12}(1)}{\alpha_{11}(1)}. \tag{8}$$

For money to be neutral, $\gamma_{ym} = 0$ and $\alpha_{11}(1) \neq 0$; In estimating γ_{ym}, we construct the standard errors using the delta method. The polynomials in the matrix $\beta_{ij}(L)$ are in the bivariate vector moving average process, where the neutrality measure is defined, and $\alpha_{ij}(L)$ are the parameters to be estimated. The standard errors are constructed by computing the matrix multiplication $V'\Sigma\, V$, where Σ is the parameter covariance matrix and V is the vector

$$V = -\begin{bmatrix} \dfrac{\partial \gamma_{ym}}{\partial \alpha_{11}} \\ \dfrac{\partial \gamma_{ym}}{\partial \alpha_{12}} \end{bmatrix}. \tag{9}$$

[4] Following Nelson and Plosser (1982), we find macro-variables to be I(1).

RESEARCH NOTES

THE ECONOMIC IMPACT OF WORKPLACE SAFETY PROGRAMS: A COST-BENEFIT ANALYSIS OF INVESTING IN OCCUPATIONAL SAFETY VS. COSTS OF WORKPLACE INJURIES

Mehdi Khalighi, Millersville University

ABSTRACT

Workplace safety is a critical driver of operational efficiency and employee well-being, yet organizations often struggle to quantify its economic value. This study examines the cost–benefit dynamics of investing in occupational safety programs relative to the financial burdens associated with workplace injuries. Using quantitative analysis, industry benchmarks, and an applied case study, the research evaluates the return on investment (ROI) of proactive safety interventions. Findings demonstrate that comprehensive safety programs reduce the frequency and severity of workplace injuries while generating measurable financial benefits, including lower insurance costs, reduced absenteeism, and improved operational efficiency. Although safety initiatives require upfront investment, the results show that proactive interventions consistently outweigh injury-related costs and deliver substantial long-term economic advantages. The study further highlights the strategic importance of integrating safety management into organizational decision-making, emphasizing its role in enhancing economic performance, workforce sustainability, and organizational resilience.

INTRODUCTION

Workplace safety has long been recognized as a core component of organizational performance, yet its economic implications are frequently misunderstood or undervalued. Employers have traditionally viewed safety primarily as a regulatory or compliance-driven function; however, a growing body of evidence demonstrates that workplace safety functions as a fundamental economic driver influencing productivity, operational stability, and competitive performance (Tompa et al., 2021; Mustard & Yanar, 2023). The costs associated with workplace injuries extend well beyond medical payments and workers' compensation claims, affecting productivity, employee retention, supply chain continuity, quality control, insurance premiums, and organizational reputation (Grimani et al., 2018; van Dongen & van der Beek, 2022).

According to the U.S. Bureau of Labor Statistics (BLS, 2023), private industry employers reported approximately 2.8 million nonfatal workplace injuries and illnesses in 2022, highlighting the scale of occupational injury exposure across industries. Complementing these findings, the National Safety Council (NSC, 2023) estimates the economic cost of preventable workplace injuries at $167 billion, including wage losses, medical expenses, administrative overhead, and employer-incurred costs. Together, these national estimates underscore the substantial and persistent financial burden organizations face when injuries occur, even before accounting for longer-term indirect and reputational impacts.

Despite mounting empirical evidence that investments in workplace safety generate measurable economic returns, many organizations continue to operate reactively, addressing hazards only after incidents occur (Lee et al., 2018; Mustard & Yanar, 2023). Executive decision-makers frequently perceive safety expenditures as unavoidable costs rather than strategic assets, limiting integration of safety into core business decision-making (Fernández-Muñiz et al., 2014; van Dongen & van der Beek, 2022). This study addresses this gap by evaluating workplace safety programs through a financial and risk-management lens, demonstrating that proactive investment can yield strong return on investment (ROI), mitigate operational risk, and enhance long-term organizational competitiveness (Tompa et al., 2021; Mustard & Yanar, 2023).

LITERATURE REVIEW

Direct and Indirect Costs of Workplace Injuries

Contemporary occupational safety research consistently distinguishes between direct costs, such as medical treatment, workers' compensation, legal fees, and insurance premiums, and indirect costs, including productivity losses, workforce disruption, retraining expenses, quality defects, and administrative burden (Tompa et al., 2021). Recent economic analyses show that indirect costs frequently exceed direct costs, particularly in industries characterized by complex operations and tight production schedules (Grimani et al., 2018; van Dongen & van der Beek, 2022).

Updated empirical evidence indicates that indirect costs often account for a substantial majority of total injury-related losses, reflecting the cumulative impact of lost workdays, supervisory time, schedule disruptions, and reduced employee engagement following incidents (Hassard et al., 2018; Tompa et al., 2021). These findings emphasize that injury costs extend far beyond immediately visible expenses, reinforcing the importance of proactive safety investment as an effective organizational risk-mitigation strategy (Mustard & Yanar, 2023).

Recent studies further demonstrate that indirect injury costs are magnified in organizations with lean staffing models and just-in-time production systems, where even short disruptions can cascade into significant operational and financial losses due to reduced buffers and high process interdependence (van Dongen & van der Beek, 2022).

Economic Burden Across Industry Sectors

Industry-level research highlights substantial variation in injury-related economic burden across sectors, reflecting differences in work processes, hazard profiles, labor structures, and operational complexity (Tompa et al., 2021; Mustard & Yanar, 2023). Liberty Mutual's 2023 Workplace Safety Index identifies overexertion, falls, and contact with objects and equipment as the leading causes of disabling workplace injuries, collectively accounting for tens of billions of dollars in annual employer costs (Liberty Mutual Research Institute for Safety, 2023). Overexertion alone continues to represent the single most expensive injury category, exceeding $12 billion per year in direct workers' compensation costs, even before accounting for associated indirect losses such as productivity disruption and workforce instability (Liberty Mutual Research Institute for Safety, 2023).

Recent sector-specific empirical analyses reinforce these findings. In the transportation and warehousing sector, injury-related costs are strongly associated with elevated turnover rates, increased overtime expenditures, and recruitment challenges, effects that are particularly pronounced among aging workforces and physically demanding job roles (Mitropoulos et al., 2019; Schwatka et al., 2020; Wadsworth et al., 2023). In manufacturing, studies consistently show that injury-related downtime disrupts production flow, degrades product quality, and increases unit costs across interconnected supply chains, amplifying the total economic impact well beyond the initial incident (Mustard & Yanar, 2023).

Collectively, these studies demonstrate that injury costs are not isolated operational issues but systemic economic stressors, affecting organizational resilience, supply-chain stability, and long-term competitiveness across industries (Tompa et al., 2021; van Dongen & van der Beek, 2022; Mustard & Yanar, 2023).

Return on Investment (ROI) of Safety Programs

A growing body of contemporary literature provides strong empirical support for the positive return on investment (ROI) associated with occupational safety and health programs (Lee et al., 2018; Mustard & Yanar, 2023). Authoritative analyses from the Occupational Safety and Health Administration (OSHA, 2022) and the National Institute for Occupational Safety and Health (NIOSH, 2021) report that employers commonly realize returns ranging from approximately $2 to $6 for every $1 invested in comprehensive safety and health initiatives, driven by reductions in injury costs, productivity losses, and insurance expenditures.

Recent longitudinal and quasi-experimental studies confirm that organizations with mature safety management systems experience sustained reductions in injury incidence, workers' compensation costs, insurance premiums, and absenteeism, while simultaneously demonstrating measurable gains in productivity, workforce stability, and operational reliability (Tompa et al., 2021; Mustard & Yanar, 2023). Importantly, the economic return on safety investments is consistently strongest when interventions prioritize engineering controls, ergonomic design, and systems-based risk management, rather than relying exclusively on training or administrative controls (van der Molen et al., 2018; Schwatka et al., 2020; van Dongen & van der Beek, 2022).

Recent meta-analyses and systematic reviews further indicate that safety investments generate compounded economic benefits over time, particularly when integrated with organizational performance metrics, continuous improvement frameworks, and enterprise risk-management strategies (Grimani et al., 2018; Tompa et al., 2021; van Dongen & van der Beek, 2022).

Safety as a Competitive and Strategic Advantage

Recent strategic management and occupational safety literature increasingly recognizes occupational safety as a source of competitive advantage, rather than merely a regulatory obligation (Podgórski et al., 2017; Mustard & Yanar, 2023). High-performing organizations integrate safety into broader Environmental, Social, and Governance (ESG) strategies, explicitly linking worker protection to corporate reputation, investor confidence, operational continuity, and supply-chain reliability (de Villiers et al., 2021; Christensen et al., 2022; Flammer, 2021).

Empirical studies indicate that employees particularly early career and highly skilled workers place significant value on workplace safety and health practices when evaluating employers, directly influencing recruitment, retention, organizational commitment, and employer attractiveness (Schwatka et al., 2020). In parallel, strong safety performance increasingly affects eligibility for contracts and partnerships in construction, energy, logistics, and public infrastructure sectors, where lagging indicators (e.g., injury rates) and leading indicators (e.g., safety management systems) are routinely incorporated into procurement and prequalification decisions (Lingard et al., 2022; Zou et al., 2023).

Collectively, these findings support the conclusion that occupational safety functions as a strategic organizational asset, enhancing resilience, strengthening human capital, improving stakeholder trust, and contributing directly to long-term economic and competitive performance (Flammer, 2021; Tompa et al., 2021; Mustard & Yanar, 2023).

RESEARCH OBJECTIVES

The primary objective of this research is to quantify the economic impact of workplace safety programs by systematically comparing the costs of occupational injuries with the financial returns generated by proactive safety investments. Specifically, this study examines both direct costs, such as medical expenses, workers' compensation, and insurance premiums, and indirect costs, including productivity losses, operational disruptions, workforce turnover, and administrative burden. By evaluating these cost categories in parallel, the research aims to clarify the full economic burden associated with workplace injuries and incidents.

A second objective is to assess the return on investment (ROI) associated with proactive safety interventions, with emphasis on engineering controls and hazard elimination strategies. This analysis considers short-term and long-term financial outcomes, including payback periods, discounted cost savings, and cumulative net benefits.

Finally, the study seeks to demonstrate how workplace safety functions as a strategic and competitive organizational asset rather than a compliance-driven expense. By linking safety performance to economic resilience, workforce sustainability, and operational reliability, the research contributes to the growing body of evidence supporting the integration of safety management into broader organizational and strategic decision-making. To ground these objectives in practice, the study applies structured cost–benefit analysis within a manufacturing case study, illustrating how data-driven safety investment decisions can be evaluated and justified using standard financial metrics.

METHODOLOGY

This study employs a quantitative, data-driven analytical approach supplemented by an applied case study to evaluate the economic impact of workplace safety programs. The methodology is designed to align with contemporary economic and occupational safety research by integrating national injury data, industry benchmarks, and established financial evaluation techniques.

Data Sources

Multiple authoritative data sources were used to support the analysis. National occupational injury and illness statistics were obtained from the U.S. Bureau of Labor Statistics, providing baseline incident frequency and severity data across industry sectors. Injury cost estimates and indirect cost ratios were derived from National Safety Council datasets and the Liberty Mutual Workplace Safety Index, which identifies leading causes of disabling workplace injuries and associated employer costs. In addition, the OSHA Safety Pays Estimator was used to estimate injury-related costs and calculate the level of business activity required to offset incident losses through revenue. Industry-specific ROI benchmarks from published studies and professional safety organizations were incorporated to validate and contextualize the findings.

Analytical Framework

The economic evaluation applies a cost–benefit and ROI framework commonly used in occupational safety and business analytics research. Financial analysis techniques include present value (PV) calculations, discounted cash flow (DCF) modeling, and payback period analysis to assess both immediate and long-term economic impacts of safety investments. Direct and indirect cost multipliers were applied to capture the full scope of injury-related losses, consistent with contemporary economic literature. The framework also incorporates OSHA Safety Pays output to illustrate the relationship between injury costs and required sales volume, reinforcing the business relevance of safety decisions for managerial audiences.

Case Study Design

To illustrate the application of the analytical framework, a manufacturing case study was conducted at a packaging facility where employees routinely accessed elevated work areas using an unstable A-frame ladder. This condition presented a significant fall hazard with potential for serious injury. The average cost of a fall-related injury ($62,521), based on national cost estimates, was used as a baseline for evaluating potential savings. Three alternative engineering control solutions were analyzed: a prefabricated stairway, a fortress stair system, and a rolling safety ladder. Each option was evaluated based on initial implementation cost, degree of hazard reduction, long-term injury prevention potential, projected ten-year net savings, and calculated payback period. This structured comparison demonstrates how financial analysis can be used to inform evidence-based safety investment decisions in manufacturing environments.

COST EFFECTIVENESS COMPARISON TABLE/FIGURE REFERENCES

Figure 1 illustrates the *economic costs of preventable workplace injuries by sector.* **Figure 2** presents the *Liberty Mutual Workplace Safety Index,* highlighting leading causes of disabling injuries. **Figure 3** depicts the *distribution of work-related injury and illness costs by category,* distinguishing between direct and indirect cost components. **Figure 4** illustrates the *return on investment (ROI) of workplace safety programs by industry.*

Table 1 presents a *cost-effectiveness comparison for a prefabricated stairway system.* **Table 2** provides a *cost-effectiveness comparison for the fortress stair system,* and **Table 3** presents a *cost-effectiveness comparison for a rolling safety ladder.*

All figures and tables are presented at the end of the manuscript in the Appendix for reference.

FINDINGS

National and Industry-Level Economic Trends

Analysis of national occupational injury data confirms that workplace injuries impose a substantial and persistent economic burden on U.S. employers. Recent estimates from the National Safety Council indicate that total employer costs associated with work-related injuries exceed $167 billion annually, encompassing medical expenses, workers' compensation, wage losses, productivity disruptions, and administrative costs. Consistent with prior economic research, indirect costs including lost output, supervisor time, schedule disruptions, and workforce instability represent a majority share of these losses, reinforcing that injury-related financial impacts extend well beyond direct compensation expenses.

Industry-level data reveal that sectors characterized by hazardous tasks and physically demanding work environments particularly manufacturing, warehousing, transportation, and utilities experience disproportionately high injury-related costs. These industries face compounded economic effects due to operational interdependencies, lean staffing models, and time-sensitive production systems. As documented in industry benchmarks and national datasets, even single serious incidents can produce cascading effects across production cycles, logistics flows, and workforce availability, amplifying total economic loss.

Return-on-investment analyses derived from national benchmarks and prior studies demonstrate that workplace safety programs consistently produce positive and measurable financial returns. Across multiple industries, safety investments yield returns typically ranging from 20 to 24 percent, driven by reductions in injury frequency and severity, lower workers' compensation premiums, reduced absenteeism, and improved operational continuity. These findings align with contemporary literature emphasizing that engineering controls and systems-based interventions generate the strongest and most durable economic benefits.

Case Study Findings: Cost–Benefit Analysis of Fall-Hazard Interventions

The manufacturing case study provides applied evidence supporting the national and industry-level trends. Economic evaluation of three engineering control options implemented to address fall hazards in a packaging facility demonstrated that all intervention alternatives achieved a payback period of approximately one year. When projected over a ten-year horizon using discounted cash flow analysis, each option generated net savings exceeding $300,000, substantially surpassing initial capital investment costs.

The prefabricated stairway option effectively eliminated repeated ladder use, significantly reducing fall exposure by replacing unstable access methods with fixed, engineered access. The fortress stair system addressed multiple elevated hazard points, providing comprehensive risk reduction across the work area. The rolling safety ladder option, while representing the lowest initial investment, reduced ergonomic strain and improved task accessibility, offering rapid implementation and favorable short-term returns.

Across all scenarios, the present value of avoided injury costs based on an average fall-injury cost of $62,521 exceeded implementation costs within the first year of operation. These findings demonstrate that even modest engineering investments can yield substantial long-term financial benefits when evaluated using structured cost–benefit and ROI frameworks. Collectively, the case study results reinforce the conclusion that proactive safety investments function not only as risk mitigation measures but also as economically justified strategic decisions.

DISCUSSION

Economic Value Beyond Compliance

The findings of this study reinforce the conclusion that workplace safety investments extend well beyond regulatory compliance and function as value-generating strategic business decisions. National and industry-level data demonstrate that injury-related costs are substantial, recurring, and largely indirect, amplifying their financial impact on organizational performance. The case study results further illustrate that targeted engineering interventions addressing fall hazards can generate rapid payback and substantial long-term savings. These outcomes confirm that safety investments directly contribute to cost containment through reductions in medical expenditures, workers' compensation claims, insurance premiums, and productivity losses.

Importantly, the financial benefits of safety interventions compound over time. Discounted cash flow analysis shows that avoided injury costs accumulate quickly, allowing organizations to recover initial capital investments within a short horizon while continuing to capture economic benefits in subsequent years. When evaluated through standard financial metrics such as net present value, payback period, and return on investment, safety initiatives perform comparably to, and in many cases exceed, traditional capital investment projects. This positions occupational safety as an economically rational component of financial risk management rather than a discretionary expense.

Strategic and Competitive Advantages

Beyond direct cost savings, the findings highlight the role of workplace safety as a strategic driver of organizational resilience and competitive advantage. Organizations operating in high-risk sectors such as manufacturing, logistics, and utilities face heightened exposure to operational disruption from injuries. Proactive safety investments reduce this exposure by stabilizing production capacity, minimizing downtime, and preserving workforce availability.

From a human capital perspective, strong safety performance supports employee retention, engagement, and attraction in increasingly competitive labor markets. Moreover, safety outcomes are increasingly visible to external stakeholders, including customers, insurers, investors, and regulators. As occupational safety metrics are more frequently incorporated into ESG evaluations, safety performance contributes directly to corporate reputation, access to capital, and supply-chain partnerships. The results of this study support emerging ESG frameworks that treat worker safety and health as core components of long-term value creation and risk governance.

MANAGERIAL IMPLICATIONS

The findings of this research provide several actionable implications for managers, executives, and financial decision-makers. First, managers can use ROI-based safety analyses to communicate the economic value of safety investments in terms familiar to executive leadership, boards of directors, and finance functions. Presenting safety initiatives using financial indicators such as net present value, payback period, and avoided cost strengthens the business case for proactive investment.

Second, the results suggest that organizations should prioritize engineering controls and system-level interventions, which consistently demonstrate strong long-term financial returns and durable risk reduction. These investments should be evaluated alongside other capital projects during budgeting and capital allocation processes, rather than being treated as discretionary operational expenses.

Third, integrating safety metrics such as injury rates, cost of incidents, and avoided losses into organizational performance dashboards enables leaders to monitor safety as a financial and operational indicator, reinforcing accountability and continuous improvement.

Finally, the alignment between safety performance, financial outcomes, and ESG objectives indicates that managers can leverage occupational safety initiatives to support broader sustainability and governance strategies. Incorporating safety outcomes into ESG reporting, enterprise risk management frameworks, and strategic planning processes enhances organizational credibility while reducing long-term economic and operational risk. In this context, workplace safety emerges as a measurable contributor to sustainable value creation, organizational resilience, and competitive positioning.

LIMITATIONS AND FUTURE RESEARCH

This study is subject to several limitations that should be considered when interpreting the findings. First, the analysis relies primarily on national and industry-level cost estimates rather than organization-specific injury and financial data. While these datasets provide robust and widely accepted benchmarks, actual injury costs and returns on safety investments may vary based on organizational size, geographic location, workforce composition, and operational complexity. Second, the manufacturing case study focuses on fall-hazard interventions in a single operational context. Although the findings illustrate the economic rationale for safety investments, the results may not be directly generalizable to all industries or hazard types.

Future research should extend this work in several important directions. Longitudinal, multi-industry studies using firm-level data would allow for more precise estimation of the long-term financial performance and return on investment associated with safety programs across diverse organizational contexts. Additionally, the integration of AI-driven predictive analytics offers significant potential for advancing safety economics by enabling early identification of high-risk conditions, dynamic risk modeling, and real-time evaluation of intervention effectiveness. Further research should also examine how safety investment outcomes are incorporated into ESG reporting and enterprise risk management frameworks, helping organizations better align worker protection with financial performance and sustainability objectives.

CONCLUSION

This study demonstrates that workplace safety programs represent sound economic investments that generate measurable financial returns while supporting long-term organizational resilience. Using national and industry-level data together with an applied manufacturing case study, the analysis confirms that proactive safety interventions reduce the frequency and severity of workplace injuries and significantly lower associated direct and indirect costs, including medical expenses, insurance premiums, productivity losses, and operational disruptions.

When evaluated using standard financial metrics such as return on investment, payback period, and discounted cost savings, safety investments consistently outperform their initial expenditures and produce sustained economic benefits over time. Beyond these quantifiable outcomes, the findings show that integrating safety management into organizational strategy strengthens workforce sustainability, operational efficiency, and risk management practices.

By framing occupational safety as a strategic component of business decision-making rather than a compliance obligation, this research contributes to the growing evidence base supporting the economic rationale for workplace safety. The results underscore the role of safety programs in enhancing both organizational performance and broader ESG objectives, reinforcing safety's importance in building resilient, responsible, and competitive organizations.

REFERENCES

Christensen, D. M., Serafeim, G., & Sikochi, A. (2022). Why is corporate virtue in the eye of the beholder? The case of ESG ratings. *The Accounting Review, 97*(1), 147–175. https://doi.org/10.2308/TAR-2019-0506

de Villiers, C., Kuruppu, S., & Dissanayake, D. (2021). A (new) role for business—Promoting the United Nations' Sustainable Development Goals through the internet-of-things and blockchain technology. *Journal of Business Research, 131*, 598–609. https://doi.org/10.1016/j.jbusres.2020.06.066

Flammer, C. (2021). Corporate green bonds. *Journal of Financial Economics, 142*(2), 499–516. https://doi.org/10.1016/j.jfineco.2021.01.010

Grimani, A., Bergström, G., Riaño-Casallas, M. I., Aboagye, E., Jensen, I., & Lohela-Karlsson, M. (2018). Economic evaluation of occupational safety and health interventions from the employer perspective: A systematic review. *Journal of Occupational and Environmental Medicine, 60*(2), 147–166. https://doi.org/10.1097/JOM.0000000000001224

Lee, G., et al. (2018). A systematic review of occupational health and safety interventions with economic evaluations. *Journal of Occupational and Environmental Medicine, 60*(2), 127–146. https://doi.org/10.1097/JOM.0000000000001223

Liberty Mutual Research Institute for Safety. (2023). *2023 Workplace Safety Index*. https://business.libertymutual.com/safety/workplace-safety-index/

Lingard, H., Cooke, T., & Blismas, N. (2022). Do leading indicators improve safety outcomes in construction supply chains? *Safety Science, 149*, 105681. https://doi.org/10.1016/j.ssci.2021.105681

Mitropoulos, P., et al. (2019). Age, safety attitudes, and injury rates in the transportation industry. *Journal of Safety Research, 69*, 1–9. https://doi.org/10.1016/j.jsr.2019.01.005

Mustard, C. A., & Yanar, B. (2023). Estimating the financial benefits of employers' occupational health and safety expenditures. *Safety Science, 159*, 106008. https://doi.org/10.1016/j.ssci.2022.106008

National Institute for Occupational Safety and Health. (2021). *Business value of safety and health*. https://www.cdc.gov/niosh/business

National Safety Council. (2023). *Injury facts®: The economic impact of preventable injuries*. https://injuryfacts.nsc.org/

Occupational Safety and Health Administration. (2022). *Safety and health management systems: Return on investment*. https://www.osha.gov/safety-management

Podgórski, D., et al. (2017). A conceptual framework for OSH performance measurement. *Safety Science, 93*, 1–15. https://doi.org/10.1016/j.ssci.2016.11.003

Schwatka, N. V., et al. (2020). A Total Worker Health® approach to workforce well-being. *Journal of Occupational and Environmental Medicine, 62*(4), 252–258. https://doi.org/10.1097/JOM.0000000000001818

Tompa, E., Mofidi, A., van den Heuvel, S., van Bree, T., Michaelsen, F., Jung, Y., Porsch, L., & van Emmerik, M. (2021). Economic burden of work injuries and diseases: A framework and application in five European Union countries. *BMC Public Health, 21*, 49. https://doi.org/10.1186/s12889-020-10050-7

U.S. Bureau of Labor Statistics. (2023). *Employer-reported workplace injuries and illnesses—2022*. https://www.bls.gov/iif/

van der Molen, H. F., et al. (2018). Interventions to prevent injuries in construction workers: A systematic review. *Scandinavian Journal of Work, Environment & Health, 44*(2), 123–136. https://doi.org/10.5271/sjweh.3691

van Dongen, J. M., & van der Beek, A. J. (2022). Economic evaluations in occupational health: What brings the best bang for the buck? *Scandinavian Journal of Work, Environment & Health, 48*(4), 249–252. https://doi.org/10.5271/sjweh.4026

Wadsworth, E., Walters, D., & Quinlan, M. (2023). Employment arrangements, vulnerability, and occupational health outcomes in logistics and transport. *Safety Science, 165*, 106198. https://doi.org/10.1016/j.ssci.2023.106198

Zou, P. X. W., Sunindijo, R. Y., & Dainty, A. R. J. (2023). Safety culture maturity and its impact on project performance. *Engineering, Construction and Architectural Management, 30*(3), 1021–1039. https://doi.org/10.1108/ECAM-05-2021-0462

Dr. Mehdi Khalighi is an Associate Professor of Occupational Safety and Environmental Health at Millersville University of Pennsylvania. His research interests include occupational safety, environmental health, cost-benefit analysis of safety interventions, and the integration of emerging technologies and AI into safety and business education.

APPENDIX

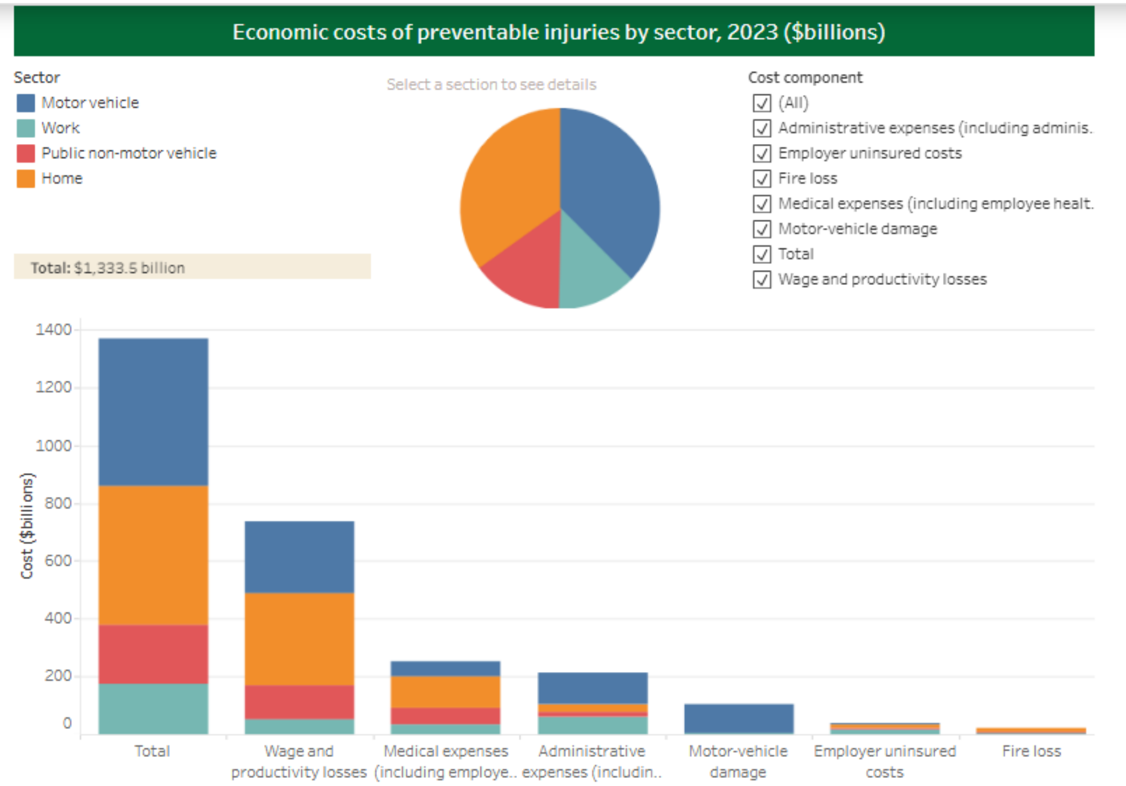


Figure 1: Economic Costs of Preventable Injuries by Sector (2023)

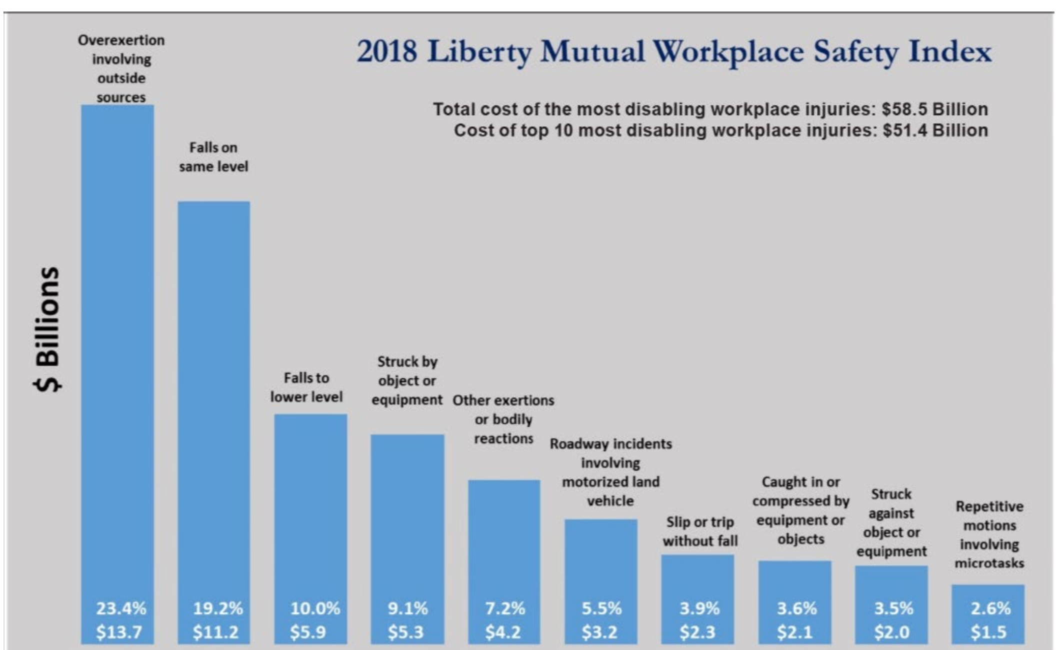


Figure 2: Liberty Mutual Workplace Safety Index (2018)

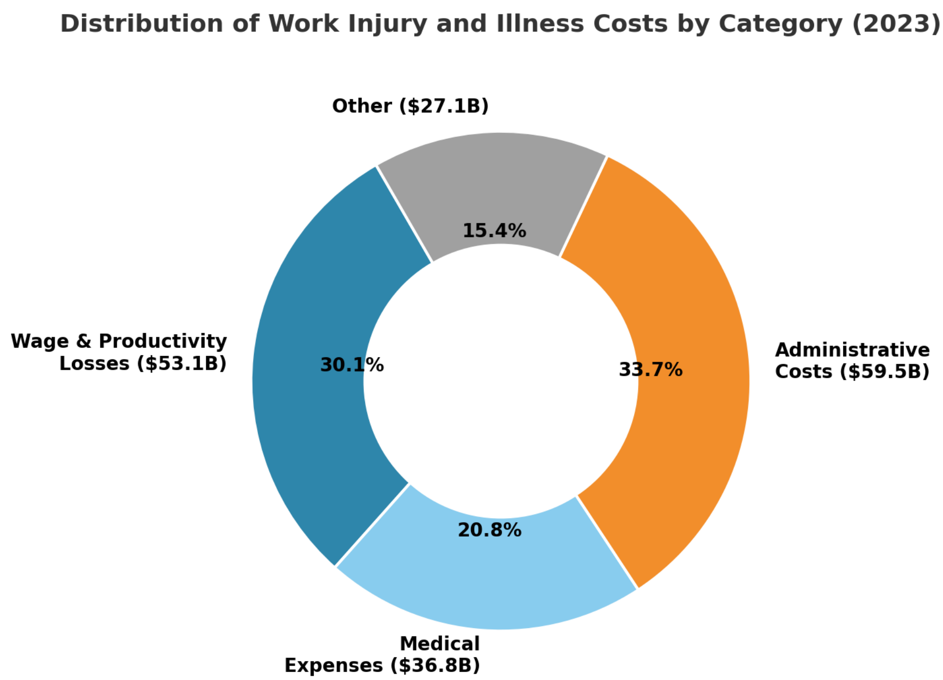


Figure 3: Distribution of Work Injury and Illness Costs by Category (2023)

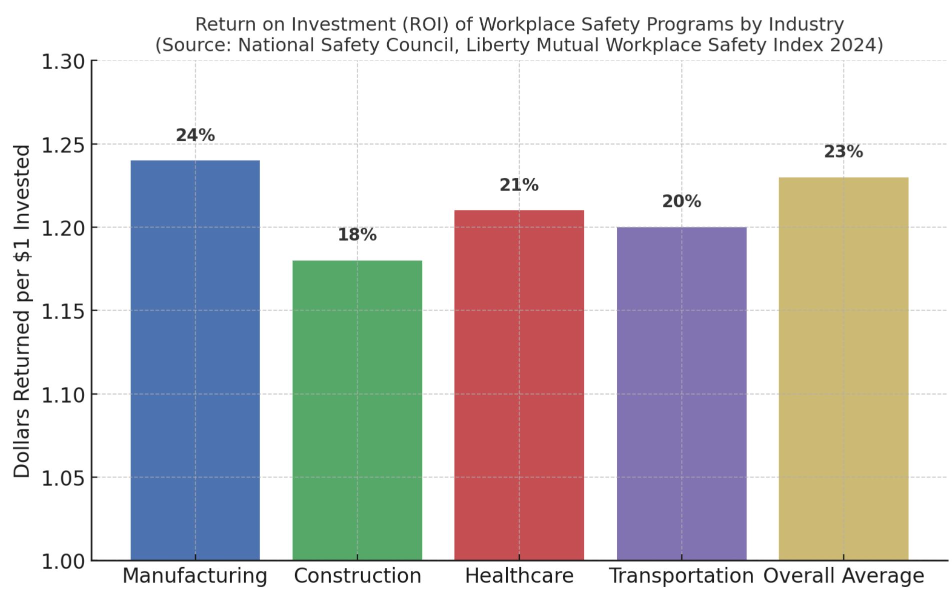


Figure 4: Return on Investment (ROI) of Workplace Safety Programs by Industry (2024)

Table 1: Cost Effectiveness Comparison - Prefabricated Stairway

A	3	4	5	6	7	8
	(1x2)		(3x4)		(5/6)	
(n) year	Annual Cost Savings in Current Dollars	Average Inflation Factor (1 + Inflation rate)$^{n-1}$	Actual Savings in Future Dollars	Discount Factor (1 + Minimum Desired Rate of Return on Investment)$^{n-1}$	Present Value of Future Savings	Cumulative Future Savings
1	$62,521.00	1.000	$62,521.00	1.000	$62,521.00	***$62,521.00***
2	$62,521.00	1.030	$64,396.63	1.200	$53,663.86	***$116,184.86***
3	$62,521.00	1.061	$66,328.53	1.440	$46,061.48	***$162,246.34***
4	$62,521.00	1.093	$68,318.38	1.728	$39,536.10	***$201,782.44***
5	$62,521.00	1.126	$70,367.94	2.074	$33,935.15	***$235,717.59***
6	$62,521.00	1.159	$72,478.97	2.488	$29,127.67	***$264,845.27***
7	$62,521.00	1.194	$74,653.34	2.986	$25,001.25	***$289,846.52***
8	$62,521.00	1.230	$76,892.94	3.583	$21,459.41	***$311,305.93***
9	$62,521.00	1.267	$79,199.73	4.300	$18,419.33	***$329,725.26***
10	$62,521.00	1.305	$81,575.72	5.160	$15,809.92	***$345,535.18***

Required Data

Average Cost of FALL Injury =	(1)	$62,521.00	9. Cumulative Future Savings (Max value Col. 8)	$345,535.18
Goal: Number of Injuries to be Prevented/Year=	(2)	1.00	10.(Subtract) Cost of Controls	$5,886.62
Annual Cost Savings in Current Dollars (1 X 2)=	(3)	$62,521.00	11. Present Value of Savings for Program (9-10)	$339,648.56
Cost of Controls (Original Investment)=	(10)	$5,886.62		
Life Expectancy of Control (Write-Off Period in Yrs)=	(A)	10		
Average Inflation Rate over Write-Off Period=	(4)	3.00%	Payback Period	1 Year
Company's Opportunity Cost of Capital (Minumum Desired Rate of Return on Investment=	(6)	20.00%	(Determined from Column 8.) the payback period is the time	

Table 2: Cost Effectiveness Comparison – Fortress Stair System

A	3	4	5	6	7	8
	(1x2)		(3x4)		(5/6)	
(n) year	Annual Cost Savings in Current Dollars	Average Inflation Factor (1 + Inflation rate)$^{n-1}$	Actual Savings in Future Dollars	Discount Factor (1 + Minimum Desired Rate of Return on Investment)$^{n-1}$	Present Value of Future Savings	Cumulative Future Savings
1	$62,521.00	1.000	$62,521.00	1.000	$62,521.00	***$62,521.00***
2	$62,521.00	1.030	$64,396.63	1.200	$53,663.86	***$116,184.86***
3	$62,521.00	1.061	$66,328.53	1.440	$46,061.48	***$162,246.34***
4	$62,521.00	1.093	$68,318.38	1.728	$39,536.10	***$201,782.44***
5	$62,521.00	1.126	$70,367.94	2.074	$33,935.15	***$235,717.59***
6	$62,521.00	1.159	$72,478.97	2.488	$29,127.67	***$264,845.27***
7	$62,521.00	1.194	$74,653.34	2.986	$25,001.25	***$289,846.52***
8	$62,521.00	1.230	$76,892.94	3.583	$21,459.41	***$311,305.93***
9	$62,521.00	1.267	$79,199.73	4.300	$18,419.33	***$329,725.26***
10	$62,521.00	1.305	$81,575.72	5.160	$15,809.92	***$345,535.18***

Required Data

Average Cost of _________ Injury =	(1)	$62,521.00	9. Cumulative Future Savings (Max value Col. 8)	$345,535.18
Goal: Number of Injuries to be Prevented/Year=	(2)	1.00	10.(Subtract) Cost of Controls	$1,494.99
Annual Cost Savings in Current Dollars (1 X 2)=	(3)	$62,521.00	11. Present Value of Savings for Program (9-10)	$344,040.19
Cost of Controls (Original Investment)=	(10)	$1,494.99		
Life Expectancy of Control (Write-Off Period in Yrs)=	(A)	10		
Average Inflation Rate over Write-Off Period=	(4)	3.00%	Payback Period	1 Year
Company's Opportunity Cost of Capital (Minumum Desired Rate of Return on Investment=	(6)	20.00%	(Determined from Column 8.) the payback period is the time	

Table 3: Cost Effectiveness Comparison – Rolling Safety Ladder

A	3	4	5	6	7	8
	(1x2)		(3x4)		(5/6)	
(n) year	Annual Cost Savings in Current Dollars	Average Inflation Factor (1 + Inflation rate)$^{n-1}$	Actual Savings in Future Dollars	Discount Factor (1 + Minimum Desired Rate of Return on Investment)$^{n-1}$	Present Value of Future Savings	Cumulative Future Savings
1	$62,521.00	1.000	$62,521.00	1.000	$62,521.00	***$62,521.00***
2	$62,521.00	1.030	$64,396.63	1.200	$53,663.86	***$116,184.86***
3	$62,521.00	1.061	$66,328.53	1.440	$46,061.48	***$162,246.34***
4	$62,521.00	1.093	$68,318.38	1.728	$39,536.10	***$201,782.44***
5	$62,521.00	1.126	$70,367.94	2.074	$33,935.15	***$235,717.59***
6	$62,521.00	1.159	$72,478.97	2.488	$29,127.67	***$264,845.27***
7	$62,521.00	1.194	$74,653.34	2.986	$25,001.25	***$289,846.52***
8	$62,521.00	1.230	$76,892.94	3.583	$21,459.41	***$311,305.93***
9	$62,521.00	1.267	$79,199.73	4.300	$18,419.33	***$329,725.26***
10	$62,521.00	1.305	$81,575.72	5.160	$15,809.92	***$345,535.18***

Required Data				
Average Cost of FALL Injury =	(1)	$62,521.00	9. Cumulative Future Savings (Max value Col. 8)	$345,535.18
Goal: Number of Injuries to be Prevented/Year=	(2)	1.00	10.(Subtract) Cost of Controls	$947.59
Annual Cost Savings in Current Dollars (1 X 2)=	(3)	$62,521.00	11. Present Value of Savings for Program (9-10)	$344,587.59
Cost of Controls (Original Investment)=	(10)	$947.59		
Life Expectancy of Control (Write-Off Period in Yrs)=	(A)	10		
Average Inflation Rate over Write-Off Period=	(4)	3.00%	Payback Period	1 Year
Company's Opportunity Cost of Capital (Minumum Desired Rate of Return on Investment=	(6)	20.00%	(Determined from Column 8.) the payback period is the time	

A MODEL OF EMINENT DOMAIN TAKINGS ON PROPERTY INVESTORS UTILITY

Tianning Li, Hood College
Ramon DeGennaro, University of Tennessee

ABSTRACT

Eminent domain grants governments the power to seize private property for public use with compensation, yet the fairness and efficiency of this practice remain contested. Landmark cases such as *Kelo v. City of New London* and subsequent state-level reforms illustrate deep tensions between economic development goals and property rights protections.

This project models how government takings, compensation levels, and public use benefits influence property investment decisions. It predicts three outcomes: (1) fair compensation keeps investment aligned with market incentives; (2) under compensation discourages new investment unless returns are unusually high; and (3) overcompensation encourages excessive investment, creating potential taxpayer losses. The project bridges legal theory, economics, and public policy to offer evidence-based guidance on balancing property rights with community development needs. Findings will inform ongoing debates over reforming eminent domain laws to better align private incentives with genuine public benefits.

INTRODUCTION

Eminent domain—the government's authority to take private property for public use with just compensation—is a complex and controversial issue at the intersection of law, economics, and public policy. The landmark Supreme Court case *Kelo v. City of New London* (2005) held that the Public Use Clause of the Fifth Amendment permits governmental authorities to exercise eminent domain to transfer private property to other private entities for the purpose of advancing economic development objectives. The *Kelo* ruling provoked one of the most intense public and scholarly controversies in recent judicial history and sparked national debate over the balance between individual property rights and public interest. The 2017 film *Little Pink House* dramatized this case, underscoring its human and societal impact.

In the aftermath of *Kelo*, more than forty states enacted statutory or constitutional reforms aimed at strengthening private property protections and constraining the scope of eminent domain powers. Nonetheless, the depth and effectiveness of these reforms have varied substantially across jurisdictions. Some states adopted only symbolic or nominal restrictions that left broad discretion intact at the local level. López, Jewell, and Campbell (2007) highlight West Virginia as a notable example, where the legislature continued to permit the condemnation of "blighted" properties. The statutory definition of "blight" in the West Virginia Code encompasses approximately 200 words and includes a wide range of qualifying conditions—such as irregular street or lot layouts, deteriorated improvements, defective titles, or any state of disrepair that might pose a hazard due to fire or "other causes" (Code of West Virginia, 1931/2006, Chapter 16, Article 18, Section 16-8-3). López, Jewell, and Campbell (2009) contend that such expansive definitions effectively delegate the determination of property rights protections to municipal authorities, thereby rendering these legislative reforms largely ineffectual in constraining eminent domain abuse.

Do property investors alter their investment decisions when the likelihood of losing their assets through an eminent domain taking increases? This project models how the probability of government takings, the level of compensation, and the distribution of public use benefits shape property investors' decisions. The framework predicts three key outcomes. First, when investors are fairly compensated, investment follows market incentives, and the risk of taking does not deter entry. Second, when investors are undercompensated, new investments are discouraged unless returns exceed the market rate, leading to fewer but riskier investment. Third, when overcompensation occurs, investors may actually prefer to be taken, leading to excessive investment and potential taxpayer losses.

LITERATURE REVIEW

Blume, Rubinfeld, and Shapiro (1984) demonstrated that providing full market-value compensation is inefficient, as it generates a moral hazard by encouraging property investors to over-invest in properties that may be subject to government regulation or expropriation.

Miceli (2007) simplified Blume, Rubinfeld, and Shapiro's model and derived the following results. When assuming a benevolent government, compensation should equal the full market value of the land at its efficient investment level. Providing full compensation prevents a secondary moral hazard arising from the fact that the probability of a taking depends on the landowner's investment decisions. Specifically, if property investors anticipate being undercompensated, they can reduce the likelihood of a taking by overinvesting. Conversely, if they expect to be overcompensated, they have an incentive to underinvest. When a government is not benevolent, their model shows that the government takes property if public value of the land is greater than compensation owed. Takings are efficient only if full compensation equals the land's market value. However, this creates a trade-off. Full compensation could avoid fiscal illusion but causes moral hazard (owners over-invest expecting payment). On the other hand, zero compensation could avoid moral hazard but lead to too many takings, since government undervalues private costs.

The basic BRS model that Miceli simplified is based on the assumption that landowner's utility being social optimal. We argue that private property investors do not necessarily care about optimizing social benefits. They only care about maximizing their own utilities. We propose our own model that reflects this assumption in the next section of the paper.

MODEL

We model eminent domain takings by beginning with the property investor's utility function as a function of his/her investment in his/her property, *x*:

$$U(x) = (1 - p) * v(x) + p * [\gamma * v(x) + B] - x \tag{1}$$

where: p = probability of the government taking the property,
$v(x)$ = value of the property if it is not taken,
$\gamma * v(x)$ = compensation for the property if it is taken, as a function of $v(x)$. γ is a scalar. If γ=1 then the property investor is exactly compensated for the market value of the property he/she loses because of the taking. If γ exceeds (is less than) unity, then he/she is overcompensated (undercompensated).
B = the property investor's share of the benefits of the public project resulting from the taking, which we call the *public use benefits,*
x = investment in the property

We identify the circumstances under which *Kelo* and subsequent state laws affect property investment. We show that property investment can be encouraged, unaffected, or discouraged as the probability of takings increases, depending on the level of compensation for the takings and the magnitude of the investors' public use benefits.

Comparative statics show that:

$$\frac{\partial u(x)}{\partial p} = -v(x) + [\,\gamma * v(x) + B] = (\gamma - 1) * v(x) + B$$

This equation offers several insights:

A. When $\gamma < 1$, and if $(\gamma - 1) * v(x) + B \geq 0$, then $\frac{\partial u(x)}{\partial p} \geq 0$.

When governments insufficiently compensate the property investor and the investor's loss is no more than his/her public use benefits, then the investor's utility increases or remains unchanged as the probability of taking increases. Most people would consider Case *A* to be a desirable outcome. Not only does society enjoy the public benefits of the government project, but these benefits are so great that the property investor's share of them overcomes the loss he/she suffers when the government underpays for the property.

B. When $\gamma < 1$, and if $(\gamma - 1) * v(x) + B < 0$, then $\frac{\partial u(x)}{\partial p} < 0$.

When governments insufficiently compensate the property investor, and property investor's loss is more than his/her public use benefits, utility decreases as the probability of taking increases. Property rights proponents fear this case. The property investor is worse off if his/her property is taken. This happens if governments honestly underestimate the value of the property. But it could also be the case that a corrupt government use its' power to benefit politically connected firms at the expense of other property investors.

C. When $\gamma \geq 1$, $\frac{\partial u(x)}{\partial p} > 0$.

When governments overcompensate for the taking, utility increases as the probability of taking increases. The property investor misses the public use benefits and possibly a premium over the market value of his/her property if it is not taken. Taxpayers fear this case. The property investor is better off if his/her property is taken, but taxpayers are harmed even if governments act honestly and simply overestimate the value of the property. Worse still, corruption is also possible. Interest groups might make large campaign contributions to help corrupt politicians win reelection and then be repaid with taxpayer dollars by having their properties taken at a premium.

Taking the first derivative of utility *u(x)* with respect to other parameters also yields insights. The first derivative with respect to γ is:

$$\frac{\partial u(x)}{\partial \gamma} = v(x) * p \geq 0$$

D. The property investor's utility increases as governments pay a higher price for the property.

Taking the first derivative of utility *u(x)* with respect to B, we obtain:

$$\frac{\partial u(x)}{\partial B} = p \geq 0$$

E. The property investor's utility increases as his/her public use benefits increase.

Maximizing the property investor's utility (Equation 1) with respect to his/her investment, *x*, we obtain:

$$(1 - p) * v^{'}(x) + p * \gamma * v^{'}(x) = 1\,, \quad (2)$$

which implies:

$$v^{'}(x) = \frac{1}{1 - p + p * \gamma} = \frac{1}{1 + (\gamma - 1) * p} \quad (3)$$

Equation (3) indicates that to maximize his/her utility, a property investor chooses a property which satisfies the following:

where $\beta = \frac{1}{1 + (\gamma - 1) * p}$ and C is a constant.

Comparative statics show that $\frac{\partial \beta}{\partial p} = \frac{1 - \gamma}{[1 + (\gamma - 1) * p]^2}$

When $\gamma = 1$, meaning that a property investor is fairly compensated, then $v'(x) = 1$. Thus $v(x) = x + C$. This indicates that when the property investor is fairly compensated, he/she can maximize his/her utility by investing in a property whose return on equity is the market rate. In addition, because $\frac{\partial \beta}{\partial p} = 0$, the return on equity that the property investor seeks is unaffected by the probability of taking if he is fairly compensated.

When $\gamma > 1$, suggesting that a property investor is over compensated relative to fair market value, then $v'(x) < 1$. This indicates that when the property investor is over compensated, he can increase his/her utility by growing the property even if the return on his/her marginal investment is below the market rate.

Overcompensating the property investor creates incentives to grow the property beyond what is socially optimal. In addition, because $\frac{\partial \beta}{\partial p} < 0$, the return on equity that the investor requires decreases as the probability of taking increases, given that he is overcompensated if his/her property is taken. The more likely his/her property is taken, the lower return on equity the property investor requires.

When $\gamma < 1$, suggesting that a property investor is under-compensated relative to fair market value, then $v'(x) > 1$. When the property investor is under compensated, he must invest in a property which offers a return on equity that is greater than the market rate. Because $\frac{\partial \beta}{\partial p} > 0$, the higher the probability of taking, the higher return on equity the property investor requires.

Comparative statics also show that $\frac{\partial \beta}{\partial \gamma} = \frac{-p}{(1-p+p*\gamma)^2}$

When $p = 0$ (the probability of taking is zero) the return on equity that the property investor seeks is not affected by the compensation.

When $p > 0$ (the probability of taking is greater than zero) the return on equity that the property investor requires is a decreasing function of the proportion of compensation received, γ.

Equation (1) writes utility in terms of *x*, the property investor's investment in his/her property.

Our model suggests that the property investor's utility can be increased, unaffected, or decreased as the probability of taking his/her property increases.

REFERENCES

Blume, L., & Rubinfeld, D. (1984). Compensation for takings: An economic analysis. *California Law Review, 72*, 569–628.

Blume, L., Rubinfeld, D., & Shapiro, P. (1984). The taking of law: When should compensation be paid? *Quarterly Journal of Economics, 99*, 71–92.

DeGennaro, R. P., & Li, T. (2013). Business formation in the wake of states' responses to Kelo. *Journal of Accounting and Finance, 13*(5).

DeGennaro, R. P., & Li, T. (2014). Do states with stronger property rights protection encourage more entrepreneurship? Revisiting states' responses to Kelo. *Journal of Business and Economics*.

López, E. J., Jewell, R. T., & Campbell, 2009. Pass a law, any law, fast! State legislative response to the Kelo backlash. *Review of Law & Economics*, 2009.

López, E. J., & Totah, S. M. (2007). Kelo and its discontents: The worst (or best?) thing to happen to property rights. *The Independent Review, 11*(3), 397–416.

Miceli, T. J. (2008). Public goods, taxes, and takings. *International Review of Law and Economics, 28*(4), 287–2

Dr. Tianning Li has been a member of Hood College's faculty for more than 15 years. He is also the current director of Hood's MBA program and faculty advisor of the Hood College Blazing Alpha Fund (student-managed endowment). He teaches courses in financial management, corporate finance, and investment analysis. His research interest includes capital structure, ownership structure, information asymmetry, and entrepreneurship.

Dr. Ramon P. DeGennaro is the Haslam College of Business Professor in Banking and Finance at the University of Tennessee. He previously served as a Visiting Scholar at the Federal Reserve Bank of Atlanta, as the Tennessee Bankers Association Scholar, and as a William B. Stokely Scholar at Tennessee. Before joining the university, he was a visiting scholar at the Federal Reserve Bank of Cleveland. His current research involves financial markets and institutions, financial regulation, small-firm finance, and public policy.

BUSINESS CONFIDENCE, 1946-2025

Clifford F. Thies, Shenandoah University of Virginia

ABSTRACT

This paper presents a new time series concerning business confidence from 1946 to 2025. The time series is based on an increasing number of underlying time series; initially, four; increasing to 17. The resulting index toughly reflects the movements of the economy during these eighty-years as indicated by NBER Business Cycle turning points. However, there are many deviations. The most obvious exception is the growth recession of 1967, where this index predicted a recession as did consumer confidence, the index of leading indicators and the yield curve, and yet the economy merely slowed a bit. Other exceptions apparently reflect "jobless recoveries," as well as geopolitical and other uncertainties. In general, business confidence is more volatile than the economy.

INTRODUCTION

This business confidence series joins with surveys of consumers, initiated by the Survey Research Center of the University of Michigan in 1946; and, surveys of economists, initiated by the Joseph Livingston of the *Philadelphia Enquirer* also in 1946; to give us a roughly complete picture of sentiment of the major actors of the economy beginning soon after World War II.

The work for this paper began with the development of business price expectations (Thies, 1986), which included a mix of library work and private correspondence. The work thereafter continued in real time, mostly based on the release of survey data.

The underlying data vary significantly in scale, among other ways. For example, in some surveys 100 is at least nominally neutral; and, in others, zero is at least nominally neutral. In this study, all scales are transformed into percentile ranks so that zero is extremely pessimistic, 50 neutral and 100 extremely optimistic. Subperiods are identified of at least a half business cycle in length, which subperiods overlap by a year. The available surveys that span the subperiod are identified and averaged. Finally, at the years of overlap, the indexes of the adjacent subperiods are melded.

1946-1967

Since 1946, there has been a growing number of sources of information concerning business confidence. See Table 1 for the first several sets of sources. Two sources carryover from before WWII: Metropolitan Life Insurance Co.'s index of help wanted advertising and Railroad Shippers' forecasts of car loadings. In addition, the U.S. Securities & Exchange Commission (SEC) began a quarterly survey of corporate plans for capital expenditures and *Fortune* magazine began a semiannual readers' poll of business conditions.

Table 1
Components of Business Confidence, first four subperiods

	1946-49	1949-57	1957-62	1962-74
Met Life help wanted	ü	Ü		
Railroad Shippers	ü	Ü		
Fortune magazine	ü			ü
SEC capital expenditures	ü	Ü		
Dun & Bradstreet		Ü	ü	ü
NAPM / ISM manufacturers		Ü	Ü	ü
Bankers' magazine			Ü	ü
US Chamber of Commerce			Ü	ü
Conf Board capital expenditures			Ü	ü
Conf Board help wanted			Ü	Ü
Purchasing magazine			Ü	Ü

The Metropolitan Life help wanted index was eventually replaced by a similar index compiled by the Conference Board. In both cases, the index is based on inches of help wanted advertising in major newspapers. The SEC survey of capital expenditure was also eventually replaced by a similar survey conducted by the Conference Board.

The original *Fortune* magazine survey was short-lived; but, was effectively replaced by a quarterly survey conducted by Dun & Bradstreet (D&B) that continued for many years. The National Association of Purchasing Managers (NAPM) survey, begun in 1948, continues to today as surveys of the Institute for Supply Managers (ISM) for manufacturing, services and health care companies.

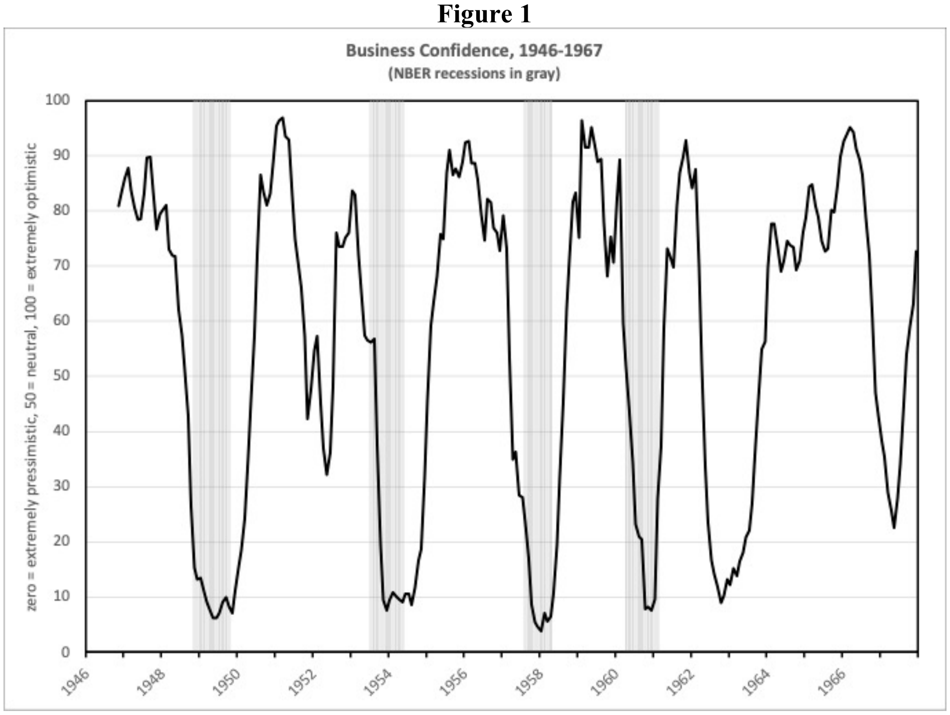
The *Fortune* magazine survey was a readers' poll. Many thousands of responses were received, but only the first several thousand were tabulated. The poll asked questions such as do you expect business in general to be better, to be worse, or to be about the same. There were similar questions regarding sales, income, employment and capital expenditure. From the responses to these kinds of questions, an optimism ratio can be calculated, equal to the percent responding positively minus the percent responding negatively. For example, for November 1947:

Optimism Ratio = % saying "improve" – % saying "worsen" = 22 – 49

To form an index, this number, -27, was converted into a percentile rank over the period of this survey. The percentile ranks of this and the other pertinent questions of the *Fortune* magazine survey were then averaged to give this component's index over the period of the survey. Finally, the percentile ranks of the several surveys of the subperiod were averaged to give the (overall) business confidence index.

The thus-constructed business confidence index turned down prior to the beginning of each of the recessions and turned up at or near the beginning of each recovery of the time segment from 1946 to 1970. See Figure 1. (The complete time series is available on request from the author.) The index also turned down two other times during this time segment, in 1951 and in 1963.

Figure 1



The 1951 and 1963 episodes might involve something other than the anticipation of a slowdown. The 1951 false signal of a recession might have reflected geopolitical considerations due to the Korean War. That is, the index might have been revealing uncertainty about the near-term future, as opposed to a forecast of a slowdown. And, the 1963 false signal might likewise have reflected geopolitical considerations in conjunction with the assassination of President John F. Kennedy and the build-up of American military forces in South Viet Nam.

Whatever the explanation of these and other false signals of recession since 1946, business confidence appears to be quite volatile at times, careening back and forth between very optimistic and very pessimistic.

1966-1987

During the next time segment, 1966 to 1987, the number of surveys useful for inferring business confidence was relatively steady; nevertheless, there was some turnover in composition. Notably, the National Federation of Independent Business (NFIB) began its survey of small business. The NFIB survey has been a membership survey, initially on a quarterly basis and now on a monthly basis.

Just as notable, the Conference Board began a quarterly survey of large corporation CEOs that continues to today. The second Fortune magazine survey ran its course, as did the Dun & Bradstreet survey. The first senior loan officer survey of the Federal Reserve Board (FRB) also ran its course, and a second such survey was started. The National Association of Home Builders (NAHB) began its survey. The complete list of surveys from 1974 to 2004 used to infer business confidence is shown in Table 2.

Table 2
Components of Business Confidence, second four subperiods

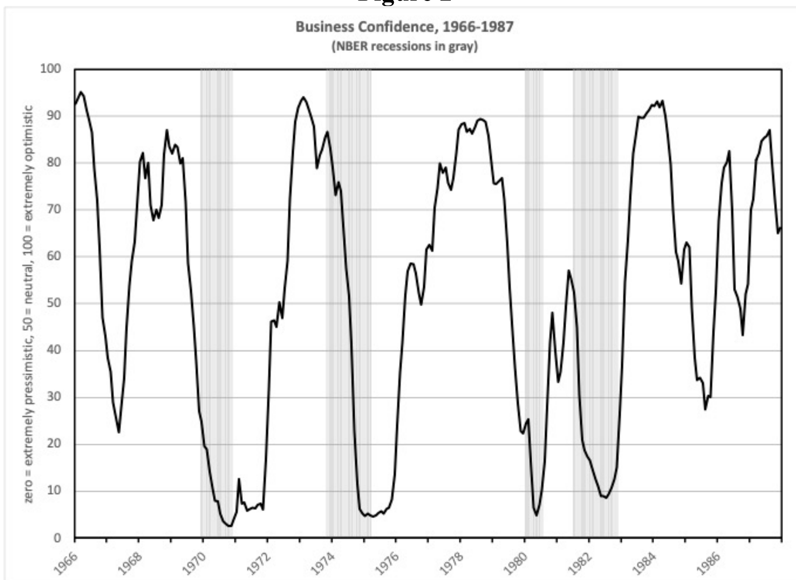
	1974-78	1978-83	1983-91	1991-04
Fortune Magazine	ü	Ü	Ü	
Dun & Bradstreet	ü	Ü	Ü	ü
NAPM / ISM manufacturing	ü	Ü	Ü	ü
Bankers' magazine	ü	Ü		
US Chamber of Commerce	ü	Ü	Ü	
Conf Board capital expenditures	ü	Ü		
Conf Board help wanted	ü	Ü	Ü	ü
Conf Board CEO		Ü	Ü	ü
Conf Board CFO		Ü		
Natl Fed of Independent Business	ü	Ü	Ü	ü
FRB senior loan officer	ü	Ü	Ü	

As Figure 2 shows, business confidence turned down before each NBER recession and turned up at or near each recovery of the next time segment (1966 to 1987). Business confidence also turned down in 1967. Not only did business confidence signal a recession in 1967, so did consumer confidence, the yield curve and the index of leading indicators.

While GDP growth did not turn negative in 1967, it slowed, and the episode is widely recognized as a "growth recession." Alfred Hayes (1967), President of the New York Fed, thought the economy would slow down during 1967 as policy aimed to restrain the imbalanced growth of the mid '60s. The 1967 *Economic Report of the President* (pp. 46-48) talked of the need for moderation following a spurt of economic growth. The 1968 *Economic Report of the Pre*sident (p. 6), posited that the year 1967 as a whole would be viewed as satisfactory; but, began with a "brief pause" following a long period of strong growth.

There is also a ragged period during this time segment. From about 1984 to about 1987, business confidence seems to fluctuate without connection to business cycle turning points. With regard to the turbulent 1984-1987 episode, any number of possibilities for fluctuations in business confidence independent of the economy might be mentioned: the Space Shuttle Challenger disaster of January 1986, the explosion of the Chernobyl nuclear power plant in April 1986, the Iran-Iraq War of 1980-1986, and the Iran-Contra Affair of 1985-1987. The Tax Reform Act of October 1986 was only enacted after several fits and starts. The Stock Market Crash of October 1987 seems sufficient to explain the gyration in business confidence of that year. The fact is, in spite of the events of 1984-1987, GDP growth was robust through the episode. The economy only faltered in 1990-1991.

Figure 2



1987-2007

During the recent several decades, the list of surveys useful for inferring business confidence has grown large. As shown in Table 3, there are now three surveys of small business (NFIB, WSJ/Vistage and

Table 3
Components of Business Confidence, last three subperiods

	2004-12	2012-18	2018-25
Inst for Supply Mgt (ISM) manufacturing	ü	Ü	ü
Inst for Supply Mgt (ISM) services		Ü	ü
Conf Board on-line help wanted	ü		
Conf Board CEO	ü	ü	Ü
Natl Fed of Independent Business (NFIB)	ü	ü	Ü
FRB senior loan officer	ü	ü	Ü
PriceWaterhouseCoopers	ü	ü	
US Chamber of Commerce mid-market			Ü
Manpower Employment Outlook	ü	ü	Ü
Natl Assn of Home Builders (NAHB)	ü	ü	Ü
Natl Assn of Manufacturers (NAM)	ü	ü	Ü
Natl Assn of Credit Managers (NACM)	ü	ü	Ü
Business Roundtable CEO	ü	ü	Ü
CFO Magazine / Richmond Fed	ü	ü	Ü
CEO Magazine	ü	ü	Ü
WSJ / Vistage	ü	ü	Ü
Wells Fargo small business	ü	ü	
MSNBC SurveyMonkey small business			Ü
JOLTS help wanted		Ü	Ü
Amer Inst of CPAs (AICPA)		Ü	Ü

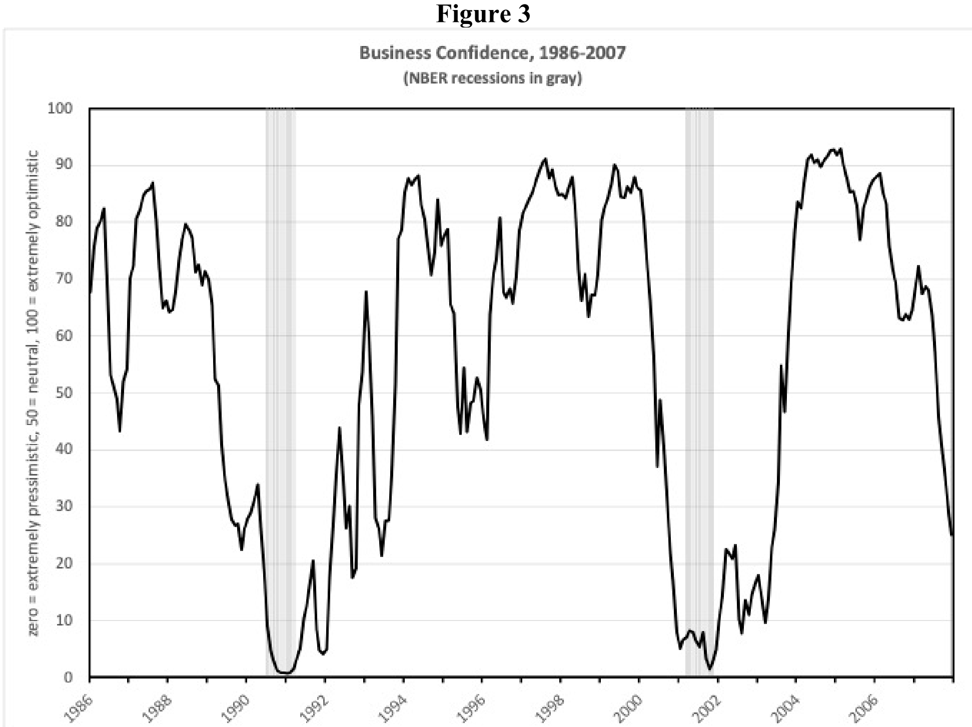
MSNBC/SurveyMonkey surveys) included in the list. There are also three surveys of medium to large business (Conference Board, Business Roundtable and U.S. Chamber of Commerce mid-market surveys).

There are three surveys of production managers (ISM manufacturing, National Association of Home Builders [NAHB] and National Association of Manufacturers [NAM]); and, three surveys of people in banking and finance (FRB senior loan officer, National Association of Credit Managers [NACM] and *CFO* Magazine / Richmond Fed CFO).

The diversity, nowadays, of business surveys invites tracking business confidence of specific segments of the business community.

Figure 3 shows business confidence over the time segment 1987 to 2007. Business confidence turned down before each of the two recessions of this time segment. However, business confidence did not sharply recover following the subsequent recoveries of the economy. Business confidence, instead, remained tentative during the early 1990s and the early 2000s. The fluctuations in business confidence during both the early 1990s and early 2000s might be attributed to the jobless recoveries from the recessions of 1990-91 and 2001. Among roughly contemporary discussions of these jobless recoveries were: Bernanke (2003), Gordon & Bailey (1993), Groshen & Potter (2003), and Schreft & Singh (2003). In real time, Alan Greenspan referred to the initially faltering recovery from the 2001 recession as a "soft patch." (Wolk, 2002)

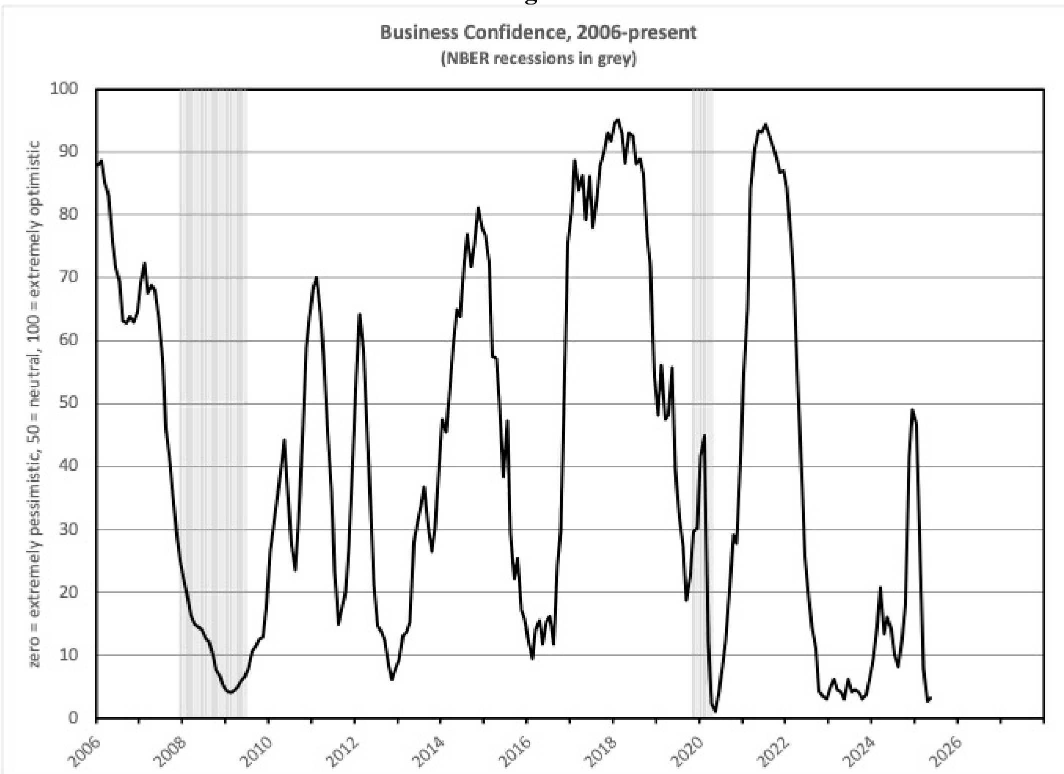
Figure 3



2006-25

Figure 4 shows the course of business confidence during the time period 2006 to 2025. With respect to the two recessions of this time period, business confidence started falling before the start of the recession and business confidence starting rising at or prior to the start of the recovery. But, aside from these movements, there appears to be little cohesion between business confidence and the business cycle.

Figure 4



The fall and then rise of business confidence during 2016 might be related to the Presidential election of that year (Davidson, 2016). Larry Summers famously described the jump in confidence following Trump's election as a 'sugar high.' (Belvidere, 2017) Similar gyrations in business confidence about 2020 and 2024 might reflect similar developments.

Pierre Lemieux (2020, p. 20) connects the volatility of business confidence during Trump's first administration with uncertainties regarding trade and an economy increasingly subject to executive action. Early during Trump's second administration, surveys seemed to indicate that uncertainties regarding trade again undermined business confidence. (Crumly, 2025; Thomas, 2025) Whatever explains the wild fluctuations in business confidence during recent years, business confidence was much more volatile than the economy.

BIG VERSUS SMALL BUSINESS

Figure 5 contrasts Big Business versus Small Business Confidence. Each is based on three underlying surveys as discussed above. While the time series are correlated, there are some notable differences: Big Business Confidence fell prior to the 2020 COVID- and shutdown-related recession, whereas Small Business Confidence did not fall by as much prior to the start of the recession.

Big Business Confidence rose more than Small Business Confidence during the initial months of the recoveries both from the 2007-2009 recession and the 2020 recession. These rises in Big Business Confidence relative to Small Business Confidence may have reflected that federal bailout programs benefited big business more than small, whether explicitly as in the policy of "Too Big To Fail," or otherwise. (Katz, 2011; Meier and Smith, 2021)

Figure 5

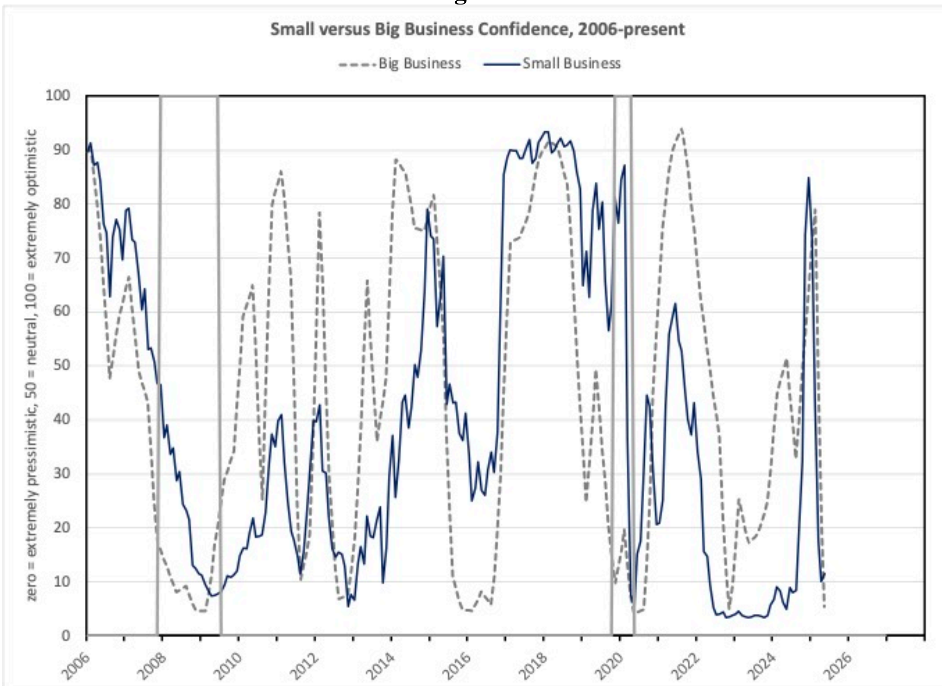
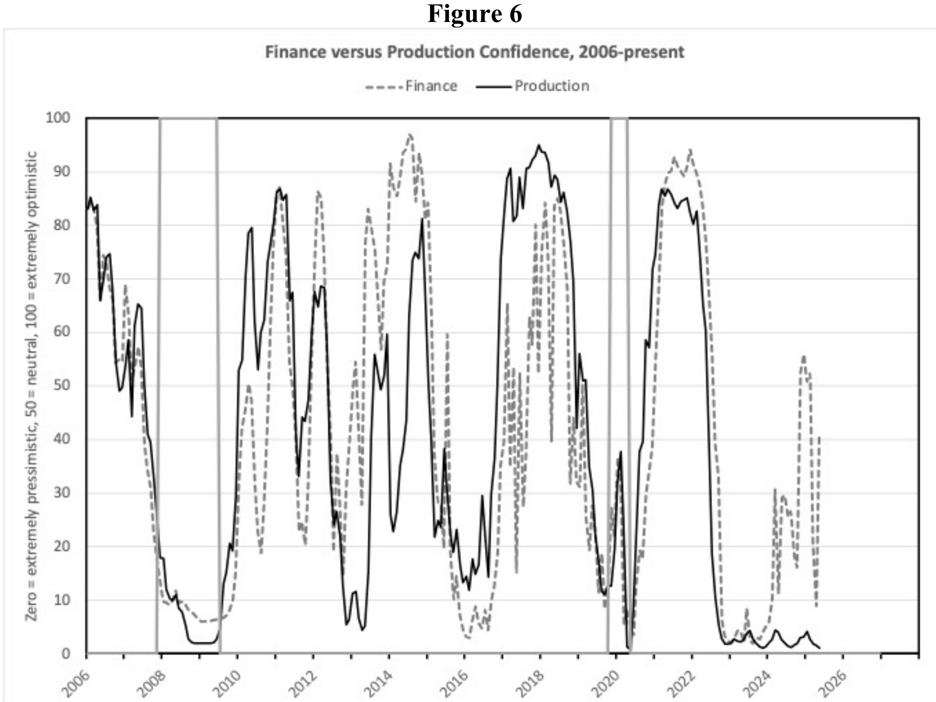


Figure 6



PRODUCTION VERSUS FINANCE

Figure 6 contrasts Production Confidence with Finance Confidence. As with Big and Small Business Confidence, each is based on three underlying surveys as discussed above. With these time series, there is one notable divergence: Finance Confidence exhibited something of a rebound during 2024-2025 while Production Confidence remained flat-lined.

The finding that Finance Confidence rose in the surveys included in this study is confirmed by other surveys. A recent survey by Deloitte indicated that CFOs have turned optimistic about their companies but have not turned optimistic about the economy as a whole. (Poinski, 2025) Another recent survey by CFO Leadership and Blackstone showed that CFO Confidence reached a high early this year, on the hope that the incoming Trump administration would be "friendlier to business, particularly domestic companies, mainly due to the promises of lower corporate taxes and deregulation." (Nolen, 2025)

SUMMARY AND CONCLUSION

It is possible to compile an internally consistent time series of business confidence since 1946. However, to do this, it is necessary to use the information in an evolving set of surveys.

The resulting time series, while roughly correlated with the course of the economy as indicated by NBER business cycle turning points, has exhibited some notable deviations from the course of the economy. It is clear that business confidence has been more volatile than the economy. While this excess volatility might be considered frustrating, business confidence might be interesting specifically because of this excess volatility. It may be useful to investigate what information businesspeople take into account when they form their forecasts, sentiments and plans. It appears that businesspeople consider information concerning geopolitical uncertainties and the challenges associated with the jobless recoveries of the recent past, among other things.

REFERENCES

Belvedere, M.J. (2017). Optimism over Trump a 'sugar high' with no signs of 3-4% economic growth, Larry Summers warns. *CNBC*, March 30. https://www.nbcnews.com/id/wbna3073337

Bernanke, B.S. (2003). The jobless recovery. http://www.federalreserve.gov/boarddocs/speeches/2003/200311062/default.htm

Crumley, B. (2025). New data confirms plunging small-business confidence from trump tariffs. Inc, March 25. https://www.inc.com/bruce-crumley/new-data-confirms-plunging-small-business-confidence-from-trump-tariffs/91165728

Davidson, P. (2016). Trump is juicing business confidence. *USA Today*, December 13. https://www.usatoday.com/story/money/2016/12/13/trump-juicing-business confidence/95372718/

Economic Report of the President, 1967.

Economic Report of the President, 1968.

Gordon, R.J., & Baily, M.N. (1993). The jobless recovery: Does it signal a new era of productivity-led growth? *Brookings Papers on Economic Activity* (1), 271–316.

Groshen, E.L., & Potter, S. (2003). Has structural change contributed to a jobless recovery. *Federal Reserve Bank of New York Current Issues in Economics and Finance* 9:8 (*August*), 1-7.

Hayes, A. (1967). Some economic problems of 1967. *Federal Reserve Bank of New York Monthly Review*, *February*, 23-26.

Katz, J.G. (2011). Who benefited from the bailout? *Minnesota Law Review 95*(5), 1568-1613.

Lemieux, P. (2020). The trump economy: Three years of volatile continuity. *Regulation*, *Summer,* 36-44.

Meier, J.M., & Smith, J. (2021). The COVID-19 bailouts. *COVID Economics, 82*(July): 54-98.

Nolen, M.C. (2025). CFO Optimism hists multi-year high. *CFO Leadership*. https://cfoleadership.com/cfo-optimism-hits-multi-year-high-trump-revenues-capex/

Poinski, M. (2025). At the dawn of a new era, CFOs are optimistic. *Forbes*, January 21. https://www.forbes.com/sites/cfo/2025/01/21/at-the-dawn-of-a-new-era-cfos-are-optimistic/

Schreft, S.L. & Singh, A. (2003). A closer look at jobless recoveries. *Kansas City Federal Reserve Bank Review* (2), 45-73.

Thomas, I. (2025). America's small businesses preparing for recession with Trump tariffs: CNBC SurveyMonkey survey. CNBC, May 9. https://www.cnbc.com/2025/05/09/small-businesses-expect-recession-with-trump-tariffs-poll.html

Thies, C.F. (1986). Business price expectations, 1947-1983. *Journal of Money, Credit and Banking 18*(2): 336-354.

Wolk, M. (2002). Greenspan: Economy in soft patch. *NBC News*, (13). https://www.nbcnews.com/id/wbna3073337

Clifford F. Thies is Professor Emeritus of Economics and Finance at Shenandoah University of Virginia. He received his Ph.D. in economics from Boston College in 1982. Dr. Thies is the author of *Global Economics: A Holistic Approach* (Lexington Books, 2018) and of 104 articles in scholarly journals.

APPENDIX

Table 4
Business Confidence Index, 1946-1985

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1946											80.9	83.7
1947	85.9	87.8	83.7	80.7	78.5	78.6	83.1	89.7	89.7	83.8	76.7	79.5
1948	80.2	81.0	73.1	72.0	71.7	62.2	57.4	51.1	42.8	26.5	15.4	13.2
1949	13.4	11.0	9.1	7.5	6.2	6.3	7.1	9.0	10.0	8.3	7.1	11.1
1950	15.3	18.5	24.3	35.4	45.6	57.3	71.8	86.6	83.5	81.0	83.2	88.4
1951	95.4	96.5	96.9	93.5	92.9	82.7	75.0	70.0	66.1	57.4	42.3	47.6
1952	54.7	57.2	45.3	37.1	32.2	36.2	48.1	76.1	73.5	73.6	75.1	76.0
1953	83.7	82.9	72.1	65.6	57.5	56.6	56.2	56.8	38.0	19.6	9.5	7.6
1954	9.5	10.9	10.2	9.6	9.1	10.5	10.6	8.6	11.7	16.6	18.7	31.8
1955	45.7	59.2	63.5	68.1	75.8	75.0	86.9	91.1	86.5	87.7	86.1	88.9
1956	92.4	92.7	88.6	88.6	85.6	79.6	74.6	82.1	81.6	77.0	76.1	72.8
1957	79.1	73.2	53.5	35.0	36.4	28.5	28.0	23.4	17.0	8.7	5.5	4.6
1958	3.8	7.1	5.6	6.5	10.7	19.4	32.0	45.3	62.3	71.2	81.6	83.3
1959	75.2	96.5	91.6	91.6	95.1	91.8	88.9	89.4	77.9	68.2	75.3	70.6
1960	82.2	89.3	59.7	51.9	43.5	34.3	23.2	21.0	20.4	7.9	8.3	7.6
1961	9.6	27.8	37.3	58.6	73.2	71.8	69.8	81.0	86.9	89.8	92.8	86.9
1962	84.2	87.6	70.2	52.1	33.8	23.5	16.8	14.2	11.8	8.9	10.2	13.2
1963	12.3	15.2	13.9	16.5	18.1	20.8	22.1	26.9	37.5	47.4	54.9	56.3
1964	69.8	77.6	77.7	73.5	69.0	70.9	74.5	73.8	73.4	69.3	71.0	76.2
1965	78.7	84.4	84.7	80.9	78.9	74.5	72.7	73.2	80.1	79.8	84.6	89.7
1966	92.7	94.0	95.2	94.3	91.2	89.3	86.5	78.7	72.0	62.0	47.0	43.0
1967	38.4	35.4	28.9	25.8	22.6	27.9	34.1	44.8	54.0	59.1	63.2	72.7
1968	80.3	82.1	76.8	80.0	71.1	67.8	70.1	68.3	71.0	81.8	87.0	83.5
1969	82.0	83.9	83.4	79.9	81.0	71.7	58.7	52.8	45.1	37.7	27.1	24.8
1970	19.6	19.0	14.2	10.6	7.9	7.8	5.3	3.6	3.1	2.6	2.5	4.0
1971	5.6	12.6	7.4	7.6	5.8	6.2	6.4	6.3	7.1	7.4	6.1	16.7
1972	31.6	46.2	46.4	45.1	50.3	46.9	53.2	59.1	72.4	82.3	88.8	91.8
1973	93.2	94.0	93.0	91.5	89.5	87.7	78.9	81.7	82.9	85.4	86.7	83.1
1974	79.1	73.2	75.9	74.1	64.9	57.6	51.6	41.2	22.8	11.8	6.3	5.2
1975	4.7	5.2	4.8	4.6	4.8	5.4	5.8	5.2	6.2	6.5	8.3	13.6
1976	23.9	35.3	41.7	51.8	56.9	58.5	58.4	56.4	52.3	49.9	53.6	61.6
1977	62.6	61.3	70.4	74.7	79.9	78.0	79.0	75.8	74.3	76.8	82.0	87.1
1978	88.3	88.6	86.7	87.3	86.2	87.6	89.0	89.3	89.2	88.7	85.8	81.2
1979	75.6	75.6	76.1	76.8	72.1	62.9	53.5	44.1	36.7	28.9	22.9	22.4
1980	24.5	25.3	14.4	6.5	4.9	7.1	10.2	16.4	27.9	41.5	48.1	39.9
1981	33.4	35.4	41.7	49.2	57.1	55.3	52.4	44.9	30.6	21.0	18.7	17.4
1982	16.6	14.4	12.5	11.2	9.0	8.9	8.5	9.3	10.7	12.5	15.1	26.4
1983	36.6	54.7	63.5	72.8	82.0	85.4	89.9	89.7	89.6	90.7	91.3	92.4
1984	92.2	93.2	91.8	93.3	90.0	85.9	79.6	70.0	61.1	59.0	54.3	61.5
1985	63.0	62.0	49.3	38.6	33.7	34.2	33.1	27.4	30.3	30.0	43.9	52.3

Table 5

Business Confidence Index, 1986-2025												
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1986	67.7	75.6	79.0	80.2	82.5	69.8	53.1	51.4	48.9	43.3	52.0	54.2
1987	70.0	72.3	80.6	82.2	84.6	85.4	85.8	87.0	80.2	72.2	65.0	66.2
1988	64.1	64.7	67.7	73.6	77.0	79.6	78.6	77.2	71.2	72.5	68.9	71.3
1989	69.9	65.3	52.4	51.3	40.9	34.6	31.4	27.8	26.8	27.1	22.4	26.1
1990	28.0	29.0	31.2	33.9	26.1	18.1	9.2	4.6	2.7	1.2	0.8	0.8
1991	0.7	0.8	1.7	3.3	5.1	10.4	12.6	16.9	20.5	8.4	4.7	4.1
1992	5.1	17.5	26.7	35.5	43.9	35.0	26.2	30.2	17.5	19.2	47.7	53.6
1993	67.7	60.4	45.6	28.0	26.2	21.5	27.4	27.7	35.6	51.4	77.2	78.6
1994	85.2	87.7	86.5	87.4	88.2	83.1	80.5	76.0	70.7	74.7	83.9	75.9
1995	77.5	78.7	65.5	64.0	47.7	42.9	54.3	43.2	48.2	48.6	52.6	50.4
1996	45.7	41.8	63.7	70.8	73.4	80.7	67.7	66.8	68.4	65.7	70.5	78.4
1997	81.5	82.9	84.1	85.4	87.0	88.9	90.5	91.1	87.7	89.2	86.2	84.8
1998	84.9	84.2	85.9	87.9	82.8	71.6	66.2	70.9	63.4	67.2	67.2	70.8
1999	80.3	82.5	84.2	86.8	90.1	89.0	84.4	84.3	86.3	85.1	88.0	86.1
2000	85.6	80.7	72.2	66.5	56.2	37.1	48.7	40.5	32.5	22.2	16.7	8.1
2001	5.2	6.7	7.1	8.3	7.9	6.4	5.4	7.9	3.4	1.5	2.4	4.8
2002	10.0	14.4	22.5	21.8	20.8	23.2	10.3	7.8	13.5	11.0	14.4	16.5
2003	18.0	13.9	9.6	13.8	22.9	25.9	34.5	54.7	46.7	60.4	69.1	78.2
2004	83.5	82.4	86.6	90.9	91.8	90.5	91.0	89.7	90.8	91.5	92.6	92.7
2005	91.8	92.8	90.1	87.7	85.3	85.4	82.9	76.9	82.6	84.0	86.3	87.3
2006	88.0	88.6	85.3	83.1	75.9	71.6	69.4	63.2	62.8	63.8	62.9	64.6
2007	69.0	72.3	67.5	68.8	68.1	63.5	57.1	46.0	40.9	35.5	29.1	25.1
2008	22.0	19.0	16.2	14.9	14.5	14.0	12.7	12.1	10.0	7.6	6.6	5.1
2009	4.3	4.0	4.4	4.9	5.9	6.6	7.7	10.5	11.5	12.5	12.9	17.5
2010	26.6	30.2	34.9	40.0	44.2	34.7	27.6	23.5	30.8	44.2	59.0	63.7
2011	68.5	70.1	64.4	57.3	45.7	36.3	23.4	14.8	17.2	20.0	27.1	40.3
2012	54.9	64.2	58.5	47.5	33.3	21.4	14.6	13.6	12.2	8.4	6.2	7.9
2013	9.1	13.0	13.7	15.2	27.8	30.7	33.8	36.8	30.6	26.5	30.3	40.0
2014	47.5	45.4	51.3	59.0	64.9	63.8	72.2	76.9	71.8	75.3	81.2	77.9
2015	76.8	72.4	57.5	57.2	50.1	38.4	47.3	29.3	22.2	25.4	17.2	15.8
2016	12.1	9.4	14.0	15.4	11.7	15.4	16.1	11.8	24.6	29.8	54.8	75.5
2017	80.5	88.6	83.9	86.2	79.3	86.1	78.0	82.5	87.8	89.9	93.0	91.8
2018	94.6	95.1	92.9	88.3	93.0	92.6	88.1	89.0	86.6	76.7	72.1	54.5
2019	48.2	56.1	47.5	48.2	55.7	39.6	32.0	27.3	18.7	22.6	29.6	30.2
2020	41.6	44.8	12.1	2.3	1.0	3.6	8.3	12.3	20.1	29.1	27.8	40.8
2021	54.7	65.3	84.2	90.5	93.4	93.2	94.4	92.9	91.1	89.0	86.8	87.0
2022	84.5	76.8	69.2	53.2	38.5	25.8	19.0	14.4	11.0	4.3	3.3	3.0
2023	4.6	6.2	4.5	4.2	3.0	6.1	4.1	4.6	4.1	3.0	3.7	6.1
2024	9.3	14.5	20.6	13.4	16.0	14.1	10.1	8.1	12.6	18.1	41.4	48.9
2025	46.6	29.8	7.9	2.6	3.1							

JOURNAL
OF
BUSINESS, ECONOMICS AND TECHNOLOGY