

ABSTRACT

Title of Dissertation: PROMOTING SAFETY IN FOOD SUPPLY CHAINS: NAVIGATING REGULATIONS, INSPECTOR SCHEDULES, AND INCENTIVE STRUCTURES

Abhay Kumar Grover, Doctor of Philosophy,
2024

Dissertation directed by: Professor Martin E. Dresner, Department of Logistics, Business and Public Policy

The dissertation examines different aspects around safety in U.S. food supply chains, using the context of regulatory policy implementation and inspections.

The first essay explores the impact of supply chain accountability regulations on firm level inventory performance within the context of global sourcing. Using the case of the Food Safety Modernization Act, the study suggests that regulatory policies have the potential to negatively impact a firm's inventory performance by increasing regulatory stress. Using the stress-coping theory, the study finds that sourcing from developed markets exacerbates the regulatory stress, while sourcing from emerging markets alleviates it, thus altering the firm's coping response as reflected by its inventory leanness performance. The essay has implications for safety in food supply chains.

The second essay investigates the impact of work-break schedules on task performance of field staff. Using the context of the U.S. Food and Drug Administration's regulatory inspections, the study explores different work-break regimes and their impact on food inspectors' quality assessments of food facilities. The study finds that temporal pacing of inspections increases task performance, but at a diminishing rate. Multi-tasking and non-standard schedules negatively affect performance, while intermittent breaks and start-day of inspections may have a positive effect on inspection outcomes. Strategic scheduling of inspections may increase violation detection.

The third essay investigates the impact of incentive design on task performance of field staff. Using the context of the U.S. Food and Drug Administration's regulatory inspections, the study explores how salary differences relative to multiple referent groups impact food inspectors' quality assessment of food facilities. Grounded in equity theory, the study finds that a higher salary as compared to previous year, internal peers, and operational interface referents, leads to work withdrawals due to complacency. Conversely, a higher salary as compared to industry referents enhances inspector task performance. A strategic incentive design for inspectors may increase the detection of violations.

These studies contribute to the literature at the intersection of supply chain and operations management, public policy, and public sector operations. They do so by advancing our understanding of the factors affecting safety in food supply chains.

PROMOTING SAFETY IN FOOD SUPPLY CHAINS: NAVIGATING
REGULATIONS, INSPECTOR SCHEDULES, AND INCENTIVE
STRUCTURES

by

Abhay Kumar Grover

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2024

Advisory Committee:

Professor Martin E. Dresner, Chair

Professor Thomas M. Corsi

Assistant Professor Adams B. Steven

Professor David E. Cantor

Professor Ali Haghani, Dean's Representative

© Copyright by
Abhay Kumar Grover
2024

Dedication

In honor of my grandparents, who laid the foundation.

Late Smt. Shimla Wanti & Late Sh. Kharayati Lal Grover
Late Smt. Sudarshan & Dr. Om Prakash Malhotra

Acknowledgements

First and foremost, I would like to thank my family, without whose support I could never have dreamed of leaving my career in industry to follow my passion for research. I could not have done this without my pillar of strength, my wife, Isha. I thank her for leaping into the unknown with me and for her faith in my decisions and capabilities. I thank her for giving me the greatest gift of my life, my daughter Amyra, whose smiles have brightened my life. I also want to thank my parents, Drs. Kiran, and Davinder Grover, who inspired me to dream big. I would not be here without them.

Next, I extend my heartfelt thanks to my advisor, Dr. Martin Dresner. My gratitude towards him is immense, for he has inspired me in several ways. Beyond the realm of research, he has instilled in me the values of humility and integrity. I owe special thanks to Dr. Dave Cantor for his guidance throughout my Ph.D. journey. I would not have made it without his guidance and support. I would also like to thank Dr. Tom Corsi. He is a true inspiration and has fueled my passion for research with his enthusiasm and dedication. I also want to express my appreciation to the other committee members, including Dr. Adams Steven and Dr. Ali Haghani, whose invaluable insights have greatly enhanced my dissertation.

A heartfelt thank you to my in-laws, Sangeeta, and Mahendra Chawla, for their support during challenging times. I also want to express my gratitude to my family, friends, mentors, colleagues, and well-wishers, listed in no particular order: Dr. Philip Evers, Dr. Cristian Dezso, Sudiksha and Akshay Grover, Pulkit Chawla, Brett Wang, Lin Lei, Andy Li, Yunfei Wang, Matthew Sabol, Oguz Gencay,

Weihong Zhao, Aleia Elnaeim, Humberto, Coronado, Dr. Laharish Guntuka, Dr. Muhammad Hasan Ashraf, Dr. Kevin Park, Dr. Hao Su, Dr. Xiaodan Pan, Dr. Clare Narrod, Dr. Oliver Yao, Dr. George Chondrakis, Justina Blanco, Miloyka Wilson, Dr. Prabhjot Mukandwal, Dr. Sining Song, Jon Crocker, Dr. Yasmine Deol, Jasdeep Sidhu, Drs. Charu and Neeraj Gandotra, Drs. Ekta and Rajesh Gupta, Hongyi Zhang, Naveen Mangla, Bhavit Chhabra, Anshul Rana, Gurtej Dhillon, and Jaskaran Singh. Thank you to all others not mentioned who have contributed to my journey. I am forever grateful.

Finally, I cannot express my gratitude enough towards that invisible power, for “*Their*” blessings are infinite.

Table of Contents

Dedication	ii
Acknowledgements	iii
Table of Contents	v
Chapter 1: Overview	1
Chapter 2: Unintended Consequences of Regulations	4
1. Introduction	4
2. Literature Review and Theoretical Foundation	8
3. Hypothesis Development	13
3.1. Regulations and Inventory Performance	13
3.2. Moderators of Regulatory Stressors-Inventory Performance Relationship	17
4. Setting and Data	27
4.1. Research Setting: The Food Safety Modernization Act	27
4.2. Data Sources and Collection Process	29
4.3. Variables and Variable Construction	33
5. Methods	39
5.1. Coarsen Exact Matching	40
5.2. Parallel Trend Assumption Checks for DID Analysis	42
5.3. Model Specification	44
6. Results	45
6.1. Analysis of Results	45
6.2. Robustness Checks	46
7. Discussion	48
7.1. Theory Contributions	54
7.2. Managerial and Policy Implications	57
7.3. Limitations and Future Research	58
Acknowledgments	59
Chapter 3: Work-Break Schedules and Quality Assessment	60
1. Introduction	60
2. Related Literature	66
2.1. Breaks and Productivity	66
2.2. Quality Management and Inspections	70
2.3. Scheduling and Operational Outcomes	72
3. Hypothesis Development	74
3.1. Pseudo Work-Break Schedule	74
3.2. Intermittent Work-Break Schedule	77
3.3. Work-week Effects	79
3.4. Non-standard Work-Break Schedule	82
3.5. Concurrent Work Schedule	84
4. Methods	86
4.1. Research Setting: Food Safety Inspections	86
4.2. Data	89
4.3. Measures	91
4.4. Empirical Specification	98

5. Results.....	101
5.1. Model Results	101
5.2. Identification	106
5.3. Robustness Tests.....	111
5.4. Further Analysis.....	114
6. Discussion	117
6.1. Contributions.....	121
6.2. Managerial Implications	123
6.3. Limitations and Future Research	124
Acknowledgments.....	125
Chapter 4: Impact of Salary Differences on Quality Assessment.....	126
1. Introduction.....	126
2. Related Literature.....	134
2.1. Incentive Design and Operational Outcomes	134
2.2. Field Staff Motivation and Operational Performance.....	138
3. Hypothesis Development	141
3.1. Self Referent	143
3.2. Industry Referent	144
3.3. Internal Peer Referent	145
3.4. Operational Interface Referent.....	147
4. Methods.....	149
4.1. Research Setting: Food Safety Inspections.....	149
4.2. Data	152
4.3. Measures	154
4.4. Empirical Specification.....	164
5. Results.....	166
5.1. Model Results	166
5.2. Econometric Issues	170
5.3. Robustness Tests.....	175
5.4. Further Analysis.....	177
6. Discussion	182
6.1. Contributions.....	183
6.2. Managerial Implications	185
6.3. Limitations and Future Research	187
Acknowledgments.....	188
Chapter 5: Conclusions and Future Extensions	190
Appendices.....	192
Bibliography	206

Chapter 1: Overview

The dissertation comprises three essays that systematically examine different aspects around safety in the U.S. food supply chains. Specifically, the first essay explores the impact of food safety regulations on the inventory performance of affected firms. The second essay investigates the effect of work-break regimes, and the third essay studies the impact of incentive design on the quality assessment of regulatory food inspectors.

The first essay examines the impact of regulatory stress on the inventory performance of firms. This study uses the theoretical foundations of stress and coping theory to examine the Food Safety Modernization Act (FSMA), a regulation focused on supply chain accountability and product safety. It aims to understand the unintended consequences of regulatory stress on the inventory performance of firms. Although regulations such as the FSMA are intended to safeguard consumer safety by mandating supply chain accountability for a firm, they may also impact a firm's inventory performance. The situation may be further influenced by the global sourcing behavior of the firms.

To address the research question, the study uses a comprehensive archival panel dataset comprising 4,119 firm-year observations, combined from multiple sources including a proprietary database on global sourcing, PIERS, and a database on financial and market information of publicly listed firms, COMPUSTAT. Using a difference-in-difference methodology, the study finds strong evidence of unintended impact of the FSMA policy on a firm's inventory leanness. Contrary to existing beliefs, it finds that sourcing from developed markets exacerbates the negative impact of regulations on inventory performance and sourcing from emerging markets alleviates its impact. This research contributes to the literature at the

intersection of supply chain and public policy by expanding the understanding of how public policies impact firm behavior and performance in the context of food safety.

The second essay examines the impact of different work-break schedules on task performance of field staff. The study uses the context of regulatory food safety inspections conducted by the U.S. Food and Drug Administration to explore the research question. The study posits that field staff, such as food inspectors, experience schedule-induced workload fatigue that impacts their quality assessment of food facilities. The fatigue manifests in the form of under detection of violations at non-compliant facilities which could increase food safety incidents and jeopardize consumer safety. The existing literature provides limited evidence into how the timing of breaks relative to inspection tasks affect their outcomes. Accordingly, this study explores how strategic work-break scheduling can improve violation detection.

The study utilizes a proprietary dataset of 7,464 FDA inspections across 3,291 food facility inspections from January 2013 to March 2020. Using fixed-effects Poisson Pseudo-Maximum Likelihood (PPML) regression the study finds that pseudo breaks, i.e., time between two successive inspections, improves violation detection but at a diminishing rate. Further, concurrent, and non-standard work-break schedules are counterproductive. However, intermittent breaks, i.e., breaks in-between multi-day inspections, help improve inspection outcomes. Lastly, the study finds that timing of inspections within the workweek significantly influences inspection effectiveness. This study makes several significant contributions to the literature on breaks, productivity, quality management, and inspections in the context of food safety.

The third essay explores the impact of incentive design on task performance of field staff. This study uses the context of the U.S. Food and Drug Administration's food facility inspections

to understand how salary differences with respect to multiple referent groups affects work withdrawal behavior of food inspectors. We posit that salary differences with respect to prominent referent groups, including self-referents, internal peer referents, industry referents, and the operational interface referents, could impact operational performance of employees. In the context of food inspections, it could lead to undetected violations at food facilities which can affect safety in the food supply chains.

To examine the research question, this study uses proprietary data from 594 FDA food inspectors conducting 16,382 inspections across 14,441 food facilities, along with the salary information of individual food inspectors from Jan. 2015 to Dec. 2019. The study finds that higher salary in comparison with self-referent, internal peers and operational referents results in a decrease in quality assessment of food inspectors, likely due to complacency. Further, a higher salary in comparison with industry referent groups results in an improved quality assessment of food inspectors, likely due to increased motivation. The study contributes by introducing the salary variance aspect of incentive design to operations management literature. The study finds that referent groups have differential impacts on work withdrawal behavior in the context of food safety inspections.

Chapter 2: Unintended Consequences of Regulations

1. Introduction

Public policies that promote supply chain accountability – the regulatory mandate that firms actively manage their global supply chains (Parmigiani et al., 2011) – are increasingly becoming an important component of the firm's corporate strategy. Many firms are under increased pressure to implement supply chain accountability practices because consumers, investors, and regulatory bodies demand greater transparency, ownership, and commitment to promote responsible behavior across the supply chain (Gualandris et al., 2015). Several government agencies are considering or implementing rules to enhance supply chain accountability because many firms remain very dependent upon foreign suppliers for critical materials to support the production process (Stanczyk et al., 2017). For instance, McDonald's faced reputational harm when one of its meat suppliers in China was shut down for food safety violations, exposing several customers to foodborne illnesses (Jones, 2020). However, regulations designed with specific objectives may impact other aspects of firm performance.

Undoubtedly, public policies that promote supply chain accountability offer many benefits to firms, shareholders, and society at large. Indeed, there has been a broader stream of supply chain research that has examined the impact of mandatory government policies on several dimensions of firm performance such as inventory (Darby et al. 2020), safety (Cantor et al. 2009; McFadden & Hosmane 2001), buyer-supplier conflict, and supply chain compliance (Gundlach et al., 2019). Some of this literature has demonstrated that public policies can have a positive effect on firm performance (Cantor et al., 2009; Gundlach et al., 2019). For instance, Cantor et al. (2009) demonstrate how electronic logbook adoption improves safety performance. In contrast, other research such as Miller et al. (2020) demonstrate that government policies have a

tapering effect on safety performance because some firms will eventually recede in their poor safety practices. However, there is a potential for unintended consequences of a mandatory govt policy on an aspect of supply chain performance not directly designed/ intended by the policy.

Surprisingly, prior research has not explored the impact of a mandatory government's supply chain accountability mandatory rules on inventory performance. While Parmigiani et al. (2011) approached this topic by developing an integrative supply chain accountability model, they did not empirically test their propositions. Swift et al. (2019) evaluated the impact of the Dodd-Frank legislation, a conflict mineral disclosure act, on the firm's profitability and market performance. However, the focus of their paper was not on supply chain performance (such as inventory). Thus, there is a serious gap in the literature regarding how mandated govt. regulations impact supply chain performance of firms.

The purpose of this study is to examine how a mandate requiring firms to adopt supply chain accountability measures to improve inventory performance. To investigate our theoretical model, we use stress and coping theory from the organizational psychology literature (Biggs et al., 2017; Folkman, 1984; Somerfield & McCrae, 2000) as a basis for our hypothesized relationships. Using this theoretical lens, we suggest that the regulatory environment that mandates supply chain accountability can act as an external stressor for firms operating in our setting (Dobrzykowski, 2019; Fugate et al., 2019). Because firms must navigate the rules and norms of the accountability regulation, this stress or burden caused by the new requirements creates a new form of risk due to the increased uncertainty associated with the mandate. Based on the firm's appraisal of this new stress, the firm's managers must consider employing varying degrees of coping strategies to minimize the level of stress (Scott et al., 2021). This stress is further exacerbated for firms that significantly utilize a global sourcing strategy, hence, making

compliance more complicated. Put simply, these firms must monitor global suppliers and ensure that the suppliers adhere to the specific requirements of supply chain accountability regulations. In addition, this research examines the moderating role of the firm's dependence on global sourcing on this relationship.

We test our hypotheses in the context of the Food Safety Modernization Act (FSMA), a mandatory regulation rooted in product safety (USFDA, 2024). We use the introduction of FSMA to study the research question. We created a unique archival panel dataset of 4,119 publicly traded firms that stored, transported, handled, or manufactured food products from 2010 to 2019. We evaluate inventory performance before and after the regulation. Using a difference-in-difference strategy, we find that the greater stress due to implementation of FSMA led to a significant decrease in inventory performance. We find that a firm's overall global sourcing strategy doesn't enhance the main effect relationship. However, we further explore the role of global sourcing and find that a firm's dependence on developed markets for sourcing exacerbates the regulatory stress, worsening the inventory performance. We find that a firm's dependence on emerging markets for sourcing alleviates the regulatory stress enhancing inventory performance. However, we find no evidence of the effect of sourcing from frontier markets. Put simply, we find no evidence that global sourcing moderates this relationship; however, it depends on the market from which firms' source. We use the Morgan Stanley Capital International (MSCI, 2024) classification to differentiate and categorize different markets as developed, emerging, and frontier (such as Han et al., 2012).

Our research makes several valuable contributions to literature and provides managerial and public policy insights. First, we add to the debate in the literature by demonstrating that introduction of supply chain accountability regulations impact supply chain performance

(inventory leanness). While some past research has shown that public policies can positively impact firm performance (Cantor et al., 2009), our study shows that firms are reacting to the mandatory safety regulation by pursuing an inventory buffering strategy to guard against risk associated with having to comply with the safety regulation. This reaction may have negative consequences. Relatedly, we contribute to the supply chain public policy literature by demonstrating how this regulation influences inventory leanness. Inventory leanness, a practice to minimize inventory levels to reduce costs and increase efficiency, is important for firms striving for operational excellence (Eroglu & Hofer, 2011). This understanding is important for enhancing supply chain resilience and responsiveness of firms to regulatory demands and changes.

Third, this is one of the few studies that examine how a mandatory policy can have unintended consequences on an important dimension of firm performance, namely, inventory leanness. It is important to do so because the government regulation did not explicitly consider the mandatory rules could impact an important measure of a firm's internal inventory efficiency. Fourth, we also contribute to literature on international purchasing and lean management by exploring the role of global sourcing as a moderator. Prior "lean management" literature has not studied how dependence on global sourcing can either enhance or weaken the impact on a firm's inventory leanness (Eroglu & Hofer, 2011; Han et al., 2012; Hofer et al., 2012). Lastly, we extend the stress and coping theory and establish its utility in a public policy supply chain context (Folkman, 1984). The academic literature on how firms respond to regulations and employ coping mechanisms in the face of stressors is limited. This theory has found applications in supply chain and operations management research to explore the impact of emotions on relational outcomes (Thomas et al., 2021; Urda & Loch, 2013), understand customer revenge

behavior (Zourrig et al., 2009), examine employees' response to workload (Kuntz et al., 2015), study customer reactions to delivery driver behavior (Masorgo et al., 2023), assess the effects of online order fulfillment on customer outcomes (Griffis et al., 2012; Ta et al., 2023), and investigate the relationship between personnel satisfaction, service quality, and profitability (Yee et al., 2008). However, its use in understanding a firm's behavior under stress is limited. Drawing from these applications, we extend the stress and coping theory to analyze firm-level behavior in response to regulatory stressors.

Further, our research has managerial and public policy implications for firms that participate in different markets for sourcing of products. Our findings suggest that under regulatory stress, especially supply chain accountability, a firm's inventory performance may be improved or less impacted by operating in emerging markets as compared to developed markets. From a public policy perspective, we suggest that regulations may have unintended consequences that need to be considered for an in-depth impact analysis of such regulations.

In the next sections we present the following: section 2 presents theoretical foundation of the study along with conceptual model. Section 3 presents the hypothesis development. Section 4, 5 and 6 provide insights into the detailed research setting, methods, and results. Finally, section 7 discusses the implications, limitations, and contributions of the study.

2. Literature review and Theoretical foundation

We use the stress and coping theory from the organizational psychology literature as the theoretical basis for this paper. The theory was proposed by Lazarus and Folkman (1984) and has been extensively used in psychology literature (Biggs et al., 2017; Folkman, 1984; Lazarus 1998; Somerfield & McCrae, 2000). This theory is primarily used to examine how entities experience and respond to stress that arises from either the internal or external environment.

The two main elements of the theory include stress and coping. The theory defines stress as the interaction between the entity and the environment, where the entity evaluates environmental demands as challenging and threatening to the status quo (Biggs et al., 2017; Somerfield & McCrae, 2000). Coping is the effort made by the entity to manage these environmental demands (Biggs et al., 2017). Entities conduct appraisals to assess the significance of stressors in their environment. Further, the theory delineates coping efforts as either externally or internally focused (Compas et al., 2001). Externally focused strategies are actions initiated by the entity to manage or alter the external environment causing the stress, while internally focused coping involves actions targeted to regulate internal responses to stressors (Biggs et al., 2017; Herman & Tetrick, 2009). The outcome is adaptation which can be positive, leading to growth and resilience or negative resulting in maladaptation or dysfunction. The effectiveness of coping strategies is gauged by the entities ability to reduce the level of stress.

This theory has been predominantly used to explain how an entity attempts to cope with adverse events. For instance, Rafferty and Griffin (2006) used this theory to study how different characteristics of job changes affect employee well-being. They suggested that the frequency of change and the potential disruptive impact of change influence individuals' appraisal of uncertainty, resulting in job withdrawal and turnover. Trougakos et al. (2020) applied stress and coping theory (or transactional stress theory) to understand how COVID-19-related health anxiety impacts critical work outcomes of employees. They found that fears of contracting COVID-19 impaired employees' goal progress and other aspects of their lives, including social engagement. In another study, Stahl and Caligiuri (2005) focused on the coping aspect of the theory, examining how different coping strategies used by expatriate managers help them

manage problems related to cross-cultural adjustments during international assignments. They found that externally focused coping strategies were more effective at influencing expatriates' intention to remain on the assignment.

There is increased interest within the supply chain and operations management literature to use this theory to study individual coping behaviors. Masorgo et al. (2023), used the appraisal aspect of the theory to understand how customers assess the inappropriate behavior of last mile delivery drivers affecting their repurchase intentions. Autry and Daugherty (2003) used the stress and coping theory to study person-organization fit, job satisfaction, and coping responses among warehouse operations employees. Morrison (2015) used the concept of stress and coping in a supply chain setting to understand how resource shortages act as stressors or triggers of workplace pressures for employees that may instigate front line coping mechanisms called workarounds. They highlight the role of managers, frontline employees, and physical characteristics of workplaces to alleviate stress.

While many scholars have used the stress and coping theory and related research to examine several factors that can affect the performance of individuals (Autry & Daugherty, 2003; Masorgo et al., 2023; Rafferty & Griffin, 2006; Stahl & Caligiuri, 2005; Trougakos et al., 2020), some scholars have demonstrated the utility of this theory at the firm level. Thomas et al. (2011) examined how time pressure (stress) from buyers prompt coping behaviors in suppliers, such as working faster or avoiding risks. In another study, Zhang et al. (2009) explored how relational stress between buyers and suppliers affects the latter's investment in technology to cope with their demands. We demonstrate further utility of this theory by theorizing how firms behave under stressful conditions such as changes to the regulatory environment.

We suggest that the regulatory environment in which a firm operates can act as an external stressor for them (Dobrzykowski, 2019; Fugate et al., 2019). Many firms are often the subject of regulatory agencies in the firm's operating environment; hence firms must navigate the rules and norms that are enacted by regulations (Cantor et al., 2022). The stress comes from the regulatory burden i.e., the cost and efforts required to meet regulatory standards and requirements to gain legitimacy (Grover & Dresner, 2022). The introduction or changes to a regulation could threaten the existing norms of the operation, becoming a source of external stress. Depending on a firm's appraisal of stress, they would engage in different degrees of coping efforts to reduce the level of stress (Scott et al., 2021).

Supply chain literature recognizes that regulations are a significant stressor in supply chain risk management (Manuj & Mentzer, 2008; Wagner & Bode, 2008). The change in regulations triggers stress responses by firms, prompting them to implement coping strategies such as inventory management (e.g., Dooley et al., 2010; Elking et al., 2017) to buffer against potential disruptions. Historically, firms have used inventory strategies as defense mechanisms to manage stress induced by regulatory risks, such as increased tariffs (Darby et al., 2020). For example, the high-tech sector experienced an increase in 11% inventory between 2018 and 2021, a strategic response to mitigate the stress of uncertainties related to regulatory compliance and supply chain disruptions (McKinsey & Co., 2022). Darby et al. (2020) suggested that government policies, acting as external stressors, can alter the flow of crucial resources between firms through such policies and regulations. This leads to stress-coping behavior where firms safeguard against the risk. We propose that firms cope by building inventory as it allows them to buffer against potential supply chain disruptions caused by regulations.

We study how firms, when faced with the stress of mandatory regulatory compliance, utilize inventory buffering—a targeted coping mechanism—to navigate these pressures. We also study how a firm’s appraisal of stress of regulatory changes based on their sourcing from different markets affect the relationship between the stress from regulatory changes and coping strategy of inventory management i.e., how resource dependence on other markets based in different regulatory environments can alleviate or worsen the impact of stress of regulatory compliance. We develop our theoretical model by examining how different markets where firms source their resources from i.e., developed, emerging, and frontier markets influence the firm’s coping strategy (refer figure 1).

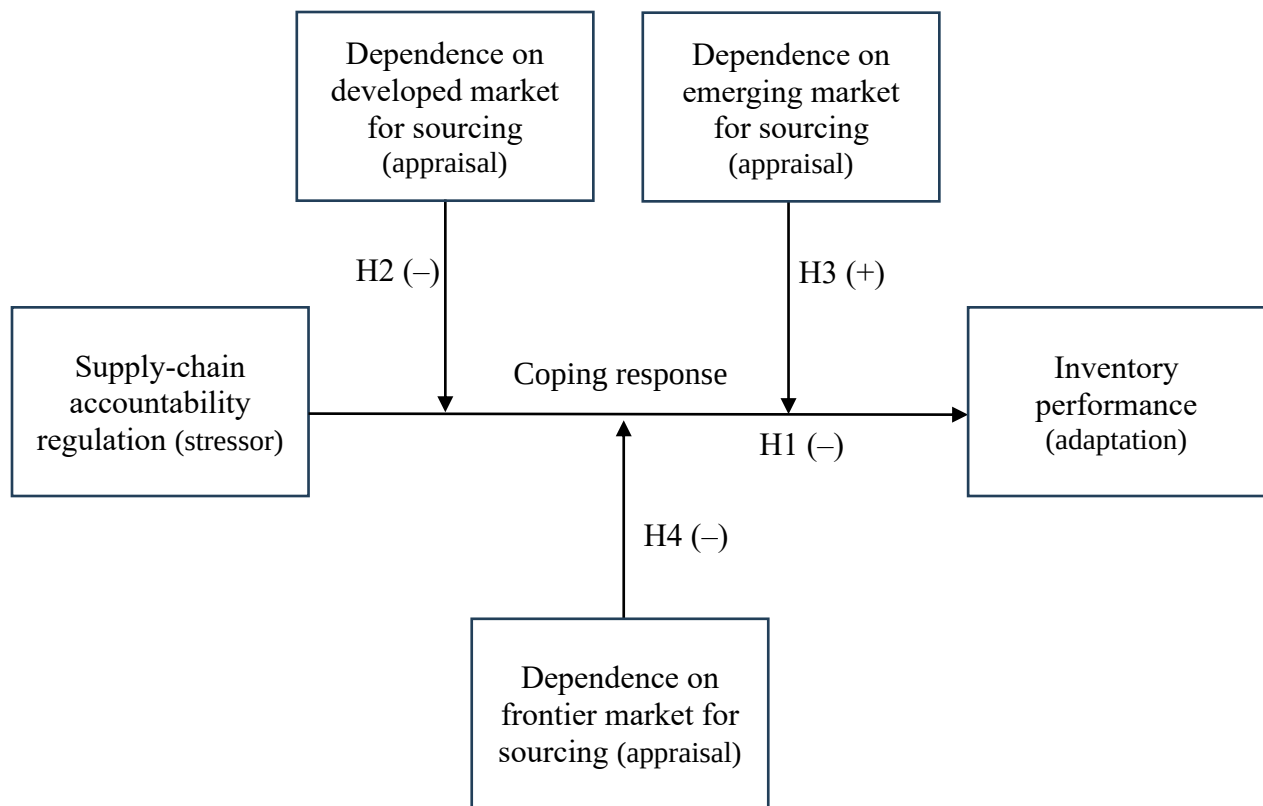


Figure 1. Hypothesized model

In this paper, our baseline hypothesis (H1) posits that introduction of such regulations act as stressors by imposing stricter safety standards and firms' buffer against the stressor leading to decreased inventory leanness as a coping response. We then theorize that when firms appraise the stress from the regulation based on sourcing from developed markets (H2) it worsens the stress of regulatory requirements on inventory leanness due to high rigidity of such markets to adapt to new requirements. We further suggest that when firms appraise the regulatory stress based on sourcing from emerging markets (H3), it reduces the stress impact of regulation on inventory leanness due to higher flexibility and adaptability of the market. Also, when firms evaluate the regulatory stress based on sourcing on frontier markets (H4), it exacerbates the stress impact of regulatory requirements on the inventory performance due to high rigidity from lack of infrastructure and high market distance.

3. Hypothesis development

3.1. Regulations and inventory performance

Our baseline hypothesis is that introduction of supply chain accountability regulatory stressors triggers an inventory buffering coping response. Government regulations create policies and constraints concerning the firm's operations such as production, distribution, and consumption of critical resources (Fugate et al., 2019). DiMaggio and Powell, (1983) and Oliver (1991) suggest that firms perceive pressures and restrictions imposed upon them by regulations with the goal of motivating firms to adjust their boundaries. Accordingly, the introduction of new regulations disrupts the status quo. It introduces new ways or modifies existing ways of conducting business by changing the rules or expectations inducing stress targeted to certain firms (Tokar & Swink, 2019).

In general, there is tension in the literature on how some firms may cope with this external stress. Some studies suggest that firms cope positively with this stress. In the transportation context, Cantor et al. (2009), Miller et al. (2020) and McFadden and Hosmane (2001) demonstrate how introduction of government transportation policies reduce a firm's involvement in crashes, thereby promoting safety. On the other hand, there are studies that suggest that firms may cope negatively with regulatory stress. Previous research has acknowledged that government policies can increase supply chain risk, causing firms to restrict their growth strategies (Kleindorfer & Saad, 2005; Tang & Tomlin, 2008). We suggest that firms hedge against the regulatory stress from supply chain accountability regulations by hoarding inventory leading to decreased inventory leanness.

Levie and Autio (2011) introduced the idea of regulatory burden where they suggest that newer stringent policies might be sending signals of stressors. Regulatory burden is defined as direct or indirect penalties that government bodies impose on the conduct of economic activities in their jurisdiction. The goal of regulations is to ensure fairness of business transactions, promote responsible conduct of business, and protect the interest of consumers. Such restrictions require firms to build the capacity to comply with rules (Easley, & O'Hara, 2009). The first source of stress due to the regulatory burden is the cost and effort required to meet current supply chain regulatory expectations. Some supply chain accountability regulations require firms to increase traceability and accountability thus increasing barriers to entry to the industry and sustenance of business operations. For example, Swift et al. (2019) suggested that compliance with regulations mandating supply chain visibility, such as the Dodd-Frank Act of 2010, typically requires firms to invest in IT tools and external consulting, leading to additional costs, which create operational barriers for firms.

The second source of stress is the risk and uncertainty arising from the introduction of new supply chain regulations. The introduction of new regulations involves a multi-stage process including input from key stakeholders that may be impacted by the implementation of such policies (Kim et al., 2012; Fowler 2006). Fowler (2006) suggests that it may be difficult to predict whether the agencies would implement a moderate or very constraining version of the regulatory proposal. Moreover, there might be uncertainty related to its uniform implementation across different jurisdictions and the effectiveness of regulating agencies implementing such policies. This uncertainty is due to the dependency on several unpredictable factors, such as logistics competence. The alignment between different agencies including federal, state, local administrations may affect the nature of execution of such regulations (Kim et al., 2012). Such risks from uncertain execution can make it difficult for firms to forecast costs and predict demand, leading to inefficiencies in supply chains. It might reduce investment in supply chain innovation and increase stockpiling of resources.

However, not every firm reacts to regulatory stress in the same way. The stress and coping theory emphasize the role of appraisal in perception of stress. Firms within the supply chain context that encounter such stressors first evaluate and assess the impact on supply chain operations to determine their level of severity. This appraisal process is critical as it influences coping strategies related to supply chain management (Lazarus & Folkman, 1984). Appraisal of stress can be affected by the coping resources available within the supply chains of the firms i.e., firms with more supply chain resources to cope with stress would still experience stress but to a lower degree. Furthermore, different supply chain industries might experience stress differently based on their dependence on such regulations or political ties through lobbying. We focus on the firms dealing with production, storage, and distribution of food products.

According to the stress-coping theory, stress induces a process that involves the production of stress that originates from the external environment (e.g., introduction of a regulation) and the response of an entity that is subject to these stressors (supply chain coping mechanism) (Lazarus & Folkman, 1984). Many firms will try to manage the environmental risk such as building inventory to mitigate uncertainty. Inventory buffering is a risk mitigation strategy used by many firms to safeguard themselves from the potential negative effects of such stressors (Darby et al., 2020). Accumulating inventory enables firms to guard against the potential limited access to critical resources in the short-term allowing them to reduce the risk that could arise from the uncertainty of policy implementation (Pfeffer & Salancik, 1978). Surplus inventory reduces the vulnerability of firms allowing them access to critical resources (Anand et al., 2008). Existing supply chain literature on risk management suggests the use of inventory as buffers to hedge against the risk of future sustenance of order fulfillment (e.g., Chopra & Sodhi, 2004; Kull & Closs, 2008; Talluri et al., 2013). For instance, Fan et al. (2022) suggest that firms increased the days of supply of goods during the 2018 U.S.-China trade war and the regulatory sanctions that followed.

The regulatory burden (stressor) may also lead to firms increasing their inventory as a coping mechanism. Stringent regulatory requirements, such as supply chain accountability regulations, could increase the burden of testing, quality inspections, certifications, documentation, and record keeping (Lee et al., 1997). For example, the Consumer Product Safety Improvement Act (CPSIA) of 2008 mandates testing requirements for supplier of toys to ensure they meet safety standards for lead content and phthalates (US Consumer Product Safety, 2024). These requirements may increase lead times for approval necessitating larger inventories. Firms might be forced to hold larger safety stock to fulfill their future customer commitments. A higher

regulatory burden may slow down the supply chain making it difficult for firms to respond to market demands. Holding more inventory enables the firm to improve their response time to customer orders, mitigating the impact of a slower supply chain and building a competitive advantage (Chopra & Sodhi, 2004). Further, increased regulatory scrutiny and inspections adds to the existing uncertainties. For instance, firms face greater risks of product rejections at the port of entry or greater risks of product recalls during inspections, forcing them to build more inventory both in the short and long-term.

However, firms that use the inventory buffering practice are doing so to the detriment of inventory leanness (Eroglu & Hofer, 2011). Many firms aspire to achieving their inventory leanness goals by reducing or eliminating wasteful manufacturing and distribution practices such as Just-In-Time (JIT) inventory management. The ideas of stockless production have been promoted for a long time. Managing higher levels of inventory is commonly viewed as wasteful practice that drains scarce and valuable resources of firms by blocking capital and warehouse space. We propose that under regulatory stressors firms cope by buffering inventory resulting in a negative adaptation or (maladaptation) of decreased inventory leanness.

Hypothesis 1: An increase in mandatory supply chain regulatory stressors will motivate firms to increase their inventory buffering as a coping strategy.

3.2. The moderating role of global markets on the regulatory stressors–inventory performance relationship

We now turn to examining how global sourcing further exposes many firms to additional risks by complying with a mandatory supply chain accountability regulation (Bozarth et al., 1998). Put simply, the global sourcing literature suggests that the confluence of regulatory stressors and global sourcing may create additional concerns such as opportunism and bounded

rationality for many firms leading to increased regulatory stress (Han et al., 2013; Kusaba et al., 2011). Firms involved in global sourcing must navigate the multiple regulatory environments in both buying and selling markets. Indeed, it is challenging to ensure that product quality standards are met due to information asymmetry that arises in foreign markets (Steven & Britto, 2016). Global sourcing gives rise to the uncertainty of access to critical resources compounding existing regulatory stress.

The multiple challenges with global sourcing may not be uniform across all markets. In fact, sourcing from different markets may have varying associated costs and benefits, potentially affecting the relationship between regulatory stress and inventory performance in distinct ways. For example, Hitt et al. (2000) suggest that the nature of formal and informal institutions in developed and emerging markets affect the degree of resource uncertainty differently. Bozarth et al. (1998) suggest that deployment of supply chain strategies (e.g., contracts, multiple sourcing) could vary by market depending on the threat for the buyers. Hitt et al. (2000), Kusaba et al. (2011), and Shou et al. (2016) further highlighted the unique supply chain challenges faced by firms operating in emerging and developed markets, suggesting that these markets have the potential to compound existing regulatory stress differently.

There is debate in the literature about how various types of markets could impact inventory performance of firms. For instance, Steven and Britto (2016) suggest that emerging and developed markets can be differentiated based on contract enforcement and institutional maturities where transactions in emerging markets negatively affects inventory performance due to market inefficiencies as compared to developed markets. Whereas Han et al. (2013) suggests that firms engaging with emerging markets may experience lower inventories as compared to developed markets because contract manufacturing—a primary method of engagement with

these markets—allows multinational firms to operate at a steady state. This highlights a tension in the literature about how sourcing arrangements in different markets could impact a buyer's performance differently.

Although the existing supply chain literature clearly establishes the fundamental need to evaluate these markets differently with respect to inventory performance, there is a lack of consensus on the impact of different markets on firms' inventory performance (coping response), especially under regulatory stress. We suggest that each type of foreign market exposes the firm to different levels of uncertainty and stress. We propose that market rigidity – referring to inflexibility and strictness of regulations (Shi et al., 2017) and market distance – the degree of regulatory difference from the buyers' market (Wu & Salomon, 2016) act as stress regulating factors within the stress and coping framework. These factors modulate how firms appraise and respond to the stress of regulatory compliance when sourcing from foreign markets (Belderbos et al., 2017; Wu & Salomon, 2016).

According to MSCI (2024) classification, developed markets, such as Australia, United Kingdom, Germany, or Japan, might have higher rigidity with stringent laws and regulations making it challenging for local businesses to adapt or modify their operations for regulatory compliance (Li et al., 2020). In contrast, sourcing from frontier markets such as Kenya, Jamaica, or Croatia, which exhibit greater regulatory distance and less stability, likely intensifies the stress response compliance as it is more challenging for firms to navigate and comply with regulations in such markets (Witt & Lewin, 2007). These factors predominantly shape how firms respond to the regulatory stressors when sourcing globally. The dependence on a particular foreign market can be viewed as a stressor that affects how firms cope with the mandatory regulatory induced

burden. Each foreign market presents unique combinations of rigidity and distance, altering the firm's intensity of coping strategy i.e., managing inventory leanness.

As noted, each type of global market presents unique regulatory, political, and economic challenges that the buyer firm must navigate. Developed markets, as an example, offer stable regulatory, political, and economic conditions to buyer firms (Hitt et al., 2000; Kusaba et al., 2011; Shou et al., 2016). Hitt et al. (2000) suggests that the relative economic and social stability in developed markets provides robust institutional frameworks. Overall, developed markets are characterized by stringent regulatory requirements, mature infrastructure, and capital markets (Han et al., 2013).

The supply chain literature suggests that due to political and economic stability, operating in developed markets decreases supply chain uncertainty (Hitt et al., 2000; Wiedmer et al., 2021). This further reduces ambiguity in resource decisions of firms in that market including demand and supply forecast etc. This leads to a reduced risk for buyer firms when faced with regulatory stress. However, a different stream of supply chain research suggests that the well-defined formal structures and a strong inclination towards adherence to existing regulations within the markets lead to a reluctance to change exacerbating risk (Gilbert 2005; Helpman & Itskhoki, 2010). This highlights the tension in the literature regarding the impact of developed markets.

We propose that institutional rigidities in developed markets create inertia, making it difficult for firms operating in these markets to adapt or innovate (Gilbert 2005). Further, developed markets require significant time and effort to achieve regulatory convergence, a process through which the regulatory requirements across different countries or regions become more aligned due to the gradual adoption of globally accepted technical guidelines or standards

(Kauffmann, 2013; Kell et al., 2005; Rodrik, 2008). For example, in order to do business in European Union, U.S. based firms need to comply with General Data Protection Regulation (GDPR) which affects how they can use, process and store user data. In addition to the GDPR, U.S. based firms are already grappling with various data privacy laws across different states such as the California Consumer Privacy Act (CCPA).

The core idea of supply chain accountability regulations is that buyers are responsible for ensuring that their foreign suppliers comply with certain standards and regulations, making them accountable for the entire supply chain (Parmigiani et al., 2011). This implies that buyers must monitor and manage their suppliers to ensure adherence to legal and ethical practices. Accordingly, when regulatory changes in the buyer's market necessitate supplier adaptation to maintain supply chain accountability, we suggest that institutional rigidities in developed markets pose barriers, requiring significant effort and cost to comply with these requirements. This decreases the ability of the buyers that depend on suppliers in developed markets to be compliant with regulatory requirements.

The inability of the buyer firms to comply with local regulations increases supply chain risk as it leads to legal consequences, financial penalties, and reputational loss. Noncompliance can disrupt the supply chain leading to increased lead times and costs, and strained stakeholder relationships affecting their ability to operate in that market and increasing regulatory stress multifold. Moreover, the cost of compliance is transferred to buyers as they are responsible for ensuring that their supplier adheres to the regulations. This adds to the cost of production leading to a further increase in stress. Buyer firms might also need to focus on proactive planning and contract renegotiations with suppliers to navigate the rigid regulatory environment. Consequently, the mandated policy in the buyer's market triggers a coping response, where

buyers accumulate inventory to buffer against the risk of relying on supplier in developed markets.

Higher dependence on suppliers in developed markets thus magnifies a firm's vulnerability to the regulatory stressor because of enhanced market rigidities in the form of inflexibility and stringencies. To reduce such vulnerabilities, we believe that firms will build up more inventory.

Hypothesis 2: An increase in mandatory supply chain regulatory stress that motivates firms to increase their inventory buffering as a coping strategy will have a stronger effect for firms sourcing from developed markets relative to other global markets.

The second moderating factor that we explore is the buyer firm's dependence on suppliers located in emerging markets. According to MSCI (2024) classification, emerging markets are based in developing nations that have high growth rates, strong economic performance, accessibility by foreign investors, high per capita income, and a reliable regulatory system such as India, Mexico, China, or Brazil (Hitt et al., 2000; Han et al., 2013; Steven & Britto, 2016). Emerging markets tend to be more flexible with lesser regulatory burden and monitoring as compared to developed markets (Aguilar & Gopinath, 2007; Huq et al., 2016; Kusaba et al., 2011).

However, the literature presents mixed evidence about the impact of sourcing from emerging markets. One perspective in the supply chain literature suggests that firms that participate in emerging markets as compared to other markets experience higher levels of risk that can negatively impact their supply chain performance. For instance, Kouvelis and Tang (2012) suggest that emerging markets have inadequate production infrastructure which increases the chances of supply chain disruptions and unexpected delays. Lai (2005) suggests that

emerging market suppliers might be less reliable due to differences in business norms, standards, and culture which may add to the riskiness and regulatory stress. Further, the lack of physical and institutional infrastructure tends to increase transportation costs, uncertainty, and lead times. This pressurizes the firm buying from emerging markets to maintain higher safety stocks as a coping response (Golini & Kalchschmidt, 2011).

In contrast, the literature suggests that sourcing from emerging markets alleviates the risk on a firm's supply chain performance. For instance, Hitt et al. (2000) argues that firms in the emerging markets have less capacity and experience for exploiting partnership to their advantage and exhibit less opportunistic behavior which may decrease regulatory stress. Han et al. (2012) suggest that firms commonly use contract manufacturing as a sourcing strategy in emerging markets allowing buyers to operate their own supply chain and operations (including inventory management) at a steady state. Contract manufacturers may carry buffer inventory on behalf of buyers that would otherwise be on the books of the buyers leading to decreased stress. Shou et al. (2016) suggests that contract ineffectiveness, due limited information transparency and legal enforceability in emerging economies, forces firms to seek social, business, and political ties to overcome "institutionally induced" contract limitations. Such ties lessen the stress of regulatory burdens on firms in emerging economies more than in developed economies, leading firms to maintain a lower inventory.

Further, suppliers based in emerging economies can inexpensively adapt their business processes and meet quality standards because they have access to multiple qualified and motivated suppliers with competitive bargaining (Calvo et al., 2006). For instance, suppliers based in India have access to qualified technical labor at inexpensive rates that can help firms transition to newer requirements economically. Accordingly, regulatory changes in the buyer's

market, especially related to supply chain accountability, do not require significant effort and cost to comply when sourcing from emerging markets as compared to developed markets. The improved or sustained ability of buyers to comply with regulatory changes decreases supply chain risk and regulatory stress. Moreover, the cost of compliance may be transferred to suppliers due to a contract manufacturing sourcing strategy. Buyer firms might also not have to engage in contract renegotiations with suppliers due to relational ties and limited contract enforcement. Consequently, the mandated policy in the buyer's market triggers a lower coping response leading to improved inventory performance.

Such flexibilities may fasten the adaptation process, further decreasing the risk for U.S. based firms engaging with foreign suppliers in emerging countries. So, firms sourcing from emerging markets will buffer less inventory as leveraging the market's flexibility to implement changes efficiently can enhance supply chain agility. This will alleviate the negative impact on inventory leanness. The flexibility in these markets can partially offset the challenges posed by the institutional differences. The adaptability of the market can help in promoting inventory leanness.

Overall, we propose that higher dependence on emerging markets alleviates the regulatory burden requiring firms to maintain lower inventories leading to increased inventory leanness.

Hypothesis 3: An increase in mandatory supply chain regulatory stress that motivates firms to increase their inventory buffering as a coping strategy will have a weaker effect for firms sourcing from emerging markets relative to other global markets.

The third moderating factor that we explore is the buyer firm's dependence on suppliers located in frontier markets. As per MSCI, Frontier markets are characterized by political

instability, inadequate regulations, large currency fluctuations, and insufficient financial reporting such as in Myanmar, Mozambique, Vietnam, or Rwanda (Berger et al., 2011; Marshall et al., 2015). Many countries in these markets lack a formal institutional or regulatory framework. Likewise, these markets tend to have a higher institutional distance relative to the buyer firm due to limited similarity to the developed regulatory institutions leading to a much higher risk.

A limited stream of supply chain and international business research has identified that many firms, including suppliers that participate in frontier markets may have found it challenging to comply with mandatory regulatory requirements. Suppliers based in frontier markets face significant political, economic, and operational risks due to institutional immaturities leading enhance regulatory stress for buyers (Berger et al., 2011; Marshall et al., 2015). For instance, in Kenya, political stability and investor rights might be a significant challenge affecting firm's ability to operate effectively (FINRA, 2015). The variable economic growth rates in such markets make the demand and supply difficult to forecast.

The rigidity due to lack of infrastructure makes it difficult for buyer firms sourcing from frontier markets to be compliant with newer regulations. Further, many buyers use local intermediaries to engage with frontier markets, further increasing the supply chain complexity. The increase in distance among the supply chain partners lead to longer lead times and higher delivery variability (Kusaba, 2011). Likewise, because of the risks that exist in these markets, buyers may need to institute increased monitoring and oversight to ensure that their suppliers comply with the mandatory regulations from the buyer's home country (Berger et al., 2011). For instance, suppliers based in Myanmar might not have access to capacity building in the form of training which can make it difficult for such firms to transition to newer requirements.

Accordingly, firms experience an increase in regulatory stress, and to guard against the risks of sourcing from frontier markets, they exhibit increased coping responses, leading to inventory overstocking.

Further, due to less developed legal and regulatory systems, firms operating in frontier markets will find it difficult to enforce contractual agreements and protect their intellectual property rights. Shou et al. (2016) suggests that due to limited legal enforceability, firms perceive a higher risk in safeguarding their interests. In addition, due to limited business norms and unstable political regimes, the firm's pursuit of social and political ties imposes an extra burden on their operational efficiency. This exacerbates the negative effects of regulatory burden for buyers based in developed markets.

Inflexibilities due to infrastructural limitations may slow the adaptation process of newer regulatory requirements, further increasing the risk for U.S. based firms engaging with foreign suppliers in frontier countries. So, firms sourcing from frontier markets will buffer more inventory as market inflexibility can decrease supply chain agility. This will worsen the negative impact on inventory leanness. Firms might need to hold more inventory due to more delays in compliance and higher need for monitoring, increasing the regulatory risk. The significant differences and rigidity in the institutional environment can lead to substantial challenges in achieving compliance, necessitating higher inventory levels as a risk mitigation strategy. Firms may need to reevaluate their sourcing strategies in these markets to balance compliance and inventory efficiency. This scenario likely intensifies the negative impact on inventory leanness.

Although firms based in frontier markets might offer a lower cost of production, the lack of robust infrastructure increases the logistics and operational risk. When sourcing from frontier markets, the cost of compliance falls on the buyers, which increases their production costs. This

situation often leads them to build up their inventory as a buffer. Under conditions of higher sourcing from frontier markets, there is an increased potential for regulatory burden. Higher dependence on frontier markets increases a firm's vulnerability to the regulatory stressor because of higher market rigidities and distance with respect to the U.S. market. We believe that buyer firms will build more inventory when dealing with frontier market-based suppliers. In contrast, firms might experience lesser regulatory burden when they appraise the stress of regulatory changes under conditions of lower dependence on frontier markets.

Overall, we propose that higher dependence on frontier markets exacerbates the regulatory burden requiring firms to maintain higher inventories leading to decreased inventory leanness.

Hypothesis 4: An increase in mandatory supply chain regulatory stress that motivates firms to increase their inventory buffering as a coping strategy will have a stronger effect for firms sourcing from frontier markets relative to other global markets.

4. Setting and data

We tested our hypotheses in the context of the Food and Drug Administration's (FDA's) Food Safety Modernization Act (FSMA). We analyzed the effect of FSMA mandate on the inventory management decisions of firms in the U.S. food industry. FSMA, which was mandated for firms in 2016, represents an exogenous shock to the industry. We compared changes in the inventory leanness of food firms that were affected by FSMA (treatment group) to firms that were also regulated by the FDA but were not affected by FSMA (control group) using difference-in-difference models (Imbens & Wooldridge, 2009; Scott et al., 2023). In doing so, we adopted an approach like other scholars in operations management (OM) literature. (e.g., Kistler et al., 2021; Scott et al., 2023; Wiedmer et al., 2021).

4.1. Research setting: The Food Safety Modernization Act (FSMA)

FSMA was enacted as Public Law on January 4, 2011, amending Title 21 of the United States Code (FDA, 2023). However, it was not until 2015 that the FDA finalized the regulations under the FSMA, specifying that large firms were required to achieve compliance by 2016, followed by small firms in 2017 and very small firms in 2018.

FSMA aimed to improve food safety by shifting the focus to preventive controls via enhanced supply chain accountability. The legislation is divided into four titles, covering various aspects of food safety (FDA, 2022). We focus on Section 103 under Title 1, i.e., the rule for Preventive Controls for Human Food for facilities handling or processing food (FDA, 2023). FSMA is considered one of the major federal legislations addressing food safety in more than 80 years. It significantly impacted the businesses which processed, held, packed, or distributed food. It represents a fundamental change from the existing food safety management system (FSMS).

Additionally, the legislation provides the FDA with enhanced authority. This includes the authority to mandate product recalls, conduct inspections and enforce compliance, and set a required frequency for these inspections. The FDA also gains complete access to a firm's safety records and the ability to conduct sample testing through accredited laboratories. This expansion of powers significantly strengthens the FDA's role in ensuring food safety.

For firms, FSMA regulation is an exogenous shock out of their direct control. However, the FSMA mandated supply chain accountability measures affected one important factor that is both endogenous and influential to a firm's FSMA compliance, that is inventory management. The need for enhanced traceability, rigorous supplier verification, and adherence to stringent safety standards leads to these firms being more cautious about inventory management. FSMA

indirectly incentivizes firms to alter their inventory management strategy towards less leanness to buffer against the risk of defaulting suppliers and FDA mandated recalls and inspections.

4.2.Data sources and collection process

For our main analysis, we compiled secondary data from the Compustat annual fundamentals database for North American firms, Compustat financial ratios suite, Port Import and Export Reporting Service (PIERS), economic uncertainty database (Baker et al., 2016), firm level risk database (Hassan et al., 2019) and FDA recalls database (FDA, 2023). We collected financial data at the firm level for publicly traded firms from the Compustat Fundamental Annual and Compustat financial ratio databases. They were matched together based on a firm level unique identifier, i.e., GVKey. Only active firms were retained. We collected firm-year level trade activity data from PIERS database, which is widely adopted in supply chain research studies including trade partnership participation (Tong et al., 2022), analysis of containerized ports (Steven & Corsi, 2012), intermodal transportation (Hayuth, 1994), import services (Anderson et al., 2009), containerization of bulk trades (Kawasaki & Matsuda, 2015) and import routes (Martinez et al., 2016) (refer figure 2).

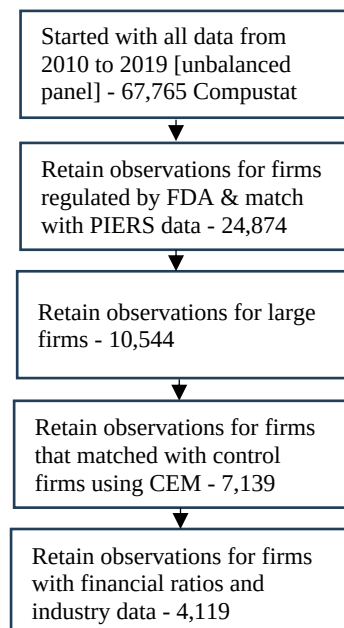


Figure 2. Flowchart of data collection steps

We also collected FDA recalls data for all firms and compiled it at a firm-year level. Then, we matched it to Compustat using the firm names as the identifiers.¹ Further, we collected a firm specific political risk index (Hassan et al., 2019) and policy uncertainty index (Baker et al., 2016), data from the firm level risk and Economic Uncertainty databases respectively. We matched the firm specific policy risk with Compustat annual fundamentals database using the GVKeys. The policy uncertainty index was matched based on the year. We collected data from 2010 to 2019. We started with the year 2010 because FSMA was signed into law in 2011 but did not get implemented until 2016. Firstly, we wanted to cover about six years prior to the implementation of the preventive control rule of FSMA and secondly, we wanted to observe any change in trade behavior due to the announcement of a policy that has not been implemented yet. To minimize potential confounds caused by the pandemic such as economic uncertainty, our data collection and analysis ended in the year 2019.

Table 1a summarizes the data collection process which is based on the FSMA implementation that started in 2016. Since FSMA is a U.S. based domestic regulation, we focused on publicly listed firms based in the United States. Our data collection started with 67,765 firm-year observations, an unbalanced panel of firms ranging from 5213 firms in 2010 to 8499 firms in 2019 (refer to Table 1a and Table 1b). We needed a few years prior and a few years after FSMA implementation to examine the effect.

Given our focus on firms regulated by the FDA, we removed firms with no products or services in the following industries including food, drugs, medical devices, biologics, cosmetics,

¹ we used fuzzy-logic matching algorithm to match names.

tobacco, and veterinary based on three-digit NAICS code. For example, we focused on NAICS codes including 311 – Food Manufacturing, 312 – Beverage and Tobacco Product Manufacturing etc. We corroborated the NAICS code identification of industries by using the FDA data dashboard that lists information related to all FDA regulated firms that were inspected, had a compliance action, recalled products, had imported products between 2009 to 2023. We believe that this comprehensive FDA database provides an exhaustive list of firms that were regulated by FDA between 2010 to 2019.

We matched Compustat firm names with the FDA regulated firm names using the fuzzy matching (probabilistic data matching) in Python (Deep, 2020). We were able to

Table 1a. Data collection steps based on firm-year observations

	Data screening steps	Number of observation reduction	Number of firm-year observations	Database used
1	Started with all data from 2010 to 2019 [unbalanced panel]		67,765	Compustat
2	Remove observations for firms not regulated by FDA	42,891	24,874	Compustat
3	Remove observations for small and very small firms	14,330	10,544	Compustat
4	Remove observations for firms without financial ratios and industry data	6,425	4,119	Compustat

Table 1b. Data collection steps based on an unbalanced panel of firms by year

Year	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	Total
------	------	------	------	------	------	------	------	------	------	------	-------

Number of firms in each year in step1	5213	5529	6060	6407	6607	6872	7186	7465	7927	8499	67,765
Number of firms regulated by FDA	1952	2060	2257	2390	2451	2527	2633	2715	2852	3037	24,874
Number of large firms regulated by FDA	947	965	989	1024	1041	1063	1082	1112	1144	1177	10,544
Number of large firms with financial ratios and industry data	364	377	384	391	403	418	424	431	457	470	4,119
FSMA	125	128	134	132	135	139	138	142	163	165	1,401
Non-FSMA	239	249	250	259	268	279	286	289	294	305	2,718

identify other NAICS codes that were regulated by FDA but were missed in our initial classification based on the three-digit NAICS code. For example, NAICS code 481 – airport transportation, 482 – rail transportation etc., which hold or distribute food that is intended for consumption in the U.S. At this stage we were left with 24,874 firm-year observations. Since our focus is on large firms (more than 500 full time employees) only, we removed small (less than 500 full time employees) and very small sized firms (less than 500 full time employees and less than \$1 million sales) leaving us with 10,544 firm-year observations (refer figure 2). We use the FSMA definition to classify firms (Grover et al., 2016). We further removed the firm-year

observations that had no financial data. Our final panel dataset uses 4,119 firm-year (unit of analysis) observations from 2010 to 2019 with firms from 20 different industries (as defined by the three-digit NAICS codes) which are regulated by FDA.

Regardless, using standard practice (Darby et al., 2020), we tested for bias comparing our final data set with the data at step 3 (Table 1a). We created a dummy variable using data at step 3 that was coded as 1 if the observation was included in the final data and zero otherwise (Kim & Zhu, 2018). We then estimated a probit model with the dummy variable as the dependent and our firm-level variables of interest as the independent variable. The results indicated that the firms included in the sample tend to have lower sales and numbers of recalls but a higher perception of policy risk. No significant difference was found in the advertisement to revenue ratio.

4.3. Variables and variable construction

4.3.1. Dependent variable

Inventory leanness. Inventory leanness is defined as a firm's relative level of inventory as compared to its peers in the firm's industry. We used the Empirical Leanness Indicator (ELI) to measure inventory leanness of a firm. ELI is a better measure of inventory leanness than the traditional inventory turnover as it considers "economies of scale in inventory management and industry-specific relationships between sales and inventory" (Eroglu et al., 2015). Instead of simply comparing a firm's inventory to its sales, ELI compares inventories to a benchmark level depending on the firm's size (sales) and industry. We used the sales and inventory data for firm i in a specific industry j at a given time t (Eroglu & Hofer, 2011; Bendig et al., 2017). We determine the ELI by implementing a regression of log-transformed sales against the log-transformed total inventory. This model is estimated for each industry categorized by a three-digit NAICS code and for each year as follows:

$$\ln(\text{Inventory}_{jt}) = \alpha_{jt} + \beta_1 \times \ln(\text{Sales}_{jt}) + \varepsilon_{jt} \quad (1)$$

We calculated the inventory leanness for every firm in each year by keeping the studentized residuals ε_{jt} and multiplying them by -1. A positive ELI value suggests a greater level of inventory leanness compared to the industry norm and firm size. A higher inventory leanness suggests lower inventory.

4.3.2. Independent variables

Government policy mandate (FSMA). Our main independent variable is the impact of FSMA on firms, i.e., the government policy mandate. We coded the treatment variable, FSMA_{it} as an indicator variable equal to 1 if firm i was impacted by FSMA after its implementation in 2016 (large firms) and zero for firms that were in operation before FSMA's implementation and zero in all years for firms that were not affected by FSMA from 2010 through 2019 (i.e., control group) (Scott et al., 2023; Tan & Netessine, 2020). For example, large firms that were impacted by FSMA were coded 1 in the year 2016 and after. They were coded zero before 2016 along with firms that were not impacted by FSMA until that time.

Firm sourcing reliance on different markets. We measured a firm's sourcing reliance on different markets as shipment volume of imports for a firm i in year t from that market by its total shipment volume of imports in year t (as in equation 2). A firm's sourcing reliance on a particular market suggests how much a company relies on sourcing activities from those markets as a business model (Zhang et al., 2022). We measured the import shipment volume by aggregating the weight of shipments measured in kilo tons (kt.) imported per firm i in a year t (Wiedmer et al., 2021). The volume of import shipments also reflects a firm's capacity to procure the necessary quantity of raw material or other products from a certain market, which in turn influences their ability to manufacture and sell products. It can help firms take advantage of

lower costs, higher quality, or exclusive products. It holds significant relevance to the operational performance of firms.

For example, a firm's higher reliance on a developed market would suggest that a firm depends greatly on foreign suppliers from developed markets for its raw materials or goods. This could be because of inexpensive goods abroad or nonavailability of raw materials domestically. However, it could also indicate a firm's risk exposure to policies, regulations, and foreign relations between countries. We differentiate markets as developed, emerging, and frontier markets.

$$\text{Firm sourcing reliance on a market}_{it} = \frac{\text{Shipment volume of imports from that market (kt)}_{it}}{\text{Total shipment volume of imports from all markets}_{it}} \quad (2)$$

4.3.3. Control variables

Shipment per year. We controlled shipments per year, measured as the natural log of frequency of imports that a buyer firm receives per year. This helps us to account for annual variations in the import shipment volume of firms, i.e., helps control changes in shipment volume due to more (fewer) imports per year (Wiedmer et al., 2021). By controlling for this variable, it allows the model to adjust for differences in the number of import shipments per year.

Firm specific policy risk. Following Darby et al. (2020), we utilized the PRisk Index by Hassan et al. (2019) to assess firm-specific policy risk. We transformed this quarterly data into annual data by averaging the four quarters' values and then applying a natural logarithm (as in Darby et al., 2020). It is derived from the Street Events database (Thomson Reuters) and is based on quarterly earnings conference calls, which are platforms for discussing potential firm risks (Hollander et al., 2010). Hassan et al. (2020) used computational linguistics techniques to analyze these calls, focusing on policy risk (Schutze et al., 2008; Song & Wu, 2009). This method of measuring policy risk helps control for firms' perceptions of government policy risks

which might be varied (Hassan et al., 2019). This index has been used in various studies (Bloom et al., 2018; Darby et al., 2020; Ellis et al., 2011; Hankins et al., 2020; Nagar et al., 2018; Pagell & Krause, 2004), offering controls beyond time and industry fixed effects (Hassan et al., 2019).

Policy uncertainty (MPU). We utilized the Monetary Policy Uncertainty (MPU) Index by Baker-Bloom-Davis (2016), derived from over 2000 U.S. newspapers in the “Access World News database.” This index, broader than the Economic Policy Uncertainty (EPU) Index, comprehensively measures policy uncertainty affecting trade, considering monetary policy's significant role. It helps control for the uncertainty in the policy environment. The index is measured by identifying articles with specific terms related to monetary policy, economics, politics, and uncertainty, standardizing their frequency to maintain an average value of 100 from January 1985 to December 2010. The data, compiled annually, captures nuances in policy decision-making, implementation, and economic impacts (Baker et al., 2016). Validated in various studies (Bonaime et al., 2018; Brogaard & Detzel, 2015), the MPU Index offers a comprehensive control of policy uncertainty, extending beyond the scope of election-focused proxies (Julio & Yook, 2012).

Industry concentration (HHI). Market structure affects firm conduct, with firms in concentrated industries having more market power and a different perception of policy risk and trade patterns. We control for this by using the Herfindahl-Hirschman Index (HHI) to measure industry concentration, calculated by summing the squared market shares (Wiengarten et al., 2017).

$$HHI = \sum (s_{it})^2 \quad (3)$$

where s is the market share of firm i in the year t in the same three-digit NAICS code. A low HHI suggests a highly competitive and fragmented industry with many small firms and a high HHI suggests a monopolistic industry where a few firms control the whole market.

Advertisement intensity ($Adv\ intensity_{it}$). This ratio is a financial metric that helps measure how much of a firm's revenue is invested in marketing and advertising to attract new and retain existing customers. It is the ratio of a firm i 's expenditures on advertising to its total revenue in year t .

$$Adv\ intensity_{it} = \frac{Total\ advertising\ expenditure_{it}}{Total\ revenue_{it}} \quad (4)$$

This ratio assesses the efficiency of marketing and advertising investments in generating revenue, controlling for factors like market expansion and penetration. Higher advertising spend may indicate efforts to enhance market presence, influencing imports and potentially bolster resilience to policy changes like FSMA by shaping firm reputation and risk perception (Beattie et al., 2021). It also affects product demand, trade, and inventory levels, as seen in successful campaigns increasing product demand and imports. This approach, accounting for strategic variations among firms, has been utilized in studies by Beattie et al. (2021), Fainmesser et al. (2021), Lu et al. (2022), and Zantedeschi et al. (2017).

Other controls. We controlled for global sourcing behavior of a firm in a particular year by adding a dummy variable importer which is equal to 1 if a focal firm i sourced globally in year t . We controlled for the size of the firms by using the natural log of sales of the firm i in year t in USD. We controlled for the size of the firms because previous scholars identified it as one of the major sources of heterogeneity that would confound firm level inventory management decisions (Darby et al., 2020; Fan et al., 2022). We also controlled for profitability of the firm using return on equity (RoE) which accounts for the firm's ability to generate profits from its

equity, which might confound the inventory management practices and decisions (Azadegan et al., 2013). Further we control for quality performance of firms as measured by the number of recalls for firm i in year t . This helps control for the effectiveness of a firm's quality control processes and effects of risk management practices on inventory management (Steven & Britto, 2016). Lastly, we control for financial health (financial stability) of the firms using the Altman Z-scores. Firms with lower scores may be at greater risk of bankruptcy and accordingly might manage their inventories more conservatively. Conversely, firms with higher scores might be more willing to invest in larger inventories (Chen et al., 2024).

Table 2. Variable definitions and summary statistics

Variable	Description	Mean	SD
Inventory leanness	Regress a firm's log-transformed sales against its log-transformed inventory, adjusted for industry and size (Eroglu and Hofer, 2011)	-0.152	1.142
FSMA	For firms affected by FSMA: 1 for the years 2016 through 2019 and 0 otherwise; 0 for all years for other firms	0.147	0.355
Reliance on developed markets*	% of imports from developed markets	0.435	0.390
Reliance on emerging markets*	% of imports from emerging markets	0.534	0.385
Reliance on frontier markets*	% of imports from frontier markets	0.019	0.094
Shipment per year	Natural log of frequency of shipments that a buyer firm receives per year	1.121	2.250
Firm specific policy risk	Natural log of yearly frequency of policy-related terms in earnings conference call transcripts, normalized by length of the call (Hassan et al., 2019)	3.882	1.700
Policy uncertainty	Count newspaper articles that mention terms related to monetary policy, economy, and uncertainty, then normalizing by the total	4.210	0.350

	number of articles each month, and finally standardizing it to maintain an average value of 100 and further averaging these monthly values and taking their natural log (Baker et al., 2016)		
Industry concentration	Herfindahl-Hirschman Index (HHI): square and then sum the market share of each firm competing in the market within 3-digit NAICS codes	0.110	0.209
Importer	Dummy variable equal 1 if a firm source globally	0.284	0.451
Firm size	Natural log of sales	7.241	1.506
Firm profitability	Measured as return on equity (RoE) as net-income by average shareholders' equity	0.081	0.258
Firm adv intensity	Ratio of total advertising expenditure to total revenue	0.008	0.025
Firm quality performance	Number of recalls	0.001	0.031
Firm financial health	Altman Z-score calculated using different financial ratios	1.600	1.415

*For firms in engaged in global sourcing

5. Methods

To assess the impact of the FSMA policy on the inventory leanness of affected firms, it is necessary to compare their performance before and after the policy implementation. However, this is challenging because counterfactual outcomes are not directly observable (Fan et al., 2022). To address this, we employed a treatment-control matching approach to construct a natural experiment that accounts for unobservable outcomes (Heckman et al., 1998). The strategy helps estimate the changes in the performance difference between the treated firms and the control after the implementation of exogenous shock, i.e., FSMA (Barrios et al., 2020). This helps in distinguishing the true effects of the FSMA implementation from factors that may affect trade behavior of both treated and control firms over time (e.g., demand shocks, technological advances, competition, infrastructure etc.). It is a method for assessing the influence of public

policies, as highlighted by studies such as Card and Krueger (2000). It has also been effectively applied in investigating matters related to operations management, as evidenced by the works of Wiedmer et al. (2021), and Wang et al. (2024).

Out of a sample of 570 firms, we identified 186 treatment firms and 384 control firms for our analysis. We also utilized the FDA's public database on inspections, recalls, compliance actions, and import refusals to validate the accuracy of our databases in distinguishing between FSMA and non-FSMA firms. The FDA's classification of these firms as either dealing with food or other product categories confirmed the precision of our identification.

Large firms had to comply with FSMA's preventative control rule for human food by September 2016. While FSMA was implemented in the third quarter of 2016, we considered the entirety of the year as the FSMA implementation period. This approach was taken because operational changes, like preventive control measures arising from FSMA, require time to be effectively incorporated. And FSMA's final rules were announced in 2015 giving firms early 2016 to comply with the regulations. Furthermore, the impact on a firm's trading operations may not be immediately evident, given that firms, their suppliers, and channel partners typically maintain inventory stocks that were procured prior to the FSMA implementation (Wiedmer et al., 2021).

5.1.Coarsen exact matching (CEM)

Differences between treatment and control firms could potentially confound the effects of FSMA. For instance, if a treatment firm's return on equity (profitability) is significantly lower than that of a control firm, any further negative impact of FSMA on the treatment firm might be attributed to its limited resources for adapting to policy changes, rather than its involvement in food products. To address this, we employed Coarsened Exact Matching (CEM) (Iacus et al.,

2012), a widely used approach that ensures high similarity between treatment and control firms and improves the balance of covariates. CEM is favored for its ability to create a robust balance in matched covariates, eliminating the iterative process of matching and balance checking that is typical in other matching methods, such as propensity score and nearest-neighbor matching (Iacus et al., 2012).

The CEM algorithm creates bins based on the matched covariates and places treated and control observations into the same bins for all matching covariates. We matched the following covariates – firm size, profitability, region, quality performance and adv intensity – that are commonly used in similar policy impact analysis studies (Cantor et al., 2017; Li et al., 2022; Miller et al., 2020). Firm size, profitability, quality performance, and adv intensity are measured the same as above. Region is the state in which the firm is located. We use exact matching on all covariates. Firms that dealt with food were in the treatment group for the matched sample and the matched firms that were regulated by FDA but not affected by FSMA in the control group. Table 3 presents the summary statistics for the sample after matching.

After the matching process, there are no significant statistical differences between the treated and untreated groups regarding the variables we matched, as the t-statistics in Table 3 show. We calculated these statistics by performing a regression of the variables we matched, like firm size, against the treatment variable, using the weights from the Coarsened Exact Matching. The t-statistics help us understand if the variables we matched can predict which firms received treatment. According to Scott et al. (2023), who used a similar approach, none of our matched variables significantly predict treatment, with all p-values above 0.10. This suggests that the treated and control groups are quite similar in terms of the variables we matched.

Table 3. Covariate balance in the matched sample

	FSMA firms		Non-FSMA firms		
Variable	Mean	SD	Mean	SD	t-stat
Reliance on developed markets	0.423	0.396	0.444	0.384	
Reliance on emerging markets	0.540	0.392	0.529	0.379	
Reliance on frontier markets	0.026	0.122	0.013	0.064	
Shipment per year	1.063	2.213	0.994	2.170	
Firm specific policy risk	3.810	1.728	3.717	1.951	
Policy uncertainty	4.214	0.349	4.210	0.349	
Industry concentration	0.216	0.317	0.081	0.168	
Importer	0.293	0.455	0.277	0.448	
Firm size	7.949	1.788	7.613	1.719	0.26
Firm profitability	0.096	0.284	0.086	0.288	1.34
Firm adv intensity	0.008	0.023	0.008	0.027	-1.30
Firm quality performance	0.002	0.044	0.001	0.029	-0.01
Firm financial health	2.029	1.503	1.505	1.316	
Firm region	20.801	18.494	18.920	18.375	-0.05

Note: The last column reports the t-statistics from CEM weighted OLS regression of the variable on the FSMA affected firm dummy. t statistics significance *** p<0.01, ** p<0.05, * p<0.1

5.2.Parallel trend assumption checks for DID analysis

For the Difference-in-Differences (DID) analysis to accurately capture any effects of the treatment, it assumes parallel trends. This means that, before the impact of the FSMA is felt, there should be a similarity in the patterns of the dependent variables, such as inventory leanness, between the treatment and control groups. To verify this assumption and ensure the validity of

our DID analysis, we implemented an analysis based on a model that specifically examines the relative time trends between these groups. (Angrist & Pischke., 2008; Song et al., 2020):

$$Inventory\ leanness_{it} = \delta_0 + \sum_{t=2010}^{2019} k_t Treat_i X D_t + X_{it}\delta + \tau_t + \varepsilon_{it} \quad (5)$$

Where D_t are dummy variables that indicate the study period of our analysis, i.e., years from 2010 to 2019. $Treat_i$ is the treatment indicator equal 1 if firm is affected by FSMA. We omitted the interaction between the treatment indicator and the dummy variable for 2016 (i.e., $Treat_i X D_{2016}$) because it serves as the reference group (Dhanorkar, 2019; Fan et al., 2022). The other variables remain the same as in our main analysis in equations 6 and 7 (described below). The analysis output is presented in Table 6 and plotted the estimated coefficients and confidence intervals of k_t in figure 3.

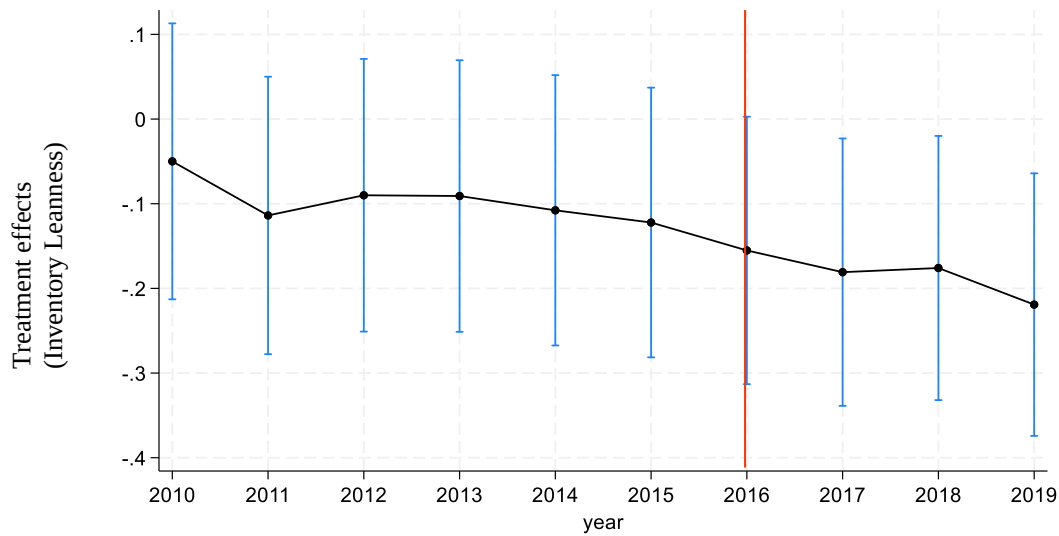


Figure 3. The effects across periods of FSMA

Table 6 shows that the coefficients for the years 2011 to 2015 are non-significant which suggests that the differences between the treatment and control groups in the 5 -year pretreatment period are not significant (Fan et al., 2022). This result is illustrated in figure 3, which shows that the limits of the 95% confidence intervals till 2015 include zero (Fan et al., 2022). However, the

coefficients for the year 2017 ($b = -0.115$, $p < 0.10$), 2018 ($b = -0.153$, $p < 0.05$) and 2019 ($b = -0.137$, $p < 0.05$) are significantly negative in the inventory leanness model. Figure 3 also illustrates that the same (Fan et al., 2022). Hence, the results support the parallel-trends assumption of our analysis.

5.3. Model specification

We created a 10-year panel dataset from 2010 to 2019. This time window was constructed based on the year FSMA was implemented for large firms (i.e., 2016). We used 6 years before the FSMA implementation at the benchmark years and then subsequently examined the impacts of the FSMA in 2017, 2018, and 2019. Our event window cutoff was 2019 because the COVID-19 global pandemic affected the global supply chains, and this might confound the firms' behavior.

To test our hypotheses, all our models use ordinary least square (OLS) regression. We test for normality and homoscedasticity along with linearity through histograms and residuals vs. fitted plot (Appendix A). The spread of residuals seems to be constant across the bounds of fitted values and the residuals seem to be centered around zero. We estimate a series of models to evaluate the effect of FSMA on inventory leanness of firms. In the first set of models, we compared inventory leanness between the control and treatment group before and after FSMA implementation to determine the impact of the policy. We use the following difference-in-difference specification:

$$\text{Inventory leanness}_{it} = \delta_0 + \delta_1 \text{FSMA}_{it} + \mathbf{X}_{it}\boldsymbol{\delta} + \tau_t + \varepsilon_{it} \quad (6)$$

where ε_{it} is the error term and $\mathbf{X}$ is a vector of control variables and τ_t are year dummy variables². A negative and significant coefficient of FSMA_{it} , i.e., δ_1 would suggest a decrease in

² We don't include an "after" dummy since it would be colinear with the time effects.

inventory leanness of firms affected by FSMA (treatment group) as compared to firms not affected by FSMA (control group) after FSMA implementation. And a positive and significant coefficient of $FSMA_{it}$ would suggest otherwise.

For H2 to H4, we examine whether firms with higher dependence for global sourcing on developed, emerging and frontier markets suffered from the FSMA. To test the moderating effect of global sourcing from three different regions, we estimated three models. The three models test the moderating effect of global sourcing of treatment group firms with respect to control group firms after FSMA was affected for firms that engage in global sourcing (Fan et al., 2022). We use the following specification.

$$Inventory\ leanness_{it} = \delta_0 + \delta_1 FSMA_{it} X Moderator_{it} X Importer_{it} + X_{it}\delta + \tau_t + \varepsilon_{it} \quad (7)$$

Where $Moderator_{it}$ refers to the three moderators of firm i in the year t for testing hypothesis H2, H3, and H4 respectively. Specifically, we used moderators of Reliance on developed markets, reliance on emerging markets and reliance on frontier markets to create interaction terms (Shen et al., 2023). $Importer_{it}$ refers to the dummy variable which is equal to 1 if a firm engaged in global sourcing. The other variables of $\delta_1 FSMA_{it}$, X_{it} , τ_t , and ε_{it} are defined the same as those in Equation 6.

6. Results

6.1. Analysis of results

Table 2 summarizes the descriptive statistics for our dependent, independent, and control variables. Table 4 summarizes the correlation matrix for our main independent variables, while Table 5 presents the results for examining H1, H2, H3 and H4. Examination of the variance inflation factors (VIFs) indicated that the VIF values are significantly below the recommended

threshold of 10 (Cohen et al., 2003, p. 423-424), indicating that there is no significant issue of multicollinearity³.

With regards to H1, we find a significant decrease in the inventory leanness ($b = -0.170$, $p < 0.01$) post implementation of FSMA for affected firms with respect to the control group by considering the coefficient of $FSMA_{it}$. It suggests that the implementation of FSMA (after 2016 for affected firms) is associated with a decrease in inventory leanness compared to the control group. This implies that after FSMA was implemented, affected firms, on average, tend to manage their inventory less efficiently relative to the control group. This result presents strong support for hypothesis H1.

Furthermore, we tested the moderating effects of firms' reliance on global sourcing from different markets and its effects on the impact of FSMA policy on trade behavior of firms. With regards to inventory leanness, we find that firm's reliance on developed markets has a significant negative moderation affect ($b = -0.245$, $p < 0.01$), i.e., firms that global source and have a higher reliance on developed market experience a larger decrease in inventory leanness as compared to firms with lower reliance (Table 5). This suggests strong support for hypothesis H2. We also tested the firm's reliance on emerging markets, and we found a significant positive moderation affect ($b = 0.157$, $p < 0.10$), i.e., firms that global source and have a higher reliance on emerging market experience a smaller decrease in inventory leanness as compared to firms with lower reliance (Table 5). This suggests that hypothesis H3 is supported. With regards to frontier markets, we found no significant outcome suggesting that we do not have evidence to support hypothesis H4.

6.2. Robustness checks

³ The mean VIF of main variables is 1.34.

We performed multiple robustness checks to further validate our findings and minimize potential alternative explanations. A key principle in Difference-in-Differences (DID) analysis is the assumption of parallel trends. This assumption posits that, in the absence of intervention (in this case, FSMA policy), the outcomes for both the treated and control groups would have been similar. Essentially, the control group serves as a stand-in to estimate what the treated group's inventory management practices would have looked like had FSMA not been introduced. We tested the validity of this assumption in two ways. First, we employed another test to identify any outcome differences before the intervention, or the external shock, occurred, i.e., the slope test (Kube et al., 2019). We found that the difference in slopes from 2010 to 2016 between the control and treated groups is not significant (p value 0.121). And the difference in slopes from 2016 to 2019 between the control and treated groups is statistically different (p value 0.054). Second, we performed a falsification test by making the years 2013 and 2014 the treatment and removing data after 2016 (as in Scott et al., 2023). Using the DID model on this data set, the effect of FSMA is insignificant for the matched sample (p = 0.212 for 2013 and p = 0.116 for 2014).

Other robustness checks include changing the dependent variable to another measure, i.e., inventory (Darby et al., 2022) and inventory days (Fan et al., 2022) instead of inventory leanness. Inventory was calculated using inventory turnover which focuses on how often a firm can sell its inventory. It has been extensively used in literature to measure operational slack (Darby et al., 2020; Weingarten et al., 2017). As in Darby et al. (2020), we calculated inventory using Compustat financial annual database as:

$$Inventory_{it} = (-1) \times \ln \left(\frac{COGS_{it}}{Average\ Inventory_{it}} \right) \quad (8)$$

where the Cost of Goods Sold (COGS) for firm i in year t is divided by the average inventory of the same firm i during the end of fiscal years $t-1$ and t . This average inventory acts as a balancing figure between two fiscal years. We calculated inventory days as in Weingarten et al., 2017 which helps understand the firm's days of supply in a year.

$$Inventory\ days_{it} = 365 \times \left(\frac{Average\ Inventory_{it}}{COGS_{it}} \right) \quad (9)$$

The results suggest that after FSMA implementation, the inventory for food firms increased compared to the control group ($b = 0.142$, $p < 0.01$) and inventory days for such firms increased by 98 days when the mean inventory days is approx. 216 days ($b = 98.12$, $p < 0.10$). This corroborates our findings. Further, we find that firms that have a higher reliance on developed market experience a higher increase in inventory as compared to firms with lower reliance ($b = 0.177$, $p < 0.01$) but sourcing from emerging and frontier market were not significant.

7. Discussion

The purpose of this study is to examine how the introduction of supply chain accountability regulations impacts a firm's inventory performance. In addition, this research examines the moderating role of the firm's dependence on global sourcing in this relationship. This study is important because it helps to elucidate the complex dynamics between regulatory stress and inventory performance (Pagell et al., 2018). Although supply accountability regulations aim to promote transparency and consumer safety, they may inadvertently impose stress and operational challenges for firms and their supply chains (Scott et al., 2021). This study makes a strong contribution to the literature by integrating stress and coping theory to the study on how an important public policy impacts supply chain management performance, offering a new perspective on how firms behave under regulatory stress.

Table 4. Correlation matrix for main independent variables

	1	2	3	4	5	6	7	8	9	10	11	12	13
1 – FSMA	1.00												
2 – Shipment per year	0.03***	1.00											
3 – Firm specific policy risk	0.06***	0.06***	1.00										
4 – Policy uncertainty	-	-0.01	0.01	1.00									
5 – Industry concentration	0.06***	0.12***	-0.02**	0.03***	1.00								
6 – Firm size	0.06***	0.12***	0.27***	-0.01	0.18***	1.00							
7 – Firm profitability	-0.02**	0.02**	0.02	-0.01	0.02*	0.12***	1.00						
8 – Firm adv intensity	-0.01	-0.01	-	-0.01	0.04***	-	0.03**	1.00					
9 – Firm quality performance	0.01	-0.01	0.01	-0.02**	0.03***	0.01	-0.01	0.01	1.00				
10 – Firm financial health	0.06***	0.09***	-	0.01	0.19***	0.18***	0.04***	-0.01	0.01	1.00			
11 – Reliance on developed market	0.04***	0.51***	0.06***	-0.01*	-0.02*	0.05***	0.01	-0.01	0.01	0.06***	1.00		
12- Reliance on emerging market	0.02**	0.68***	0.05***	0.01	0.01	0.03***	0.03***	-0.01	-0.01	0.07***	0.12***	1.00	
13 – Reliance on frontier market	0.01	0.18***	0.01	-0.01	-0.01	0.02**	-0.01	0.01	-0.01	0.01	0.02	0.08***	1.00

Table 5. Effect of FSMA on inventory leanness using difference-in-differences

VARIABLES	(1) DV= Inventory leanness	(2) DV= Inventory leanness	(3) DV= Inventory leanness	(4) DV= Inventory leanness	(5) DV= Inventory leanness
FSMA	-0.170*** (0.036)	-0.131*** (0.039)	-0.199*** (0.039)	-0.170*** (0.036)	-0.159*** (0.042)
FSMA X Developed markets X Importer	-	-0.245*** (0.086)	-	-	-0.238*** (0.086)
FSMA X Emerging markets X Importer	-	-	0.157* (0.083)	-	0.147* (0.083)
FSMA X Frontier markets X Importer	-	-	-	-0.011 (0.355)	-0.052 (0.355)
Shipment per year	-0.013 (0.015)	-0.013 (0.015)	-0.014 (0.015)	-0.013 (0.015)	-0.014 (0.015)
Firm specific policy risk	-0.008 (0.008)	-0.007 (0.008)	-0.008 (0.008)	-0.008 (0.008)	-0.008 (0.008)
Policy uncertainty	1.065*** (0.401)	1.040*** (0.401)	1.056*** (0.401)	1.065*** (0.401)	1.032** (0.401)
Industry concentration	-0.985*** (0.314)	-0.988*** (0.314)	-1.017*** (0.315)	-0.985*** (0.314)	-1.017*** (0.314)
Importer	-0.063 (0.214)	-0.089 (0.214)	-0.067 (0.214)	-0.063 (0.214)	-0.082 (0.214)
Firm size	-0.230*** (0.031)	-0.225*** (0.031)	-0.229*** (0.031)	-0.230*** (0.031)	-0.223*** (0.031)
Firm profitability	0.059 (0.044)	0.059 (0.044)	0.058 (0.044)	0.059 (0.044)	0.059 (0.044)
Firm adv intensity	-1.145 (1.866)	-0.944 (1.866)	-1.156 (1.866)	-1.144 (1.866)	-0.955 (1.866)
Firm quality performance	-0.181 (0.295)	-0.158 (0.294)	-0.198 (0.295)	-0.181 (0.295)	-0.174 (0.295)
Firm financial health	0.109*** (0.021)	0.108*** (0.021)	0.109*** (0.021)	0.109*** (0.021)	0.109*** (0.021)
Reliance on developed market X Importer	0.055 (0.214)	0.124 (0.215)	0.065 (0.214)	0.055 (0.214)	0.131 (0.215)
Reliance on emerging market X Importer	0.063 (0.213)	0.083 (0.213)	0.044 (0.214)	0.062 (0.213)	0.064 (0.214)
Reliance on frontier market X Importer	0.309 (0.276)	0.317 (0.276)	0.320 (0.276)	0.313 (0.313)	0.349 (0.313)
Year fixed effects	yes	yes	yes	yes	yes
Firm fixed effects	yes	yes	yes	yes	yes
Constant	-3.384* (1.784)	-3.319* (1.783)	-3.345* (1.784)	-3.384* (1.785)	-3.286* (1.783)
Observations	4,119	4,119	4,119	4,119	4,119

Number of firms	570	570	570	570	570
R-squared	0.034	0.037	0.035	0.034	0.037

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 6. Results of DID analysis on years for parallel trend assumption

VARIABLES	INVENTORY LEANNESS
Year 2010*Treated	0.192*** (0.069)
Year 2011*Treated	0.076 (0.068)
Year 2012*Treated	0.081 (0.067)
Year 2013*Treated	0.021 (0.067)
Year 2014*Treated	0.002 (0.066)
Year 2015*Treated	-0.020 (0.065)
Year 2017*Treated	-0.115* (0.065)
Year 2018*Treated	-0.153**(0.064)
Year 2019*Treated	-0.137**(0.064)
Treated	0.167*(0.100)
Shipment per year	0.002 (0.012)
Firm specific risk	-0.011 (0.006)
Industry concentration	-0.280** (0.126)
Importer	0.068 (0.184)
Firm size	-0.142*** (0.018)
Firm profitability	0.001 (0.015)
Firm add intensity	-0.338 (1.056)
Firm quality performance	0.017 (0.092)
Firm financial health	0.110*** (0.145)
Reliance on developed market X Importer	-0.077 (0.185)
Reliance on emerging market X Importer	-0.056 (0.184)
Reliance on frontier market X Importer	0.213 (0.251)
Year fixed effects	yes
Firm fixed effects	yes
Constant	0.618*** (0.137)
Observations	5,683
Number of firms	722
R-squared	0.051

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

This research also offers practical managerial and public policy insights into how firms can navigate challenges of regulatory stress, especially when firms are dependent on foreign markets for global sourcing.

Using the stress and coping theory, our analysis demonstrates that supply chain accountability regulations impact inventory performance of focal firms. Further, when a firm is more dependent on sourcing from some foreign market segments, this stress might further increase, leading to further deterioration of inventory performance (Shi et al., 2017).

The first hypothesis suggests that a mandatory regulatory stressor will incentivize firms to increase inventory buffering as a coping strategy. In doing so, firms decrease their inventory leanness in response to the mandatory regulations. We find strong support for our hypothesis suggesting that firms that are affected by FSMA experience a significant decrease in inventory leanness performance after the implementation of the FSMA regulation. The average inventory leanness of firms under FSMA regulation prior to the implementation is -0.152 units which is measured as a deviation of a firm's inventory levels from size adjusted within-industry average inventory levels. This is consistent with other studies including Eroglu and Hofer (2014). We find that firms that are affected by FSMA as compared to other firms experience 0.159 units decrease in inventory leanness post FSMA implementation. This represents a 111.84% decrease in inventory leanness relative to the average inventory leanness of -0.152. This finding suggests that FSMA regulation has a major impact on inventory leanness of affected firms making firms considerably less lean compared to average before implementation of the regulation.

The second hypothesis suggests that a firm's reliance on developed market for sourcing critical materials exacerbates the relationship between regulatory stress and inventory leanness. This finding is important as it shows how strategic sourcing from different markets can affect a

firm's response to regulatory stress. Figure 4 demonstrates the moderating role of firm's reliance on developed markets for sourcing indicates the presence of a moderation affect. Firms that have high reliance on developed markets appear to be more significantly affected by the FSMA regulatory stress, leading to a decrease in their inventory leanness as a response to regulatory stress. Likewise, the third hypothesis suggests that a firm's reliance on emerging markets for sourcing critical materials alleviates the effect of regulatory stress leading to more inventory leanness. Figure 5 demonstrates the moderating role of firm's reliance on emerging markets for sourcing; it shows that the impact of FSMA on inventory leanness is more significant for firms that are less likely to source from emerging markets compared to those that do.

Thus, our study offers important theoretical, managerial, and policy implications.

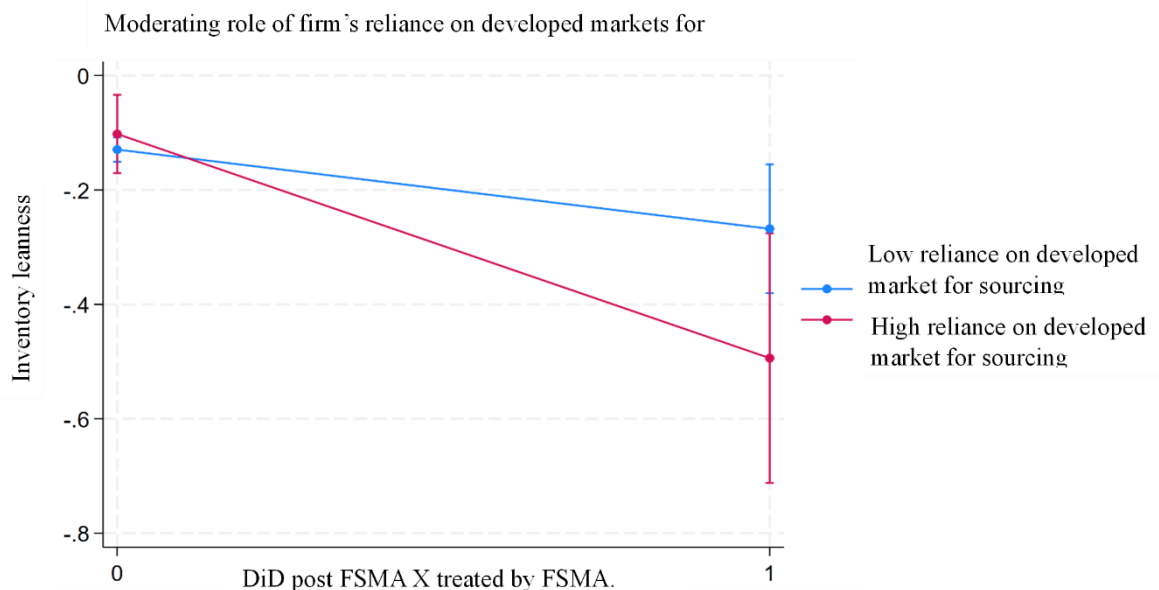


Figure 4. Moderating role of firm's reliance on developed markets for sourcing

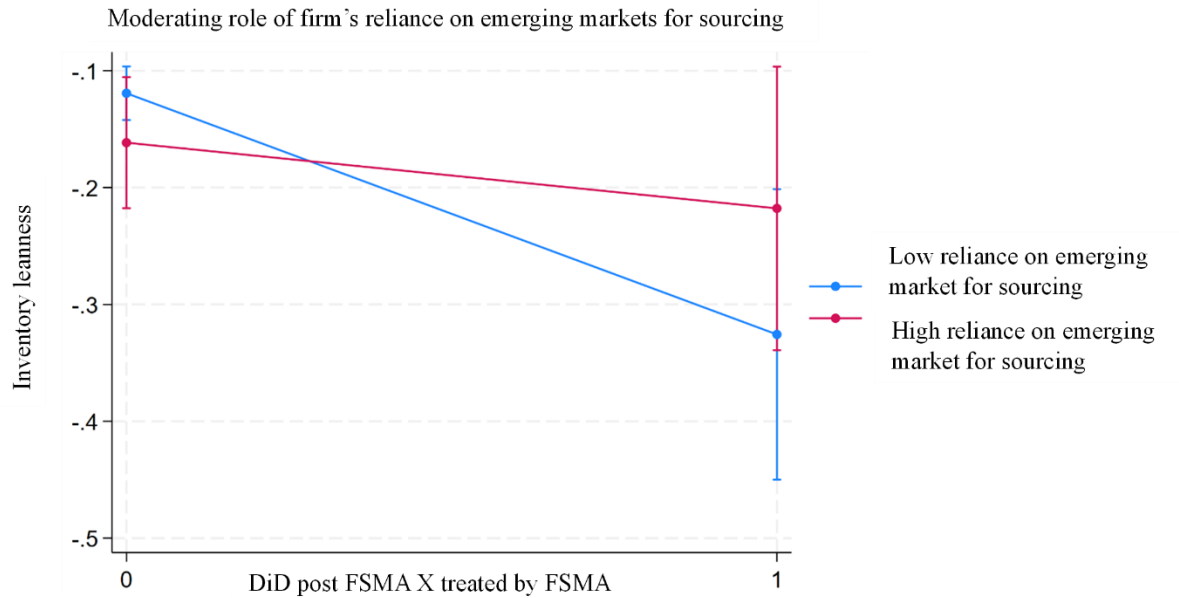


Figure 5. Moderating role of firm's reliance on emerging markets for sourcing

7.1.Theoretical contributions

Our study makes several contributions to supply chain literature. First, we theorize and provide empirical evidence that introduction of public policies, such as supply chain accountability regulations, result in a decrease in inventory performance (Pagell et al., 2018). Previous research provides mixed evidence about the relationship between public policies and firm performance. Some studies suggest that public policies positively impact firm performance (Cantor et al., 2009; Gundlach et al., 2019) and others suggest that they have a negative effect on firm performance (Darby et al., 2020; Miller et al., 2017). Our research adds to this debate in the literature because it demonstrates that there are unintended consequences associated with public policies on performance of supply chains (Scott et al., 2021). The inadvertent impact of this public policy is much more difficult to predict and can have significant impact. We show that a policy intended to improve a firm's supply chain accountability might change their inventory management behavior.

Second, we contribute to the literature at the intersection of supply chain management and public policy and government regulation (PPGR) (Dobrzykowski, 2019; Fugate et al., 2019; Pagell et al., 2018). There have been several recent calls in the supply chain literature to expand the understanding of this intersection (Pagell et al., 2018; Helper et al., 2021). Research at this intersection has the potential to inform policymakers and help them consider a broader supply chain perspective. It can indirectly affect policy making by sensitizing stakeholders about the intended and unintended consequences of regulations on supply chains (Fugate et al., 2019). For example, Pagell et al., (2018) suggests that while extensive research has informed stakeholders of public policy about the costs and benefits of regulations, this analysis is largely missing the systemic implications for supply chains. Answering the call, we contribute to this literature by demonstrating how firms “adapt practices in response to a new or proposed SCM-related PPGRs”. Further, we contribute to this literature by uncovering the impact of public policy on the inventory performance of firms impacted by regulations (Helper et al., 2021). Although the policy was introduced as supply accountability regulation to promote consumer safety, it might have far reaching consequences for firms.

Thirdly, we expand the literature in supply chain management by theorizing government policies as supply chain risk. Building on the existing literature, we delve deeper into how supply chain accountability regulations manifest as critical risk factors affecting inventory performance. Prior studies have identified the changes in the regulations as critical supply chain risk factors (Darby et al., 2020; Manuj & Mentzer, 2008; Wagner & Bode, 2008). It suggests that government policies control the flow of critical resources that firms need for production and distribution (Cohen, 2008). Unfavorable policies can cause disruption in supply chains and affect the flow. A few other studies suggest that such policies can reduce the risk in supply chains such

as by facilitating conflict resolution among buyers and suppliers. We extend the prior literature by studying a unique context of the introduction of a supply chain accountability regulation, rooted in product safety.

Fourth, we enumerate the impact of public policies on inventory leanness. We show that public policies make firms less internally efficient as measured by inventory leanness. Prior literature suggests that inventory leanness is critical as it helps in improving cash flow as less capital is tied up in inventory. Maintaining lean inventory, especially for firms handling food, can reduce the risk of spoilage and goods depreciation (Eroglu & Hofer, 2014). However, increased risk and inefficiencies due to a regulatory mandate leads to decreased inventory leanness. Eroglu and Hofer (2014) studied how a higher inventory leanness impacts financial performance of firms. Barker et al. (2022) studied how higher inventory leanness of supplier's impact financial performance and Chuang et al. (2019) studied how higher inventory leanness affects operational efficiency. We are among the first to study inventory leanness as an outcome. We contribute to the inventory leanness literature by identifying the antecedents of inventory leanness, i.e., factors affecting firms inventory leanness performance.

Lastly, we make a theoretical contribution by demonstrating the utility of stress and coping theory in public policy and supply chain contexts. Stress and coping theory have been largely used in the organizational psychology literature to study individual behavior and responses under stressful circumstances such as work pressure, marital abuse, peer pressure etc. (Lazarus & Folkman, 1984). Lately, supply chain and operations management literature has used the theory to explain individual level behavior of several stakeholders such as consumers and employees (Autry & Daugherty, 2003; Masorgo et al., 2023; Morrison, 2015). We suggest that one of the primary ways that firms cope with the supply accountability regulation is by adjusting

their inventory behavior. Further, their appraisal of sourcing from different global markets can alleviate or exacerbate their stress and coping response. We contribute to the theory by applying it to a unique phenomenon and introducing a new unit of analysis (Makadok et al., 2018). We further introduce a new casual mechanism of regulatory burden and global sourcing to identify the conditions under which a firm's stress increases or decreases (Makadok et al., 2018).

7.2. Managerial and policy implications

Our study offers valuable managerial and policy implications. First, global sourcing is a double-edged sword. The decision to source from different markets requires careful consideration, recognizing the risks and uncertainties associated with each market. While global sourcing can offer cost advantages by providing access to different markets, our findings highlight the potential for increasing risk and burden on inventory performance from different regulatory environments. Further, our study helps managers re-think their preexisting beliefs about sourcing from different markets. For instance, we show that reliance on developed markets may prove counterproductive under regulatory stress when compared to emerging markets, contrary to popular belief. Balancing the costs advantages and added risks and burden from global sourcing from different markets becomes essential for inventory management.

Moreover, our research documents the adverse impact of government regulations on how a firm maintains its internal efficiency. Regulations can inadvertently incentivize firms to accumulate higher levels of inventory to buffer against the regulatory burden and risk (Darby et al., 2020). This undermines a firm's ability to operate with lean inventory, affecting its inventory performance. For competitive reasons, many firms maintain lean inventory to minimize waste and improve efficiency, cash flow, and profit margins (Eroglu & Hofer, 2014). Unfortunately, the regulatory environment encourages managers to adhere to proactive regulatory compliance to

minimize the negative impact of such regulations on their operations. Higher regulatory intelligence – a process of proactively gathering and analyzing the impact of regulations on a firm’s operations – may safeguard firms against disruptions.

Lastly, our findings emphasize the role of public policies in shaping inventory performance of firms. It highlights the unintended consequences of public policies and the role of policy makers. Such an unintentional impact increases the regulatory burden on firms affecting their ease of doing business. Policy makers must balance the regulatory burden on firms while safeguarding the interest of consumers. Our study highlights the importance of staggered implementation of regulatory policies or pilot programs to test the policy on a smaller scale before full scale implementation. Further, impact assessments of such policies are required before and after the policy implementation to understand its economic, social, and environmental impact.

7.3.Limitations and future research

Although our paper presents a comprehensive analysis of the impact of supply chain accountability regulation on a focal firm’s inventory performance, there are a few limitations that future research can address. Firstly, we measure inventory performance by only using inventory leanness. Future research could explore additional inventory performance measures including inventory days, inventory turnover, stock outs etc. These outcomes can help provide a more comprehensive understanding of the effect of such regulations on a firm’s inventory performance. Secondly, we use the specific context of Food Safety Modernization Act to observe this phenomenon. Future research can explore if the findings remain consistent under different contexts of supply chain accountability regulations. For example, the German Supply Chain Due Diligence Act or EU’s Corporate Sustainability Due Diligence Directive.

Thirdly, we use the context of regulatory compliance in a developed market where firms based in such markets engage with the global market. Future research may expand the context to emerging or frontier markets as well. They could study how a change in such policy in an emerging market affects buyers' behavior based in emerging markets. Lastly, we assume that firms are at the receiving end of the unintended consequences of regulations. However, if the political environment is unfavorable to a firm's supply chain performance, it can seek to engage with its political environment to shape it favorably (Grover & Martin, 2022). For instance, firms can lobby to save their competitive advantage. Similarly, future research can explore different alternatives where firms can use their regulatory environment to their supply chain's advantage.

Acknowledgements

I would like to express my gratitude to my mentors of this essay Dr. David Cantor (Iowa State University) and Dr. Thomas M. Corsi (University of Maryland). I would also like to extend my appreciation towards the Center for Global Business at the Robert H. Smith School of Business, University of Maryland, for its financial support provided via the Centers for International Business Education and Research (CIBER) grant. I would also like to extend my appreciation to Dr. Martin Dresner for his invaluable feedback on the manuscript and to Dr. Adams Steven for his assistance with the PIERS database, offering insight into maritime shipping and international trade.

Chapter 3: Work-Break Schedules and Quality Assessment

“Field work is understaffed. Expect overtime that is not clearly stated in the job description... If you don't cover everything, your performance rating and the consumer suffer the consequences ... overtime and directed assignments will strain relationships if you have a family, and if you don't have a family, you won't have a social life.” - (Consumer Safety Officer, FDA)

1. Introduction

Several studies have highlighted the vital role of plant inspections in managing quality across increasingly complex supply chains (Ibanez & Toffel, 2020; Power & Terziovski, 2007; Short et al., 2016). These inspections serve dual purposes: they gauge the existing level of compliance and pinpoint areas of improvement in stakeholders' facilities. However, inspection schedules can influence the outcomes of inspections (Ibanez & Toffel, 2020), potentially affecting the reliability of managers' decisions based on those results. For example, Ibanez and Toffel (2020) found that sequence is critical such as positive outcomes from previous inspections lead to fewer violations in subsequent inspections. Similarly, Ball et al., (2017) found that scheduling inspections with experienced inspectors can be counterproductive for example inspections carried out by inspectors familiar with the site tend to be less effective in reducing recall hazards due to relational ties. Thus, the quality and reliability of inspection outcomes are directly contingent upon the inspection schedules, as well as the inspectors assigned.

Literature on inspection scheduling has primarily focused on inspector experience (familiarity with the inspected site) and inspection sequencing (the order in which inspections are conducted) (Ball et al., 2017; Ibanez & Toffel, 2020). For example, inspectors get tired from mental and physical work during the day, so inspections done later in the day are not as effective as those done earlier (Ibanez & Toffel, 2020). However, the literature offers limited insight into

how the timing of breaks relative to inspection tasks intersects with inspector workload management and performance. For example, an increased number of back-to-back inspections with limited breaks can lead to inspector burnout, impairing their ability to detect violations, resulting in scheduling-induced workload fatigue (Ibanez & Toffel, 2020). The issue raises significant concern for essential workers who are on the front lines of the workforce, for example regulatory food inspectors (Mayo et al., 2024). According to Bachman and Stokes (2019) U.S. Dept. of Agriculture (USDA) inspectors can be “scheduled for five straight days of 16.5-hour shifts and adequate oversight of meat processing has become impossible.”

Breaks, in operations management, are defined as intervals of rest and downtime scheduled within and around tasks to provide employees time to recover from work-related mental and physical fatigue (Bard & Wan, 2008; Schultz et al., 2003; Upton, 1995). Breaks can improve an employee’s efficiency by helping them reduce burnout resulting in improved job satisfaction. The timing and length of breaks can be designed by the managers to benefit their employee’s productivity (Bard & Wan, 2008; Hunter & Wu 2016). A growing body of research examines various types of breaks focusing on their benefits (Leroy et al., 2020; Pendem et al., 2022). However, breaks could have downsides, for example, longer breaks between tasks can make it hard for workers to regain momentum and continue the workflow (Cai et al., 2018). Hence, an important question that field staff managers face is how they should effectively schedule field tasks to balance breaks and work intensity.

Balance between break and work intensity for field staff can take myriad forms. On one end of the broad spectrum, work-break scheduling can result in multiple days of time-off within a task assignment offering several days to recharge to enhance productivity (Fritz & Sonnentag 2006, Fritz et al. 2010). Under such conditions, employees are completely disengaged from

work, improving the quality and effectiveness of breaks. For example, a food inspector that gets idle time between inspections can improve their quality assessment of the facility. On the other end of the spectrum, work-break scheduling can result in job creeps, i.e., when work encroaches on personal time or scheduled breaks, leading to a blurred line between work and rest, diminishing the quality of time-off (Bolino et al., 2013). Under such conditions employees are not completely disengaged from work during their scheduled time off, reducing the quality and effectiveness of breaks, for example, a food inspector is required to perform an inspection on a weekend due to food safety emergency. We introduce a novel conceptualization of work-break scheduling structures based on the timing of breaks relative to a work task, the length of the breaks, varied levels of disengagement from work, and their predictability. Using the following dimensions, we capture a broad range of work-break schedule regimes in operational setting that may impact the effectiveness of work performance.

We examine the impact of five work-break schedule regimes of field staff (FDA Consumer Safety Officers or Food Inspectors) on their quality assessment of food facilities. Although there can be several work-break regimes, we consider five based on our identified dimensions and empirically test their effects in the field context of food safety inspections. First, we focus on the pacing of the work-break schedule, i.e., the time between two inspection assignments where inspectors are working on secondary tasks. This time is used by employees to recover from the last inspection and prepare for the next assignment, but it is not a complete break as they work on secondary tasks - we call it pseudo break. Second, we study an intermittent work-break schedule where breaks are scheduled during an ongoing task (Pendem et al., 2022). Third, we study fixed workweek schedules, where workweeks are sandwiched between fixed breaks (Yao et al., 2019). Fourth, we study inspections on non-standard days or

job creeps i.e., where tasks are assigned during (normal) scheduled time off and may encroach on personal time (Bolino et al., 2013). Lastly, concurrent task assignments are considered where employees are assigned multiple tasks simultaneously. Instead of breaks between tasks assignment, this is a situation of task overlap or tasks intruding on other tasks (Goes et al., 2018). In such schedules, work from one task may affect the performance of the other task. The impact may be related to attention residue when part of one's focus is on another task instead of fully being devoted to the current (focal) task. We hypothesize how these factors can influence the likelihood of detection of violations using the case of FDA food inspectors.

Our thematic analysis of job feedback from over 80 Food and Drug Administration (FDA) inspectors indicates that workload related to scheduling is among the top five concerns impacting their job performance and morale. Some of the inspector comments include: "Extremely high workload"; "Tons of traveling that may exhaust you, especially if you have a family"; "Expect burnout"; "Expect overtime that is not clearly stated in the job description"; "Additional overtime is required to meet assignment due dates." Such an escalated workload could negatively affect the inspectors' work performance. It is important to note, despite the food inspections, foodborne illnesses affect approximately 48 million people in the U.S. each year, resulting in 128,000 hospitalizations and 3,000 deaths (CDC, 2018). In addition, the comments from the inspectors draw attention to the issue of responsible scheduling: prioritizing employees' wellbeing while optimizing productivity through scheduling.

A 2017 report by the Department of Health and Human Services (DHHS) highlighted multiple challenges about "how" FDA schedules food inspections. For example, DHHS suggested that the FDA needs to re-consider "how it handles attempted inspections to ensure better use of resources." Specifically, it needs to focus on augmenting its workforce and re-

evaluating inspection scheduling strategy (Laack, 2017; Levinson, 2017). Accordingly, using the FDA context, our research aims to provide insights about how different aspects of work-break schedules affect inspection outcomes that may improve performance by changing the proposed inspection scheduling. It is essential to determine how scheduling can be made more effective to enhance inspection outcomes.

Existing literature has established and used violations as a tangible indicator of regulatory compliance and effectiveness of quality assessments (Short et al., 2016; Ibanez & Toffel, 2020; Ibanez et al., 2023). Violations reflect the thoroughness of the inspection process and the inspector's ability to identify and document non-conformance against a code of conduct. The number of violations cited is influenced by various factors, including the facility's adherence to quality standards and the inspector's rigor. Since inspectors are required to provide evidence for violations they cite, this may lead the inspectors to underreport issues, a phenomenon noted by Ibanez and Toffel in 2020. Furthermore, although FDA inspectors initially document their observations on-site, these issues undergo subsequent review and vetting by the FDA's regional office. Consequently, there is a tendency among inspectors to be overly careful in citing violations, a scenario described by Ibanez and Toffel (2020) as "inspector bias in the form of underreporting violations."

To address our research objective, we utilize fine-grained data from FDALytics, a regulatory compliance database firm that extracts information from enforcement reports of food inspections via Freedom of Information Act requests to FDA. Based on our analysis of FDA inspections of food facilities from January 2013 to March 2020, we find evidence that inspection work-break schedules significantly impact inspection outcomes. First, we observe that pseudo breaks are productive for inspections. Increasing the interval between successive assignments

using pseudo breaks leads to a rise in the number of violations detected, with inspectors having more time for preparation, analysis, or recovery. However, this effect exhibits diminishing returns; specifically, beyond a 3-day interval, the number of detected violations begins to decline. This reduction could be due to a decline in inspectors' proficiency as break length increases. Second, we find that conducting inspections simultaneously, when inspections intrude on other inspections, reduces citations by 14.7%, suggesting that concurrent schedules diminish inspectors' effectiveness in identifying issues. Third, intermittent breaks during ongoing inspections lead to a 78.8% increase in citations compared to continuous inspections, although, the duration of these intermittent breaks' exhibits diminishing returns. Fourth, inspectors detect 21.9% fewer violations during inspections when working on non-standard schedules compared to standard Monday to Friday workweeks. Lastly, we find that inspections starting on Monday (generally directly following a weekend break) result in 18.4% more citations than those initiated on Friday (generally, directly before a scheduled break) suggesting that the timing within the workweek significantly influences inspection effectiveness. These workweek results provide evidence for fade-out effects and the anticipatory effects from breaks.

Our work contributes to both theory and practice in several ways. First, our study introduces an operational perspective to the research on breaks and its impact on food inspectors' performance, a relatively unexplored area. We analyze how different work-break arrangements, including their timing and length, influence the effectiveness of food safety inspections. This research shifts the focus from general break benefits to their strategic scheduling in the context of food facility inspections, examining the outcomes of different break regimes. Second, we explore new concepts, such as pseudo breaks, temporal pacing of inspections, and the effects of break anticipation, on task performance. Our findings suggest that while breaks can improve

inspector performance, their effects vary. Additionally, we highlight breaks as strategic elements that can be managed to enhance work performance.

Thirdly, we also contribute to the quality assurance and inspections literature by identifying factors that affect inspection effectiveness (e.g., Ball et al., 2017; Ibanez & Toffel, 2020; Mani & Muthulingam, 2018). Our focus on how inspection break scheduling affects inspection outcomes brings attention to the inspection process contributing to the literature on inspection design and implementation. This literature has largely been focused on inspector's biases, experiences, or other sociological and economic factors (e.g., Ball et al., 2017; Short et al., 2016). Our focus on how inspection break scheduling facilitates or impede inspectors' ability to conduct inspections also contribute to the literature on behavioral operations, which focuses on the intersection of factors affecting human behavior in operations management (Bendoly et al., 2006). Finally, results uncover actionable insights for federal regulatory agencies and other private or non-profit entities that engage in inspections. As managers in various industries strive to enhance their monitoring abilities, our research indicates several effective strategies: utilizing the break scheduling for effective task implementation.

2. Related literature

We build on the following streams of operations management research: (i) breaks and productivity, (ii) quality management and inspections, and (iii) scheduling and operational outcomes.

2.1. Breaks and productivity

An extensive body of research explores the impact of breaks on employee productivity, covering weekend breaks, vacations, and end-of-day breaks. For instance, Fritz et al. (2010) suggest that weekend breaks reduce work demands, allowing for recovery from stress induced by

such demands. However, they also note that not all weekend breaks lead to disengagement. The choice of specific recovery activities during the weekend, like relaxation and detachment, influences employees' emotional states throughout the workweek. In another study, Fritz and Sonnentag (2006) investigate the effect of extended breaks (vacations) on work performance. They find that health complaints and burnout reports from employees decrease immediately after a vacation, resulting in improved self-reported task performance. However, the positive effects of vacation breaks are temporary and tend to fade over time, leading to a return to pre-vacation stress levels and impacting employee performance. Binnewies et al. (2009) examine the impact of overnight or end-of-day breaks on employees' work performance the following day. Their findings indicate that effective overnight recovery is positively linked to daily task performance, personal initiative, and organizational citizenship behavior the next day.

Cai et al. (2018) distinguished between breaks based on their predictability, categorizing them as either expected or unexpected. They introduced a crucial theoretical distinction between scheduled breaks (such as weekends and lunchtimes) and interruptions (like equipment breakdowns and unscheduled meetings), exploring how unplanned breaks (interruptions) affect workers' productivity on the following day. Their findings indicate that such interruptions lead to a decrease in worker productivity, with a machine breakdown, for instance, associated with a 3.3 percent drop in productivity the next day. Similarly, Froehle and White (2014) observed that interruptions can lead to errors and the need for rework in production, thereby diminishing worker performance. Gurvich et al. (2020) discovered that interruptions due to task changeovers result in a 20% increase in service-processing time. The consensus in the literature is that interruptions are counterproductive, reducing worker productivity by causing distractions, frustration, stress, and anxiety among employees in relation to their work tasks.

In contrast, Pendem et al. (2022) suggested that not all interruptions result in negative outcomes. They described interruptions as events within the employees' working hours that divert their attention from the primary task and break the flow of work. Unlike previous literature, Pendem et al. (2022) proposed that standard scheduled breaks, which offer a chance for respite and renewal, can act as a type of interruption likely to shift employees' focus away from their primary work tasks. They also introduced the concept of 'pauses' i.e., work interruptions that provide a brief period of respite, allowing employees to maintain their focus on the primary task, such as when an employee waits for another to complete a task due to process interdependencies. However, their study focused on farm workers harvesting fruit, whose shifts involve a series of simple and short tasks that demand exhaustive physical effort but minimal cognitive work where errors do not directly lead to immediate, significant adverse outcomes.

Our research specifically focuses on work-break schedules, that is, the timing of breaks in relation to work tasks. We define breaks as intervals of rest and downtime scheduled within and around tasks to allow employees to recover from work-related mental and physical fatigue. Using the existing literature, we theorize that breaks are fundamentally different from interruptions. Firstly, breaks, though not guaranteed, can be predicted, controlled, and managed to a greater extent by employees and their managers, unlike interruptions, which are generally unpredictable and uncontrollable.

Secondly, the mechanisms through which breaks and interruptions influence work performance may differ. For instance, short interruptions, delays due to task interdependencies, might offer physical respite to employees in the short term but result in increased mental stress due to reduced task productivity overtime. This could lead to more physically demanding work as employees may need to compensate for lost productivity by working faster. While this might,

in some cases, lead to increased productivity, it does so through heightened task-related stress and anxiety, unlike breaks, which pave the way to productivity by promoting positive recovery through rest and rejuvenation. Third, for breaks, the objective of managers-employees is to improve effectiveness by increasing the level of disengagement from work. Interruptions, on the other hand, reduce effectiveness, so companies seek to minimize their occurrences, intensity, and the degree to which they shift attention of employees from their tasks.

Our study focuses on the timing of breaks in relation to specific work tasks. Our research conducts an analysis of five distinct classes of work-break schedules that result in both positive and negative outcomes. We explore complex facility inspections requiring both physical (e.g., collecting samples, navigating production facilities) and mental (e.g., conducting science-based evaluations, interpreting regulatory standards) efforts, where mistakes can lead to immediate and significant adverse outcomes. These tasks typically extend over multiple days (averaging three days in our sample) and aim for enhanced quality assessments. Moreover, we systematically develop a theoretical framework to categorize work-break schedule structures by considering a combination of a) the timing of breaks relative to a work task, b) the length of the breaks, c) the degree of disengagement from work, and d) predictability (refer Table 1). Our study is among the first to thoroughly examine the impact of work-break schedules on task outcomes.

Table 1. Categorization of work-break schedules

Types of work-break schedules	Break timing in relation to a task	Length of breaks	Degree of disengagement from work	Level of predictability
Pseudo work-break schedule	Before the start of a new task	Varies (short to long)	Partial	Low, variable
Intermittent work-break schedule	During an ongoing task	Short	Moderate to complete	Moderate, as breaks depend on

Work-week schedule	Series of tasks are sandwiched by breaks	Short and fixed	Complete	re-start date of task High, with scheduled weekends
Non-standard work-break schedules	Task intrudes into scheduled break periods	Reduced	Low to none	Low to high
Concurrent work schedules	Task intrudes into other tasks without a formal break between them	None, tasks continue without formal breaks	None, continuous engagement in work tasks	Low, formal break periods not specified

2.2. Quality management and inspections

Previous research has delved into a variety of methods for monitoring operational tasks to ensure compliance with required quality standards, industry norms, regulatory requirements, and customer expectations (Corbett, 2006; Gray et al., 2015; Levine & Toffel, 2010). The body of research on inspections has considered who conducts these inspections; for instance, Short et al. (2016) investigated how an inspector's ability to detect non-conformance is influenced by factors such as site familiarity, experience, and the funding source of the inspection. This research additionally delved into the dynamics of inspection teams and their gender compositions. Palmarozzo et al. (2021) focused on the sourcing of inspections, analyzing how second-party inspections compare to third-party ones in assessing supplier compliance with buyers' codes of conduct.

Further studies have focused on the methods and formats for executing inspections. Scott et al. (2021) revealed that electronic monitoring of truck drivers' speeding behaviors, while improving compliance in monitored aspects, inadvertently increases non-compliance in unmonitored areas. Staats et al. (2017) found that individual electronic monitoring significantly

boosts hand hygiene among healthcare workers. Palmarozzo (2021) identified that remote inspections generally yield lower quality results than in-person inspections, a trend that worsens with an increase in the number of inspectors in a remote or virtual team. Nonetheless, in-person inspections remain the preferred execution mode in scenarios such as internal quality control within manufacturing processes (Shah et al., 2016), inventory accuracy assessments (Kok & Shang, 2007), and evaluations of supplier operations by buyers' inspectors (Handley & Gray, 2013; Short & Toffel, 2016).

Other studies explore how inspections can promote organizational learning opportunities (Short et al. 2019; Mani & Muthulingam 2018) and act as external shocks prompting organizations to renew their commitment to operational routines, as noted by Anand et al. (2012). These routines may decay unless rejuvenated by external shocks, though not all shocks, such as mergers, have a revitalizing effect. Research has also investigated how the accuracy of outsourced inspections is influenced by the financial source of the third-party inspections, be it buyers or suppliers (Duflo et al. 2013; Ronen 2010; Short & Toffel 2016), and the competitive dynamics among inspection firms and their cross-selling abilities (Bennett et al. 2013; Koh et al. 2013; Pierce & Toffel 2013).

Overall, previous research indicates that the number of issues detected at a facility is influenced by various factors including the firm's inspection history (e.g., number and outcomes of prior inspections, time since the last inspection), characteristics of the inspector (e.g., training, experience, familiarity with the site), composition of the inspection team (e.g., number of inspectors, gender makeup), method of inspection (e.g., virtual, electronic monitoring, in-person), how inspections are sourced (e.g., in-house, second or third-party, outsourced), and the scheduling of inspections (e.g., order, timing during the day) (Ko et al., 2010; Ibanez & Toffel,

2020; Toffel et al., 2015). Expanding on research related to inspection scheduling, our study explores how the outcomes of facility inspections are affected by the work-break schedules of inspectors.

2.3.Scheduling and operational outcomes

Our research contributes to the existing body of knowledge on "how work schedules affect task performance." In the manufacturing sector, scheduling covers a broad spectrum of activities, including sourcing raw materials, establishing production schedules, and conducting machine maintenance. Specifically, our study delves into the timing aspect of scheduling for quality assessments, examining how break timing relative to work tasks (inspections) affects task performance. The impact of inspection scheduling on quality assessments has not been extensively explored within the operations management literature.

Ibanez and Toffel (2019) investigated how inspector schedules could introduce biases that adversely affect inspection outcomes, particularly regarding inspection stringency. They argue that accuracy and consistency are essential for ensuring that quality assessments are both effective and equitable. Through their analysis of food inspections, they highlighted how sequencing biases could compromise an inspector's objectivity. Firstly, they explored inspection sequencing, noting that the outcomes of previously inspected facilities could impair judgments about subsequent facilities due to a negativity bias. Secondly, they examined the sequencing of inspections within an inspector's daily schedule, finding that inspectors report fewer violations in inspections conducted later in the day compared to those scheduled earlier. This reduction in citations is attributed to fatigue, which diminishes inspectors' capacity and willingness to identify violations as the day progresses. Their research underscores the influence of sequence-induced

inspector bias, suggesting that inspections scheduled out of sync with inspectors' natural daily rhythms can lead to biased outcomes.

Mayo et al. (2024) examined the influence of scheduling decisions on the turnover of frontline workers, such as part-time certified nursing assistants (CNAs). They investigated the effects of two scheduling aspects: the number of hours nurses worked and the variability of their coworkers. They proposed that both the hours worked by the nurses and the variability of their coworkers play a significant role in nurse turnover. Their findings suggest that managers can mitigate nurse turnover by increasing the hours worked and consistently scheduling coworkers together. They emphasized how the quantity and quality of work allocation could enhance nurses' willingness to remain longer with a caregiving facility, thereby improving both the quality and cost of care.

Kesavan et al. (2022) introduced the concept of responsible scheduling, which prioritizes worker well-being. They analyzed how various dimensions of work schedules such as inconsistency, unpredictability, inadequacy, and lack of employee control, affect employee well-being and performance. Their research indicated that implementing responsible scheduling practices could increase retail store throughput and sales by motivating employees to put in extra effort and adhere more closely to labor plans. Similarly, Lu et al. (2022) explored the effect of schedule consistency on worker productivity, specifically examining its impact on cashier productivity in a grocery retail setting. They discovered that more consistent work schedules significantly boost productivity, especially among less experienced workers, such as new hires. Moreover, Goes et al. (2018) investigated the effects of multitasking in customer service by studying agents handling multiple online service chats simultaneously. They found that agents

assigned multiple chats at once experienced longer service delays, which decreased problem resolution rates and adversely affected customer satisfaction.

Previous research has delved into various aspects of scheduling and its impact on task outcomes, including sequencing, number of hours worked, coworker variability, consistency, predictability, adequacy, employee schedule control, and the simultaneous allocation of tasks. Building on this body of work, our research contributes to the literature by examining how the work-break schedules of inspectors' influence quality outcomes. We focus on how workload-induced fatigue affects inspectors' ability to detect violations.

3. Hypotheses development

3.1 Pseudo work-break schedule

First, we focus on the pseudo work-break schedule. We propose that work-break schedules can generally be categorized based on when the break occurs in relation to a work task, how long the break lasts, and different degrees of disengagement from work tasks. Unlike other literature that examines the effects of complete time-off between tasks, our study focuses on the pacing of pseudo breaks—a period that serves as a break from a primary task but does not completely disconnect employees from work (Eisenhardt & Brown, 1998). Rather, workers perform secondary tasks such as completing documentation from previous tasks, filing travel reimbursement reports, answering emails, preparing for the next primary task, and attending training, among other activities.

Psychology literature suggests that such task transitions, i.e., a mental and physical shift from one task to another, play a very important role in task performance (Fritz & Sonnentag, 2006). Firstly, a work-break schedule structure with transition times before the start of the next tasks (inspections in our case) provides a cooling period for employees to recover from the

mental and physical fatigue from the last inspection. We suggest that pseudo breaks enable employees to disengage from their primary tasks while simultaneously helping them maintain momentum. Secondary tasks require less cognitive and physical effort as compared to primary tasks, ensuring that inspectors are not completely disconnected from work. After taking a brief break from the primary task, employees return with a refreshed mind and enhanced vigilance, which can lead to a more thorough inspection process. This time-off from the primary task leads to mental refreshment improving their attention to details.

In our context of FDA inspections, Ibanez and Toffel (2020) suggest that inspectors are affected by the outcomes of prior inspections. The number of violations from prior inspections affects inspectors' sentiments and perceptions about the state of compliance of food facilities in general. This impacts their approach to subsequent evaluations. For example, a prior inspection with higher facility compliance may lower their vigilance in the next inspection leading to leniency and fewer citations. In another example, negative experiences from prior inspections can lead to a tunnel vision effect where inspectors focus too narrowly on specific issues encountered previously, missing the broader compliance issues. This suggests that inspector's attitudes and expectations may be shaped by the impressions from the recent past. We suggest that incorporating effective transition times between two inspections can act as a critical buffer. These pseudo breaks allow for disengagement from the prior primary task which can prevent perception spillovers from the last inspection, reducing bias. By breaking the inertia of negativity or positivity bias, transition times reset inspector's cognitive and emotional states enabling them to approach inspections with renewed objectivity. Breaks also help dispel emotional residues, reducing reactionary approaches (Fritz & Sonnentag, 2006). This disengagement results in a more vigilant and unbiased inspection resulting in inspectors detecting a greater number of

violations. Transition times are not merely a pause from primary task but a vital reset mechanism that ensures that each inspection is conducted with a fresh perspective resulting in an increase in number of citations due to increased thoroughness and reduced bias.

Further, the transition time allows for the inspectors to prepare for the next assignment. The preparation time helps them understand the compliance history of facilities and specific unresolved problems highlighted by prior inspections. It also helps devise a targeted inspection strategy and plan their next inspection in advance. Better preparation could lead to more efficient inspection execution. FDA's Investigator Operations Manual (2024), suggest that prior to inspections inspectors should *"Acquaint themselves with the firm's inspectional, compliance, personal safety, and recall history; open complaints (see IOM 5.2.3); related firms; responsible persons; trademarks; practices; and products. This review may help reveal or identify products [inspectors] will want to cover... Determine the type of operation (manufacturer, warehouse, own-label distributor, etc.) to be inspected to ensure [inspectors] apply the appropriate regulations."* Transition times allow for the preparation time enabling inspectors to conduct a thorough inspection resulting in a greater number of violations.

However, very long transition times could prove counterproductive. After a certain point of disengagement from a primary task, such transition periods can give rise to "productivity-stealing temptations." Beyond the optimal recovery period, longer breaks away from primary tasks may lead to a decline in sharpness of inspection skills – also known as skill atrophy. The inspectors might not spend as many days on the field practicing their primary tasks of conducting inspections. According to psychological distance theory, people feel more removed from an event as time progresses, leading them to view their experiences of the distant event in a more abstract rather than concrete manner. So, after a longer gap from a prior inspection, inspectors

might underestimate the efforts to conduct an inspection resulting in overconfidence. Further, inspectors returning to their primary task after a longer pseudo break might require a period of readjustment affecting inspectors' ability to perform at their peak.

Therefore, we suggest that the relationship between pacing of pseudo work-break schedule and the ability to find issues in an inspection has an inverted U-shaped relationship underscoring the importance of balancing rest and work to maintain optimum levels of performance (Haans, et al., 2016). We suggest that there is an optimal interval between tasks (in this case, inspections) that maximizes performance and anything above or below that could lead to under-stimulation or over-relaxation respectively. Hence, we propose that:

Hypothesis 1a: Longer pseudo break intervals between successive tasks will increase the task performance; **Hypothesis 1b:** However, it will occur at a diminishing rate.

3.2 Intermittent work-break schedule

We define intermittent work-break schedules are short scheduled pauses taken in between an ongoing task offering greater disengagement from work tasks as compared to pseudo breaks. They are unpredictable and depend on the start day of the inspections. Existing literature presents conflicting viewpoints on the effect of intermittent breaks on task performance. The first set of literature suggests that intermittent breaks act as interruptions that break the workflow. This literature indicates that intermittent breaks, while offering a pause from primary tasks, act as cognitive restarts and necessitate additional warm-up time afterward. Such breaks can be counterproductive, leading to increased distractions and negatively impacting the performance outcomes of the primary task (inspections) resulting in lesser number of citations. A significant amount of research is dedicated to minimizing the detrimental effects of these interruptions.

According to Pendem et al., (2022), intermittent breaks “*divert attention from the primary work task*” and they may negatively impact productivity immediately after the break. However, Pendem et al. (2022) focused on the intermittent work-break schedules of farm workers, whose shifts consist of a series of simple and short tasks requiring exhaustive physical effort but minimal cognitive work with low margin of error. In contrast, our study investigates complex facility inspections that demand both physical and mental effort with a high margin of error. These tasks generally span several days (averaging three days in our sample) with the objective of improved quality assessment.

However, the second set of literature suggests that intermittent breaks provide a period of rest for employees (Fritz & Sonnentag, 2006). This period can be used to revitalize and refresh from the exhaustion of work. This break can help re-focus attention leading to more thorough inspections. Further, conducting inspections includes both physical and cognitive load. Inspectors usually get a limited time to familiarize themselves with inspected facilities and assess them. They encounter an influx of information which they must process during that limited time. Intermittent breaks may provide them with some cooling period between work tasks to process that information better.

Further, such breaks improve inspector readiness for the inspections. It provides them with time to refresh and re-confirm their knowledge about regulatory standards. It also provides them with time to connect with their supervisors or other senior colleagues to seek clarifications and corroborate their own opinions. Overall, it provides them with much needed time to be more thorough resulting in higher citations.

Hence, we take a side and propose that:

Hypothesis 2: Intermittent breaks during tasks will increase task performance.

3.3 Work-week effects

Now, we focus on the impact of workweek effects on inspection outcomes. We examine the work-break schedule, where workweeks are sandwiched by fixed breaks, i.e., weekends. Our attention is on the start day of inspection tasks during the workweek. The probability and duration of time off during weekends are predictable throughout the workweek cycle. Weekends provide complete disengagement from work tasks, except for overtime work during weekends, which we exclude from our analysis to examine the workweek effect. Workweeks have proponents on both sides, with several industries in Japan, Singapore, and India still adhering to 6-day workweeks, with industrialists demanding 70-hour workweeks (Varanasi, 2023). Meanwhile, firms in the UK, Australia, and Belgium are piloting 4-day workweeks. We explore inspectors' performance throughout the workweek by examining how the outcomes of inspections vary with the start day of the inspection during a regular five-day workweek. We provide theoretical justification for why inspections conducted earlier (versus later) in the 5-day workweek lead to variations in inspector stringency and differences in violations detected.

Previous research has recognized the impact of task timing during workweeks on operational outcomes. Dai et al. (2014) and Karimi et al. (2020) suggest that the start of the week acts as a temporal landmark, signifying a 'fresh start' and new 'mental accounting periods' for inspectors, resulting in increased throughput and motivation levels. Dai et al. (2014) argue that organizing operational tasks closer to the beginning of the week, rather than the end, can enhance people's aspirations and motivation, thereby boosting operational performance (Karimi et al., 2020). In our context, this 'fresh start' effect could mean that inspectors are more motivated and productive at the beginning of the week, potentially leading to more thorough inspections, a phenomenon we refer to as post-break productivity.

Psychology literature supports the idea that breaks, such as vacations, can positively influence employees' emotions and mood, which is reflected in their performance after returning to work (Fritz & Sonnentag, 2006). This literature also introduces the concept of a "break fade-out effect," indicating that the positive impacts of breaks diminish shortly after employees return to work, leading their performance-related outcomes to revert to pre-break levels (Westman & Eden, 1997). We propose that regular days off during workweek cycles have a positive initial impact on employee performance upon returning to work, but this effect gradually fades out over the course of the workweek.

Moreover, the psychology literature indicates that employees' anticipation of upcoming breaks at the end of the workweek cycle can further impact their work performance negatively towards the end of the workweek (Writz et al., 2003). For example, inspections initiated immediately post-break, i.e., on Monday, will result in a higher number of citations. However, as the impact of the break fades over the week and anticipation of the upcoming break begins to build, the number of citations is likely to decrease (Fritz et al., 2010). Employees may experience decreased focus and motivation as the week progresses, and anticipation for the next break intensifies. They may become distracted or less engaged, looking forward to their time off. In some instances, employees might rush to complete tasks before their break, prioritizing speed over quality, which could lead to errors or incomplete work, ultimately affecting overall efficiency.

Further, specific to our context, according to Ibanez and Toffel (2020), the work of inspectors involves series of assessment tasks that include both physical activities (such as lifting, bending, or reaching into equipment or storage areas for sample collection or extensive walking and standing while examining various parts of the facility) and mental tasks (such as

reviewing documentation, interviewing facility staff, or applying complex food safety regulations to the facility context). As the week progresses both the physical and mental fatigue tends to increase (Brachet et al., 2012). Mental fatigue is further known to increase physical fatigue (Ibanez & Toffel, 2020; Wright et al., 2007). Throughout the week, this accruing fatigue can greatly diminish inspector's physical and cognitive abilities impacting their diligence for tasks (Ibanez & Toffel, 2020); for example, equipment and process verification which involves physical movement around the facility and the cognitive efforts of confirming that it is operating in compliance with regulatory standards and making extemporaneous decisions. As the inspectors become more fatigued over the week, they might be more liable to accept the status quo (Danziger et al., 2011; Muraven & Baumeister, 2000) resulting in overlooking certain violations.

Further, inspection tasks that are very procedural might result in disengagement, negatively impacting inspectors' cognitive arousal as the week progresses (Hendey et al. 2005; Karimi et al., 2020). Inspectors might also face resistance from facility managers who intend to protect their firm's reputation. Regulatory violations might impact the perception of their stakeholders including customers and partners. Such backlash can increase mental fatigue. As the week progresses, fatigued inspectors may prefer to avoid confrontations, potentially leading them to a more lenient approach. (Ibanez & Tofell, 2020).

In addition, food regulatory bodies are severely resource constrained resulting in a shortage of food inspectors. Shortages in staff might lead to longer hours, more work, and frequent overtime, increasing the risk of fatigue and burnout among inspectors. This can negatively affect their performance (Yestiana et al. 2019). For instance, Bradley and McAuliffe (2009) found that staff shortages and high workloads can harm public health supply chains. In

our case, inspectors are likely to become more fatigued as the week goes on, which may reduce their thoroughness. This implies that inspections done earlier in the week might yield better results.

However, according to Yao et al. (2019), weekend breaks are viewed as downtime in operational systems, effectively acting as "bottlenecks" that interrupt the continuous flow of operational routines and create pauses in operations. As a result, weekends might slow down operational flow for two days, potentially affecting performance in the early part of the subsequent week, a phenomenon often referred to as the "Monday Effect." Additionally, behavioral science research suggests that individuals' (inspectors') moods fluctuate throughout the week (Pettengill, 1993). For those who work from Monday to Friday, the start of the week can seem particularly daunting (Fritz & Sonnentag, 2005; Yao et al., 2019). For instance, Pettengill (1993) observed that investors tend to be more cautious on Mondays, favoring safer investments. Similarly, Fritz and Sonnentag (2005) identified that stress experienced over the weekend can impact work performance at the beginning of the week. Furthermore, the early week is often considered a "warm-up" period for inspectors, influenced by both their routines and mood (Yao et al., 2019). Performance tends to stabilize as the week progresses (Yao et al., 2019). The disruption caused by weekends often leads to less efficient inspections early in the week, resulting in lengthier inspections and fewer completed checks as operations attempt to regain momentum post-weekend (Yao et al., 2019).

However, we take the positive side from scheduled breaks and propose:

Hypothesis 3: Task performance will be highest at the beginning of the workweek, decreasing during the workweek (from Monday to Friday).

3.4 Non-standard work-break schedules

Next, we discuss the effects of non-standard work-break schedules on inspection outcomes. The Office of Personnel Management (OPM), an independent U.S. agency that oversees various human resource functions of federal agencies such as FDA, defines a "nonstandard work-break schedule" as *“any schedule in which full-time employees work other than the standard schedule of 8 hours per day and 5 days per week in an administrative workweek [Monday to Friday].”* Under the Fair Labor Standards Act (FLSA), which establishes regulations for wages, overtime pay, and working hours, the FDA's consumer safety officers (the official title for food inspectors) are classified as exempt employees. This means they are not entitled to overtime compensation for working more than 40 hours per week and are usually paid a salary. However, they may receive compensatory time off for additional hours worked, including weekends, with their supervisor's approval. Given the workload of field staff, obtaining such compensatory time can be challenging.

In non-standard work-break schedules, work encroaches on breaks, reducing their length or completely overtaking them. This results in employees' engagement with work being renewed as work infringes upon their personal time. Van Dyne and Ellis (2004) conceptualized this phenomenon as “job creep” to describe such situations where “employees feel ongoing pressure to do more than the requirements of their jobs” and escalating citizenship behaviors informally become a part of employees’ formal duties (Bolino et al., 2013). This phenomenon often affects employees who work outside of traditional work hours, including evenings, weekends, or during what was previously considered personal time, potentially leading to a decrease in number of citations. In the analysis of the job concerns of FDA’s food inspectors one of the comments included:

“Field work is understaffed. Expect overtime that is not clearly stated in the job description...Additional overtime is required to meet assignment due dates or short reporting timeframes for violative inspections. The routine overtime...causes people to breakdown with fatigue or leave without even having a job to go to.”

Non-standard work-break schedules can blur the boundaries between work and personal time, leading to an increased workload that may become unsustainable. Bolino et al., (2013) suggest that job creep makes employees believe that their personal freedom is compromised. Such factors negatively impact employees' motivation and engagement. Consequently, this diminished motivation and engagement can lead to fewer citations being issued during inspections conducted under non-standard work schedules.

Secondly, working over non-standard work-break schedules can lead to higher error rates leading to missed violations. Thoughts of missing out on social obligations, coupled with guilt over not being with family, contribute to social isolation. This underscores how one's work schedule is out of sync with the conventional work-break patterns observed by other people, including family and friends, along with general social norms. The discord between an individual's work responsibilities and the usual leisure times of their social circle, especially such distractions combined with an increased workload, can result in inspectors missing violations and demonstrating a lack of thoroughness, ultimately finding fewer issues.

Hence, we propose:

Hypothesis 4: Work tasks on non-standard days (i.e., weekends and holidays) will decrease task performance.

3.5 Concurrent work schedules

Concurrent work schedules entail managing multiple tasks (inspections) at the same time within a specific timeframe (Goes et al., 2018). These schedules necessitate sharing time among various tasks, requiring employees (inspectors) to split their focus, time, and efforts across multiple tasks (inspections). In such work-break schedules, tasks overlap without formal breaks in between Kc (2014). Concurrent work schedules are common for field staff due to increased workload. However, the literature on the effects of concurrent schedules on employee productivity is mixed.

One part of the literature suggests that multi-tasking has a positive impact and leads to an increased productivity and quality of task performance (Altman et al., 2021). Employees may be able to utilize the downtime resulting from delays or interruptions within an ongoing task. This limits the idle time between tasks. This not only improves the employee utilization rate but also keeps the inspectors actively engaged, enhancing their alertness and attentiveness to details.

Such an active engagement strategy results in a dual benefit: it boosts productivity by ensuring that time is utilized efficiently, and it improves the quality of inspections due to the capacity to identify issues. Prior literature suggests that flexibility has made employees more efficient, suggesting that concurrent schedules can lead to more thorough inspections and a greater number of identified citations.

Another set of literature suggests that multi-tasking negatively impacts task performance (Kc, 2014). Theoretically, during multitasking, individuals focus on one task at a specific time but switch between tasks without completing the previous task. This switching between tasks has a switching cost in terms of time and efforts of an employee as there is a warm-up period for every task. Even when individuals switch to the new task, part of their focus remains on the previous task. Moreover, when individuals shift between tasks without completing the previous

one it results in a feeling of dissatisfaction which affects their mood. This negatively affects task performance.

Several studies have established that individuals struggle to execute concurrent tasks effectively. The lack of attention on a single task result in an increased error rate. Multitasking results in cognitive blocks. In essence, switching between tasks results in a series of sequential cognitive tasks. Charron and Koechlin (2010) discovered that individuals slow down and make numerous errors while multitasking in laboratory settings.

Hence, we take a side and propose that:

Hypothesis 5: Engaging in concurrent tasks will reduce task performance.

4. Methods

4.1. Research Setting: food-safety inspections

We tested our hypotheses in the setting of food-safety inspections conducted by the Food and Drug Administration (FDA). FDA is responsible for overseeing about 80% of food supply chains in the U.S. It does so by conducting food safety inspections of food facilities that manufacture, process, pack, or hold food for consumption in the United States (FDA, 2023). These inspections are conducted by individual inspectors or a team of food inspectors. However, in our study, we focus on inspectors who work individually. This is because in team inspections it is challenging to discern the individual behaviors from those of their colleagues (Ibanez & Toffel, 2020). These food inspectors, also called the consumer safety officers, monitor the compliance of the food facilities in their assigned geographic areas to minimize food borne illnesses. The effectiveness of these inspections depends on the accuracy and completeness of these inspectors (Ibanez & Toffel, 2020).

Inspection inaccuracies can lead to significant health, social, and economic impacts on both consumers and businesses. Despite the monitoring, foodborne illnesses affect approximately 48 million people each year in the U.S., resulting in 128,000 hospitalizations and 3,000 deaths, according to the CDC's 2018 report. Additionally, the economic burden of foodborne illnesses in the United States, including medical expenses, legal fees, public health investigations, product recalls, loss of firm reputation and income, is estimated to be between \$20 and \$43 billion annually, as reported by Hoffman and Ahn (2020).

Unlike, prior literature on local health department inspections which focus on food retailers, we study FDA inspections as they focus on food processing facilities, including the manufacturing processes, hygiene of production lines, and warehouse storage conditions – collectively termed *manufacturing quality management*. On the other hand, food retail inspections focus on restaurants and grocery stores investigating conditions at the point of sales and quality of the retail environment. Food processing facilities offer several advantages for assessing the impact of inspector scheduling on inspection outcomes such as a.) more controlled environment compared to retail settings which have a much broader scope (Ibanez & Toffel, 2020), b.) wider range of quality issues providing more nuanced understanding, c.) standardized procedures and protocols providing consistency, and d.) more opportunities to proactively influence product quality and minimize risk in the upstream food supply chain.

Several studies have examined FDA inspections of manufacturing facilities. For example, Ball et al., (2017) found that FDA medical device inspectors with more site-specific experience become less effective at finding violations over time because of familiarity and relational ties with the facilities. They suggest investigator rotation and effective scheduling as policies to make inspections more effective. Anand et al., (2011) used FDA pharmaceutical inspections to

illustrate how these inspections act as external renewals that halt the decay in operational routines used to maintain product quality. They found that external shocks such as mergers lead to decreased inspection performance. Further, Gray et al., (2015) used FDA pharmaceutical inspections to identify how R&D and manufacturing co-location of firms helps them with better conformance to quality standards as determined through inspection outcomes.

Other researchers have specifically focused on inspections of food processing facilities. For example, Powell et al., (2013) compared food safety inspections of processing facilities and suggested how regulatory inspections need to be augmented by regular internal audits to maintain a strong food safety culture within processing facilities. Starbird (2005) examined several food safety inspection policies and suggests that there is uncertainty inherent to sampling inspections that undermines confidence in regulatory inspection outcomes. Buckley (2015) studied the impact of food safety inspections on small food processors by qualitatively studying the interactions between inspectors and food processors. They found that if inspectors used a collaborative-assistive approach, the results would include better compliance by small food facilities.

However, inspection outcomes can be affected by several factors. Since inspectors need to provide evidence for violations they find, they often report fewer actual problems to avoid, potentially, over reporting, as noted by Ibanez and Toffel in 2020. Several studies on inspection bias (e.g., Bennett et al., 2013; Duflo et al., 2013; Short et al., 2016) assume that “deviations from the true number of violations are due mainly to under detection and that bias does not generally lead inspectors to cite nonexistent violations” (as in Ibanez & Toffel, 2020). Also, while FDA inspectors initially note their observations on-site, these notes are reviewed and approved by the FDA's regional office (as mentioned in FDA's Investigator Operations Manual 2024). Further,

companies have the right to dispute the findings of the inspectors, which makes inspectors hesitant to report issues, as mentioned in the FDA's 2022 report. As a result, there's a tendency for inspectors to under-report violations – “inspector bias takes the form of underreporting the violations” (Ibanez & Toffel, 2020).

4.2.Data

Our data are derived from FDAlytics, a company that gathers facility level food-safety inspection reports from FDA. FDAlytics is a regulatory compliance database firm that extracts information from enforcement reports of food inspections via Freedom of Information Act (FOIA) requests to FDA. It uses this information to create data dashboards and inspection intelligence platforms, selling them to food manufacturers that are interested in gaining insights into regulatory compliance. The FDAlytics dataset includes information about the inspected food facility (name, a unique number, address with city, state, zip code), the inspector, and the start and end date of inspections. We augment this database by gathering information from the publicly available dataset from FDA inspection dashboard using unique FEI identifiers to include number of citations, inspection depth and severity of citations. Further, we use publicly available information on FDA inspectors, including their start date with FDA, age, and position grade. We also use the Namsor database to determine the gender of inspectors. This approach employs artificial intelligence (AI) to analyze name structures to determine factors related to inspectors, such as origin, ethnicity, gender, and residence country. We corroborate the veracity of results via LinkedIn.

The FDAlytics' food inspection datasets include inspection start and end times, as well as names of inspectors. Our database includes all recorded food facility inspections in the U.S. from January 2013 to March 2020 (right before COVID-19 impacted inspections). From 67,854

inspections in the raw dataset, we exclude 116 duplicates attributed to data errors. Further, we drop 12,262 inspections that were conducted by more than one inspector, i.e., inspections conducted by a team. We then exclude 10,305 inspections for which we were not able to calculate relevant variables (for example, inspector experience) or were missing lagged data necessary to operationalize prior inspection outcomes and prior inspection facilities outcomes. This yields a sample of 45,171 inspections. Poisson Pseudo-Maximum Likelihood (PPML) fixed effects dropped 37,707 inspections based on facilities that had (a) just a single inspection, called singletons (Correia, 2015), or (b) multiple inspections all citing zero violations allowing no variation to study within-firm changes. This resulted in a sample containing 7,464 inspections of 3,291 food facilities (of firms) conducted by 1,070 inspectors.

To get a better understanding of the context, we a.) interviewed FDA officials including the senior advisors and directors in the office of food safety, b.) studied the FDA's Investigations Operational Manual 2024 and c.) analyzed the feedback of 80 FDA Inspectors on Glass Door⁴ from 2010 to 2023. Our interviews with FDA officials augmented by the Glass Door data suggest that FDA Inspectors have limited control over scheduling. We conducted a thematic analysis of the feedback from these 80 FDA inspectors by first categorizing each mentioned feedback around scheduling and then calculating the frequency of mentions for each category.

Some of the comments of FDA inspectors included *"Not a lot of control over your schedule"; "decisions come down from above, no ability to negotiate (it is how it is)"; "Expect to be asked to travel [for assignments] with as little as 24 hours' notice"; "busy workload with inflexible deadlines"; "assignments seem to be also doled out willy-nilly."* Despite these

⁴ Glass Door is a website based in the U.S. It is used by current and former employees of an organization to anonymously provide reviews as a service to those who are making career decisions. Glass Door "collects company reviews and real salaries from employees of large companies and displays them anonymously for all members to see."

comments, it is acknowledged that inspections assignment could still be non-random, and thus we test for bias in the identification section of the paper.

FDA follows a risk-based inspection schedule where high risk facilities are required to be inspected at least every three years and non-high-risk food facilities are required to be inspected at least every five years. So, most inspections are scheduled based on the facilities' due dates and compliance histories. As mentioned in the FDA Operations Manual 2024, "firm's inspectional, compliance, and recall history" is considered before scheduling inspections. We control for a firm's prior quality performance as a proxy for its risk profile.

Although inspectors conduct administrative tasks, for example compiling reports, most of the time is involved in conducting inspections and travelling to inspection sites. Per the FDA Operations Manual, during the start of the inspections, inspectors engage with the person in-charge at the facility and encourage them to accompany the inspector during the inspection. They walk through the facility acquainting themselves with the layout and setup. Throughout the inspection, they observe employee practices, interview personnel, collect samples and make notes of any violations. Upon completion of the inspection, inspectors review their findings with the facility in-charge and highlight violations by issuing a form 483. They also provide an opportunity for clarifications and discuss potential corrective actions along with the next steps. Thus, underreporting of violations can have significant ramifications for the safety of the food supply chain.

4.3.Measures

4.3.1. Dependent and independent variables

Our main dependent variable is *violations*. For our analysis, we measure it as the number of violations cited in each inspection as used by similar studies (e.g., Ibanez & Toffel, 2020; Scott et al., 2021; Short et al; 2016).

Pseudo is measured as the time interval in days between current inspection and the inspector's earlier assigned inspection or in other words it is the time since the last inspection ended. We winsorized the variable at 0.1th and 99.9th percentile to reduce the effect of potential outliers. Since the data was right skewed, we take a natural log by adding one to our data to accommodate the zero values, i.e., no days between inspections. Further, we mean center the *Pseudo* variable to minimize collinearity issues with the quadratic term of *Pseudo*² in our model (Mayo et al., 2024). *Pseudo* is the linear part of the independent variable for Hypothesis 1. *Pseudo*² is the square of mean-centered *Pseudo* (winsorized and logged as above) which is the quadratic portion of the independent variable for Hypothesis 1.

Concurrent scheduling of inspections is measured as the number of inspections assigned to an inspector at the time of the current inspection, i.e., number of inspections an inspector will oversee simultaneously, or when a second assignment is made before the first is completed. We further winsorize the variable at 0.1th and 99.9th percentile and log it to deal with the right skewness. We code the *Start-day* as the day of the week the current inspection started. We focus on the start-day and not the end-day of the inspection because the initial few days are used mainly to inspect, and the end-days of the inspection are used by inspectors to compile their reports and share their findings with the facility managers to help them understand the next steps. An inspector's regular work week is from Monday to Friday, and they are discouraged to work overtime. We code the *Start-day* as a dummy variable for the day of the week when the inspection started.

We measure *Intermittent* inspections as inspections that are affected by a weekend or holiday, i.e., inspections that span across weekends and federal holidays. These inspections start on a regular or standard workday and are interrupted by a weekend or holiday only to be completed after the weekend or holiday. We code *Intermittent* inspections as 1 if the inspections span across weekends/holidays and zero otherwise. However, these spanning inspections could also suggest that inspectors worked on the weekend. To rule out that possibility we confirmed the operational hours of the sub-sample of the holiday/weekend spanning inspected facilities from their respective websites and Google Business Profile.

A *Non-standard* work schedule includes inspections in which the inspectors worked other than the standard schedule of an administrative workweek of Monday to Friday (U.S. Office of Personnel Management, 2024). We code *Non-standard* work schedule as 1 if the inspection was started on a non-standard work schedule, i.e., weekends (Saturday or Sunday), Federal Holidays (including Christmas and Thanksgiving breaks), or during a government shutdown. Otherwise, it is coded as zero. We identify *Non-standard* inspections by the start day of the inspections.

4.3.2. Control Variables

We control for inspector *experience* as it can affect the skill, diligence, or efficiency with which inspections are conducted (Short et al., 2016). This control helps reduce bias in the estimated effect of inspection schedule on the inspection outcomes accounting for the heterogeneity among inspectors. It also helps address endogeneity concerns of non-random inspection assignments where more experienced inspectors might be assigned to more complex inspections. We measure it as the number of inspections an inspector had conducted at any food facility from the beginning of our sample period till the start of the current inspection (as in Ibanez & Toffel, 2020; Lapre & Tsikriktsis, 2006; Lapre & Scudder, 2004).

We control for *average time* taken by an inspector to complete inspections prior to the current inspection. Despite the training, inspector's own personalities impact inspections. Different inspectors might have different styles and pace of conducting inspections. Some might inherently take longer or be slower. It might also reflect the depth and thoroughness with which an inspector conducts an inspection. We measure it as an average of all the durations of inspections performed by the inspector from the beginning of our sample period prior to conducting the current inspection (in days).

We control for the *gender* of the inspectors. Prior research suggests that there can be behavioral differences among male and female inspectors during inspections which affect their decision making (Short et al., 2016). For example, females might be more nuanced whereas males might be more strategic in their approach towards inspections (Short et al., 2016). Further, including gender helps control for differences in training, mentorship or variations in inspection assignment based on gender biases in the organization. We create a *Gender* dummy and code it as 1 for males and 0 for females.

As in prior papers on inspections such as Ball et al., (2017) and Ibanez and Toffel (2020), we control for the *site-specific experience* of the inspectors. This is measured by the number of inspections the inspector has conducted in the current facility prior to the present inspection. It helps control the familiarity of the inspectors with the current facility and its employees. Prior research suggests that the greater the familiarity of the inspectors with the inspected facility, the less they cite the facility because of their relational ties that develop over time.

We also control for the inspection outcomes of the last facility the inspectors visited – *prior inspection outcomes*. Prior research suggests that, owing to negativity bias, inspectors are affected by the outcomes of their last inspection citing more violations at their current facility

after they inspected a facility that had a higher number of violations (Ibanez and Toffel, 2020). We measure it as the number of violations the inspectors found at their last inspections prior to the current inspections.

Our interviews with the FDA officials suggested that the main criteria leveraged by inspector supervisors to decide which facility to inspect is the past inspection performance of these facilities – *past inspection outcomes*. We measure it as the number of citations the current facility had in their last inspection conducted by any inspector. This helps to control for endogeneity concerns of non-random inspection assignments where worse performing facilities are inspected more often. We also control for the number of recalls the firm had in the year prior to the current inspection – *recalls last year*. A recall, as defined by FDA, is a way of removing or correcting products that violate FDA laws and regulations. This helps us to control for the quality performance of facilities in the recent past. A greater number of recalls can shape the perception of these inspectors that might affect the inspection outcomes. Recalls in the recent past might also suggest that firms might have undertaken improvement efforts to address the identified issues.

We control for the *inspection depth* of the current inspection. This variable reflects the intensity and thoroughness with which the current inspection is conducted. Inspections that are more in-depth are likely to reveal more issues, leading to a greater number of citations. Further, FDA utilizes risk-based strategies, where the intensity of an inspection is influenced by the perceived risk level of the facility. High-risk facilities might get more in-depth inspections. Adding *inspection depth* allows the model to adjust for such risk-based strategies. FDA suggests that there are 6 possible areas that a food inspection can cover, i.e., pesticides and chemical contaminants, biological hazard, natural toxins, food composition, standards and labelling, food

and color additives, and food handling and storage practices. This variable is measured as the number of areas that were covered in the current inspection by the inspectors, e.g., if the inspector covers all 6 areas in their inspection it is coded as 6.

Further, we control for the *inspection duration* of the current inspection. Duration is the number of days the inspector takes to complete the current inspection. Duration reflects the complexity of the facility, or the operation being inspected. More complex facilities might require longer inspections due to the number of processes, the size of the operation, or the type of products being manufactured. Controlling for duration of the inspection allows the model to account for this complexity, ensuring that outcomes are not biased by the innate challenges of inspecting more complex facilities.

Lastly, we control for the *political affiliation* of the Commissioner of FDA. Since the FDA commissioners are political appointees, their strategical approach to food safety is shaped by the philosophy of their political parties. They are also driven by the mandate of their political parties. For example, when it comes to policy restraints on the industry, the Republican party generally has a more moderate stance than the Democratic party which could affect the outcomes of the inspections. We measure it as a political affiliation dummy to indicate whether the FDA commissioner was a Republican or Democrat appointee (Republican (0) or Democrat (1) appointee).

Table 2 presents summary statistics of the dependent, independent and control variables. On average, the food facilities undergo 3.12 inspections in the sample. These inspections result in an average of 2.59 violations per inspection.

Table 2. Summary Statistics

Variable	Description	Mean	Standard deviation	Minimum	Maximum
----------	-------------	------	--------------------	---------	---------

<i>Violations</i>	Number of violations cited in each inspection	2.59	2.94	0	26
<i>Pseudo</i> (Hypothesis 1)	Time since the last inspection was allocated to the inspector (days)	68.78	131.05	0	379
<i>Concurrent</i> (Hypothesis 2)	Number of inspections assigned to an inspector at the time of current inspection	1.28	0.87	1	38
<i>Intermittent</i> (Hypothesis 3)	1 if the inspections spans across holidays and zero otherwise	0.32	0.47	0	1
<i>Start-day</i> (Hypothesis 4)	Variable for the day of the week when the inspection started	2.57	1.23	1	5
<i>Non-standard</i> (Hypothesis 5)	1 if the inspection was started on a non-standard work schedule and zero otherwise	0.03	0.15	0	1
<i>Experience</i>	Number of inspections an inspector conducts at any food facility from the start of our sample period till the current inspection	56.52	49.66	1	390
<i>Average time</i>	Average of all the durations of inspections performed by the inspector from the start of our sample period till the current inspection (in days)	3.69	2.41	0	21.57
<i>Gender</i>	Gender dummies coded as 1 for males and 0 for females	0.52	0.50	0	1
<i>Site-specific experience</i>	Number of inspections the inspector has conducted in the current facility till the current inspection	0.23	1.18	0	47
<i>Prior inspection outcomes</i>	Number of violations the inspectors found at their last inspection prior to the current inspection	1.59	2.71	0	30

<i>Past inspection outcomes</i>	Number of citations the current facility had in their last inspection conducted by any inspector	1.71	2.93	0	26
<i>Recalls last year</i>	Number of recalls the firm had in the year prior to the current inspection	0.02	0.14	0	3
<i>Inspection depth</i>	Number of areas that were covered in the current inspection by the inspectors	1.24	0.51	1	5
<i>Inspection duration</i>	Number of days the inspector takes to complete the current inspection	3.12	3.35	0	10
<i>Political affiliation</i>	Dummies to indicate whether the FDA commissioner was a Republican (0) or Democrat (1) appointee	0.54	0.50	0	1

Note. $N = 7,464$

4.4. Empirical specification

We test our hypotheses by estimating (via quasi-maximum likelihood) Poisson Pseudo-Maximum Likelihood (PPML) while accommodating high-dimensional fixed effects which predict the number of violations cited in an inspection. PPML is designed to estimate count models useful for dealing with large panel datasets with high dimensional fixed effects, heteroskedasticity, and potential overdispersion. It is especially helpful in contexts when we need to control for unobserved heterogeneity across several dimensions, i.e., multi-way fixed effects (e.g., month, years, FDA divisions, facilities' nth inspection, facilities). PPML is used in econometrics when dependent variables are count data. It is designed to handle high dimension fixed effects efficiently where traditional Poisson or Negative Binomial models may require approximations to handle fixed effects leading to biases and computational inefficiencies. PPML

accommodates overdispersion through robust standard errors where traditional negative binomial models are computationally inefficient while dealing with many fixed effects. Our unit of analysis is an inspection of a facility started on a particular day (a facility inspection). Our specification exploits the panel structure of data where inspection can be viewed as the n^{th} inspection of the facility i in our sample.

Our independent variables include (1) *Pseudo* and *Pseudo*² work-break schedule – temporal spread or distribution of inspections; (2) *Concurrent* schedule – number of simultaneous inspections; (3) *Intermittent schedule* – holiday spanning inspections (4) *Start-day* of inspections – the day of the week the inspection started; (5) *Non-standard* work schedule inspections – inspections started during weekends or holidays.

We include a series of fixed effects to control for *facilities' nth inspection*. Prior research suggests that facilities improve their compliance over their successive inspections (Ko et al., 2010; Ibanez and Toffel, 2020; Toffel et al., 2015) We also created a set of indicator variables to account for the order of facilities inspections, varying from the 2nd to the 10th or beyond. Each variable identifies if an inspection is the facilities second, third, and so on, inspection within our study period. The first inspection serves as the baseline category from which comparisons are made.

Further, we include fixed effects for the FDA divisions. FDA's Office of Human and Animal food Operations (OHAFO), responsible for executing inspections, has twelve divisional offices. These twelve divisions are regional offices that oversee the operations of their divisions. A division consists of 3-4 states, each and has their own divisional leadership, i.e., Division Director, Director of Compliance Branch and Director of Investigations Branch, to which the FDA inspectors (Consumer Safety Officers) report. In our analysis of FDA Inspector's

comments, one of them indicated “*The workload disparity between western and eastern divisions is unbelievable.*” Since the divisions may have variations in approaches to scheduling and inspections, we control time-invariant characteristics of these divisions by including *FDA division* fixed effects.

We also include fixed effects for *months* in which the inspection was conducted to control for seasonality. We include fixed effects for the *years* in which the inspections were conducted to control for the unobserved heterogeneity and probable confounding variables that differ over time but are constant through entities each year. Further, we include the facility fixed effects to account for differences between facilities such as management, company culture, operational practices, quality management systems, or size to ensure that the estimated effects of interest are not confounded by these unobserved factors. We include several inspector level controls to ensure that we capture the effects of inspector experience, gender, efficiency, facility familiarity, and other biases that might affect their average stringency. We also cluster the standard errors by inspectors to account for the possibility that inspection outcomes might not be independent within the same inspectors. It also helps to ensure that we account for correlation in the error terms for inspections conducted by same inspector – overdispersion and serial correlation. Clustering by inspectors also helps to mitigate the effects of heterogeneity on the standard errors. Moreover, inference is more robust to heteroskedasticity and autocorrelation within clusters (inspectors).

We estimate the following model to evaluate the effect of different scheduling regimes on the inspection outcomes of the facilities.

$$\begin{aligned}
Violations_{ni} = & \alpha_0 + \alpha_1 Pseudo_{ni} + \alpha_2 Pseudo^2_{ni} + \alpha_3 Concurrent_{ni} + \\
& \alpha_4 Interrmitent_{ni} + \alpha_5 Startday_{ni} + \alpha_6 Nonstandard_{ni} + X_{ni}\alpha + \gamma_t + \theta_{m(t)} + \tau_{l(ni)} + \\
& \mu_i + \rho_{i(n)} + \varepsilon_{it}
\end{aligned} \tag{1}$$

Here, *Violations* is the outcome variable, with n signifying the inspection and i indicating the facility. The X vector encompasses a range of controls variables. The term ε_{it} captures the error terms, and α stands for the parameters being estimated.

γ_t – are the year fixed effects

$\theta_{m(t)}$ – are the month fixed effects for the month m in year t in which the observation was made

$\tau_{l(i)}$ – are the FDA division fixed effects for FDA division l in whose jurisdiction facility i located

μ_i – are the facility fixed effects

$\rho_{i(n)}$ – are the facilities' n th inspection fixed effects

5. Results

5.1. Model results

We estimate the count model using fixed-effects Poisson Pseudo-Maximum Likelihood (PPML) regressions where the dependent variable is the number of citations during each inspection. We report the correlations among variables and primary results with standard errors clustered by inspectors in Table 2 and Table 3 respectively. PPML extends the Poisson regression allowing for consistent estimation accounting for presence of overdispersion using the multiway fixed effects (Chang et al., 2022; Elfenbein et al., 2023). Our results are robust to different alternatives: clustering standard errors by facilities including inspector fixed effects,

estimating the model negative binomial regression with conditional fixed effects, and using ordinary least squares (OLS) regression to predict log violations to estimate the model.

Further, multi-collinearity is not a concern given that the variance inflation factors (VIF) are less than 1.82 (mean VIF = 1.34) for all hypothesized variables and other control variables. As with other studies, such as Chen et al. (2016), Ibanez and Toffel (2020), and Short et al., (2016), we interpret the coefficients of the hypothesized variables as “evidence of bias”, as our model specification controls for a several factors that affect the number of violations detected. Further, we assume that the deviations from the actual number of violations are due to under detection resulting from the scheduling bias. We interpret negative coefficients as indicating the level of under detection and positive coefficients as the degree to which under-detection is minimized. We infer the effect sizes using the incidence rate ratios (IRR).

Table 3 shows the results for the empirical tests exploring hypotheses 1-5. We first present the results of Model 1, which is a control-only model. We focus on the inspector control variables. The estimated coefficient of inspector *experience* is not significant and is consistent with prior literature which finds very small changes if any. The *average time* that an inspector spends conducting inspections prior to the current inspection is positive and significant ($\beta = 0.045$ $p < 0.001$, IRR = 1.046) indicating that inspectors who have a history of spending more time at a facility conducting inspections, issue 4.6% more citations in their current inspection as compared to their colleagues. When it comes to *gender*, the coefficient is negative and significant ($\beta = -0.133$ $p < 0.001$, IRR = 0.875) suggesting that males inspectors detect 12.5% fewer violations than female inspectors which is consistent with prior literature (Short et al., 2016). For site-specific experience, we find that the coefficient is negative and significant ($\beta = -0.067$ $p < 0.05$, IRR = 0.935) suggesting that inspectors who are more familiar with the current

inspection site tend to find 6.5% fewer violations as compared to their colleagues, which is consistent with prior literature (Ball et al., 2017; Ibanez & Toffel, 2020). Further, the coefficient of the outcome of an inspector's *prior inspection* is positive and significant ($\beta = 0.034$ $p < 0.001$, IRR = 1.035) suggesting that each additional citation at a facility inspected immediately prior to the current inspection increases the number of violations cited in the current inspection by 3.5%, consistent with prior literature (Ibanez & Toffel, 2020).

Next, we focus on the facility and inspection control variables. For *past inspection outcomes*, i.e., the number of citations for a facility in their immediate last inspection, there is a negative and a significant coefficient ($\beta = -0.082$ $p < 0.001$, IRR = 0.921). This suggests that each additional citation at the current facility's immediate last inspection leads to a 7.9% decrease in citations in their current inspection, likely from a learning effect. Each inspection acts as an external shock that halts the decay in operational routine of these facilities (Anand et al., 2012). However, we find no evidence that the *recalls* from last year have an effect on the current inspection's outcomes. Further, we find that the *inspection depth* of the current inspection has a positive and significant coefficient ($\beta = 0.173$ $p < 0.001$, IRR = 1.189) suggesting that with a unit increase in inspection depth, inspectors find 18.9% more violations in the current inspection. We also find that the *inspection duration* of the current inspection has a positive and significant coefficient ($\beta = 0.024$ $p < 0.001$, IRR = 1.024) suggesting that which each additional day of the inspection the number of violations cited in the current inspection increase by 2.4%. The *political affiliation* of the FDA commissioner is not significant.

We use the results of model 2 from Table 3 for interpreting Hypothesis 1. *Pseudo* is positive and significant ($\beta = 0.028$ $p < 0.001$, IRR = 1.028), whereas *Pseudo*² is negative and significant ($\beta = -0.012$ $p < 0.001$, IRR = 0.988). This implies an inverted U-shaped relationship

between the pacing of inspections and the number of citations at a current inspection, supporting Hypothesis 1. Thus, extending the interval between consecutive inspections initially leads to an increase in the number of violations detected by an inspector at a facility, but eventually results in a decrease in the number of citations, indicating diminishing returns. Refer to figure 1.

For the rest of Hypotheses 2 through 5 we use model 6 to interpret results. For concurrent inspections we find that the coefficient is negative and significant ($\beta = -0.159$ $p < 0.001$, IRR = 0.853). This indicates that for every additional inspection an inspector conducts simultaneously with the current inspection, the number of citations at a focal firm decreases by 14.7%. This provides strong support for Hypothesis 2 and indicates that multitasking with inspections reduces the efficacy of the inspectors in detecting issues.

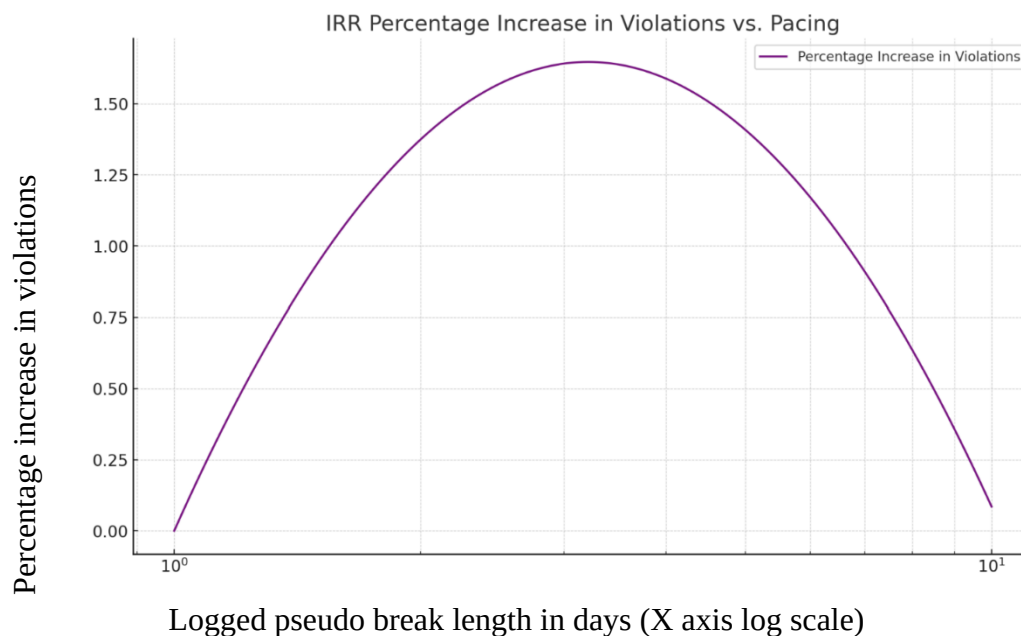


Figure 1. Percentage increase in the number of violations detected by length of pseudo breaks (in days where x axis is a log scale)

Applying this 14.7% effect to the 2.59 average violations per inspection results in an average marginal effect of 0.38 fewer violations being cited for each inspection in addition to the focal inspection. This amounts to about 380 violations not cited for every 1000 inspections when an inspector is conducting 1 additional inspection simultaneously with current inspection and 761 violations not being cited for every 1000 inspections conducted by inspectors who are working on 2 additional inspections.

For Hypothesis 3 we find the coefficient of intermittent inspections to be positive and significant ($\beta = 0.581$ $p < 0.001$, IRR = 1.788). This suggests that the inspections that are interrupted by holidays, forcing inspectors to take a schedule pause during an inspection. Results in a 78.8% increase in the number of citations, compared to multiday inspections with no scheduled pauses or inspections completed in a single day. This result is achieved after controlling for the duration of the inspection. This result supports Hypothesis 3 that multiday inspections with scheduled pauses are much more effective in helping inspectors detect violations. However, we find that the coefficient of *non-standard* work week inspections is negative and significant ($\beta = -0.247$ $p < 0.001$, IRR = 0.781). This suggests that inspections that are scheduled during a *non-standard* work week, i.e., during a weekend or a holiday, result in a 21.9% decrease in the number of citations as compared to inspections that are conducted on a standard administrative work week from Monday through Friday. This results in an average marginal effect of 0.57 fewer violations cited per inspection i.e., 570 violations not cited for every 1000 inspections when an inspector is conducting the inspection during a non-standard work week. This provides strong support for Hypothesis 5.

Further, for *start-day* we find that the coefficient of inspections conducted on a Tuesday, Wednesday, Thursday, and Friday are negative and significant as compared to a Monday.

Tuesday ($\beta = -0.076$ $p < 0.05$, IRR = 0.927), Wednesday ($\beta = -0.152$ $p < 0.001$, IRR = 0.860), Thursday ($\beta = -0.221$ $p < 0.001$, IRR = 0.802), and Friday ($\beta = -0.312$ $p < 0.001$, IRR = 0.732). This suggests that inspections conducted by inspectors starting on a Tuesday result in 7.3% fewer citations (compared to inspections started on Monday). The citations decrease as the week progresses, with 14% fewer citations on Wednesday, 19.85% fewer on Thursday and 26.8% fewer on Friday, all compared to Monday's count. This strongly suggests that the inspections conducted early in the week result in a greater number of citations as compared to later in the week providing support for Hypothesis 4. This result is achieved after controlling for the duration of the inspection. So, for an inspection started on a Friday, 694 violations are missed for every 1000 inspections as compared to an inspection started on Monday.

5.2. Identification

We examine each hypothesized variable in the food inspection context to identify potential sources of endogeneity. We believe that we control for several relevant variables significantly decreasing the chances of omitted variable bias. Further, we believe that it is unlikely that our variables have issues of simultaneity or reverse causality. For example, it is unlikely that the number of citations found in a inspector's current inspection influences how much time has passed since the inspector's last inspection. Further, pacing is determined by the timing of the inspection rather than the inspection outcomes. However, selection bias might be a concern as the scheduling of inspections might not be random and could be influenced by firm characteristics such as firm's compliance history. We use several inspector level controls and facility fixed effects to control for time-invariant and unobservable characteristics of the facility. Below, we discuss further empirical approaches to address this endogeneity concern.

We took several steps to ensure that our empirical approach examines our hypothesized relationships by either controlling for or ruling out other possible explanations. For example, the temporal distribution of inspections for an inspector (pacing) could affect the inspection outcomes, i.e., for an inspector, the inspections scheduled further apart from each other may yield a greater number of violations up to a certain point beyond which there could be diminishing returns (Hypothesis 1). But this could also be possible if inspectors clustered their schedules by the facilities that expect to yield fewer violations and inspectors schedule challenging inspections further apart to get more time to complete paperwork and work on other assignments. However, our discussions with FDA officials suggested that such decisions are based on geographical proximity rather than prior compliance history of the facility. While violations might be spatially correlated due to geographic clustering, our inclusion of facility fixed effects controls for such facility characteristics (applies to other hypotheses as well). Further, we performed a Poisson Pseudo-Maximum Likelihood analysis, using inspector-days along with other control variables and fixed effects. We predict the time between inspections using the prior compliance history of the firms (prior violations $\beta = -0.03$ $p = 0.007$, prior recalls $\beta = 0.14$ $p = 0.321$) and find no support. On the contrary, we find that as the prior number of violations increases, inspectors reduce the time between the inspections suggesting a bias against our hypothesis.

There could be other explanations for the hypothesis that as the number of concurrent inspections increases for the inspector, the number of violations detected by the inspectors decreases (Hypothesis 2). There could be a possibility that inspectors schedule facilities with better compliance history together assuming that there is less work, and they could be managed simultaneously. Regardless, using the Poisson Pseudo-Maximum Likelihood model and using

inspector-days along with other control variables we predict the number of concurrent inspections using the prior compliance history of the firms (prior violations $\beta = 0.03$ $p = 0.013$, prior recalls $\beta = -0.01$ $p = 0.972$) and find no supporting evidence. Rather, we find that as the prior number of violations increases inspectors schedule a greater number of concurrent inspections suggesting a bias against our hypothesis. This could be because inspectors are on strict deadlines and if they feel it is going to take more time at a current facility, they might multitask during these inspections to complete their work on time.

We believe that there could be other alternative explanations for the hypothesis that inspections that span over a holiday/weekend lead to a greater number of violations (hypothesis 3). One could assume that only inspections of challenging facilities might span over longer periods of time and thus are more likely to be interrupted by a weekend. However, based on our discussions with FDA officials and per the FDA Investigations Operations Manual 2024, inspectors have limited control over their schedule. The duration of the inspection is determined by the facility's current inspection status and is not impacted by previous performance. Moreover, we performed a logistic regression to predict the likelihood of a facility's inspection being interrupted based on its prior quality performance and find limited evidence for prior violations $\beta = -0.03$ $p = 0.18$ or prior recalls $\beta = -0.20$ $p = 0.23$. We find that prior severity of violations ($\beta = -0.16$ $p = 0.002$) predicts the likelihood of interruptions, but its coefficient suggests that a poorer performing facility's inspection is less likely to be interrupted or span over holidays, which is biased against our hypothesis.

There could be alternative explanations for the hypothesis that inspector's psychological factors and fatigue accumulation over the week affects their ability to find more violations (Hypothesis 4). The first alternative explanation could be that the facilities improve their quality

performance as the week progresses since the beginning of the work week are generally slower and facilities tend to gain momentum over the week leading to fewer quality issues later in the week. However, the literature provides evidence that the employees of the inspected facilities could also be fatigued as the week progresses leading to more oversights later in the week. Further, existing literature also suggests that staffing levels generally reduce later in the week as the employees tend to take vacations during the end of the week, leading to reduced quality performance of the facility towards the late week. In addition, our analysis of detailed descriptions of these violations from the FDA data suggest that the majority of violations reflect longer-term issues where their intensity does not change through the week (e.g., developing a supplier verification plan) which would be biased against our hypothesis.

Second, there may be likely an assumption that inspectors intentionally plan their week to inspect poorer facilities – those with a history of quality issues – earlier in the week as compared to better facilities later in the week. However, additional analyses find little evidence to support this supposition. We conduct a Pearson’s pairwise correlation analysis between the past quality performance of the facility and the start day of its current inspection (as in Ibanez & Toffel, 2022). We use the facility’s prior recalls (Pearson’s $r = 0.0005$, $p = 0.90$), prior citations (Pearson’s $r = -0.008$, $p = 0.06$), and prior severity of citations (Pearson’s $r = -0.0023$, $p = 0.59$) to find no evidence. Additionally, we run a Poisson regression to determine if there is an intentional sequencing of inspections within the week based on the facility’s past quality performance. We run a regression, predicting the days an inspector has been inspecting in each week before starting the inspection of the current facility (time inspecting earlier in the week), based on the current facility’s number of recalls in the last year (while keeping the rest of the model same). We find that a higher number of recalls from last year does not predict when a

facility's subsequent inspection will be scheduled in the week [$\beta=-0.003$, standard error (S.E.) = 0.023, and $p = 0.90$ with standard errors clustered by inspectors]. This finding suggests that the history of quality issues of the firm does not affect its schedule earlier or later in the week.

Our final hypothesis suggests that inspections conducted over a non-standard work week, i.e., weekends or federal holidays, reduce inspector stringency (Hypothesis 5) leading to lower numbers of violations. However, non-standard work week inspections could also yield fewer violations if inspectors choose to inspect more compliant facilities that yield fewer violations to minimize their workloads over the holidays, as these facilities could be inspected more quickly with lesser cognitive load. Our evidence suggests otherwise. First, our interviews with FDA officials suggest that inspectors are discouraged from working overtime. They work on non-standard schedules, i.e., weekends, federal holidays, or government shutdowns, to either accommodate facilities that are working on non-standard schedules, or if these facilities are critical – with a history of quality issues or handling vital / risky products such as baby food. We control for facility fixed effects to capture constant and unobserved heterogeneity, such as inherent risk levels of products handled, quality control standards, and operational practices.

Second, facilities with a non-standard work week inspection had no significant difference in the prior quality performance (prior violations $t = -0.63$ $p=0.52$, prior recalls $t=1.01$ $p=0.31$, prior severity of violations $t=-2.18$ $p=0.03$) compared to facilities that were inspected during the regular work week. When it comes to prior severity of citations, the difference was positive and significant, i.e., facilities with non-standard work week inspections averaged higher severity of violations in the previous inspections as compared to regular work week facilities suggesting bias against our hypothesis. Further, a logistics regression with dependent as a binary non-standard work week (1 or 0) with prior inspection outcomes predicting the inspection schedule

yields similar results (prior violations $\beta = -0.03$ $p = 0.15$, prior recalls $\beta = -0.17$ $p = 0.46$, prior severity of violations $\beta = 0.14$ $p = 0.006$). Third, another argument could be that facilities have reduced staffing over the non-standard work weeks, however that would lead to more violations not less.

5.3. Robustness tests

We conduct several supplemental analyses to determine the robustness of our results (Appendix B, Table B.1 and B.2). Our primary results are founded on a conservative approach that includes several fixed effects. We find very similar results if we include inspector-facility dyad fixed effects. We include inspector fixed effects with standard errors clustered by facilities and facilities fixed effects with standard errors clustered by inspectors and results remain robust. Further, we re-run our model using negative binomial conditional fixed effects and ordinary least squares regression using logged number of citations suggesting that our results are robust. Our results are robust to omitting facilities' fixed effects as well.

We changed how we measured the outcome of the inspections using a different approach. First, we used the average severity of citations. The outcome of these inspections could be no action indicated (NAI), voluntary action indicated (VAI), and official action indicated (OAI) in the order of increasing severity. Since there could be as many outcomes as the number of citations, we code the outcomes as 1 – NAI, 2 – VAI, and 3 – OAI. We then take an average of these outcomes and use it as a facility inspection performance criterion. A higher severity suggests that inspectors are able to find issues of higher intensity. Our results are robust using severity instead of number of citations (Appendix B, Table B.1., Model 2). Second, we use a binary outcome of whether the inspector is able to detect any issues (or not). We code it as 1 if an inspector found at least one issue at a facility and zero otherwise. We implement a logistic

regression model using the binary outcome of issues detected or not and found our primary results to be robust (Appendix B, Table B.1., Model 1).

To assess whether unusually busy weeks may make inspectors more fatigued and confound our results of schedule-induced fatigue that can affect the pacing of inspections (Hypothesis 1), concurrent inspections (Hypothesis 2), interrupted inspections (Hypothesis 3), start-day affect (Hypothesis 4), and non-standard work week inspections (Hypothesis 5), we re-estimate our model on the sub-sample of inspector-weeks with no more than five inspections (the 99th percentile). Our hypothesized results are robust to this subsample test (Appendix B, Table B.1., Model 3).

We also perform a falsification test. We have 40 federal holidays which fell on a Monday. Hence, Tuesdays following the federal holidays on a Monday turn into the “New Mondays” (Yao et al., 2019). We focus on a subsample of inspections conducted on Tuesdays and conduct an analysis by comparing the performance of the post-holiday Tuesdays with regular Tuesdays. The coefficient of post-holiday Tuesdays is positive and significant ($\beta = 0.204$ $p < 0.05$, IRR = 1.226). The inspection performance of post-holiday Tuesday is better than regular Tuesdays such that the number of citations increases, although some of our other hypothesized results become insignificant (Appendix B, Table B.2., Model 3).

For Hypothesis 4, we measure the independent variables in a few different ways, however our results regarding the effects of schedule-induced fatigue over the week are robust. In the first alternative we re-code the days of the week as early (Monday and Tuesday), mid (Wednesday), and late week (Thursday and Friday). Although this takes away the nuance of the days of the week, it captures the essence of the time of the week the inspection is conducted. We find that inspections conducted mid and late week are less effective than inspections conducted

early in the week (Appendix B, Table B.2., Model 1). In our second alternative approach, we construct another variable as the *days inspecting earlier in the week* (as in Ibanez and Toffel, 2020). This variable is the count of number of days already spent conducting inspections in the same week before the current inspection (week of Monday through Sunday). We find that in the same PPML model we use for our primary model that the coefficient for the variable *days inspecting earlier in the week* is negative and significant ($\beta = -0.075$ $p < 0.001$, IRR = 0.927). This suggests that with each additional day spent inspecting earlier in the week prior to the current inspection, the number of citations decrease by 7.3% supporting our hypothesis 4 (Appendix B, Table B.2., Model 2).

Our analysis is also robust in including additional inspector control variables in our model. Our results are robust after including the salary of the inspectors in the year of the inspection. We also include the time since the inspector's last salary hike (in years), time since last promotion (in years), age of the inspectors at the time of inspection (in years), tenure of the inspector with FDA at the time of inspection (in years) and federal grade of the inspector at the time of inspection. Our results are robust against all these control variables. However, the sample size dropped to about 3051 inspections and a few variables including the pacing and non-standard work week scheduling were now significant at 10% instead of 1% as in our primary results.

Further, our results are robust to controlling for various aspects of the facilities' past quality performance. We use the average severity of past inspection performance as discussed above. We control for the average of NAI, VAI and OAI and find our results to be robust. Furthermore, we use the recalls in the same year as the year of inspection and find that our

results are robust. We also control for the severity of recalls as in the average of recall severity from 1 - minor issues such as mis-labels to 3 – life threatening products with robust outcomes.

Lastly, our results remain consistent when we account for the weekly or monthly workload, defined as the number of inspections an inspector completed during the week or month of the focal inspection. Incidentally, the estimated coefficient for workload in a week is negative and significant ($\beta = -0.069$ $p < 0.001$, IRR = 0.933), implying that, with each additional inspection increase in the workload of the inspector in a week of the current inspection, inspectors detect 6.7% fewer violations (Appendix B, Table B.1., Model 4). Similarly, the estimated coefficient for workload in a month is negative and significant ($\beta = -0.031$ $p < 0.001$, IRR = 0.969) suggesting that with each additional inspection in a month, the number of violations detected decrease by 3.1% (Appendix B, Table B.1., Model 5). Hence, inspectors in our sample do not appear to be immune to the workload which forms the basic premise of our story. However, this also demonstrates that our outcomes are effective and robust signifying the significance of the effects we have identified (Prentice and Miller 1992).

5.4.Further Analysis

Extensions can be found in Appendix C (Table C.1 and Table C.2). Our primary results suggest that interrupted inspections (or holiday spanning inspections) result in a greater number of citations (Hypothesis 3). To augment this, we also investigate the length of interruption or holiday span, i.e., does the length of interruption affect the inspections outcome? We find that it does affect the number of inspections. We introduce the term log length of the interruption and its quadratic term, i.e., square of log length of the interruption. We find that the term of length of interruption is positive and significant ($\beta = 3.566$ $p < 0.001$), whereas its square term is negative and significant ($\beta = -1.055$ $p < 0.001$) (Appendix C, Table C.1., Model 1). This implies an

inverted U-shaped relationship between the length of the interruption and the number of citations at a current inspection; that is, increasing the interval of interruption initially increases but eventually decreases the number of citations. Refer figure 2.

In addition, we explore different inspector characteristics that can moderate the scheduling variables. This helps us understand how inspector characteristics such as experience, gender, site familiarity or inspection duration can help offset some of the scheduling induced biases. We performed PPML analysis using the number of citations as the outcome with month, year, facility, FDA division, and facility's nth inspection fixed effects with standard errors clustered by inspectors, as in our primary models.

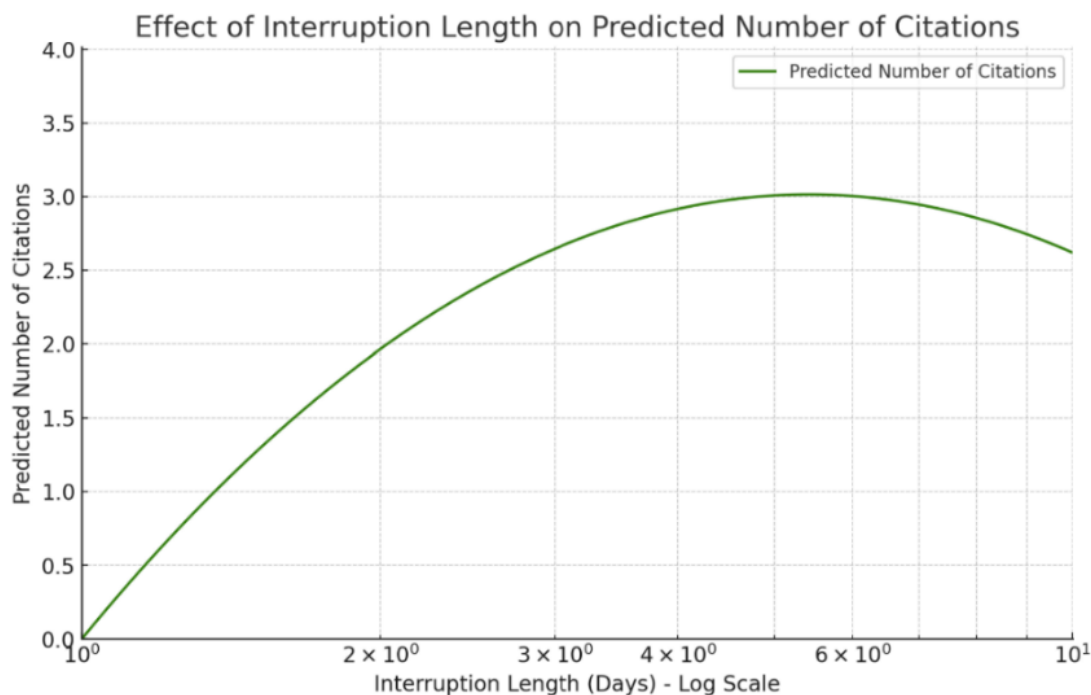


Figure 2. Number of citations by the length of interruption

First, we find that inspector experience, i.e., number of inspections the inspector has conducted prior to current inspection within our sample period, positively moderates the relationship between interrupted inspection and number of citations (Appendix C, Table C.2.,

Model 1). This suggests that inspectors with more experience benefit more from an interruption (or a scheduled pause) as compared to less experienced colleagues. This is further corroborated by the moderation results of site-specific experience which suggests that inspectors more familiar with the inspection site benefit more from interruption (or a scheduled pause) as compared to less their familiar colleagues (Appendix C, Table C.2., Model 3). More experienced inspectors become desensitized or immune to routine issues due to their familiarity with them. This becomes worse with increased familiarity with the inspection site. Routine and familiarity lead to observational fatigue where inspectors overlook issues because they become part of the “background.” An interruption (or a scheduled pause) can break their routine.

Further, more experienced inspectors are immune to the start day effect as compared to their less experienced colleagues suggesting that the number of citations over the week remain consistent. Second, we find evidence for “Friday Effect” among males. Our analysis suggests that male inspectors tend to perform worse than female inspectors on a Friday leading to fewer citations (Appendix C, Table C.2., Model 2). We find the variable gender (male =1) negative moderates the relationship between inspections scheduled on a Friday and the number of violations.

Third, we look at the average duration of past inspections of an inspector (Appendix C, Table C.2., Model 4). This variable informs us about the average number of days an inspector takes to conduct an inspection. We perform PPML analysis using the number of citations as the outcome, however, we restrict the sample to the 25th percentile of the inspector’s experience, i.e., we use observations where an inspector has performed a minimum 20 inspections prior to the current inspection. This is because the average duration of the inspection may be affected by inexperience and learning effects. New inspectors may take more time to conduct an inspection.

By focusing on this sub-sample, we ensure that inspector duration is related to their inspection approach rather than their inexperience. We also control for facility complexity and facility fixed effects for unobserved time-invariant facility characteristics.

We find that inspectors who, on average, spend more time at a facility for an inspection have their ability to find violations decrease if the inspection is interrupted by a holiday. We also find that duration of the inspection is positively and significantly linked to the number of violations found at a facility while controlling for an exhaustive list of confounders. So, the inspectors with a slower (or more thorough) approach to inspections are affected more if they are further slowed down by a holiday interruption. Further, such inspectors tend to find a greater number of violations as compared to their faster (less thorough) colleagues as the week progresses. This suggests that cognitive fatigue can have greater impacts on those inspectors who are accustomed to working faster, affecting their ability to process information and make decisions over the week. This also suggests that while moving fast is associated with a greater number of inspections, there might be diminishing returns to spending less time at a facility for an inspection.

6. Discussion

Our analysis of FDA food facility inspections demonstrates that work-break schedules significantly impact inspection outcomes. First, we discovered that allowing inspectors more time between inspections, during which they engage in secondary tasks, leads to an increase in the number of detected violations, but only to a certain extent.

Table 3. Correlation matrix for main independent variables

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1-Pseudo	1.00															
2-Pseudo ²	0.60***	1.00														
3-Concurrent	-0.13***	0.03***	1.00													
4-Intermittent	0.02	0.02*	0.08***	1.00												
5-Start-day	-0.01	0.02**	0.00	-0.05***	1.00											
6-Non-standard	0.02*	0.01	-0.02	0.10***	-0.05***	1.00										
7-Experience	0.16	0.08***	0.01	-0.01	-0.04***	0.06***	1.00									
8-Average time	0.02	0.03**	0.10***	0.30***	0.01	0.01	-0.10***	1.00								
9-Gender	0.00	-0.00	-0.05***	-0.07***	0.01	0.00	0.03***	-0.18***	1.00							
10-Site specific experience	0.00	0.00	0.04***	0.03***	-0.02	0.00	0.13***	0.03***	0.02	1.00						
11-Prior inspection outcomes	-0.04***	0.06***	0.03**	0.09***	-0.02	-0.01	-0.03***	0.18***	-0.08***	-0.01	1.00					
12-Past inspection outcomes	0.12***	0.06***	-0.02	0.10***	-0.02**	0.03**	0.21***	0.11***	-0.01	0.06***	0.03***	1.00				
13-Recall last year	-0.00	-0.01	-0.00	-0.01	0.01	0.00	-0.00	0.00	-0.02	0.00	0.01	0.00	1.00			
14-Inspection depth	0.02*	-0.00	-0.02	0.05***	0.00	0.01	0.02**	0.03***	-0.04***	-0.01	0.04***	0.03***	-0.00	1.00		
15-Inspection duration	0.01	0.01	0.16***	0.54***	-0.01	0.04***	-0.02**	0.28***	-0.06***	0.03***	0.05***	0.09***	-0.01	0.04***	1.00	
16-Political affiliation	0.33***	0.17***	-0.02	-0.01	-0.02	0.03***	0.51***	0.06***	-0.01	0.02	-0.03***	0.23***	-0.00	-0.03**	-0.01	1.00

*** p<0.01, ** p<0.05, *p<0.10

Table 4. How inspector's work-break schedules affect inspection outcomes

Independent variables	(1) Violations	(2) Violations	(3) Violations	(4) Violations	(5) Violations	(6) Violations
<i>Pseudo (H1)</i>	-	0.028*** (0.008)	0.024*** (0.008)	0.025*** (0.008)	0.019** (0.008)	0.019** (0.008)
<i>Pseudo² (H1)</i>	-	-0.012*** (0.003)	-0.011*** (0.004)	-0.012*** (0.003)	-0.010*** (0.003)	-0.010*** (0.003)
<i>Concurrent (H2)</i>	-	-	-0.157*** (0.057)	-0.150*** (0.055)	-0.159*** (0.055)	-0.159*** (0.055)
<i>Intermittent (H3)</i>	-	-	-	0.493*** (0.040)	0.567*** (0.043)	0.581*** (0.044)
<i>Start-day (Monday) (H4)</i>	-	-	-	-		
<i>Tuesday</i>					-0.074** (0.035)	-0.076** (0.035)
<i>Wednesday</i>					-0.143*** (0.038)	-0.152*** (0.037)
<i>Thursday</i>					-0.212*** (0.042)	-0.221*** (0.042)
<i>Friday</i>					-0.298*** (0.058)	-0.312*** (0.059)
<i>Non-standard (H5)</i>	-	-	-	-	-	-0.247*** (0.089)
<i>Experience</i>	-0.001 (0.001)	-0.001 (0.001)	-0.001 (0.001)	-0.001 (0.001)	-0.001 (0.001)	-0.001 (0.001)
<i>Average time</i>	0.045*** (0.009)	0.045*** (0.009)	0.046*** (0.008)	0.040*** (0.008)	0.038*** (0.008)	0.039*** (0.008)
<i>Gender</i>						
<i>Male (1)</i>	-0.133*** (0.033)	-0.132*** (0.033)	-0.135*** (0.033)	-0.135*** (0.033)	-0.133*** (0.032)	-0.132*** (0.032)
<i>Site-specific experience</i>	-0.067** (0.033)	-0.069** (0.033)	-0.071** (0.033)	-0.071** (0.032)	-0.073** (0.032)	-0.075** (0.032)
<i>Prior inspection outcomes</i>	0.034*** (0.005)	0.037*** (0.005)	0.037*** (0.005)	0.037*** (0.005)	0.037*** (0.005)	0.036*** (0.005)
<i>Past inspection outcomes</i>	-0.082*** (0.006)	-0.082*** (0.006)	-0.082*** (0.006)	-0.080*** (0.006)	-0.079*** (0.006)	-0.077*** (0.006)
<i>Recalls last year</i>	-0.122 (0.0794)	-0.124 (0.080)	-0.117 (0.080)	-0.083 (0.086)	-0.069 (0.086)	-0.066 (0.088)
<i>Inspection depth</i>	0.173*** (0.028)	0.175*** (0.029)	0.175*** (0.029)	0.171*** (0.029)	0.167*** (0.029)	0.166*** (0.029)
<i>Inspection duration</i>	0.024***	0.025***	0.026***	0.013***	0.011***	0.011***

	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
<i>Political affiliation</i>						
<i>(Democrat)</i>						
<i>Republican</i>	-0.115	-0.119	-0.124	-0.176	-0.177	-0.162
	(0.169)	(0.167)	(0.167)	(0.166)	(0.167)	(0.167)
<i>Month Fixed</i>	Included	Included	Included	Included	Included	Included
<i>Effects</i>						
<i>Year Fixed Effects</i>	Included	Included	Included	Included	Included	Included
<i>Facility Fixed</i>	Included	Included	Included	Included	Included	Included
<i>Effects</i>						
<i>FDA Division</i>	Included	Included	Included	Included	Included	Included
<i>Fixed Effects</i>						
<i>Facilities' nth</i>	Included	Included	Included	Included	Included	Included
<i>inspection Fixed</i>						
<i>Effects</i>						
<i>Constant</i>	0.984***	1.023***	1.028***	0.951***	1.032***	1.032***
	(0.092)	(0.091)	(0.091)	(0.091)	(0.094)	(0.094)
Observations	7,464	7,464	7,464	7,464	7,464	7,464
(Inspections)						
Number of	3,291	3,291	3,291	3,291	3,291	3,291
Facilities						
Number of	1,070	1,070	1,070	1,070	1,070	1,070
Inspectors						
Wald's χ^2	317.33***	353.99***	346.24***	611.92***	654.54***	653.49***
Pseudo R ²	0.366	0.367	0.367	0.375	0.377	0.377
Log	-	-	-	-	-	-
pseudolikelihood	12454.079	12442.169	12435.221	12276.184	12244.928	12238.848

Note. Poisson regression coefficients with robust standard errors clustered by Inspectors.

Robust standard errors in parentheses *** p<0.01, ** p<0.05

Beyond a 3-day interval, the number of identified violations begins to decline, possibly due to inspectors becoming less adept because of infrequent work. Second, doing inspections simultaneously leads to 14.7% fewer violations found, indicating that multitasking makes inspectors less effective.

Third, holiday spanning inspections are associated with a 78.8% jump in violations found, in contrast to those done without breaks. Fourth, inspections on non-standard work schedules, begun outside of the Monday to Friday workweek, result in 21.9% fewer violations. Lastly, the number of citations per inspection decreases over the workweek (from Monday to

Friday). This reduction is attributed to the gradual fade-out of break effects and the anticipation of upcoming breaks. Overall, we contribute to the literature by classifying and evaluating specific types of breaks within the context of inspection tasks, with a goal to understanding how these breaks directly impact inspection quality.

6.1.Contributions

Our study contributes to three streams of literature. We contribute to the literature on factors affecting the ability of inspectors to perform quality inspections (Ibanez & Toffel, 2020; Ibanez et al., 2023; Short et al., 2016). We suggest that work-break schedules may impact inspection outcomes of regulatory food inspectors. We contribute to this literature by identifying specific work-break regimes and their impact on the inspector's ability to detect violations.

We also contribute to the literature on breaks and work performance (Fritz & Sonnentag, 2006; Fritz et al., 2010). We contribute to this literature in several ways. We conceptualize breaks using the dimensions of their timing around work tasks, length, degree of disengagement from work, and their level of predictability. Further, we identify different underlying mechanisms by which breaks affect work performance. We also use contribute by measuring their impact on inspection performance as an outcome. The existing literature highlights several positive and general effects of breaks on employee productivity and work performance. Our study specifically focuses on the context of monitoring and assessment.

This study theoretically grounds the concept of work-break using the operations management literature on scheduling and psychology literature on breaks. The literature on breaks suggests they are periods of rest that positively impact individuals. Prior literature studies break in isolation. We suggest that breaks are strategic elements that can be scheduled around work tasks to improve task performance and work outcomes. Prior literature also suggests that

different types of breaks might have a similar impact and offer a uniform degree of disengagement from work tasks. We contribute to this literature by suggesting that the impact of breaks on work tasks depends on different dimensions including their timing and length.

Furthermore, we introduce, define, and analyze new forms of breaks, such as pseudo breaks, wherein an employee has time off from primary tasks but still works on secondary tasks. We introduce the concept of temporal pacing of inspections, where the time between two inspections could impact the outcome of these inspections. Additionally, we test the concept of split inspections or holiday-spanning inspections, i.e., inspections that have an intermittent break in between ongoing inspections. We also find evidence for the concept of break anticipation, wherein work tasks are impacted by the anticipation of breaks.

We are among the first empirical papers to study and find evidence for break fade-outs in an operational setting, where the possible restorative effect of breaks fades out over subsequent workdays. An earlier paper by Fritz et al., 2010, used self-reporting surveys to introduce the concept of break fade-outs and found limited evidence. We are also among the first to conceptualize and test the phenomenon of “job creep,” wherein work tasks intrude into breaks and affect their restorative abilities (Bolino et al., 2013). Furthermore, we ground multitasking in work transition and work-break scheduling literature and suggest that multitasking is a phenomenon where one task intrudes on other tasks without a formal break between them, affecting transition costs and work performance.

Existing literature has focused on the positive aspects of breaks on task performance and work outcomes. We contribute by suggesting that breaks may have a negative effect. Longer breaks can be counterproductive as increased duration of breaks can have diminishing impact on

work outcomes. Lastly, we also contribute to a new stream of literature on responsible scheduling which emphasizes employee well-being by scheduling work tasks more thoughtfully.

6.2. Managerial implications

Our study demonstrates that various work-break schedules can have both positive and negative effects on inspectors' ability to detect issues during inspections. Our research suggests that managers can strategically use the breaks to improve the work outcomes of their employees. We analyze different work-break regimes and make recommendations. We suggest that pseudo breaks may effectively help inspectors recover from inspection fatigue without requiring complete time-off from work.

Further, intermittent breaks may be beneficial for work task performance but might have diminishing returns. Our research also opens discussions about effective work-week lengths. Also, employees may experience break fadeout effect over the work week affecting their performance. In addition, assigning more than one inspection to an inspector simultaneously can detrimentally affect inspection outcomes. We challenge the belief among practitioners that multitasking is an effective strategy to utilize idle time between inspections.

From a managerial perspective we analyze how inspectors might benefit from such work-break scheduling strategies differently. Managers can use our findings to design customized policies to decrease the consequences of work-break schedules eroding quality of inspections. Our findings reveal that inspector experience positively moderates the relationship between inspections with intermittent breaks and the number of citations. This suggests that more experienced inspectors achieve greater benefits from such breaks compared to their less experienced colleagues. Moreover, inspectors familiar with the inspection sites benefit more from intermittent breaks.

Further, we analyze the average duration of past inspections conducted by an inspector. Inspectors spending more time at a facility for inspections tend to find fewer violations if the inspection is holiday spanning. Additionally, inspectors with a slower approach to inspections, who spend more time at a facility, tend to find a greater number of violations, particularly as the week progresses. We also have public policy implications. With over 48 million people falling ill, 128,000 being hospitalized, and 3,000 deaths every year in the U.S. our research may benefit policy makers. They may be able to better strategize work-break regimes to help improve inspection performance of regulatory inspectors.

6.3. Limitations and future research

This study has several limitations that could be addressed in future research. Firstly, the inspection outcomes could be impacted by several individual level inspector characteristics including inspectors' individual beliefs, culture, religious background etc. We do not control for such factors due to data limitations. Future research can collect more individual level inspector information.

Further, inspector training, including formal and informal training, could impact their inspection outcomes. Formal training such as college education (such as graduate education) could impact the inspection outcomes differently. Informal training such as on-the-job training and tacit knowledge through interactions with other experts may also impact inspections outcomes. Future research may control such factors to account for the role of training.

Lastly, there could be regulatory changes that may have impacted the process of inspections and its protocol. They might have also impacted inspection scheduling. We don't control for any regulatory changes during the period of the study. Future research may also study how the change in scheduling affects the inspection outcomes differently.

Acknowledgements

This paper would not have been possible without the support, motivation, and guidance of Dr. Martin Dresner. His direction and feedback on this paper have been invaluable. Secondly, I would like to express my gratitude to the FDA officials, including food inspectors, who took the time to provide their valuable feedback on the field operations of food safety inspections. Lastly, I would like to extend my thanks to FDAlytics for generously sharing the data on food inspections.

Chapter 4: Impact of Salary Differences on Quality Assessment

"Our pay falls well below comparable positions outside of government and there is limited promotion potential...compensation is less than the industry QA [Quality Assurance] counterparts with similar education and experience...overall morale is very low." - (Consumer Safety Officers, FDA)

1. Introduction

Essential workers on the frontline (or field staff) provide necessary services to general public including health care, education, transportation, and food safety. They make critical operational decisions in the field that have significant repercussions (Balogun et al., 2015). For example, food inspectors assess quality of food facilities and ensure they comply with regulatory standards safeguarding safety of the food supply chain. The years marked by COVID-19 have brought the efforts, importance, and plight of field staff to the forefront (Blau et al., 2022). The ever-increasing job demands are affecting worker morale and engagement at work, impacting their performance (Bork et al., 2022). For example, studies suggest that disengaged healthcare workers are prone to making more mistakes and exhibit impatience with their clients (Brower, 2022; Ulfa et al., 2022). Accordingly, the question is what affects the field staff's motivation to perform effectively at work?

Although, incentives, especially salaries, are important for all employees including field staff (Hsee et al., 2010; Skitka, 2009), evidence suggests that salary comparisons are a key factor affecting workplace outcomes such as productivity and adverse behavior (Xu et al., 2024). For example, employees often discuss and compare salaries with others, and such comparisons affect how they perceive their own compensation. Through salary comparisons, employees can determine their ratio of input (effort) to output (salary) relative to a selected referent (such as a

colleague) (Xu et al., 2024). Differences in these ratios may affect an employee's motivation, impacting their work performance. Therefore, a critical question is: How do salary referent differences affect the motivation and work performance of field staff?

Further, the choice of salary referents is crucial. Different referent comparisons may shape the job satisfaction of employees differently (Trevor & Wazeter, 2006). For instance, a U.S. Food and Drug Administration (FDA) food inspector highlighted that: "[Federal Agency] Salary is below the private sector. There is little opportunity for individual reward or advancement [compared to industry], leading to job apathy and disconnectedness." So, salary differences between the employees of the monitored firm (referent group) and the inspectors regulating the industry may affect the inspectors' work engagement, possibly more than differences with co-workers. This complexity raises several questions: How do perceived incentive differences with chosen referent groups affect the operational outcomes of the field staff performance? Do comparisons with different referent groups have distinct impacts on their performance? Moreover, what constitutes an effective incentive design (in terms of referent groups) that motivates employees to modify their behavior as desired?

According to the literature, employees commonly choose two types of groups for referent comparisons: a) similar peers within the organization (internal salary comparison) and b) individuals outside the organization performing similar work (external salary comparison) (Harris et al., 2008; Hill et al., 2017; Kacperczyk & Balachandran, 2018; Williams et al., 2006). However, these comparisons can be more nuanced than the broad understanding of internal and external comparisons. We combine these insights with existing studies on referents and suggest that referent groups can be classified based on scope of comparison and degrees of relational dependence. Accordingly, we focus on four referent groups, i.e., self-referent, internal peer

referent, industry referent, and operational interface referent (a novel referent group). We propose these classes of interface referent groups that are critical, especially for field staff, previously overlooked in the literature. For example, FDA food inspectors interact with employees at facilities they inspect. Specifically, we study how incentive differences, influenced by comparisons with multiple referent groups, affect FDA inspector's quality assessments during food facility inspections.

In the monitoring and assessment literature, several studies have highlighted the vital role of plant quality assessments (inspections) in managing quality across increasingly complex supply chains (Ibanez & Toffel, 2020; Power & Terziovski, 2007; Short et al., 2016). These inspections help determine the existing level of conformity and identify areas of improvement in those facilities. Their goals span from ensuring product quality, promoting consumer safety, and bolstering operational productivity (Ball et al., 2017). The accuracy of these inspections is key for informed managerial decisions and greatly dependent on the inspectors performing them (Ibanez & Toffel, 2020).

FDA conducts food-safety inspections for food facilities where food is processed, distributed, or stored to assess for quality against the regulatory standards on a regular basis (FDA, 2023). Such inspections are critical for public health and safety. Despite these inspections, every year, 48 million Americans get sick from foodborne diseases, resulting in 128,000 being hospitalized and 3,000 dying (CDC, 2018). A 2017 report by the Office of the Inspector General in the Department of Health and Human Services found that these inspections often overlooked problems that later led to serious health issues (Levinson, 2017). For our research, we use data from 594 FDA food inspectors conducting 16,382 inspections across 14,441 food facilities from

Jan'2015 to Dec'2019. We augment this data using the salary information of individual food inspectors for the same time period.

This context is ideal to study the impact of referent group comparisons for several reasons. First, the salary information for FDA inspectors and their referent groups is publicly available since they are federal employees, and this information can be used to make actual comparisons rather than relying on perception-based information. Second, food inspectors are essential employees at the frontline making critical decisions to safeguard the food supply chain (Ibanez & Toffel, 2020). Third, food inspectors, during their operational duties, closely interact with employees of the facilities they inspect, i.e., referent groups with direct relational dependencies. Finally, our dataset provides detailed records of all four types of referent groups, including salaries, grade levels, and inspector demographics, allowing us to evaluate salary differences for each individual inspector within their grades controlling for several inspector and facility level controls.

When it comes to quality assessments, the number of violations identified by an inspector is a common measure to understand the work performance of the inspector (Short et al., 2016). The number of violations in inspection reports reflect the inspector's stringency in identifying these violations after controlling for the facility's actual quality performance. Further, given that violations need supporting documentation, inspector work performance tends to manifest as underreporting actual issues rather than citing fictitious ones (Ibanez & Toffel, 2020). In addition, while FDA inspector(s) make initial on-site observations, these are later reviewed by the FDA's regional offices. Firms can also challenge the inspector findings, deterring inspectors from over-reporting (FDA, 2023). Therefore, inspector work performance on quality assessment takes the form of under-reporting the number of violations that are present (Short et al., 2016).

Accordingly, using data on thousands of FDA inspections of food facilities, we find strong evidence that an inspector's motivation due to incentive differences affects the inspection outcomes.

First, contrary to the existing beliefs, we observe that using self or personal growth as a referent, higher pay hikes from the previous year are associated with decreased field staff performance. Specifically, a 1% higher pay as compared to last year is associated with roughly a 0.7% decrease in the expected count of citations. Similarly, we find that a 1% higher salary compared to internal peers and operational interface referents results in roughly 0.57% and 0.16% fewer citations respectively. According to equity theory, these results suggest that when the ratio of inputs to outcomes is favorable for inspectors compared to referents, contrary to existing belief that inspectors would increase their inputs to restore differences, they adjust the value attributed to their own contributions (Adams & Freedman, 1976; Lawler, 1968; Hook & Cook, 1979).

We also find that higher pay in comparison to referents is associated with a higher number of citations. This implies that when the input-to-output ratio is less favorable for inspectors compared to the referent group, contrary to existing research that suggests inspectors may be demotivated and reduce their efforts to restore differences, they enhance their efforts. Lastly, consistent with the existing literature on equity theory, we find that a 1% higher pay in comparison with industry referents (other federal employees) leads to a 1.43% higher number of citations detected by inspectors during an inspection. So, different referent group comparisons affect inspector work performance differently. Overall, our findings indicate that currently unreported violations would be cited if inspector incentives are designed strategically.

Our study makes several theoretical and managerial contributions. First, we are among the initial studies to explore the impact of the salary differences aspect of incentive design on operational outcomes. Existing literature studies several aspects of incentive design, including performance-based compensation contracts, rewards, bonuses, and penalties, but differences in incentives relative to referent groups are less explored (DeHoratius & Raman, 2007; Villena & Dhanorkar, 2020). Second, we extend the literature on salary comparisons by conceptualizing referent groups for studying incentive differences specifically in an operations management setting. We find that not all referent groups have the same impact on work withdrawal behavior, which provides a more nuanced understanding of incentive differences effects. We propose a new class of referent groups, i.e., operational interface referent, extending the existing categorization of referent groups, especially for field staff. Third, we contribute to the literature on the performance of essential workers on the frontlines, which is a relatively newer line of investigation (Lu et al., 2022; Tucker, 2007; Pendem et al., 2024). We provide answers to important questions such as what affects the field staff's motivation to perform effectively at work? Do incentive differences affect the motivation and work performance of field staff?

Fourth, we extend the equity theory by suggesting newer mechanisms through which these biases manifest themselves on the operational outcomes. When it comes to positive differences, existing literature suggests that employees increase their efforts to match the relatively higher incentives because of guilt and indebtedness (Brockner et al., 1985). But we found that employees decrease their efforts. For negative differences, existing literature suggests that employees get demotivated and decrease their efforts to match their lower incentives (Berger et al., 2013). Instead, we find that employees may exhibit over compensatory behavior where they increase their efforts. Fifth, organizational psychology literature establishes how multiple

salary comparisons are processed to affect employee perceptions literature (Xu et al., 2024). We explore the mechanics of how employee perceptions of distributive justice affect employee work withdrawal which is less understood in the literature (Xu et al., 2024). Furthermore, existing studies on incentive differences are mostly survey-based studies that depend on asking respondents to compare their perceived inputs and outcomes to others and rate their satisfaction. We use actual pay data that could help provide a different perspective.

Sixth, by identifying the factors that affect inspection outcomes, we contribute to the literature on monitoring and quality improvement (e.g., Ibanez & Toffel, 2020; Mani & Muthulingam, 2018). Our focus on how incentive differences affects inspector stringency introduces the behavioral operations lens to the literature on inspector bias, which has otherwise largely focused on inspection scheduling (Ibanez & Toffel, 2020), inspector experience (Ball et al., 2017; Short et al., 2016), and inspector qualification (Short et al., 2016). Our study extends beyond prior research by assessing both the extent of bias and its tangible outcomes in real-world scenarios.

Lastly, while previous research has explored a few individual (Becker et al., 2012) and team-level factors (Scott-Young & Samson, 2008; Short et al., 2016) affecting inspection outcomes, our research provides a more nuanced understanding of incentive differences. Our research focuses on workforce effectiveness which directly ties into job satisfaction, motivation, and retention of field staff. By having a deep understanding of these factors, organizations can adjust their incentive design strategies for improving operational effectiveness. Many studies have identified causes of inspection oversight (Becker et al., 2012; Short et al., 2016), but there is limited knowledge about factors that affect inspector's work motivation. We explore factors

that arise at the intersection of operations and workforce dynamics, i.e., incentive design, employee motivation, and quality assessment.

From a managerial perspective, how salaries are structured and compared within and across different referent groups has a significant impact on employee motivation, their perception of fairness and eventually their operational performance. The goal of an effective incentive design in operations is to encourage employees to achieve operational excellence. Leveraging incentive designs, informed by the referent groups that matter the most, can help reduce employee work withdrawals (Xu et al., 2024). Aligning salary structures with operational goals promotes operational performance. We suggest that compensation, as a part of a well-designed incentive structure, can be leveraged as a strategy to motivate the employees for operational performance.

We explore several inspector level characteristics that may be used to minimize the incentive differences. For example, we find that higher site-specific experience for an inspector can mitigate the effects of salary differences on the number of citations issued. Inspectors earning higher salaries than their internal peers and operational interface referents, who possess greater site-specific experience, tend to identify more violations. This might be because they have a better understanding of the concerns, leading to more effective inspections. Further, providing inspectors with breaks during ongoing inspections can neutralize the negative effects of salary differences with operational interface referents on the number of citations. Also, breaks, especially in between the inspections, can help alleviate stress and negative effects of observed salary differences, letting inspectors refocus and work better. However, inspector workload exacerbates the negative impact of operational interface referent differences on inspection outcomes. Increased number of inspections in a certain period can heighten the negative

emotions related to incentive differences. Overall, in addition to designing better incentive structures, managers can moderate workload, strategize breaks, and schedule more site-specific inspections for inspectors to mitigate the negative impact of incentive disparities.

2. Related literature

We build on the following streams of operations management research: (i) incentive design and operational outcomes and (ii) field staff motivation and operational performance.

2.1. Incentive design and operational outcomes

An extensive body of literature investigates how incentive structure and design affects operational outcomes. Most studies look at either firm level incentives and a subsequent firm behavior change or incentive based contracts among buyers and suppliers. For example, Lee et al. (2018) studied incentives in operations management at a firm level. They study the role of national culture in the effectiveness of bonus and penalty incentive contracts in supply chain exchanges between suppliers and buyers. They suggest that national culture influences the operational outcomes of various contracts. For example, bonus contracts are more effective than penalty-based contracts in encouraging suppliers in China to be more operationally efficient, thus reducing shirking behavior. In another paper, Lee and Li (2018) studied how buyers can manage product quality of suppliers by incentivizing supplier quality-improvement efforts. They find that both monetary incentives and inspections may be used together to encourage suppliers to improve product quality effort – a carrot-and-stick incentive policy.

Fewer studies look at task performance and operational outcomes of employees or analyze incentives at the individual level. DeHoratius and Raman (2007) studied the effect of incentive designs on store managers' performance. They examined how performance-based contracts, where store managers are penalized more for inventory shrinkage than they are

rewarded for sales, affected the retail performance of the store. They found that the way store managers were compensated was associated with a 4.2% improvement in sales profit. In another study, Villena and Dhanorkar (2020), studied how providing financial and non-financial rewards to employees help motivate them to implement climate change goals for carbon transparency. Their study offers a detailed analysis of the drivers of organizational behavior within a firm that are crucial for achieving sustainability goals in operations.

Further, Allon et al. (2023) focused on the gig economy workers who are outside the traditional boundary of a firm. The study focuses on the relationship between financial incentives and operational decisions of whether to work and the work duration of these gig economy workers. They find that financial incentives improve operational outcomes, further confirming that when employees' incomes increase, their performance tends to improve. In another study, Hutchison-Krupat and Chao (2014), studied incentive design in terms of rewards and penalties to understand how such incentives impact the decisions of individual employees working for a collaborative innovation initiative. Using experiment and implicit incentive structure they found that when penalties are small, the risk-taking behavior of a person does not change much regardless of the rewards they might get. Further, Gu et al. (2017) studied how CEO incentives in the form of compensation affects firms' suppliers' relationship-specific investments. They looked at option grants as a form of incentives. Using archival data, they found that an increase in the risk-taking incentives of CEOs of a customer firm leads to a reduction in suppliers' investments in relationship specific initiatives.

The existing individual-level studies are mainly experiment-based or survey-driven, utilizing primary data collection methods with scant research relying on archival data, which may be less prone to response bias. Moreover, these studies generally explore similar aspects of

incentive design, namely salaries, rewards, bonuses, and penalties, albeit in different ways. The impact of the incentive differences, relative to multiple referent groups on operational outcomes, is less explored in the literature. According to the incentive design literature in organizational psychology, although financial incentives are important, incentive comparisons are critical factors affecting employees' operational outcomes (Hsee et al., 2009; Skitka, 2009; Xu et al., 2024). Employees tend to perceive their own compensation in comparison to a referent group to derive work satisfaction.

Literature on salary comparison suggests that the employees ascertain the ratio of input (effort) to output (salary) relative to a selected referent (such as a colleague) (Xu et al., 2024). The impact of such comparisons can vary widely depending on the specific referent groups. Adams (1965), through Equity Theory, proposed that such inconsistencies might evoke feelings of unfair treatment among employees towards their organization. According to management literature, incentive differences affect the perceptions of distributive justice (i.e., "the perceived fairness of decision outcomes, especially the degree to which outcomes are equitable"; Colquitt et al., 2013, p. 200) which subsequently promotes employee withdrawal behavior ("behaviors such as reduced work effort and leaving the workplace"; Xu et al., 2024). However, the choice of a specific referent group is equally important.

Based on the literature, employees generally use two types of groups for referent comparisons: a) similar peers within the organization (internal salary comparison) and b) individuals outside the organization performing similar work (external salary comparison) (Harris et al., 2008; Hill et al., 2017; Kacperczyk & Balachandran, 2018; Williams et al., 2006). However, these comparisons can be more nuanced, involving varying degrees of relational dependency on the referent groups. We suggest that an employee's functional reliance and the

salience of operational interactions with another group or other individuals could affect their perception of their own incentives. This is especially true for employees who work “in the field” and are directly involved in the operational delivery of services. For example, when it comes to external referents, there could be indirect comparisons, providing a broader market context, such as employees in different organizations. Others, though outside the organization, may have a direct and significant impact on an individual's perceptions of salary differences.

This insight, combined with existing studies on referents, suggests that actual performance of employees may be impacted by various types of salary referent group comparisons which we theorize by (i) the scope of comparison, which may be internal or external to the organization, and (ii) the degree of relational dependence on the referent group (direct or indirect). By organizing referent groups along orthogonal dimensions, we capture a broad range of referent groups in the operational setting for field staff. Though there could be several candidate referents, we identify four (one in each quadrant) and empirically test their effects in the operational context of FDA food inspections (Table 1). We focus on three previously identified types of referent groups: self-referent, internal peer referent, industry referent. In the fourth quadrant we propose a novel class of referent group, operational interface referent, which has not been studied in the previous literature. Operational interface is an external referent group with whom field staff interacts directly for their operational duties within their professional field. For example, FDA food inspectors interact with employees at facilities they inspect. This referent analysis provides an overlooked yet crucial perspective for understanding how the incentives of field staff compare with those of external clients with whom they closely interact to perform their job duties. We examine the impact of salary differences with these four referent group comparisons on the field staff's performance.

We classify the four categories of referent groups in Table 1. The self-referent category directly affects an employee's perception of personal progress within an organization, as it pertains to the individual's own contributions and performance. For internal peer comparisons, although they occur within the organization, they are categorized as indirect because they relate to understanding one's standing among peers rather than reflecting personal achievement. Considering the context of field staff, food inspectors mostly conduct inspections independently and have low relational dependence on their immediate peers for job duties. The operational interface referent involves employees of facilities that are directly inspected by the inspectors, leading to direct professional interaction and operational relevance to an individual's job function, affecting their perception of efforts and outcomes. The industry referent comparison with federal employees, while external, is indirect because it reflects a broader market-wide perspective without direct job-specific interaction with these employees.

Table 1. Categorization of salary referents

Relational Dependency	Direct	Indirect
Comparison Scope		
Internal (Organizational)	Self-referent (Pay hike from last year)	Internal peer referent (Co-workers)
External (Sectoral)	Operational interface referent (Employees of facilities inspected)	Industry referent (Federal employees)

2.2. Field staff motivation and operational performance

Field staff are employees who operate outside of the traditional office setting. They may include roles like sales representatives, service and maintenance technicians, field researchers, and food inspectors (Balogun et al., 2015). They have direct involvement in the operations of a firm and conduct essential tasks that contribute to a firm's primary objective (Tucker, 2007).

Unlike typical frontline workers, they may not be stationed at specific sites, such as retail stores, hospitals, dealerships, or factory floors. They may operate in a specific territory and travel between locations to conduct their duties. Given the complexity and uncertainty of day-to-day tasks, such roles often experience a high work withdrawal (reduced work effort) (Chong et al., 2020; Deery et al., 2002). Being field based, these positions pose challenges for managers in overseeing and standardizing work performance. However, existing studies in the operations management literature have focused on the motivation of stationed frontline workers, with less attention on the mobile field staff.

For instance, Tucker (2007), using survey data, studied the conditions under which frontline workers are motivated to take initiatives to help improve work systems for avoiding operational failures. The paper, using the context of hospital nurses, finds that managers must consciously design a work environment where such employees are motivated to talk about errors “without risk of punishment.” This helps reduce the chances of operational failures in the hospital setting. Lu et al. (2022) focuses on predictability of work schedules as a motivator for frontline workers and their operational performance. They studied the performance of cashiers in grocery stores when they have more consistent workdays in a week and work hours in day and its impact on worker productivity. They found an average increase of 1-2% in worker productivity with consistent scheduling. In another study, Pendem et al. (2024) studied agricultural workers engaged in mechanized harvesting to understand how work pauses can motivate such workers to perform better. They find that brief and unscheduled pauses improve worker productivity both in the long and short run and scheduled breaks are effective in the long run.

In addition, several of these studies focus on productivity, safety, and delivery as operational outcomes. Only a handful of studies have looked at how field staff motivation impacts monitoring, auditing, and quality assessment as an operational outcome. For instance, Ibanez and Toffel (2020) studied how designing effective work task sequences can help motivate field staff in detecting more issues during quality assessments. They studied regulatory food inspectors and found that when inspectors conduct inspections later in the day or around lunch time, they tend to find 11% fewer violations at food facilities. This is in part because of physical and mental fatigue induced due to travel between inspection sites. Short et al. (2016), studied the factors that affect the field staff's motivation to identify and report problems. They studied third party auditors who travel to supplier facilities to conduct facility inspections. They found that when auditors are paid by the audited suppliers, they tend to find fewer violations compared to when they are paid by the buyers.

There are other studies in accounting literature that look at the quality of financial audits by auditors. For instance, Hoopes et al. (2018) examined the relationship between field staff salaries and the facility-level audit quality. They used the context of financial auditors who travel to audit sites to ensure financial compliance. They used both absolute salaries and relative salaries of auditors in comparison to co-workers who have more traditional office-based roles, such as tax consultants and information systems. They found that employing auditors with lower salaries both in relative and absolute terms led to a higher percentage of client restatements (i.e., lower quality audits). Furthermore, Choi et al. (2012) studied how the geographic proximity between field staff and their clients or work sites impacts site quality assessment. They also explored the context of financial auditors and found that locally based auditors deliver higher

quality audit services than non-local auditors, due to less travel fatigue, more local knowledge, and increased productivity.

Our study focuses on regulatory food inspectors who travel within their assigned territory, monitoring the quality management practices of food facilities to ensure a safe food supply chain. Unlike existing literature, we focus on incentive design and study how multiple referent comparisons affect the quality assessments of the field staff. The goal of the study is to understand when and how leveraging incentive design, informed by multiple referent groups, can help improve violation detection by field staff employees. Next, we develop hypotheses using the four identified referent groups.

3. Hypothesis development

Equity theory suggests that employees, including field staff, compare the ratio of their input (efforts) to outcomes with those from a chosen referent group (Adams, 1965). This comparison drives their perception of fairness (distributive justice - degrees to which the outcomes are equitable) in the workplace and thus affects their motivation to perform a work task (Choi & Chen, 2007). Existing literature suggests that effects of salary comparisons on distributive justice can influence employee morale which can manifest as either enhanced work engagement or work withdrawal (Xu et al., 2024). For example, a low sense of distributive justice can motivate employees to reduce their work efforts and a higher sense can motivate them to increase work efforts (Berry et al., 2012; Podsakoff et al., 2007).

Organizational psychology literature suggests that as employees improve their external or internal pay standing, i.e., receive a higher salary compared to an internal or external referent group, they will demonstrate increased work engagement (Tröster et al., 2018). This phenomenon occurs when employees feel that their salary variance is favorable compared to a

specific referent group, potentially leading to feelings of guilt or a sense of indebtedness towards their organization (Brockner et al., 1985). According to Festinger's theory of cognitive dissonance, employees in such situations will be motivated to reduce the differences to achieve a sense of equity (Morvan & Connor, 2017). They may do so by channeling the tension to increase their inputs.

Positive differences for field staff compared to the referents can foster a greater sense of responsibility to justify their higher earnings. This sense manifests itself through the feeling of reciprocity, wanting to give back to their employer, resulting in increased operational efforts and better task performance (Baker & Bulkley, 2014). The fear of losing the advantage through job loss or salary adjustments may motivate them to work harder to maintain a favorable position (Repenning, 2000). The higher salary can also heighten a sense of pride in one's job, encouraging them to meet or exceed expectations (Tymon et al., 2010). It could also instill greater trust and confidence in one's own abilities, which is critical for field jobs where employees make high-stake decisions in external environments.

On the other hand, when employees have a lower external or internal pay standing, i.e., receive a lower salary compared to the referent groups, they will exhibit lower work engagement or work withdrawal (Li-Ping Tang et al., 2000). This phenomenon occurs when employees feel their salary is less favorable in comparison to both internal and external referent groups. This may result in feelings of resentment, frustration, and demotivation. Employees who experience negative differences may act to reduce their input (efforts). This manifests in the form of withdrawals, i.e., both work withdrawals (neglect, reduced performance, absenteeism) and job withdrawals (voluntary turnover, turnover intention).

We use the following theoretical understanding along with referent comparisons with four groups including self-referent, industry referent, internal peer referent, and operational interface referent to develop our hypotheses (as in Table 1).

3.1. Self-referent

According to Adams (1965), one of the most prominent referent groups that employees compare their salary to is self-referent, i.e., an employee's own progression internally within their own organization. It involves assessing one's own personal growth or regression over time given their input (or efforts) (Tremblay, et al., 1997). This referent group focuses on an individual's own growth rather than their relative standing among peers, industry standard, or external operational interface. It links an employee's own trajectory to their motivation and performance. The focus is on personal achievement and fulfillment of individual expectations.

This referent comparison is different than other comparisons in several ways. First, their salary raise may be tied directly to their own actions (Folger & Konovsky, 1989). Their task performance can result in a higher or lower salary in comparison to last year. This sense of agency or control may have a direct impact on their sense of satisfaction with the job and task performance in the current year. Second, employees have complete information about their salary and growth trajectory in self-referent comparison as compared to other groups such as peers where they may have incomplete information.

Third, the impact of self-referent comparisons is closely tied to an individual's identity and can have an emotional and psychological impact (Tremblay et al., 1997). Self-referent groups offer a micro-level perspective on one's own growth progression within an organization. Positive differences compared to one's own salary from last year can enhance an individual's personal sense of growth and career progression. This may encourage the individual to perform

better. Positive self-reference provides field staffers with a basis for assessing their input-to-outcome ratio, which can enhance their motivation (Cullen, 2024). Conversely, a less favorable input-to-outcome ratio can lead to dissatisfaction. It can adversely affect commitment and dedication towards an organization. Therefore, we propose the following:

Hypothesis 1: Field staff receiving a higher salary raise (cut) compared to their own salary from the previous year (a self-referent comparison) will exhibit higher (lower) task performance.

3.2. Industry referent

One of the prominent groups that employees select as a referent group for incentive comparisons can be employees external to their organization who are performing comparable work in a similar industry, called industry referents (Sieweke et al., 2017; Williams et al., 2006). This is particularly true for field staff employees, as the nature of their job involves operating in environments with high mobility, thereby providing them with opportunities to encounter employees from other organizations (Balogun et al., 2015). Field employees have frequent opportunities to interact with external referents that provide them with an awareness of their external standing.

This exposure makes external comparisons visible and accessible. Further, the operating procedures of field staff positions can be standard across the industry, for example, with respect to sales, service delivery, or inspections (Field & Dunn, 2013). This makes the comparability and assessment of incentives across external jobs possible. For example, FDA inspectors can compare incentives with Dept. of Transportation inspectors, or United States Dept. of Agriculture inspectors.

In addition, due to remote working conditions, field staff may feel isolated from their own organizational culture and practices (Mulki & Jaramillo, 2011). Moreover, field staff may have opportunities and the inclination to move within their industry making external comparisons important for career planning (Rubel & Prasad, 2016). These industry referents serve as standard benchmarks that influence employees' perceptions of their own incentives.

Industry referents offer a broad perspective on one's market value, emphasizing field staff's individual progress. Further, positive differences compared to industry referents may enhance an individual's personal sense of value and job security (Austin, 2006). This could encourage loyalty towards their organization due to favorable incentives. It can also create a strong incentive to increase input (efforts) to justify their salary. Literature suggests that it leads to a psychological contract with the employer where employees exhibit a greater sense of dedication and obligation towards their organization.

Overall, positive differences as compared to industry referents for field staff provides individuals with a macro-level evaluation of their input to outcome ratio providing a sense of motivation. The opposite is true for negative differences i.e., when field staff feels that their ratio of input to outcome is less favorable than the industry referent resulting in dissatisfaction and, potentially, anger. It could negatively impact field workers' satisfaction, making them less faithful towards their organization and its mission. So, we propose the following:

Hypothesis 2: Field staff with a higher (lower) salary compared to the typical salary within the industry standard (an industry referent comparison) will exhibit higher (lower) task performance.

3.3. Internal peer referent

According to organizational psychology literature, one of the most common referent groups that employees compare their salary with is their co-workers (Xu et al., 2024). Such comparisons occur internally within the organization and shape the perception of employees about their own standing among peers (Sweeney & McFarlin, 2005). These comparisons are made with respect to colleagues within an organization or a specific department. Employees compare job roles, working conditions, and incentives within the same organizational structure. This co-worker comparison can affect an employee's immediate work environment (Sweeney & McFarlin, 2005). These comparisons may be personal and tangible, affecting the day-to-day motivation of employees.

Moreover, employees may have a detailed knowledge about their co-workers' job roles and responsibilities, capabilities, and salaries and incentive structure, especially with government positions where there is salary transparency. Employees may be able to accurately assess the input (effort) to outcome (salary) ratio of their internal peers which can impact their motivation and work performance. High or low performance of employees can be recognized and rewarded (or penalized) within an organization. Such direct comparisons also have an impact on social dynamics and professional relationships among employees. The perception of differences can result in competition and friction or cooperation and teamwork affecting the morale and work congeniality of employees (Card et al., 2012).

According to psychology proximity literature and construal level theory, since the employees interact with their internal peers, the impact of salary differences can be pronounced and concrete (Trope & Liberman, 2010). People's thinking is shaped by the psychological distance of the event, i.e., if an event is more distant it will be construed as more abstract, while a closer or frequent event will have a more significant impact on employees (Trope & Liberman,

2010). Further, comparisons with internal peers can have an immediate impact on the perception of career progression and opportunities within the organization for an employee's growth.

Internal peers offer a micro-level perspective on one's value within an organization emphasizing field staff's individual progress. As discussed above, more payment compared to internal peer referents may increase an individual's personal sense of value, creating incentives to increase input (efforts) to rationalize their salary. Comparison with internal peer referents gives field staff a basis for indirectly evaluating their input-to-outcome ratio, potentially enhancing motivation (Adams, 1965). On the other hand, negative differences, i.e., when field staff perceives their input-to-outcome ratio as less favorable than that of the operational interface referents, can lead to dissatisfaction. This perception can negatively affect motivation and task performance. Therefore, we propose the following:

Hypothesis 3: Field staff with a higher (lower) salary compared to the salary of their coworkers within the same grade (an internal peer referent comparison) will exhibit higher (lower) task performance.

3.4. Operational interface referent

One of the other prominent external groups, with which field staff may have direct relational dependence, is the operational interface referents. These referents can directly affect the field staff's perceptions of their efforts and outcomes. This is because the nature of the field staff's job involves operating in external environments (Rubel & Prasad, 2016). For example, food inspectors must visit inspection sites and interact with the employees of the inspected facilities. This regular interaction facilitates ongoing comparison of work conditions, responsibilities, and compensation, making this referent group a prominent standard for evaluating incentive differences (Trevor & Wazeter, 2006). For instance, in our context, food

inspectors spend time conducting inspections at inspection sites. During these visits, they have numerous opportunities to interact with employees of the inspected site, which gives them an understanding of their external standing.

Field staff may be able to directly relate to operational interface referents in terms of job functions, duties, and organizational context as they operate in a similar operational environment. Additionally, due to career aspirations, field staff often benchmark against employees at the operational interface. For example, federal food inspectors working for the FDA may aspire to work for the private firms they inspect to advance their careers and obtain better incentives. They may also be interested in gathering information about the incentives for employees at the firms they inspect. Federal employees, especially those trained in critical roles as food inspectors, serving as gatekeepers for regulatory standards, can be highly sought after by industry (Lucca, 2014). In some cases, these inspectors may have previously worked in the food industry before transitioning to their government positions. They may maintain close connections within the industry they inspect and possess detailed knowledge about its incentive structures. Furthermore, since employees in the food industry often share professional networks or attend the same professional conferences or training sessions as food inspectors, the inspectors have numerous opportunities for informal interactions.

Overall, perceiving positive and negative differences compared to operational interface referents gives field staff a basis for directly evaluating their input-to-outcome ratio, potentially enhancing, or adversely affecting their motivation. Therefore, we propose the following:

Hypothesis 4: Field staff with a higher (lower) salary compared to employees at operational interfaces they inspect (an operational interface referent comparison) will exhibit higher (lower) task performance.

4. Methods

4.1. Research setting: food-safety inspections

We test our hypotheses in the context of U.S. Food and Drug Administration (FDA) based food safety inspections (FDA, 2023). FDA is responsible for monitoring about 80% of the food supply chain in the United States. They assess the quality of food facilities, i.e., firms where food is processed, packed, stored, and distributed on a regular frequency. High risk facilities are inspected at least once in every three years and low risk facilities are inspected at least once in every five years (FDA, 2023). These inspections are performed by individual inspectors, or a team of inspectors based on several factors including the complexity of the facility inspection. However, we focus on inspectors who work independently because in team-based inspections it becomes difficult to discern individual behaviors from those of their colleagues (Ibanez & Toffel, 2020).

These inspectors, officially referred to as consumer safety officers, have designated geographic areas where they are responsible for identifying and controlling the risk of foodborne illnesses. The effectiveness of such an inspection relies heavily on the accuracy and thoroughness of individual inspectors. Inaccuracies can have a significant negative impact on the health of consumers. Moreover, as per prior studies, since violations require substantial documentation, and re-verification from the divisional offices, the work performance of inspectors often results in the underreporting of real problems rather than the reporting of non-existent ones (Ibanez & Toffel, 2020).

Further, FDA inspectors are General Schedule (GS) exempt employees, i.e., they are exempt from Fair Labor Standards Act's (FLSA) overtime regulations and are not eligible for overtime pay. Since these are federal employees, they follow the standard guidelines of the

Office of Personnel Management (OPM) for salary increases and promotions. Every employee generally receives a standard annual pay raise as approved by Congress which goes into effect at the start of the first full pay period of the new year. Our data set ranges from January 2015 to December 2019. Within our dataset, FDA employees received an average of 1.16% annual (standard) pay raise (FederalPay.org, 2024). In addition to standard pay raises, FDA inspectors could also receive performance-based raises based on their individual performance. This setting provides us an opportunity to compare salaries of employees across different referent groups and control for several key factors to understand the precise impact of pay disparities on employee outcomes.

Prior studies have examined salary comparisons and their effect on employee motivation in several different contexts. For example, Xu et al. (2024) conducted a scenario-based survey in a mixed professional setting with participants from different occupations, such as managers, professionals, technicians, skilled agricultural workers, and machine operators. They studied how the variance between internal and external salary comparisons affects perceptions of fairness and subsequent job withdrawal. Sieweke et al. (2017) studied the relationship between employees' actual internal and external pay standing and their work performance. They used the context of players in the National Basketball Association over 12 seasons to test their hypothesis. They find that a lower employee's objective internal and external pay standing negatively affects their work standing. Further, Torre et al. (2015) studied how the compensation design influences worker behavior at the organizational level focusing on absenteeism. They used the context of Italian manufacturing firms and blue collared workers to study their research question suggesting that lower or no differences in external pay leads to lower absenteeism.

Other authors have specifically focused on government employees for salary comparison studies because government work offers a more standardized setting and homogeneity of comparison. Trevor and Wazeter (2006) studied how social comparisons affect employees' reactions towards pay. They used data on salaries of more than 2000 public school teachers across 400 schools districted and found how pay equity perceptions resulted from the interaction of internal and external pay conditions. Krasna et al. (2024) analyzed salary comparisons of government employees using equivalent private employers in the same industry. They used the context of U.S. public health occupations to test their hypothesis and found that Inspection and Compliance roles in the government paid at least 10-19% less compared to the private sector.

Unlike prior literature, our context of FDA food inspectors offers several unique advantages to study the impact of referent group comparisons on operational outcomes. Firstly, FDA inspectors have clearly defined referent groups both for external and internal comparisons allowing for precise comparisons and assessments of salary. Second, the pay structure of FDA inspectors with annual and performance-based raises is reasonably standardized. It provides a controlled environment to study the impact of salary comparisons on employee motivation and work performance. Prior contexts often offer high variability in compensation structures. Third, studying salary dynamics in an operational context, such as regulated and mandated inspections, provides a measurable and stable performance metric.

Fourth, our context provides a unique opportunity to study the boundary spanning impact of incentive design. We use a unique government and private sector interaction to inform our study where the government employees responsible for regulating private firms are impacted by their compensation structures. Fifth, given that FDA inspectors are federal employees the data regarding their pay and promotion history is available. This provides more data transparency and

legitimacy to our study. Lastly, since FDA is a mission driven organization where employees work to safeguard the public welfare, this context provides an opportunity to explore how the inspectors' commitment to public health safety (intrinsic motivation) interacts with their extrinsic motivation of salary and job conditions affecting their performance and satisfaction.

4.2.Data

We compiled a detailed and extensive dataset from various sources. Our primary dataset comes from FDAlytics, a database company that gathers data from food inspection enforcement reports through Freedom of Information Act (FOIA) requests to the FDA. It utilizes the information to develop food inspection dashboards for food firms that purchase the database to understand regulatory compliance. The dataset encompasses information on the inspected food facility, including its name, a unique identification number, and its address with the city, state, and zip code. It also contains details about the inspector and the start and end dates of the inspections. We augmented this database by incorporating information from the publicly accessible FDA inspection dashboard, utilizing unique identifiers. This additional data includes the number of citations, the depth of inspections, and the severity of the citations.

Further, we scrape salary information of all FDA inspectors from 2013 to 2020 from a publicly available database called OpenPayrolls.com. It is one of largest public payroll transparency websites in the U.S. which allows free access to public records of government and public employees. This database was corroborated by another publicly available database called FederalPay.org. These databases provide information on inspector salary by years, their federal grade, salaries compared to other federal employees and inspector age. We use inspector names, including their middle names, to match this data to the FDAlytics database. We further use the Namsor database to determine the gender of inspectors (Sebo, 2023). Namsor employs artificial

intelligence (AI) to analyze name structures to determine factors related to inspectors, such as gender that we confirm via LinkedIn using exact inspector names.

We also use the U.S. Bureau of Labor Statistics (BLS) database to download information on average salaries in a particular county by specific industry classified using the North American Industry Classification System (NAICS) code by years. We use this information to understand the average salary of employees of the firms being inspected. We match this information to the FDAlytics database using the county and state names. Further, we downloaded information on recalls of each firm from the FDA recall database. This database provides information on the date of recall, number of products recalled, and severity of recalls. The recalls database provides firm FEI or unique ID to match them with the original database.

We also downloaded state level information from the U.S. Bureau of Economic Analysis (BEA) on state Gross Domestic Product (GDP), state population etc. to control for state characteristics. We used the state names to match this data to the original database. Lastly, we downloaded information on the change in the U.S. political landscape from the MIT election lab database. We download information on political affiliations of U.S. presidents, state governors, senators, house of representatives both at the state and federal level to calculate a political alignment index (PAI). Again, we used the state names to match this data to the original database.

Our comprehensive dataset includes all documented FDA food facility inspections in the U.S. from January 2013 to December 2019. We focus on data prior to 2020 because COVID-19 could have impacted inspections and their outcomes. From 67,854 inspections in the raw dataset, we exclude 106 duplicates due to data logging errors. Further, we dropped 12,262 team inspections that were conducted by more than one inspector as we focus on single inspectors

only. We dropped 14,782 observations from 2013 and 2014 as our salary data ranged from 2015 to 2019. Finally, we exclude 24,322 observations for which we were not able to calculate relevant variables (for example, inspector experience) or were missing lagged data necessary to operationalize prior inspection outcomes and prior inspection facilities outcomes. This yields a sample of 16,382 inspections of 14,441 food facilities conducted by 594 inspectors from Jan'2015 to Dec' 2019.

To gain a deeper insight into the context, we a) conducted interviews with FDA officials, including senior advisors and directors in the Office of Food Safety, b) reviewed the FDA's Investigations Operational Manual 2024, and c) analyzed feedback from 80 FDA Inspectors on Glassdoor from 2010 to 2023. Our discussions with FDA officials, supplemented by data from Glassdoor (a website and platform where employees and former employees anonymously review companies and their management), indicate that FDA Inspectors have several concerns with salary differences. We performed a thematic analysis of the feedback provided by 80 FDA inspectors. This involved categorizing each piece of feedback related to salary differences and then calculating the frequency of mentions for each category.

Some of the comments of FDA inspectors regarding salary differences included *"Salaries are not competitive with the private sector," "Salary less than private or academic world. Not much opportunity for advancement/promotion," "Salary - low to start and difficult career advancement. Private sector pays double or triple for the same skills. Be careful about retirement," "Pay falls well below comparable positions outside of government. Limited promotion potential," "Doesn't pay close to what our counterparts in private industry get paid."*

4.3.Measures

4.3.1. Dependent and independent variables

Our primary dependent variable is *violations*. For our study, we measure it as the number of violations cited in each inspection as in other comparable studies (e.g., Ibanez & Toffel, 2020; Scott et al., 2020; Short et al; 2016).

Self-referent is operationalized as the percentage pay raise of the inspectors from the last calendar year. FDA inspectors typically receive an annual pay raise that goes into effect at the start of the first full pay period of the new year i.e., January of every year. We subtract last year's salary (t-1) of the inspector (i) from their current salary in year (t) divided by the last year salary (t-1) of the inspector (i) multiplied by hundred. All salaries are in USD.

$$Self\ referent_{it} (\%) = \left(\frac{Salary_{it} - Salary_{it-1}}{Salary_{it-1}} \right) \times 100 \quad (1)$$

Internal peer referent is operationalized as the percentage difference of the inspector salaries (USD) from their FDA co-workers within the same federal grade. We use the salary comparison within the same federal grade. We subtract the average salary of the co-workers (c) of inspector (i) for that year (t) in the same grade (j) from the salary of the inspector (i) divided by their co-worker's salary multiplied by hundred.

$$Internal\ peer\ referent_{it} (\%) = \left(\frac{Salary_{it} - Salary_{cjt}}{Salary_{cjt}} \right) \times 100 \quad (2)$$

Industry referent is operationalized as the percentage difference of the inspector salaries (USD) from other federal employees. We subtract the average salary of the federal employees within the same grade (f) for the year (t) from the salary of the inspector (i) in year (t) divided by the average salary of the federal employees in similar position (f) for the year (t) multiplied by hundred.

$$Industry\ referent_{it} (\%) = \left(\frac{Salary_{it} - Salary_{ft}}{Salary_{ft}} \right) \times 100 \quad (3)$$

Operational interface referent is operationalized as the percentage difference of the inspector salaries (USD) from the average salary of employees in the industry within the county of the firm being inspected. We use it as a proxy for the salary of employees of the firm being inspected. We subtract the average salary of the employees in industry (d) in county (o) in year (t) from the salary of inspector (i) in year (t) divided by the average salary of the employees in industry (d) in county (o) in year (t) and multiplied by hundred.

$$\text{Operational interface referent}_{it} (\%) = \left(\frac{\text{Salary}_{it} - \text{Salary}_{dot}}{\text{Salary}_{dot}} \right) \times 100 \quad (4)$$

4.3.2. Control variables

4.3.2.1. Inspector specific control variables

We control for several inspector specific characteristics that can otherwise confound the results and affect the relationship between the hypothesized variable and the dependent variable of interest. We control for inspector *experience* as it can affect the expertise or proficiency with which inspections are implemented (Short et al., 2016). It helps reduce bias in the estimated effect of salary differences on the inspection outcomes controlling for heterogeneity amongst all inspectors. By controlling inspector experience, we can better ensure that citations are due to salary disparities rather than variations in inspector experience. We measure it by counting the number of inspections an inspector has conducted at any food facility from the start of our sample period up to the beginning of the current inspection (Ibanez & Toffel, 2020).

We control for the *average time* taken by an inspector to complete inspections prior to the present inspection. It helps in controlling for performance consistency of inspectors. If an inspector takes longer, it might mean thoroughness or inefficiency. It can also help adjust for inspector personality and training variability. Some inspectors might be inherently slower than others or may lack the necessary training to complete the inspections faster. We measure it as the

average number of days taken by the inspectors to complete inspections prior to the current inspection.

We control for the inspector *gender* as well. Prior research in operations management and organizational behavior suggests that there are inherent differences in the work style and decision-making process of males and females. While males might have a more strategic approach towards inspections, females might be more nuanced (Short et al., 2016). We use a gender dummy coded as 1 for male and 0 otherwise.

We also control for *inspector age*. Age can influence several different aspects of inspector-inspection outcomes. As per prior research, older inspectors might be slow but thorough as compared to younger inspectors who might be fast but impatient. Age might also reflect experience, judgment, and inspector's protocol adherence or deviation. We measured it as the age of inspectors (in years) at the time of inspection.

We control for inspector's *site-specific experience*. Prior research by Ibanez and Toffel (2020) and Ball et al. (2017) suggests that site-specific experience impacts the familiarity of inspectors with site employees, processes, and culture. Inspectors might also be intimately familiar with inspection issues at that particular site. We measure it as the number of inspections the inspector conducted at that site prior to the current inspection within the sample period.

We also control for *prior inspection outcomes* of the inspector. According to Ibanez and Toffel (2020), the outcomes of the inspection conducted immediately prior to the current inspection may shape an inspector's perception about the current facility. There could be a negativity or positivity bias. For example, a prior facility with poor conditions can negatively impact an inspector's beliefs about compliance of the food industry in general. We measured it

as number of citations of the facility immediately prior to the current inspection inspected by the inspector.

We also control for *inspector workload*. An increased workload may burn out inspectors leading to increased mental and physical fatigue. This may also suggest that some inspectors have an increased burden as compared to others because of experience, qualifications, or age. We measure it as the number of inspections an inspector conducted in the month of the current inspection.

Lastly, we control for *inspector designation*. We rank order the title of the inspectors conducting an inspection in the increasing order of responsibility. The designation of the inspectors could impact the number of inspections conducted by each inspector and their stringency. Supervisors might not be able to conduct as many inspections as an early career inspector. Furthermore, as we want to assess the impact of salary differences on inspection outcomes, controlling for designation helps ensure that observed effects are confounded by the grade of the inspectors.

4.3.2.2. Firm specific control variables

We control for several firm specific characteristics that may otherwise confound the results and affect the relationship between salary disparities and the number of citations. Firstly, we control for *Firm FDA distance*. This is measured as the distance in miles between the facility being inspected and the local FDA divisional headquarters. The Office of Human and Animal Food Operations (OHAFO) that manages the inspection process has 12 divisional offices across United States with each managing field operations of about 3 -5 states. The food inspectors report to their respective divisional headquarters. The distance of the facility from the FDA regional office could impact firm's compliance with regulations, networking opportunities and

training opportunities with FDA. For example, a facility closer to FDA might have better compliance as compared to others.

We control for *past inspection outcomes* of the facility. We measure it as the number of citations the current facility had in its last inspection conducted by an inspector. This helps us control for the quality performance of the firm. In our interviews with the FDA officials and inspectors, they suggested that the past quality performance of the firm is a primary criterion used to identify and prioritize firms for inspections. It can also shape the perception of the inspectors about the firm's quality performance.

We control for the recalls in the previous year of the firms. We measure it as the number of recalls the firm had in the year prior to the current inspection. Recalls are defined as a voluntary or mandatory request by the firm to return unsafe products back to them (Steven et al, 2014). Recalls can affect firm reputation, result in legal actions, leading to financial impacts. Firms with product recalls in the recent past might be more cautious about their quality performance. As per the literature, recalls might also be a proxy for firm reputation with greater numbers of recalls suggesting a lower reputation.

We also control for *inspection depth*. We measure it as the number of areas that were covered in the current inspection by the inspectors. The inspection depth helps control for firm level complexity. A firm could be inspected for a maximum of six areas, such as pesticides and chemical contaminants, biological hazards, natural toxins, food composition, standards and labelling, food and color additives, and food handling and storage practices. If the inspector covers all 6 areas in an inspection this variable is coded as 6 (the maximum value). Depth represents the intensity of an inspection. If the facility is perceived at a higher risk level or more complex it might get more in-depth inspections.

Lastly, we also control for the inspection duration. We measure it as the number of days the inspector takes to complete the current inspection. We use inspection duration to control for the operational scale of the facility. A facility with a larger operational scale might require longer inspections due to greater numbers of processes, larger size, and greater numbers of products being manufactured. This variable helps us account for the challenges of inspecting a facility with a larger, more complex operation.

4.3.2.3. State specific control variables

We control for several characteristics of the state in which the facility being inspected is located. We believe that location can confound our results, impacting the relationship between salary differences and the inspection outcomes. Firstly, we control for state Gross Domestic Product (GDP). We measure this variable as the GDP of the state as a percentage of U.S. GDP. It helps us control for the relative economic size of a state within the national context. The variable allows for standardized comparisons of economic activity across states.

We control for the state employment rate. It is measured as employment rate as a percentage of U.S. employment rate. The variable allows for comparison across different labor markets conditions. It can help to account for differences in employment levels that could affect inspection outcomes. For example, facilities in a state with a low employment rate might have easier access to qualified quality professionals, process engineers, and other expert resources.

We control for *state participation* in the federal inspection process. FDA contracts its inspections to state health and agriculture departments to assist in conducting inspections. Our study focuses on inspections conducted by FDA officials only. But we control for state participation, measured as the ratio of inspections conducted by state personnel in a year compared to the total inspections in that state-year. This variable allows us to control for the

extent of state involvement in regulatory activities. Greater state participation may impact FDA inspection frequency, quality, and outcomes.

We control for the political alignment between the state and federal government. We use a political alignment index (PAI) developed by Kim et al., (2012). It is a state-level measure of the degree of presidential party control across the political power chain. It is calculated as:

$$PAI = \frac{1}{3} \times S + \frac{1}{3} \times R + \frac{1}{3} \times G \quad (5)$$

where S is the fraction of the state's two federal senators that are members of the president's party. R is the percentage of a state's House of Representatives that belong to the president's party. G is a dummy for state governor equal to 1 if the governor belongs to the president's party. This variable helps control for the political outlook and congruence that might influence regulatory enforcement. It helps assess the impact of political factors on the effectiveness of inspections outcomes. The variable ranges between 0 and 1 where 1 would suggest a high political alignment between the state and the federal government.

Lastly, we control for the political affiliation of the FDA commissioner. FDA commissioners are political appointees, and their strategic leadership may be impacted by the philosophy of the party of their affiliation. For example, Republicans might have a more pro-industry stance, while Democrats might have a more consumer centric stance affecting the stringency of inspection and their outcomes (Kim et al., 2012). We measure this variable as a dummy coded 0 for a Republican commissioner and 1 for a Democrat appointee. Table 2 presents summary statistics of the dependent, independent and control variables.

Table 2. Summary Statistics

Variable	Description	Mean	Standard deviation	Minimum	Maximum
<i>Violations</i>	Number of violations cited in each inspection	0.86	2.01	0	28
<i>Self-referent</i> (Hypothesis 1)	Percentage pay raise of the inspectors from the last calendar year	4.50	11.68	-68.75	40.69
<i>Industry referent</i> (Hypothesis 2)	Percentage difference of the inspector salaries from other federal employees in similar positions	3.26	24.46	-53.00	82.20
<i>Internal Peer referent</i> (Hypothesis 3)	Percentage difference of the inspector salaries from their FDA co-workers within the same federal grade	-0.81	15.10	-42.46	65.68
<i>Operational interface referent</i> (Hypothesis 4)	Percentage difference of the inspector salaries from the average salary of employees in the industry within the county of the firm being inspected	96.47	80.72	-92.99	270.56
<i>Experience</i>	Number of inspections an inspector conducts at any food facility from the start of our sample period till the current inspection	72.54	52.13	1	381
<i>Average time</i>	Average of all the durations of inspections performed by the inspector from the start of our sample period till the current inspection (in days)	3.22	2.16	0	21.57
<i>Gender</i>	Gender dummies coded as 1 for males and 0 for females	0.55	0.50	0	1
<i>Inspector age</i>	Inspector age (in years) at the time of inspection	42.93	11.75	21	80

<i>Site-specific experience</i>	Number of inspections the inspector has conducted in the current facility till the current inspection	0.65	5.04	0	105
<i>Prior inspection outcomes</i>	Number of violations the inspectors found at their last inspection prior to the current inspection	1.02	2.27	0	30
<i>Inspector workload</i>	Number of inspections an inspector conducts in the month of the inspection	3.37	2.50	1	21
<i>Inspector designation</i>	Title of the inspectors in the increasing order of responsibility	5.58	2.75	1	21
<i>Firm FDA distance</i>	Distance (in miles) of the inspected firm from the FDA divisional office	231.06	183.84	13.43	604.49
<i>Past inspection outcomes</i>	Number of citations the current facility had in their last inspection conducted by any inspector	0.49	1.75	0	26
<i>Recalls last year</i>	Number of recalls the firm had in the year prior to the current inspection	0.02	0.14	0	4
<i>Inspection depth</i>	Number of areas that were covered in the current inspection by the inspectors	1.17	0.45	1	6
<i>Inspection duration</i>	Number of days the inspector takes to complete the current inspection	1.15	1.76	0	5
<i>State GDP</i>	Gross Domestic Product (GDP) of the state in which the inspected firm is located as a percentage of the U.S. GDP	4.74	4.52	0.16	14.23
<i>State Employment</i>	Employment rate of the state in which inspected firm is located as a percentage of the U.S.	100.90	5.97	81.09	132.50
<i>State participation</i>	Ratio of inspections conducted by a state in a year as compared to the	0.01	0.01	0	0.13

	total inspections in that state-year				
<i>Political alignment between state and federal govt.</i>	It is state-level measure of the degree of presidential party control across the political power chain (Kim et al., 2012)	0.49	0.33	0	1
<i>Political affiliation</i>	Dummies to indicate whether the FDA commissioner was a Republican (0) or Democrat (1) appointee	1.70	0.46	1	2

Note. $N = 16,382$

4.4. Empirical specification

We tested our hypotheses by estimating the fixed effects of negative binomial regression using a panel structure of the data (Steven et al., 2014). Our dependent variable is a count variable, i.e., the number of citations at a specific inspection. Prior literature has used a negative binomial regression when modelling count data that exhibits overdispersion (variance is greater than mean) (Hilbe, 2011). This happens when data does not fit the assumptions of the Poisson distribution due to heterogeneity in the data. Our unit of analysis is inspection of a facility. We have a double panel structure where each facility is inspected by an inspector. But our specification exploits the panel structure of our dataset where an inspection can be viewed as the n^{th} inspection of the inspector i in our sample.

Our independent variables include four referent groups for FDA food inspectors: (1) *self-referent* – pay raise of inspectors compared to last year’s pay; (2) *Internal peer referent* – salary as compared to co-workers; (3) *Industry referent* – salary compared to other federal employees; (4) *Operational interface referent* – salary of the FDA inspectors compared to employees of the facilities inspected by these inspectors.

We include several fixed effects in our model to control for unobserved heterogeneity across different entities and times. First, we control for *facilities' nth inspection*. Prior research suggests that facilities improve their compliance over successive inspections (Ibanez & Toffel, 2020; Toffel et al., 2015). Similarly, we constructed a set of indicator variables to account for the order of facilities inspections, i.e., whether the current inspection is a facility's 1st, 2nd, or 3rd inspection, etc. within our study period. The first inspection serves as the baseline category from which comparisons are made.

Further, we include fixed effects for the FDA divisions. As discussed before, FDA has 12 divisions across the United States which manage the inspection operations. Each division is responsible for 3-5 states and has a unique set of leaders and food safety challenges. In our analysis of FDA Inspector's comments one inspector indicated "*The workload disparity between western and eastern divisions is unbelievable.*" Since every division head might have variations in how they approach inspections, we control for time-invariant attributes of these divisions by including *FDA division* fixed effects.

We include *day* fixed effects, i.e., days on which the inspection was started to control for workday effect and Monday effect (Yao et al., 2020). We also include the *month* fixed effects, i.e., the month in which the inspection was conducted to control for seasonality. Further, we include fixed effects for *years* in which inspections were conducted to control for unobserved heterogeneity over time. Lastly, we include inspector fixed effects to control for unobserved attributes that might affect the inspection outcomes, although, we include several inspector level controls to ensure that we capture the effects of inspector experience, gender, age, inspection efficiency and other factors that might affect their stringency (Short et al., 2016). Furthermore, we include several facility level factors to control for the effects of facility quality performance,

reputation, operational scale, operational complexity, and its location to ensure that firm attributes are confounding the outcomes. Lastly, we control for several state-wide attributes as well.

We estimate the following model to evaluate the effect of different salary comparisons on the inspection outcomes of the facilities.

$$\begin{aligned} Violations_{ni} = & \alpha_0 + \alpha_1 Self\ referent_{ni} + \alpha_2 Internal\ peer\ referent_{ni} + \\ & \alpha_3 Industry\ referent_{ni} + \alpha_4 Operational\ interface\ referent_{ni} + X_{ni}\alpha + \beta_{w(mt)} + \gamma_t + \\ & \theta_{m(t)} + \tau_{l(ni)} + \mu_i + \rho_{i(n)} + \varepsilon_{it} \end{aligned} \quad (6)$$

Violations is the outcome variable, with *n* signifying the inspection and *i* indicating the inspector. *X* vector encompasses a range of controls variables. The term ε_{it} captures the error terms, and α stands for the parameters being estimated.

γ_t – are the year fixed effects

$\theta_{m(t)}$ – are the month fixed effects

$\beta_{w(mt)}$ – are the day of the week fixed effects

$\tau_{l(i)}$ – are the FDA division fixed effects

μ_i – are the inspector fixed effects

$\rho_{f(n)}$ – are the facilities' *n*th inspection fixed effects

5. Results

5.1. Model results

We estimate the count model using the fixed-effects negative binomial regression where the dependent variable is the number of citations during each inspection conducted by an inspector (Ibanez & Toffel, 2020). We report the correlations and primary results in Table 3 and

Table 4 respectively. Negative binomial regression enables consistent estimation by accounting for the presence of overdispersion.

Further, multicollinearity may not be a concern in our model because the variance inflation factors (VIF) are less than 10 (mean VIF = 2.22) for all hypothesized and control variables (Shrestha, 2020). We interpret the coefficients of our hypothesized variables as evidence of bias. Our model specification accounts for various factors that influence the number of detected violations. We attribute deviations from the actual number of violations to under-detection resulting from salary disparity bias. Negative coefficients are seen as indicating the level of under-detection, while positive coefficients suggest the extent to which under-detection is reduced. We use incidence rate ratios (IRR) to infer the sizes of these effects (Ibanez & Toffel, 2021).

Table 4 displays the results of the empirical tests for Hypotheses 1 through 4. Initially, we present the findings from Model 1, a model that includes only control variables, with a particular focus on the control variables related to inspectors. We only interpret coefficients with p-values of 0.05 or lower. The estimated coefficient of inspector experience is negative and significant ($\beta = -0.004$ $p < 0.001$, IRR = 0.996) indicating that with every additional inspection, inspectors detect 0.4% fewer violations. For site-specific experience, we observe that the coefficient is negative and significant ($\beta = -0.165$, $p < 0.001$, IRR = 0.848), indicating that one more prior visit by an inspector result in 15.2% fewer violations. fewer violations compared to their colleagues. This finding aligns with previous research (Ball et al., 2017; Ibanez & Toffel, 2020). Further, for inspector workload, we find that the coefficient is negative and significant ($\beta = -0.181$, $p < 0.001$, IRR = 0.834). This suggests that with each additional inspection that the inspector handles in a month the number of citations decreases by 16.6%.

Next, we interpret firm level and state level control variables. For a firm's past inspection outcomes, we find that the coefficient is positive and significant ($\beta = 0.107$, $p < 0.001$, IRR = 1.113). This suggests that a firm's prior history of poor-quality performance will influence the inspection outcomes of their current inspection. One additional citation of the current facility's immediate last inspection will lead to a 11.29% increase in citations in their current inspection. We use inspection depth as a proxy for firm level complexity. The coefficient for inspection depth is positive and significant ($\beta = 0.227$, $p < 0.001$, IRR = 1.254). This suggests that as the complexity of the firm increases by a unit, the number of citations for the facility increases by 25.48%. Further, we use inspection duration as proxy for firm's operational scale. We find the coefficient to be positive and significant ($\beta = 0.019$, $p < 0.001$, IRR = 1.019). This suggests that with each additional day of inspection the number of citations for the facility increases by roughly 2%. Lastly, for state control variables, we find that the political alignment index to be negative and significant ($\beta = -0.185$, $p < 0.001$, IRR = 0.8311). This suggests that as the political alignment of the state in which inspection is being conducted increases with the federal government, the number of citations at such facilities decreases by 16.89%.

We use Model 5 in Table 4 to interpret our main hypothesized variables. For the self-referent variable, we find that the coefficient is negative and significant ($\beta = -0.007$, $p < 0.001$, IRR = 0.993). This suggests that with every percentage point increase in the salary of inspectors compared to last year's salary, the number of citations decreases by 0.7%. We don't find support for hypothesis 1, suggesting that when inspectors benchmark their current salary against their salary from the previous year, a higher salary is associated with a decrease in effectiveness in detecting violations indicating that pay raises may not have the expected effect on performance. Rather, an increase in salary may lead to an increased sense of satisfaction or complacency, thus

reducing an inspector's motivational drive. The results also align with Herzberg's Two-Factor Theory perspective that increases in salary (a hygiene factor) do not enhance motivation or performance, as evidenced by the reduction in citations (Dartey-Baah, 2011).

When it comes to industry referents, i.e., salaries of inspectors as compared to other federal employees, we find that the coefficient is positive and significant ($\beta = 0.014$, $p < 0.001$, IRR = 1.014). This suggests that higher salaries of the inspectors as compared to their industry benchmark may positively motivate inspectors. When the salary of inspectors is one percentage point higher than federal employees, the number of citations increases by 1.41%. As per equity theory, as the inspector compares their ratio of contributions to salary outcomes with the ratio of their chosen referent (other federal employees in this case), they may try to restore differences by increasing their inputs to achieve higher outputs to justify higher salaries. This result supports our hypothesis 2.

For internal peers or co-workers as the referent group, we find that the coefficient is negative and significant ($\beta = -0.006$, $p < 0.001$, IRR = 0.994). This suggests that higher salaries of inspectors as compared to their internal peers negatively motivates inspectors. With a one percentage point higher salary of inspectors as compared to other FDA co-workers, the number of citations decreases by 0.6%. This result does not support hypothesis 3.

Lastly, for operational interface referents, i.e., salaries of inspectors as compared to the employees of facilities they inspect, we find that the coefficient is negative and significant ($\beta = -0.002$, $p < 0.001$, IRR = 0.998). This suggests that higher salaries of inspectors as compared to the county average of salaries of employees in the food industry negatively motivates inspectors. With one percentage point higher salaries of inspectors as compared to operational interface referent, the number of violations detected decreases by 0.2%. Similar to co-workers, if

inspectors feel they are overcompensated compared to the industry employees they inspect, they reduce their efforts. If they feel they are undercompensated, they increase their efforts to prove their worthiness for better pay. This result does not support hypothesis 4.

5.2. Econometric issues

We examine each hypothesized variable to identify any potential endogeneity. We control for several related variables significantly reducing the chances of endogeneity due to omitted variable bias. We control for several inspectors, firm level, and state level attributes that rule out alternative explanations for variability in the inspection outcomes. We also use several fixed effects dummies for days, months, years, FDA divisions, inspectors, facilities' nth inspection accounting for time-invariant and unobserved characteristics.

Further, we believe that it is unlikely that our variables have simultaneity or reverse causality issues. For example, the number of citations found in an inspector's current inspection cannot influence how much salary raise an inspector received last year, although inspection outcomes may have an impact on the inspector's future salary raises. Similarly, an inspector's performance today might impact their salary as compared to other federal worker and co-workers in the next calendar year as all salary raises in federal jobs are applicable at the start of a new calendar year. Also, outcomes of current inspections will not impact the current salary of inspectors compared to employees of the industry being inspected.

5.2.1. Selection bias in inspector assignment

Selection bias could be a concern if inspector assignments are not random but influenced by inspector and facility characteristics, such as a firm's compliance history. However, by controlling for inspector and facility characteristics, we try to control for these potential non-random assignment factors (Ibanez & Toffel, 2020). For example, inspectors that have a greater

salary raise from last year find fewer issues at a facility. This could be possible if inspectors with higher raises (high performing inspectors) are assigned to more compliant facilities to incentive them further, while inspectors with low raises (low performing inspectors) are assigned to less compliant and more difficult cases. Inspectors with higher salary raises might self-select into some of these facility inspections as well. Although our interviews with FDA inspectors and officials suggest that the assignment of FDA inspectors is random, and we account for inspector fixed effects and various attributes, we still empirically test for the presence of systematic non-random assignment.

5.2.2. Instrumental variable approach for addressing endogeneity

Regardless, we believe that there could still be endogeneity issues that could arise, potentially biasing the results. For example, despite a comprehensive list of control variables, we might still be missing crucial variables that influence both salary differences and the number of citations. For example, an inspector's social capital with the organization might affect their salary and inspection effectiveness. There could be other firm-related control variables that we are missing, especially when not using firm level fixed effects. Further, there might be issues of reverse causality as salary adjustments (raises) may occur in response to inspection outcomes. There is a possibility that inspectors might be rewarded with salary increases based on their performance in previous years, i.e., the causal relationship might be confounded by this feedback loop where past performance is affecting current performance - current salary is, in part, a reflection of their past inspection outcomes. To address these concerns, we use the Instrumental Variable approach.

We assume that all our hypothesized variables are endogenous. We verify and account for potential endogeneity by using the two-stage control function approach (Palit et al., 2022;

Wang et al., 2023; Woolridge, 2015). In the first stage, we regress our hypothesized variables including Self-referent, Industry referent, Internal peer referent, and Operational interface referent on instrumental variables and the control variables. The predicted residuals from this stage are used as the endogeneity correction terms in the second stage.

A valid instrumental variable needs to satisfy the conditions of relevance and exclusion (Baiochi et al., 2014). We select three instrumental variables that satisfy these requirements. For instrumental variables we use two industry level variables, the average annual salary of employees in non-food industries within the state (non-food industry pay), and the cost of living in the inspector's jurisdiction region (cost of living). We also include a firm level instrumental variable; average salary raises of other inspectors in the same grade but in other FDA divisions.

The conceptual rationale for including non-food industry pay is that it serves as a regional economic indicator unrelated to the inspected industry (Gishkayeva et al., 2019). It reflects the overall economic health of the region and is external to the operations of the FDA and its inspectors, i.e., exogenous to the inspection process itself. Average pay in sectors unrelated to the food sector could influence overall salary levels in the region without directly affecting the inspection outcomes. As per the Office of Personal Management (OPM), the regional economic conditions influence the federal wage structures broadly, but would not affect how inspectors conduct their duties, making it a suitable instrument.

Further, cost living in an area, as measured by the location quotient, affects the general salary levels of employees as their compensation accounts for the higher or lower cost of living but it is not directly related to inspector's decision-making process. Expenses in a particular area compared to the national average can affect salary negotiations and adjustments for inspectors to maintain a comparable standard of living. The average salary raises of other inspectors at the

same (federal) grade, but located in other FDA divisions, reflects broad salary trends in FDA that could influence an inspector's perceptions of fairness without being directly linked to the performance outcomes of a specific inspection.

Finally, in the first regression, we include several control variables along with fixed effects. The equations for our first-stage regression models are as follows:

$$\text{Self referent}_{ni} = \delta^{SR} + \theta^{SR} \text{nonfood industry pay}_{ni} + \gamma^{SR} \text{cost of living}_{ni} + \mu^{SR} \text{average salary raises}_{ni} + \alpha^{SR} X_{ni} + \pi_{ni}^{SR} \quad (7a)$$

$$\text{Industry referent}_{ni} = \delta^{IR} + \theta^{IR} \text{nonfood industry pay}_{ni} + \gamma^{IR} \text{cost of living}_{ni} + \alpha^{IR} X_{ni} + \pi_{ni}^{IR} \quad (7b)$$

$$\text{Internal peer referent}_{ni} = \delta^{IPR} + \theta^{IPR} \text{nonfood industry pay}_{ni} + \text{cost of living}_{ni} + \alpha^{IPR} X_{ni} + \pi_{ni}^{IPR} \quad (7c)$$

$$\text{Operational interface referent}_{ni} = \delta^{OIR} + \theta^{OIR} \text{nonfood industry pay}_{ni} + \gamma^{OIR} \text{cost of living}_{ni} + \mu^{OIR} \text{average salary raises}_{ni} + \alpha^{OIR} X_{ni} + \pi_{ni}^{OIR} \quad (7d)$$

Where X_{ni} are the set of control variables including fixed effects. The residual terms are π_{ni}^{SR} , π_{ni}^{IR} , π_{ni}^{IPR} , and π_{ni}^{OIR} , with predicted values ($\hat{\pi}_{ni}^{SR}$, $\hat{\pi}_{ni}^{IR}$, $\hat{\pi}_{ni}^{IPR}$ and $\hat{\pi}_{ni}^{OIR}$) used as correction terms in the second stage. For all models in the first stage estimations, the Cragg-Donald Wald F statistics exceed the Stock-Yogo critical value of 16.85, leading to rejecting the null hypothesis that the instrumental variables are weak. In addition, the under-identification test, conducted using the Kleibergen–Paap rk LM statistic (Kleibergen & Paap, 2006), shows significance, thereby rejecting the null hypothesis that the models are under-identified in the equation.

With number of citations as the dependent variable, we use the conditional fixed effects negative binomial regression consistent with other studies on inspections. The equation for the model after accounting for endogeneity and unobserved heterogeneity is:

$$\begin{aligned}
Violations_{ni} = & \alpha_0 + \alpha_1 Self\ referent_{ni} + \alpha_2 Internal\ peer\ referent_{ni} + \\
& \alpha_3 Industry\ referent_{ni} + \\
& \alpha_4 Operational\ interface\ referent_{ni} + X_{ni}\alpha + \eta^{SR}\hat{\pi}_{ni}^{SR} + \eta^{IR}\hat{\pi}_{ni}^{IR} + \\
& \eta^{IPR}\hat{\pi}_{ni}^{IPR} + \eta^{OIR}\hat{\pi}_{ni}^{OIR} + \beta_{w(mt)} + \gamma_t + \theta_{m(t)} + \tau_{l(ni)} + \mu_i + \rho_{i(n)} + \varepsilon_{it} \quad (8)
\end{aligned}$$

Where the variable descriptions are the same as described in equation 6 and η^{SR} , η^{IR} , η^{IPR} , η^{OIR} represent the coefficient of the endogeneity correction terms.

Table 5 presents the estimates from the first stage regression models that result in generating the endogeneity correction terms. Models 1, 2, 3 and 4 show the first stage regression results with self-referent, Industry referent, Internal peer referent, and Operational interface referent as the dependent variable. Model 1 shows that both the average annual salary of employees in non-food industries within the state ($\theta^{SR} = 0.001$ $p > 0.1$) and cost of living ($\gamma^{SR} = -0.028$ $p > 0.1$) have no relationship with an individual inspector's salary raise. However, average salary raises of other inspectors in the same grade but in other FDA divisions negatively impacts self-referent ($\mu^{SR} = -60.650$ $p < 0.001$). Similarly, model 2 shows that the average annual salary of employees in non-food industries within the state ($\theta^{IR} = -0.001$ $p < 0.001$) negatively impacts the industry referent but cost of living has no relationship with it ($\gamma^{IR} = 0.309$ $p > 0.1$). Model 3 also shows that the average annual salary of employees in non-food industries within the state ($\theta^{IPR} = -0.001$ $p < 0.001$) negatively impacts the internal peer referent but cost of living has no relationship with it ($\gamma^{IPR} = 0.248$ $p > 0.1$). We do not use the instrument average salary raises for industry referents or internal peer referents as industry raises may not be directly relevant to the specific comparisons between an inspector's salary and industry standards or co-workers. However, Model 4 shows that all the instruments including the average annual salary of employees in non-food industries within the state ($\theta^{OIR} =$

-0.002 $p < 0.001$), cost of living ($\gamma^{OIR} = -99.411$ $p > 0.001$), average salary raises ($\mu^{OIR} = -22.760$ $p < 0.001$), negatively impact the operational interface referent.

The results for the second stage negative binomial regression model accounting for endogeneity and unobserved heterogeneity are shown in Table 4 model 6. We find evidence of endogeneity for self-referent ($\eta^{SR} = 0.024$ $p < 0.001$), Industry referent ($\eta^{IR} = -0.421$ $p < 0.001$), and Internal peer referent ($\eta^{IPR} = 0.953$ $p < 0.001$) but not for Operational interface referent ($\eta^{OIR} = 0.001$ $p > 0.1$). The correction estimates are all significant at $p < 0.001$ except for operational interface referent. Importantly, however, the main hypothesized results remain robust.

5.3. Robustness tests

We conduct various supplemental analyses to determine the robustness of our results (Appendix D, Table D.1). Our main results are founded on a very conservative approach that includes several fixed effects including days, months, years, inspectors, FDA divisions, and facilities nth inspection. Our results are robust when omitting these fixed effects one by one or all at once. We also used a random effects model, and our estimated results remain robust. Finally, we re-ran our model using the Poisson partial fixed effects model, but it could not achieve convergence.

We changed how we measured inspection outcomes (our dependent variable) by using the average severity of citations. Severity outcome could be no action indicated (NAI), voluntary action indicated (VAI), and official action indicated (OAI) in the order of increasing severity. We code the outcomes as 1 – NAI, 2 – VAI, and 3 – OAI. A higher severity suggests that inspectors were able to find issues of higher intensity. We use an ordered logistic regression

model, and our results are robust using the severity instead of number of citations for all hypothesized variables except for our self-referent results (Appendix D, Table D.1., Model 4).

Further, since our dataset primarily comprises single facility inspections, that is, facilities inspected just once during the five-year period of our dataset, we do not control for facility fixed effects. Given the nature of our data, we conducted an additional robustness check. For this check, in cases of repeat observations for a facility, we randomly select one observation for a facility and drop the others. We then rerun our estimation so that each facility is represented by only one observation. This approach ensures that there will be no issues with correlated error terms across facilities. The number of observations dropped down to 11,633 but the results remain robust (Appendix D, Table D.1., model 1).

Furthermore, our analysis remains robust when controlling for the scheduling of inspections in our model, in addition to the effects of the day of the week and month that are already included in our model. For example, our results are unaffected by including the concurrent number of inspections an inspector manages simultaneously. Additionally, we account for the time between consecutive inspections. Previous studies suggest that more time between inspections could give inspectors the necessary break to recover from mental and physical fatigue leading to performance improvement. We also control for split inspections, that is, when inspectors receive a break during an ongoing inspection. Our results hold, although, the results for self-referent are now marginally significant at 10% instead of 1% (Appendix D, Table D.1., model 3).

Lastly, our findings remain stable when accounting for various attributes of a facility's past quality performance. As previously mentioned, we incorporate the average severity of past inspection outcomes. When controlling for the averages of NAI (No Action Indicated), VAI

(Voluntary Action Indicated), and OAI (Official Action Indicated), our results prove to be consistent. Furthermore, we consider the recalls occurring in the same year as the inspection year and show our findings to be unaffected. We also adjust for the severity of recalls, ranging from 1 (minor issues such as mislabeling) to 3 (life-threatening products), and observe robust outcomes (Appendix D, Table D.1., model 2).

5.4. Further analysis

Extensions to our analysis can be found in Appendix E, Table E.1. and Table B.2. Our primary results suggest that incentives in the form of pay raise from last year affect the number of citations (Hypothesis 1). To augment this, we also investigate the years since the last promotion, i.e., does the length of time since the last promotion of an inspector affect inspection outcome? We find that it does affect the number of citations. We introduce the term years since last promotion and its quadratic term, i.e., square of years since last promotion. These are mean centered to avoid collinearity between the squared term and linear term. We find that the year since last promotion is positive and significant ($\beta = 0.060$ $p < 0.001$), whereas its square term is negative and significant ($\beta = -0.007$ $p < 0.001$) (Appendix E, Table E.1., Model 1). This implies an inverted U-shaped relationship between the years since last promotion and the number of citations at a current inspection; that is, increasing the interval between promotions initially increases but eventually decreases the number of citations. Refer to Figure 1.

In addition, we explored different inspector characteristics that can moderate the salary differences. This helps us understand how inspector characteristics such as site familiarity, breaks, and workload moderation can help offset some of the salary difference induced biases. We performed partial fixed effects negative binomial regression analysis using the number of

citations as the outcome with day of the week, month, year, inspector, FDA division, and facility's nth inspection fixed effects like our primary models.

First, we find that more site-specific experience for an inspector can alleviate the effects of salary disparity on the inspection outcome. For example, inspectors earning higher salaries than their co-workers and employees of facilities they inspect, who possess more site-specific experience, tend to detect more violations. This might be for the reason that they have an improved understanding of the facility, leading to more efficient inspections. Second, providing inspectors with breaks during ongoing inspections can offset the negative effects of salary disparity with employees of the organization they inspect on the inspection outcomes. Additionally, breaks, particularly those taken between inspections, can help alleviate stress and the negative impacts of salary differences, allowing inspectors to refocus and enhance their performance.

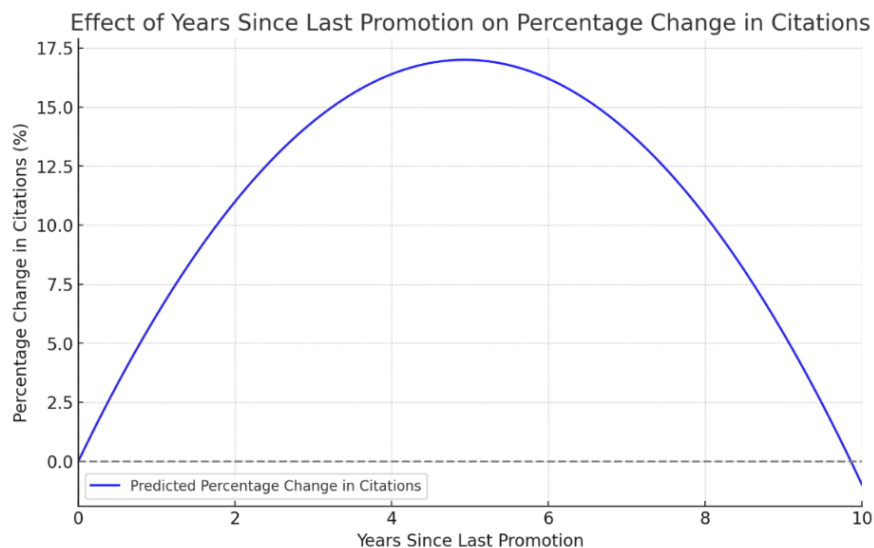


Figure 1. Relationship between years since last promotion of inspector and number of citations

Table 3. Correlation matrix for main independent variables

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1-Violations	1.00																						
2-Self referent	0.01	1.00																					
3-Internal peer	0.01	-	1.00																				
4-Industry referent	0.09***	-	0.58***	1.00																			
5-Operational interface referent	-	0.29***	0.29***	0.40***	1.00																		
6-Experience	-	-	0.23***	0.34***	0.02***	1.00																	
7-Average time	0.02***	0.11***	-	-	-	-	1.00																
8-Gender	0.17***	0.03***	0.06***	0.10***	0.12***	0.15***	0.03***	1.00															
9-Inspector age	-	0.05***	0.05***	0.05***	0.08***	0.06***	-	0.12***	1.00														
10-Site-specific	-0.01	0.02***	0.36***	0.43***	0.21***	0.19***	0.08***	-	0.01***	1.00													
11-Prior experience	-	0.14***	0.06***	0.03***	-	0.25***	0.05***	0.02***	-	-	1.00												
12-Prior inspection outcomes	0.04***	0.02***	-	0.03***	0.06***	-	0.18***	0.01***	0.02***	-	0.03***	1.00											
13-Past inspection outcomes	0.20***	0.01	-0.01*	0.03***	0.04***	0.05***	0.11***	0.00	0.01	-0.01	0.06***	0.02***	1.00										
14-Recalls last year	0.22***	-0.01	0.02***	0.03***	-	0.06***	0.04***	0.00	0.01	-	0.06***	0.03***	0.03***	1.00									
15-Inspection depth	0.00	0.01*	0.01*	0.01*	0.01	0.00	0.02***	-0.01*	0.00	0.00	0.04***	0.02***	0.00	1.00									
16-State GDP	0.23***	0.00	0.01**	0.00	-	0.23***	0.01*	0.04***	0.01***	-	0.06***	0.11***	0.00	0.03***	1.00								
17-State employment	0.06***	-	0.01	0.07***	-	-	0.07***	-	0.00	-	0.04***	0.04***	0.00	-	0.03***	1.00							
18-Firm FDA distance	-	0.04***	0.07***	0.02***	0.30***	0.03***	-	0.07***	0.01	0.00	-	0.05***	0.00	-	-	1.00							
19-State participation	0.02***	0.02***	0.13***	0.09***	0.05***	0.05***	0.05***	0.03***	0.01	0.00	0.02***	0.00	0.00	0.05***	0.01***	0.17***	1.00						
20-Political alignment index	-0.01	-0.01**	0.06***	0.09***	0.06***	0.03***	-	0.07***	0.11***	-	-0.01*	-0.01**	0.00	0.03***	-	-	0.08***	1.00					
21-Inspector designation	0.00	0.03***	-	0.01	0.07***	-	0.01**	-0.01	-0.01*	-0.01**	0.01**	-	0.00	0.02***	0.00	-	0.02***	0.05***	1.00				
22-Political affiliation	0.04***	0.03***	0.02***	0.09***	0.03***	0.08***	0.01***	0.02***	-	0.04***	0.04***	0.02***	0.01**	0.01	0.00	-	0.02***	0.03***	0.02***	1.00			
23-Inspector workload	-	0.00	0.00	-	-	0.37***	0.07***	0.01**	0.01**	-	-	0.09***	-0.01*	-	0.01***	0.05***	-	0.00***	-	0.00	-	1.00	
	0.05***	0.09***	-	0.28***	0.29***	0.13***	0.24***	-	-0.01**	0.02***	0.02***	-	0.00	0.02***	-	-	0.06***	0.08***	0.23***	0.02***	0.10***	-	1.00
	0.11***	0.02***	0.02***	0.01*	0.13***	0.24***	0.23***	0.30***	-	0.02***	0.11***	0.06***	0.00	0.04***	0.09***	0.06***	0.06***	0.08***	0.00	0.02***	0.02***	0.02***	1.00

Table 4. How incentive differences affect inspection outcomes

Type of effects	Independent variables	(1) Violations	(2) Violations	(3) Violations	(4) Violations	(5) Violations	(6) Violations
Pay raise from last year (%)	<i>Self-referent (H1)</i>	-	-0.013*** (0.002)	-0.008*** (0.003)	-0.007*** (0.003)	-0.007*** (0.003)	-0.030*** (0.010)
Pay compared to federal other employees (%)	<i>Industry referent (H2)</i>	-	-	0.010*** (0.002)	0.012*** (0.002)	0.014*** (0.001)	0.435*** (0.145)
Pay compared coworkers (%)	<i>Internal peer referent (H3)</i>	-	-	-	-0.007*** (0.002)	-0.006*** (0.002)	-0.959*** (0.276)
Pay compared to inspected industry (%)	<i>Operational interface referent (H4)</i>	-	-	-	-	-0.002*** (0.001)	-0.002*** (0.001)
Inspector control	<i>Experience</i>	-0.004*** (0.001)	-0.005*** (0.001)	-0.006*** (0.001)	-0.006*** (0.001)	-0.006*** (0.001)	-0.016*** (0.003)
Inspector control	<i>Average time</i>	-0.024* (0.013)	-0.022 (0.013)	-0.020 (0.013)	-0.021 (0.013)	-0.021 (0.013)	-0.022* (0.013)
Inspector control	<i>Gender-male</i>	0.039 (0.057)	0.031 (0.057)	0.012 (0.057)	0.008 (0.057)	0.001 (0.057)	0.002 (0.057)
Inspector control	<i>Inspector age</i>	5.36e-05 (0.002)	-0.002 (0.003)	-0.008*** (0.003)	-0.007** (0.003)	-0.007** (0.003)	0.287*** (0.091)
Inspector control	<i>Site-specific experience</i>	-0.165*** (0.026)	-0.165*** (0.026)	-0.164*** (0.026)	-0.165*** (0.026)	-0.170*** (0.027)	-0.246*** (0.038)
Inspector control	<i>Prior inspection outcomes</i>	0.011* (0.006)	0.011* (0.006)	0.010* (0.006)	0.010* (0.006)	0.010* (0.006)	0.018*** (0.006)
Inspector control	<i>Inspector designation</i>	-0.004* (0.002)	-0.005* (0.003)	-0.006** (0.003)	-0.006** (0.003)	-0.006** (0.003)	-0.006** (0.003)
Inspector control	<i>Inspector workload</i>	-0.181*** (0.010)	-0.179*** (0.010)	-0.178*** (0.010)	-0.178*** (0.010)	-0.174*** (0.010)	-0.111*** (0.025)
Firm control	<i>Past inspection outcomes</i>	0.107*** (0.005)	0.107*** (0.005)	0.105*** (0.005)	0.106*** (0.005)	0.106*** (0.005)	0.119*** (0.005)
Firm control	<i>Recalls last year</i>	-0.101 (0.122)	-0.099 (0.122)	-0.105 (0.122)	-0.104 (0.122)	-0.103 (0.123)	-0.047 (0.125)
Firm control	<i>Inspection depth</i>	0.227*** (0.035)	0.232*** (0.035)	0.240*** (0.035)	0.239*** (0.035)	0.245*** (0.035)	0.060 (0.064)
Firm control	<i>Inspection duration</i>	0.019*** (0.010)	0.019*** (0.001)	0.019*** (0.001)	0.019*** (0.001)	0.019*** (0.001)	0.018*** (0.001)
Firm Control	<i>Firm FDA distance</i>	-4.92e-05 (0.000)	-6.27e-05 (0.000)	-8.21e-05 (0.000)	-7.71e-05 (0.000)	-2.56e-05 (0.000)	0.000 (0.000)
State Control	<i>State GDP</i>	0.022* (0.011)	0.020* (0.011)	0.017 (0.011)	0.017 (0.011)	0.011 (0.011)	-0.022 (0.014)
State Control	<i>State employment</i>	-0.003 (0.005)	-0.004 (0.005)	-0.003 (0.005)	-0.002 (0.005)	-0.004 (0.005)	-0.017*** (0.006)
State Control	<i>State participation</i>	2.689 (2.186)	3.383 (2.179)	3.124 (2.177)	2.753 (2.184)	2.631 (2.200)	-25.850*** (8.406)
State Control	<i>Political alignment index</i>	-0.185*** (0.057)	-0.177*** (0.057)	-0.232*** (0.058)	-0.231*** (0.058)	-0.238*** (0.058)	0.049 (0.203)
State Control	<i>Republican</i>	-0.085 (0.149)	-0.084 (0.149)	-0.109 (0.147)	-0.102 (0.148)	-0.100 (0.148)	0.197 (0.167)
Endogeneity correction terms	<i>Self-referent</i>	-	-	-	-	-	0.024** (0.0102)
Endogeneity correction terms	<i>Industry referent</i>	-	-	-	-	-	-0.421***

Endogeneity correction terms	<i>Internal peer referent</i>	-	-	-	-	-	(0.145) 0.953***
Endogeneity correction terms	<i>Operational interface referent</i>	-	-	-	-	-	(0.276) 0.001
	<i>Day Fixed Effects</i>	Included	Included	Included	Included	Included	(0.001) Included
	<i>Month Fixed Effects</i>	Included	Included	Included	Included	Included	Included
	<i>Year Fixed Effects</i>	Included	Included	Included	Included	Included	Included
	<i>FDA Division Fixed Effects</i>	Included	Included	Included	Included	Included	Included
	<i>Facilities nth inspection Fixed Effects</i>	Included	Included	Included	Included	Included	Included
	<i>Inspector Fixed Effects</i>	Included	Included	Included	Included	Included	Included
	<i>Constant</i>	0.150 (0.528)	0.328 (0.528)	0.365 (0.527)	0.242 (0.528)	0.549 (0.534)	-15.96*** (4.829)
	Observations (Inspections)	16,382	16,382	16,382	16,382	16,382	16,382
	Number of Facilities	14,441	14,441	14,441	14,441	14,441	14,441
	Number of Inspectors	594	594	594	594	594	594
	Wald's χ^2	2346.60***	2376.45***	2419.70***	2430.90***	2458.71***	2466.16***
	Log likelihood	-14167.324	-14153.854	-14132.656	-14127.316	-14109.688	-14104.954

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 5. First stage regression results for addressing endogeneity

Type of effects	Independent variables	(1) <i>Self-referent</i>	(2) <i>Industry referent</i>	(3) <i>Internal peer referent</i>	(4) <i>Operational interface referent</i>
Instrument 1	<i>Nonfood industry pay</i>	0.001 (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.002*** (0.000)
Instrument 2	<i>Cost of living</i>	-0.028 (0.265)	0.309 (0.191)	0.248 (0.191)	-99.411*** (1.711)
Instrument 3	<i>Average salary raises</i>	-60.650*** (1.673)	-	-	-22.760** (9.930)
Inspector controls	Exp., gender, speed, age, designation, workload, site exp., prior insp. outcomes	Yes	Yes	Yes	Yes
Firm control	Prior insp. outcome, recalls last year, complexity, ops. Scale, distance from FDA	Yes	Yes	Yes	Yes
State Control	GDP, employment ratio, participation, PAI, party affiliation	Yes	Yes	Yes	Yes

<i>Day Fixed Effects</i>	Included	Included	Included	Included
<i>Month Fixed Effects</i>	Included	Included	Included	Included
<i>Year Fixed Effects</i>	Included	Included	Included	Included
<i>FDA Division Fixed Effects</i>	Included	Included	Included	Included
<i>Facilities nth inspection Fixed Effects</i>	Included	Included	Included	Included
<i>Inspector Fixed Effects</i>	Included	Included	Included	Included
<i>Constant</i>	53.916*** (7.430)	-49.352*** (6.780)	-35.78*** (5.253)	282.697*** (45.608)
F-statistics	264.04	270.99	47.13	131.70
R-squared	0.360	0.370	0.080	0.227

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Lastly, an inspector's workload exacerbates the negative effects of salary disparity between inspectors and employees of the organizations they inspect on inspection outcomes. An increased number of simultaneous inspections can amplify the negative emotions associated with perceived inequities. Overall, besides developing more equitable incentive structures, managers can alleviate the adverse effects of incentive disparities by moderating workloads, strategically planning breaks, and arranging more site-specific inspections for inspectors.

6. Discussion

Our analysis of FDA food inspections reveals that incentive differences or salary variances with respect to different referent groups significantly affect inspection outcomes. First, we found that salary raises (cuts) compared to the previous year led to a decrease (increase) in the number of violations detected. Second, a higher (lower) salary compared to coworkers also results in a decrease (increase) in the number of detected violations. Thirdly, a higher (lower) salary compared to operational interface referents leads to a decrease (increase) in the number of detected violations. This may occur because higher salaries can make inspectors complacent, although the effect of all salary referents varies. Lastly, higher (lower) salaries compared to industry referents result in an increase (decrease) in the number of detected violations. Overall,

our research contributes to the literature by classifying and assessing incentive differences with specific referent groups within the context of food inspections, aiming to understand how these disparities directly impact inspection outcomes.

6.1.Contributions

Our study contributes to three streams of literature, namely incentive design and operational outcomes, field staff motivation and operational performance, and lastly quality management and monitoring. We contribute to these literature streams in several ways. Firstly, our paper is among the very few papers that introduces a model based on referent group incentive design to the operations management literature (Anand & Gray, 2017; Bendoly et al., 2010; Browning, 2010). We explore incentive differences relative to multiple referent groups and their impact on operational performance of field staff. Existing literature in operations management has historically focused on traditional measures of incentive design such as salaries, bonuses, stock options, and allowances to study its impact on regular employees, managers, and executive staff. We contribute to the literature by studying the impact of income disparities with multiple referent groups on field staff's operational outcomes.

Further, we extend the literature on salary comparisons by theoretically grounding the classification of referent groups. We suggest that referent groups for salary comparisons be classified using the scope of comparison (internal or external) and relational dependencies of employees on referent groups (direct or indirect) (Xu et al., 2024). Grounding the referent groups using these dimensions provides a systematic way to study the impact of such referent groups, i.e., self-referents, internal peer referents, industry referents, and operational interface referents. Further, we propose a new class of referent group i.e., operational interface referent. This is especially relevant for field staff employees who engage with direct customers, business clients,

or facilities they inspect. They interact with the employees of such facilities more closely than their office-based counterparts.

Next, we contribute to the literature on field staff, essential workers on the frontlines, and their operational performance (Chong et al., 2020; Deery et al., 2002; Tucker, 2007). We explore the impact of field staff motivation on their operational performance. Existing literature studies frontline workers such as retail cashiers, agricultural workers, or nurses who are stationed in the field. However, employees such as sales representatives, maintenance technicians, last mile delivery operators, or FDA food inspectors who travel across different locations to perform their operational tasks are less studied. We study regulatory food inspectors who travel within their territory to inspect food facilities ensuring food safety in supply chains.

Further, when it comes to essential workers on the frontlines, existing literature focuses on productivity, delivery, and safety (Lu et al., 2022; Pendem et al., 2024; Tucker, 2007). There is limited literature on the quality assessment of field staff. We contribute to the literature by studying the quality aspect of decision making of field staff. This is very relevant in the context of field staff as they face uncertain and complex situations where they make high stake decisions (Balogun, 2015). We contribute by understanding the factors that affect their ability to make such decisions.

Next, we expand on equity theory by providing empirical evidence of several mechanisms proposed by the theory (Adams, 1965). The theory establishes that when an employee experiences positive differences, i.e., receives higher salary than their chosen referent, they might increase their input (efforts) to match and justify their outputs (incentives). This is because they may experience a feeling of guilt or sense of indebtedness toward their organization (Brockner et al., 1985). However, we propose that they may also experience complacency or a

sense of superiority, whereby they increase the value that they ascribe to their self-worth and decrease their inputs (efforts). Similarly, equity theory also proposes that when employees experience negative differences, i.e., receive lower salary than their chosen referent, they may decrease their inputs (efforts) due to feelings of anger or dissatisfaction (Li-Ping Tang et al., 2000). However, we propose that instead they increase their efforts and exhibit compensatory behavior, possibly to demonstrate their worth for a higher incentive.

Next, organizational psychology literature is limited in establishing how multiple salary comparisons affect employee perceptions of salary differences and ultimately performance (Xu et al., 2024). We help establish the link and explain the mechanism using the context of food inspections. Further, we contribute to the literature on salary comparisons by using actual salary comparisons instead of perception. Actual salary comparisons provide measures of employees' internal and external pay standing, compared to pay standings based on perceptions.

Lastly, we also contribute to the literature on inspections and quality assessments (Ibanez & Toffel, 2020; Short et al., 2016). The existing literature on inspections has identified several factors that affect an inspector's motivation to detect the number of violations. These factors include but are not limited to inspection scheduling, inspector teams, inspector genders, inspector qualifications and training. The effect of incentive design on inspection outcomes of inspectors is less understood. We contribute to the literature by analyzing salary inequities and its impact on inspection outcomes.

6.2. Managerial implications

We make several managerial contributions. First, we reinforce that the way salaries are structured and compared within and across different referent groups affects employee motivation. This may be due to comparisons that impact perceptions of fairness, and ultimately

performance. The aim of an effective incentive design in operations is to motivate employees to achieve operational excellence. Strategically using incentive designs, informed by the most relevant referent groups, can help mitigate employee work withdrawal. We suggest that not all referent group comparisons may affect employees the same way and that salary differences in comparison to co-workers and interface referents can be counterproductive. Aligning salary structures with operational goals may ensure better operational performance. We propose that salary differences, as part of a well-designed incentive scheme, may be used to motivate employees for enhanced operational performance.

Secondly, we investigate various inspector-level attributes that could mitigate the salary difference aspects of incentive design in an operational setting (refer Appendix E, Table E.2). We find that an inspector's site familiarity can minimize the effects of potential positive differences on their performance, as evidenced by the number of violations detected during an inspection. Inspectors who earn higher salaries than their internal peers and operational interface referents, and who possess greater site-specific experience, tend to identify more violations. This may be because they have a better understanding of the concerns, leading to more effective inspections. Their site familiarity enables them to conduct effective inspections, allowing them to detect a greater number of violations regardless of salary comparisons.

Thirdly, providing breaks during ongoing tasks can offset the negative effects of incentive disparities as compared with operational interface referents. In our case, providing intermittent breaks during inspections can help inspectors find a greater number of violations regardless of salary comparisons. This is because breaks, particularly between inspections, can help alleviate stress and the negative effects of observed salary differences, allowing inspectors to refocus and improve their performance. Lastly, an inspector's workload worsens the negative

impact of operational interface referent disparity on inspection outcomes. An increased number of inspections can increase the negative emotions associated with incentive comparisons.

6.3.Limitations and future research

This paper has various limitations, which future research could address. Firstly, regarding salary disparities, we did not examine inspector perceptions of inequity; instead, we used actual salary data to observe the differences between the inspectors and their respective referent groups. Inspectors may have different perceptions of salary equity. We assume that all inspectors have accurate input (effort) to outcome (salary) information about their referent groups. However, inspectors may not possess actual information about the efforts and salaries of others and may base their understanding of incentive inequities on perceptions that could be far from reality. Future research could extend this study by conducting surveys or semi-structured interviews with inspectors to develop a more detailed understanding of such perceptions.

Secondly, as a proxy for incentives or outcomes, we use salary comparisons of inspectors with their referent groups. Incentives can be in the form of bonuses – one time or regular reward incentives to show appreciation, extra allowances – additional money to buffer against inflation such as allowances, commissions – financial incentives for field staff working in sales to earn a percentage of profit from their sales, employee stock options – incentives to allow employees to purchase shares of a company, profit sharing – financial incentives in addition to regular salary, co-partnerships – similar to profit sharing where employees receive a share of profits, retirement and fringe benefits – providing employees with long term retirement benefits, child care, college tuition, living expenses, etc. Future research can investigate other incentive mechanisms.

Thirdly, we used a limited number of referent groups, namely self-referents, industry referents, internal peer referents, and operational interface referents. However, there could be

more nuanced referent groups that significantly impact employee motivation. For instance, within the internal peer referents, a close group of individuals with whom inspectors conduct team inspections could have a greater impact than inspectors of the same grade but outside their core network with whom they never interact for work. Furthermore, since FDA food inspection operations are divided into twelve divisions, the salaries of inspectors outside their own divisions may have different impacts, compared to those within their own divisions. Fourth, some of our variables are counted from the start of our data collection period, including facilities' nth inspection, so we have excluded prior information.

Lastly, our measure of salary comparisons with operational interface referents is limited. We calculate the salary of employees at the inspected facilities as the average salary of the specific industry to which the facility belongs and the county where the facility is located. This approach will average the salaries from other facilities in the same industry and county. Furthermore, we averaged the salary across different job roles, not just for quality assurance positions within a facility. This average will affect external salary comparisons. Future researchers can collect data on specific external referent groups for direct comparisons.

Future research could explore the impact of other inspector-level attributes on salary inequity comparisons such as individual beliefs, inspector training, cultural and religious backgrounds, formal and informal education, gender, race, and ethnicity. Further, research could be extended to other types of regulatory inspectors, such as U.S. Dept. of Agriculture, Dept. of Transportation, and pharmaceutical inspectors. Future research could also use external shocks, such as changes in compensation and benefits policies, to understand the effect of inequities in incentive design in operational settings.

Acknowledgements

I would like to thank Dr. Martin Dresner for his support, motivation, and guidance on this paper. Additionally, I would like to thank Dr. Oliver Yao for his guidance on an earlier version of the paper. I also extend my thanks to the FDA officials, including food inspectors, who provided valuable insights.

Chapter 5: Conclusions and Future Extensions

The dissertation examines different aspects around safety in the U.S. food supply chains and suggests the following findings:

First, supply chain accountability and product safety regulations such as Food Safety Modernization Act may have unintended consequences. Although, they are intended to safeguard consumers by making firms more accountable for their suppliers, they negatively impact a firm's inventory performance. The regulatory stress instigates a coping response of inventory buffering leading to decreased inventory leanness among affected firms. This stress is more pronounced when firms depend on developed markets for sourcing critical material as compared to emerging markets. Thus, the regulation may have implications for the firm's food safety compliance behavior.

Second, different work-break schedule regimes may affect the task performance in the context of food safety inspections. The study provides evidence on how the timing of breaks relative to inspection tasks for food inspectors affects quality assessments of food facilities. Pseudo-breaks, i.e., time between two inspections, leads to improved outcomes but at a diminishing rate. Intermittent breaks, i.e., time-off between ongoing inspections, improve citation detection. Non-standard and concurrent schedules lead to a decreased task performance. Finally, the study provides evidence of break fade-out effects such that task performance deteriorates during the workweek. The study has implications for the inspection accuracy of food inspections.

Third, salary differences with respect to multiple referents affect the task performance of food inspectors. The study provides evidence on how incentive design for food inspectors affects their quality assessment of food facilities. When inspectors earn, higher salaries compared to

their own salaries from the previous year, as well as compared to the salaries of their internal peers and employees of the organizations they inspect, they may cite fewer violations, likely due to complacency. On the other hand, if they earn higher salaries compared to other industry referents, they may cite more violations, likely due to a higher motivation.

The dissertation highlights several opportunities for future research: Firstly, future research can use additional contexts beyond the Food Safety Modernization Act to observe the phenomenon of regulatory burden and its impact on supply chain performance. Future research can investigate how different contexts of supply chain accountability regulations (for example, the German Supply Chain Due Diligence Act) impact various supply chain outcomes. Future research can also explore how firms engage with their political environment (such as through lobbying) to shape favorably shape supply chain performance outcomes.

Future research could also explore the impact of regulations on the scheduling of food inspections and the impact of these regulations on the quality assessment of food inspectors. Additionally, studies could extend the work from this dissertation on food inspection outcomes to other regulated industries, such as pharmaceuticals, cosmetics, and medical devices. Future research can also build on this work and explore the impact of other inspector-level attributes on inspection outcomes, such as individual beliefs, inspector training, cultural and religious backgrounds.

Appendices

Appendix A: OLS assumptions Chapter 2

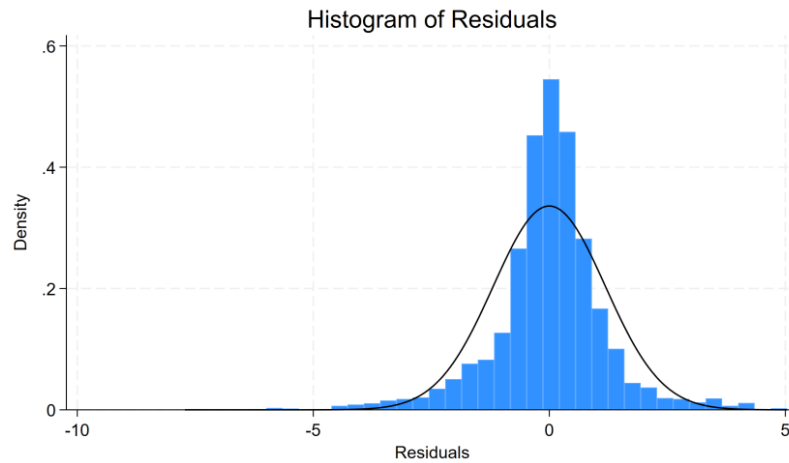


Figure A.1. Histogram to test for normality

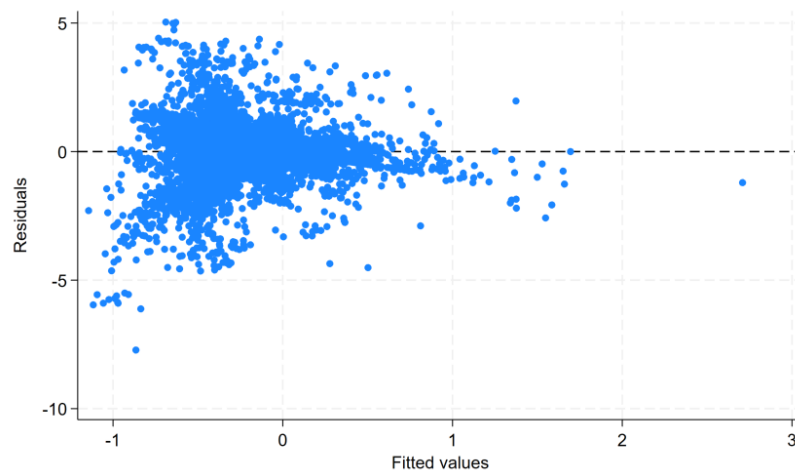


Figure A.2. Residuals vs. fitted plot

Appendix B: Robustness Checks Chapter 3

Table B.1. Combination of different dependent and independent variables for robustness

	(1)	(2)	(3)	(4)	(5)
VARIABLES	(Logistic) Citation (1/0)	(OLS) Avg. cites severity	(PPML) Number of citations	(PPML) Number of citations	(PPML) Number of citations
<i>Pseudo (H1)</i>	0.034* (0.019)	0.010** (0.004)	0.018** (0.008)	0.017** (0.008)	0.016** (0.008)
<i>Pseudo² (H1)</i>	-0.023*** (0.008)	-0.007*** (0.002)	-0.010*** (0.003)	-0.009*** (0.003)	-0.01*** (0.003)
<i>Concurrent (H2)</i>	-0.299*** (0.107)	-0.076*** (0.023)	-0.152*** (0.056)	-0.098* (0.059)	-0.118** (0.057)
<i>Intermittent (H3)</i>	0.495*** (0.131)	0.289*** (0.017)	0.581*** (0.044)	0.575*** (0.044)	0.573*** (0.044)
<i>Tuesday (H4)</i>	-0.019 (0.079)	-0.017 (0.017)	-0.078** (0.035)	-0.077** (0.035)	-0.074** (0.034)
<i>Wednesday (H4)</i>	-0.241*** (0.083)	-0.069*** (0.018)	-0.152*** (0.038)	-0.147*** (0.038)	-0.147*** (0.037)
<i>Thursday (H4)</i>	-0.428*** (0.091)	-0.130*** (0.020)	-0.222*** (0.042)	-0.212*** (0.042)	-0.212*** (0.042)
<i>Friday (H4)</i>	-0.609*** (0.133)	-0.202*** (0.027)	-0.315*** (0.060)	-0.296*** (0.060)	-0.304*** (0.059)
<i>Non-standard (H5)</i>	-0.153 (0.192)	-0.112*** (0.041)	-0.243*** (0.089)	-0.250*** (0.089)	-0.249*** (0.089)
<i>Weekly workload</i>	-	-	-	-0.069*** (0.019)	-
<i>Monthly workload</i>	-	-	-	-	-0.030*** (0.010)
<i>Experience</i>	-0.003*** (0.001)	-0.000*** (0.000)	-0.001 (0.001)	-0.001 (0.001)	-0.001 (0.001)
<i>Average time</i>	0.030** (0.015)	0.012*** (0.003)	0.039*** (0.008)	0.038*** (0.008)	0.038*** (0.008)
<i>Male (Female)</i>	-0.040 (0.057)	0.001 (0.013)	-0.136*** (0.033)	-0.131*** (0.032)	-0.133*** (0.032)
<i>Site-specific experience</i>	-0.096** (0.038)	-0.007 (0.007)	-0.074** (0.032)	-0.079** (0.032)	-0.079** (0.032)
<i>Prior inspection outcomes</i>	0.096*** (0.013)	0.016*** (0.002)	0.036*** (0.005)	0.035*** (0.005)	0.034*** (0.005)
<i>Past inspection outcomes</i>	-0.011 (0.011)	0.007*** (0.002)	-0.078*** (0.006)	-0.080*** (0.006)	-0.079*** (0.006)
<i>Recalls last year</i>	-0.200 (0.195)	-0.039 (0.044)	-0.074 (0.087)	-0.074 (0.087)	-0.073 (0.087)
<i>Inspection depth</i>	0.178*** (0.061)	-0.161*** (0.013)	0.160*** (0.029)	0.162*** (0.029)	0.162*** (0.029)

<i>Inspection duration</i>	0.170*** (0.018)	0.009*** (0.001)	0.011*** (0.003)	0.010*** (0.004)	0.011*** (0.003)
<i>Republican</i>	-0.536 (0.347)	-0.057 (0.080)	-0.170 (0.167)	-0.171 (0.168)	-0.161 (0.166)
<i>Month Fixed Effects</i>	Included	Included	Included	Included	Included
<i>Year Fixed Effects</i>	Included	Included	Included	Included	Included
<i>FDA Division Fixed Effects</i>	Included	Included	Included	Included	Included
<i>Facilities' nth inspection Fixed Effects</i>	Included	Included	Included	Included	Included
<i>Facility Fixed Effects</i>	Not included	Not included	Included	Included	Included
Constant	0.679*** (0.200)	1.925*** (0.043)	1.046*** (0.094)	1.140*** (0.100)	1.125*** (0.0955)
Observations	7,464	7,464	7351	7,464	7,464
R-squared	0.142	0.171	0.377	0.378	0.378
			Weekly workload 99 th %ile		

Robust standard errors clustered by Inspectors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table B.2. Empirical test for different ways of coding work week variable

	(1)	(2)	(3)
VARIABLES	(PPML) Number of citations	(PPML) Number of citations	(PPML) Number of citations
<i>Pseudo (H1)</i>	0.020** (0.008)	0.019** (0.008)	-
<i>Pseudo² (H1)</i>	-0.010*** (0.003)	-0.010*** (0.003)	-
<i>Concurrent (H2)</i>	-0.159*** (0.055)	-0.159*** (0.055)	-
<i>Intermittent (H3)</i>	0.565*** (0.043)	0.580*** (0.043)	-
<i>Mid-week (Early -week)</i>	-0.110*** (0.032)	-	-
<i>Late-week (Early -week)</i>	-0.203*** (0.034)	-	-
<i>Days inspecting earlier this week</i>	-	-0.076*** (0.011)	-
<i>Post-holiday Tuesday (Tuesday)</i>	-	-	0.204* (0.113)
<i>Non-standard (H5)</i>	-0.238*** (0.088)	-0.247*** (0.089)	-
<i>Experience</i>	-0.001	-0.001	0.001

	(0.001)	(0.001)	(0.001)
<i>Average time</i>	0.039***	0.039***	0.073***
	(0.008)	(0.008)	(0.018)
<i>Male (Female)</i>	-0.132***	-0.133***	-0.266***
	(0.033)	(0.032)	(0.080)
<i>Site-specific experience</i>	-0.075**	-0.075**	-0.131
	(0.032)	(0.032)	(0.106)
<i>Prior inspection outcomes</i>	0.036***	0.036***	0.016
	(0.005)	(0.005)	(0.012)
<i>Past inspection outcomes</i>	-0.079***	-0.079***	-0.072***
	(0.006)	(0.006)	(0.016)
<i>Recalls last year</i>	-0.069	-0.067	-0.127
	(0.088)	(0.088)	(0.255)
<i>Inspection depth</i>	0.166***	0.166***	0.154*
	(0.029)	(0.029)	(0.089)
<i>Inspection duration</i>	0.011***	0.011***	0.041***
	(0.004)	(0.004)	(0.011)
<i>Republican</i>	-0.160	-0.163	0.419
	(0.168)	(0.167)	(0.349)
<i>Month Fixed Effects</i>	Included	Included	Included
<i>Year Fixed Effects</i>	Included	Included	Included
<i>FDA Division Fixed Effects</i>	Included	Included	Included
<i>Facilities' nth inspection Fixed Effects</i>	Included	Included	Included
<i>Facility Fixed Effects</i>	Included	Included	Included
<i>Constant</i>	0.992***	1.033***	0.754***
	(0.091)	(0.091)	(0.211)
<i>Observations</i>	7,464	7,464	746
<i>R-squared</i>	0.377	0.377	0.382

Robust standard errors clustered by Inspectors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table B.3. Regression results without winsorization for variables: *pseudo* and *concurrent*

VARIABLES	(PPML) Number of citations
<i>Pseudo</i>	0.060*** (0.019)
<i>Pseudo²</i>	-0.010*** (0.003)
<i>Concurrent</i>	-0.100*** (0.0238)
<i>Intermittent</i>	0.556*** (0.040)
<i>Tuesday</i>	-0.073** (0.035)

<i>Wednesday</i>	-0.152*** (0.037)
<i>Thursday</i>	-0.216*** (0.042)
<i>Friday</i>	-0.304*** (0.058)
<i>Non-standard</i>	-0.238*** (0.086)
<i>Experience</i>	-0.001 (0.001)
<i>Average time</i>	0.038*** (0.008)
<i>Male (Female)</i>	-0.132*** (0.032)
<i>Site-specific experience</i>	-0.074** (0.032)
<i>Prior inspection outcomes</i>	0.036*** (0.005)
<i>Past inspection outcomes</i>	-0.079*** (0.006)
<i>Recalls last year</i>	-0.067 (0.087)
<i>Inspection depth</i>	0.164*** (0.030)
<i>Inspection duration</i>	0.015*** (0.003)
<i>Republican</i>	-0.177 (0.168)
<i>Month Fixed Effects</i>	Included
<i>Year Fixed Effects</i>	Included
<i>FDA Division Fixed Effects</i>	Included
<i>Facilities' nth inspection Fixed Effects</i>	Included
<i>Facility Fixed Effects</i>	Included
<i>Constant</i>	1.052*** (0.104)
<i>Observations</i>	7,464

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Appendix C: Further Analysis Chapter 3

Table C.1. PPML model for the length of intermittent breaks

VARIABLES	(1)
	(PPML) Number of citations
<i>Pseudo (H1)</i>	0.020** (0.008)
<i>Pseudo² (H1)</i>	-0.015*** (0.003)
<i>Concurrent (H2)</i>	-0.159*** (0.057)
<i>Intermittent (H3)</i>	-2.762*** (0.594)
<i>Length of intermittent</i>	3.566*** (0.705)
<i>Length of intermittent²</i>	-1.055*** (0.265)
<i>Tuesday (H4)</i>	-0.067* (0.036)
<i>Wednesday (H4)</i>	-0.113*** (0.043)
<i>Thursday (H4)</i>	-0.145** (0.064)
<i>Friday (H4)</i>	-0.154 (0.100)
<i>Non-standard (H5)</i>	0.061 (0.115)
<i>Experience</i>	-0.001 (0.001)
<i>Average time</i>	0.036*** (0.009)
<i>Male (Female)</i>	-0.135*** (0.032)
<i>Site-specific experience</i>	-0.058* (0.033)
<i>Prior inspection outcomes</i>	0.035*** (0.005)
<i>Past inspection outcomes</i>	-0.075*** (0.006)
<i>Recalls last year</i>	-0.079 (0.082)
<i>Inspection depth</i>	0.174*** (0.029)

<i>Inspection duration</i>	0.077*** (0.029)
<i>Republican</i>	-0.197 (0.166)
<i>Month Fixed Effects</i>	Included
<i>Year Fixed Effects</i>	Included
<i>FDA Division Fixed Effects</i>	Included
<i>Facilities' nth inspection Fixed Effects</i>	Included
<i>Facility Fixed Effects</i>	Included
Constant	0.888*** (0.108)
Observations	7,464
R-squared	0.382

Robust standard errors clustered by Inspectors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table C.2. *Combination of different moderators*

	(1)	(2)	(3)	(4)
MODERATORS	Inspector Experience	Inspector Gender	Site-specific experience	Inspector speed
VARIABLES	(PPML) Number of citations	(PPML) Number of citations	(PPML) Number of citations	(PPML) Number of citations
<i>Pseudo (H1)</i>	0.006 (0.012)	0.005 (0.011)	0.013 (0.009)	0.032 (0.023)
<i>Pseudo X Moderator</i>	0.000* (0.000)	0.028 (0.017)	0.026** (0.013)	-0.000 (0.004)
<i>Pseudo² (H1)</i>	0.001 (0.005)	-0.006 (0.005)	-0.007** (0.004)	-0.015 (0.010)
<i>Pseudo² X Moderator</i>	-0.001*** (0.000)	-0.009 (0.007)	-0.012* (0.006)	0.001 (0.002)
<i>Concurrent (H2)</i>	-0.159** (0.075)	-0.192*** (0.074)	-0.187*** (0.057)	-0.358** (0.152)
<i>Concurrent X Moderator</i>	0.000 (0.001)	0.067 (0.109)	0.071 (0.04)	0.023 (0.034)
<i>Intermittent (H3)</i>	0.497*** (0.056)	0.537*** (0.051)	0.554*** (0.045)	0.859*** (0.085)
<i>Intermittent X Moderator</i>	0.002*** (0.001)	0.091 (0.062)	0.157*** (0.058)	-0.080*** (0.018)
<i>Tuesday (H4)</i>	-0.207*** (0.052)	-0.060 (0.047)	-0.073** (0.035)	-0.260*** (0.10)
<i>Tuesday X Moderator</i>	0.002*** (0.001)	-0.033 (0.072)	-0.030 (0.064)	0.052** (0.023)

<i>Wednesday (H4)</i>	-0.179*** (0.057)	-0.135** (0.053)	-0.137*** (0.037)	-0.394*** (0.105)
<i>Wednesday X Moderator</i>	0.000 (0.001)	-0.034 (0.076)	-0.089 (0.068)	0.076*** (0.025)
<i>Thursday (H4)</i>	-0.312*** (0.059)	-0.189*** (0.059)	-0.233*** (0.042)	-0.428*** (0.108)
<i>Thursday X Moderator</i>	0.002** (0.001)	-0.070 (0.084)	0.0519 (0.0613)	0.066*** (0.021)
<i>Friday (H4)</i>	-0.356*** (0.086)	-0.183** (0.084)	-0.313*** (0.062)	-0.686*** (0.150)
<i>Friday X Moderator</i>	0.001 (0.001)	-0.257** (0.121)	0.021 (0.081)	0.124*** (0.033)
<i>Non-standard (H5)</i>	-0.200 (0.124)	-0.169 (0.121)	-0.204** (0.086)	-0.572*** (0.203)
<i>Non-standard X Moderator</i>	-0.001 (0.001)	-0.141 (0.162)	-0.151 (0.241)	0.102** (0.042)
<i>Experience</i>	-0.001* (0.001)	-0.001 (0.001)	-0.001 (0.001)	0.000 (0.001)
<i>Average time</i>	0.041*** (0.008)	0.039*** (0.008)	0.040*** (0.008)	0.033 (0.021)
<i>Male (Female)</i>	-0.132*** (0.032)	-0.100* (0.060)	-0.133*** (0.032)	-0.103** (0.043)
<i>Site-specific experience</i>	-0.081** (0.032)	-0.078** (0.032)	-0.121* (0.070)	-0.063 (0.041)
<i>Prior inspection outcomes</i>	0.035*** (0.005)	0.035*** (0.005)	0.036*** (0.005)	0.043*** (0.007)
<i>Past inspection outcomes</i>	-0.079*** (0.006)	-0.079*** (0.006)	-0.080*** (0.006)	-0.085*** (0.009)
<i>Recalls last year</i>	-0.064 (0.087)	-0.071 (0.087)	-0.072 (0.090)	-0.108 (0.103)
<i>Inspection depth</i>	0.168*** (0.029)	0.161*** (0.029)	0.169*** (0.029)	0.191*** (0.039)
<i>Inspection duration</i>	0.011*** (0.003)	0.011*** (0.003)	0.011*** (0.003)	0.015*** (0.003)
<i>Republican</i>	-0.171 (0.167)	-0.171 (0.167)	-0.167 (0.164)	-0.042 (0.265)
<i>Month, Year, FDA Div., nth inspection, facility, fixed effects</i>	Included	Included	Included	Included
<i>Constant</i>	1.077*** (0.098)	1.028*** (0.098)	1.029*** (0.093)	0.964*** (0.177)
<i>Observations</i>	7,464	7,464	7,464	4,436
<i>R-squared</i>	0.378	0.378	0.378	0.374

Robust standard errors clustered by Inspectors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table C.3. Regression results for presence of pseudo breaks versus otherwise

VARIABLES	(PPML) Number of citations
<i>Pseudo break (1)</i>	0.072** (0.030)
<i>Concurrent</i>	-0.160*** (0.056)
<i>Intermittent</i>	0.581*** (0.044)
<i>Tuesday</i>	-0.080** (0.035)
<i>Wednesday</i>	-0.157*** (0.037)
<i>Thursday</i>	-0.231*** (0.042)
<i>Friday</i>	-0.322*** (0.059)
<i>Non-standard</i>	-0.244*** (0.088)
<i>Experience</i>	-0.001 (0.001)
<i>Average time</i>	0.039*** (0.008)
<i>Male (Female)</i>	-0.133*** (0.033)
<i>Site-specific experience</i>	-0.073** (0.032)
<i>Prior inspection outcomes</i>	0.036*** (0.005)
<i>Past inspection outcomes</i>	-0.079*** (0.006)
<i>Recalls last year</i>	-0.063 (0.088)
<i>Inspection depth</i>	0.166*** (0.029)
<i>Inspection duration</i>	0.011*** (0.004)
<i>Republican</i>	-0.158 (0.171)
<i>Month, Year, Facility, FDA div., nth insp. Fixed Effects</i>	Included
Constant	0.942*** (0.100)
Observations	7,464

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Appendix D: Robustness Checks Chapter 4

Table D.1. Robustness checks

Independent variables	(1) <i>Violations</i>	(2) <i>Violations</i>	(3) <i>Violations</i>	(4- Ordered Logit) <i>Citation severity</i>
<i>Self-referent (H1)</i>	-0.006** (0.003)	-0.006** (0.003)	-0.004* (0.002)	-0.003 (0.003)
<i>Industry referent (H2)</i>	0.019*** (0.002)	0.014*** (0.002)	0.014*** (0.002)	0.016*** (0.002)
<i>Internal peer referent (H3)</i>	-0.007*** (0.002)	-0.006*** (0.002)	-0.005*** (0.002)	-0.008*** (0.002)
<i>Operational interface referent (H4)</i>	-0.002*** (0.000)	-0.002*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)
<i>Experience</i>	-0.008*** (0.001)	-0.006*** (0.001)	-0.007*** (0.001)	-0.006*** (0.001)
<i>Average time</i>	-0.034** (0.016)	-0.021 (0.013)	-0.051*** (0.014)	0.013 (0.018)
<i>Gender-male</i>	0.059 (0.068)	-0.011 (0.057)	0.042 (0.060)	0.049 (0.088)
<i>Inspector age</i>	-0.011*** (0.003)	-0.007** (0.003)	-0.008*** (0.003)	-0.013*** (0.004)
<i>Site-specific experience</i>	-0.086* (0.052)	-0.187*** (0.028)	-0.163*** (0.027)	-0.049*** (0.015)
<i>Prior inspection outcomes</i>	0.009 (0.007)	0.009 (0.006)	0.004 (0.005)	0.018** (0.008)
<i>Inspector designation</i>	-0.005 (0.003)	-0.006** (0.003)	-0.007*** (0.003)	-0.000 (0.004)
<i>Inspector workload</i>	-0.164*** (0.012)	-0.166*** (0.011)	-0.117*** (0.011)	-0.145*** (0.011)
<i>Past inspection outcomes</i>	0.112*** (0.008)	0.077*** (0.006)	0.079*** (0.005)	0.165*** (0.010)
<i>Recalls last year</i>	0.019 (0.139)	-0.085 (0.122)	-0.013 (0.120)	-0.150 (0.143)
<i>Inspection depth</i>	0.262*** (0.044)	0.251*** (0.035)	0.194*** (0.034)	0.016 (0.045)
<i>Inspection duration</i>	0.017*** (0.001)	0.018*** (0.001)	0.006*** (0.001)	0.084*** (0.004)
<i>Firm FDA distance</i>	-0.001 (0.000)	-0.001 (0.000)	-0.001 (0.000)	0.001 (0.000)
<i>State GDP</i>	0.007 (0.014)	0.009 (0.011)	0.003 (0.012)	0.009 (0.014)
<i>State employment</i>	-0.004 (0.006)	-0.004 (0.005)	-0.007 (0.005)	0.003 (0.006)
<i>State participation</i>	1.880 (2.734)	2.556 (2.208)	1.432 (2.107)	1.179 (2.836)

<i>Political alignment index</i>	-0.272*** (0.073)	-0.238*** (0.058)	-0.254*** (0.057)	-0.260*** (0.071)
<i>Republican</i>	-0.159 (0.180)	-0.120 (0.148)	-0.205 (0.157)	-0.155 (0.189)
<i>Concurrent inspections</i>	-	-	-0.223*** (0.058)	-
<i>Time between inspections</i>	-	-	-0.001** (0.000)	-
<i>Split inspections</i>	-	-	1.383*** (0.038)	-
<i>Prior severity of violations</i>	-	0.272*** (0.030)	-	-
<i>Current number of recalls</i>	-	0.002 (0.176)	-	-
<i>Recall severity</i>	-	0.127 (0.084)	-	-
<i>Day Fixed Effects</i>	Included	Included	Included	Included
<i>Month Fixed Effects</i>	Included	Included	Included	Included
<i>Year Fixed Effects</i>	Included	Included	Included	Included
<i>FDA Division Fixed Effects</i>	Included	Included	Included	Included
<i>Facilities nth inspection Fixed Effects</i>	Included	Included	Included	Included
<i>Inspector Fixed Effects</i>	Included	Included	Included	-
<i>Constant</i>	0.538 (0.659)	0.480 (0.535)	0.382 (0.531)	3.929*** (0.681)
Observations (Inspections)	11,633	16,382	16,382	19,441
Number of Facilities	11,633	14,441	14,441	16,904
Number of Inspectors	550	594	594	845
Wald's χ^2	1526.48***	2546.60***	4023.77***	1641.62***
Log likelihood	-9576.831	-14063.27	-13487.227	-12424.89

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Appendix E: Further Analysis

Table E.1. Further Analysis

Independent variables	(1) Violations
<i>Self-referent (H1)</i>	-0.006* (0.003)
<i>Industry referent (H2)</i>	0.012*** (0.002)
<i>Internal peer referent (H3)</i>	-0.005** (0.002)
<i>Operational interface referent (H4)</i>	-0.002*** (0.000)
<i>Years since last promotion</i>	0.060*** (0.019)
<i>Years since last promotion²</i>	-0.007** (0.003)
<i>Experience</i>	-0.006*** (0.001)
<i>Average time</i>	-0.022 (0.013)
<i>Gender-male</i>	0.001 (0.057)
<i>Inspector age</i>	-0.007** (0.003)
<i>Site-specific experience</i>	-0.170*** (0.026)
<i>Prior inspection outcomes</i>	0.009* (0.006)
<i>Inspector designation</i>	-0.006** (0.003)
<i>Inspector workload</i>	-0.174*** (0.010)
<i>Past inspection outcomes</i>	0.106*** (0.005)
<i>Recalls last year</i>	-0.100 (0.122)
<i>Inspection depth</i>	0.246*** (0.035)
<i>Inspection duration</i>	0.019*** (0.001)
<i>Firm FDA distance</i>	-0.001 (0.000)
<i>State GDP</i>	0.120*

	(0.071)
<i>State employment</i>	-0.005
	(0.005)
<i>State population</i>	-0.134
	(0.085)
<i>State participation</i>	3.351
	(2.225)
<i>Political alignment index</i>	-0.210***
	(0.059)
<i>Republican</i>	-0.112
	(0.148)
<i>Day, Month, FDA Div., nth insp., inspector</i>	Included
<i>Fixed Effects</i>	
<i>Constant</i>	0.505
	(0.532)
Observations (Inspections)	16,382
Number of Facilities	14,441
Number of Inspectors	594

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table E.2. Further Analysis

	(1)	(2)	(3)
	<i>Site- experience</i>	<i>Inspector workload</i>	<i>Breaks</i>
Independent variables	<i>Violations</i>	<i>Violations</i>	<i>Violations</i>
<i>Self-referent (H1)</i>	-0.007*** (0.003)	-0.005 (0.003)	-0.006** (0.003)
<i>Self-referent X Moderator</i>	0.001 (0.002)	-0.001 (0.001)	0.004 (0.003)
<i>Industry referent (H2)</i>	0.015*** (0.002)	0.012*** (0.002)	0.016*** (0.002)
<i>Industry referent X Moderator</i>	-0.003* (0.002)	0.001* (0.001)	-0.007*** (0.002)
<i>Internal peer referent (H3)</i>	-0.006*** (0.002)	-0.006* (0.003)	-0.007*** (0.002)
<i>Internal peer referent X Moderator</i>	0.003* (0.002)	-0.001 (0.001)	0.005 (0.003)
<i>Operational interface referent (H4)</i>	-0.002*** (0.000)	0.001 (0.001)	-0.002*** (0.000)
<i>Operational interface X Moderator</i>	0.001* (0.000)	-0.001*** (0.000)	0.002*** (0.001)
<i>Experience</i>	-0.006*** (0.001)	-0.006*** (0.001)	-0.006*** (0.001)
<i>Average time</i>	-0.022	-0.021	-0.052***

	(0.013)	(0.013)	(0.014)
<i>Gender-male</i>	-0.001	0.003	0.059
	(0.057)	(0.057)	(0.060)
<i>Inspector age</i>	-0.007**	-0.007**	-0.008***
	(0.003)	(0.002)	(0.003)
<i>Site-specific experience</i>	-0.205***	-0.171***	-0.158***
	(0.040)	(0.027)	(0.026)
<i>Prior inspection outcomes</i>	0.010*	0.010*	0.004
	(0.006)	(0.006)	(0.005)
<i>Inspector designation</i>	-0.006**	-0.006**	-0.008***
	(0.003)	(0.003)	(0.003)
<i>Inspector workload</i>	-0.174***	-0.117***	-0.127***
	(0.010)	(0.017)	(0.010)
<i>Past inspection outcomes</i>	0.107***	0.106***	0.100***
	(0.005)	(0.005)	(0.005)
<i>Recalls last year</i>	-0.103	-0.099	-0.023
	(0.123)	(0.122)	(0.120)
<i>Inspection depth</i>	0.243***	0.245***	0.200***
	(0.035)	(0.035)	(0.034)
<i>Inspection duration</i>	0.019***	0.019***	0.005***
	(0.001)	(0.001)	(0.001)
<i>Firm FDA distance</i>	-0.001	-0.001	0.001
	(0.000)	(0.000)	(0.000)
<i>State GDP</i>	0.011	0.011	0.004
	(0.011)	(0.011)	(0.012)
<i>State employment</i>	-0.004	-0.003	-0.006
	(0.005)	(0.005)	(0.005)
<i>State participation</i>	2.711	2.310	1.388
	(2.199)	(2.204)	(2.102)
<i>Political alignment index</i>	-0.239***	-0.247***	-0.255***
	(0.058)	(0.058)	(0.057)
<i>Republican</i>	-0.097	-0.098	-0.188
	(0.148)	(0.148)	(0.157)
<i>Day, Month, FDA Div., nth insp., inspector Fixed Effects</i>	Included	Included	Included
<i>Constant</i>	0.557	0.284	0.368
	(0.534)	(0.539)	(0.531)
Observations (Inspections)	16,382	16,382	16,382
Number of Facilities	14,441	14,441	14,441
Number of Inspectors	594	594	594
Wald's χ^2	2461.11***	2418.93***	3984.21***
Log likelihood	-14106.25	-14098.28	-13480.842

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.10

Bibliography

- Adams, J. S. (1965). Inequity in social exchange. In *Advances in experimental social psychology* (Vol. 2, pp. 267-299). Academic Press.
- Adams, J. S., & Freedman, S. (1976). Equity theory revisited: Comments and annotated bibliography. *Advances in Experimental Social Psychology*, 9, 43-90.
- Aguiar, M., & Gopinath, G. (2007). Emerging market business cycles: The cycle is the trend. *Journal of Political Economy*, 115(1), 69-102.
- Allon, G., Cohen, M. C., & Sinchaisri, W. P. (2023). The impact of behavioral and economic drivers on gig economy workers. *Manufacturing & Service Operations Management*, 25(4), 1376-1393.
- Altman, D., Yom-Tov, G. B., Olivares, M., Ashtar, S., & Rafaeli, A. (2021). Do customer emotions affect agent speed? An empirical study of emotional load in online customer contact centers. *Manufacturing & Service Operations Management*, 23(4), 854-875.
- Anand, G., & Gray, J. V. (2017). Strategy and organization research in operations management. *Journal of Operations Management*, 53, 1-8.
- Anand, G., Gray, J., & Siemsen, E. (2012). Decay, shock, and renewal: Operational routines and process entropy in the pharmaceutical industry. *Organizational Science*, 23(6), 1700-1716.
- Anand, K., Anupindi, R., & Bassok, Y. (2008). Strategic inventories in vertical contracts. *Management Science*, 54(10), 1792-1804.
- Anderson, C. M., Opaluch, J. J., & Grigalunas, T. A. (2009). The demand for import services at US container ports. *Maritime Economics & Logistics*, 11, 156-185.
- Angrist, J. D., & Pischke, J. S. (2008). *Mostly harmless econometrics: An Empiricist's Companion*. Princeton University Press.
- Austin, J. (2006). US life scientists report rising salaries and high job satisfaction. *Science*, 314(5800), 842-847.
- Autry, C. W., & Daugherty, P. J. (2003). Warehouse operations employees: linking person-organization fit, job satisfaction, and coping responses. *Journal of Business Logistics*, 24(1), 171-197.
- Azadegan, A., Patel, P. C., & Parida, V. (2013). Operational slack and venture survival. *Production and Operations Management*, 22(1), 1-18.
- Bachman, B., & Stokes, S. (2019). Federal Meat Inspector Work Overload, Burnout Has Critics Worrying About Food Safety. Retrieved from <https://investigatemedwest.org/2019/09/20/food-safety/>
- Baiocchi, M., Cheng, J., & Small, D. S. (2014). Instrumental variable methods for causal inference. *Statistics in Medicine*, 33(13), 2297-2340.
- Baker, S. R., Bloom, N., & Davis, S. J. (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics*, 131, 1593-1636.
- Baker, W. E., & Bulkley, N. (2014). Paying it forward vs. rewarding reputation: Mechanisms of generalized reciprocity. *Organization Science*, 25(5), 1493-1510.
- Ball, G., Siemsen, E., & Shah, R. (2017). Do plant inspections predict future quality? The role of investigator experience. *Manufacturing & Service Operations Management*, 19(4), 534-550.

- Balogun, J., Best, K., & Lê, J. (2015). Selling the object of strategy: How frontline workers realize strategy through their daily work. *Organization Studies*, 36(10), 1285-1313.
- Bard, J. F., & Wan, L. (2008). Workforce design with movement restrictions between workstation groups. *Manufacturing & Service Operations Management*, 10(1), 24-42.
- Barker, J. M., Hofer, C., Hoberg, K., & Eroglu, C. (2022). Supplier inventory leanness and financial performance. *Journal of Operations Management*, 68(4), 385-407.
- Barrios, J. M., Hochberg, Y. V., & Yi, H. (2020). The cost of convenience: Ridehailing and traffic fatalities. *Journal of Operations Management*.
- Beattie, G., Durante, R., Knight, B., & Sen, A. (2021). Advertising spending and media bias: Evidence from news coverage of car safety recalls. *Management Science*, 67(2), 698-719.
- Becker, D., Jin, G., & Leslie, P. (2012). Inspection design and inspector behavior.
- Belderbos, R., Du, H. S., & Goerzen, A. (2017). Global cities, connectivity, and the location choice of MNC regional headquarters. *Journal of Management Studies*, 54(8), 1271-1302.
- Bendig, D., Strese, S., & Brettel, M. (2017). The link between operational leanness and credit ratings. *Journal of Operations Management*, 52, 46-55.
- Bendoly, E., Donohue, K., & Schultz, K. L. (2006). Behavior in operations management: Assessing recent findings and revisiting old assumptions. *Journal of Operations Management*, 24(6), 737-752.
- Bendoly, E., Perry-Smith, J. E., & Bachrach, D. G. (2010). The perception of difficulty in project-work planning and its impact on resource sharing. *Journal of Operations Management*, 28(5), 385-397.
- Bennett, V. M., Pierce, L., Snyder, J. A., & Toffel, M. W. (2013). Customer-driven misconduct: How competition corrupts business practices. *Management Science*, 59(8), 1725-1742.
- Berger, D., Pukthuanthong, K., & Yang, J. J. (2011). International diversification with frontier markets. *Journal of Financial Economics*, 101(1), 227-242.
- Berger, J., Harbring, C., & Sliwka, D. (2013). Performance appraisals and the impact of forced distribution—an experimental investigation. *Management Science*, 59(1), 54-68.
- Berry, C. M., Lelchook, A. M., & Clark, M. A. (2012). A meta-analysis of the interrelationships between employee lateness, absenteeism, and turnover: Implications for models of withdrawal behavior. *Journal of Organizational Behavior*, 33(5), 678-699.
<https://doi.org/10.1002/job.778>
- Biggs, A., Brough, P., & Drummond, S. (2017). Lazarus and Folkman's psychological stress and coping theory. *The handbook of stress and health: A guide to research and practice*, 349-364.
- Binnewies, C., Sonnentag, S., & Mojza, E. J. (2009). Daily performance at work: Feeling recovered in the morning as a predictor of day-level job performance. *Journal of Organizational Behavior*, 30(1), 67-93.
- Blau, F. D., Koebe, J., & Meyerhofer, P. A. (2020). Essential and frontline workers in the COVID-19 crisis. *Econofact*, 6, 16.
- Bloom, N., Floetotto, M., Jaimovich, N., Saporta-Eksten, I., & Terry, S. J. (2018). Really uncertain business cycles. *Econometrica*, 86, 1031-1065.
- Bolino, M. C., Klotz, A. C., Turnley, W. H., & Harvey, J. (2013). Exploring the dark side of organizational citizenship behavior. *Journal of Organizational Behavior*, 34(4), 542-559.
- Bonaime, A., Gulen, H., & Ion, M. (2018). Does policy uncertainty affect mergers and acquisitions? *Journal of Financial Economics*.

- Bork, R.H., Robins, M., Schaffer, K., Leider, J. P., & Castrucci, B. C. (2022). Workplace perceptions and experiences related to COVID-19 response efforts among public health workers—Public Health Workforce Interests and Needs Survey, United States, September 2021–January 2022.
- Bozarth, C., Handfield, R., & Das, A. (1998). Stages of global sourcing strategy evolution: an exploratory study. *Journal of Operations Management*, 16(2-3), 241-255.
- Brachet, T., David, G., & Drechsler, A. M. (2012). The effect of shift structure on performance. *American Economic Journal: Applied Economics*, 4(2), 219-246.
- Bradley S, McAuliffe E (2009) Mid-level providers in emergency obstetric and newborn health care: factors affecting their performance and retention within the Malawian health system. *Hum. Resour. Health* 7(1):14.
- Brockner, J., Davy, J., & Carter, C. (1985). Layoffs, self-esteem, and survivor guilt: Motivational, affective, and attitudinal consequences. *Organizational Behavior and Human Decision Processes*, 36(2), 229-244.
- Brogaard, J., & Detzel, A. (2015). The asset-pricing implications of government economic policy uncertainty. *Management Science*, 61, 3–18.
- Brower, T. (2022). Burnout Is A Worldwide Problem: 5 Ways Work Must Change. Retrieved from <https://www.forbes.com/sites/tracybrower/2022/07/24/burnout-is-a-worldwide-problem-5-ways-work-must-change/?sh=5ef877ed6c1e>
- Browning, T. R. (2010). On the alignment of the purposes and views of process models in project management. *Journal of Operations Management*, 28(4), 316-332.
- Buckley, J. A. (2015). Food safety regulation and small processing: A case study of interactions between processors and inspectors. *Food Policy*, 51, 74-82.
- Cai, X., Gong, J., Lu, Y., & Zhong, S. (2018). Recover overnight? Work interruption and worker productivity. *Management Science*, 64(8), 3489-3500.
- Calvo, G. A., Izquierdo, A., & Talvi, E. (2006). Sudden stops and phoenix miracles in emerging markets. *American Economic Review*, 96(2), 405-410.
- Cantor, D. E., Corsi, T. M., & Grimm, C. M. (2009). Do electronic logbooks contribute to motor carrier safety performance?. *Journal of Business Logistics*, 30(1), 203-222.
- Cantor, D. E., Yan, T., Pagell, M., & Tate, W. L. (2022). From the editors: Introduction to the emerging discourse incubator on the topic of leveraging multiple types of resources within the supply network for competitive advantage. *Journal of Supply Chain Management*, 58(2), 3-7.
- Card, D., & Krueger, A. B. (2000). Minimum wages and employment: a case study of the fast-food industry in New Jersey and Pennsylvania: reply. *American Economic Review*, 90(5), 1397-1420.
- Card, D., Mas, A., Moretti, E., & Saez, E. (2012). Inequality at work: The effect of peer salaries on job satisfaction. *American Economic Review*, 102(6), 2981-3003.
- CDC. (2018). Estimates of Foodborne Illness in the United States. Retrieved from <https://www.cdc.gov/foodborneburden/index.html>
- Chang, P. L., Fujii, T., & Jin, W. (2022). Good names beget favors: The impact of country image on trade flows and welfare. *Management Science*, 68(10), 7555-7596.
- Charron, S., & Koechlin, E. (2010). Divided representation of concurrent goals in the human frontal lobes. *Science*, 328(5976), 360-363.

- Chen, D. L., Moskowitz, T. J., & Shue, K. (2016). Decision making under the gambler's fallacy: Evidence from asylum judges, loan officers, and baseball umpires. *Quarterly Journal of Economics*, 131(3), 1181-1242.
- Chen, D., Hu, N., Liang, P., & Swink, M. (2024). Understanding the impact of trade policy effect uncertainty on firm-level innovation investment. *Journal of Operations Management*.
- Choi, J. H., Kim, J. B., Qiu, A. A., & Zang, Y. (2012). Geographic proximity between auditor and client: How does it impact audit quality?. *Auditing: A Journal of Practice & Theory*, 31(2), 43-72.
- Choi, J., & Chen, C. C. (2007). The relationships of distributive justice and compensation system fairness to employee attitudes in international joint ventures. *Journal of Organizational Behavior*, 28(6), 687-703. <https://doi.org/10.1002/job.438>
- Chong, S., Huang, Y., & Chang, C. H. D. (2020). Supporting interdependent telework employees: A moderated-mediation model linking daily COVID-19 task setbacks to next-day work withdrawal. *Journal of Applied Psychology*, 105(12), 1408.
- Chopra, S., & Sodhi, M. S. (2004). Managing risk to avoid supply-chain breakdown. *MIT Sloan Management Review*.
- Chuang, H. H. C., Oliva, R., & Heim, G. R. (2019). Examining the link between retailer inventory leanness and operational efficiency: Moderating roles of firm size and demand uncertainty. *Production and Operations Management*, 28(9), 2338-2364.
- Cohen, B. J. (2008). *International political economy: An intellectual history*. Princeton University Press.
- Colquitt, J. A., Scott, B. A., Rodell, J. B., Long, D. M., Zapata, C. P., Conlon, D. E., & Wesson, M. J. (2013). Justice at the millennium, a decade later: A meta-analytic test of social exchange and affect-based perspectives. *Journal of Applied Psychology*, 98(2), 199-236. <https://doi.org/10.1037/a0031757>
- Compas, B. E., Connor-Smith, J. K., Saltzman, H., Thomsen, A. H., & Wadsworth, M. E. (2001). Coping with stress during childhood and adolescence: problems, progress, and potential in theory and research. *Psychological Bulletin*, 127(1), 87.
- Consumer Product Safety Commission. (n.d.). The Consumer Product Safety Improvement Act (CPSIA). Retrieved [April 8, 2024], from <https://www.cpsc.gov/Regulations-Laws--Standards/Statutes/The-Consumer-Product-Safety-Improvement-Act>
- Corbett, C. J. (2006). Global diffusion of ISO 9000 certification through supply chains. *Manufacturing & Service Operations Management*, 8(4), 330-350.
- Correia, S. (2015). Singletons, cluster-robust standard errors and fixed effects: A bad mix. *Technical Note, Duke University*, 7.
- Cullen, Z. (2024). Is pay transparency good?. *Journal of Economic Perspectives*, 38(1), 153-180.
- Dai, H., Milkman, K. L., Hofmann, D. A., & Staats, B. R. (2015). The impact of time at work and time off from work on rule compliance: the case of hand hygiene in health care. *Journal of Applied Psychology*, 100(3), 846.
- Danziger, S., Levav, J., & Avnaim-Pesso, L. (2011). Extraneous factors in judicial decisions. *Proceedings of the National Academy of Sciences*, 108(17), 6889-6892.
- Darby, J. L., Ketchen Jr, D. J., Williams, B. D., & Tokar, T. (2020). The implications of firm-specific policy risk, policy uncertainty, and industry factors for inventory: A resource dependence perspective. *Journal of Supply Chain Management*, 56(4), 3-24.

- Dartey-Baah, K., & Amoako, G. K. (2011). Application of Frederick Herzberg's Two-Factor theory in assessing and understanding employee motivation at work: a Ghanaian Perspective. *European Journal of Business and Management*, 3(9), 1-8.
- Deep, M. (2020). Surprisingly Effective Way To Name Matching In Python. Retrieved from <https://towardsdatascience.com/surprisingly-effective-way-to-name-matching-in-python-1a67328e670e>
- Deery, S., Iverson, R., & Walsh, J. (2002). Work relationships in telephone call centres: Understanding emotional exhaustion and employee withdrawal. *Journal of Management Studies*, 39(4), 471-496.
- DeHoratius, N., & Raman, A. (2007). Store manager incentive design and retail performance: An exploratory investigation. *Manufacturing & Service Operations Management*, 9(4), 518-534.
- Dhanorkar, S. (2019). Environmental benefits of internet-enabled C2C closed-loop supply chains: A quasi-experimental study of Craigslist. *Management Science*, 65(2), 660-680.
- DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, 48(2), 147-160.
- Dobrzykowski, D. (2019). Understanding the downstream healthcare supply chain: Unpacking regulatory and industry characteristics. *Journal of Supply Chain Management*, 55(2), 26-46.
- Dooley, K. J., Yan, T., Mohan, S., & Gopalakrishnan, M. (2010). Inventory management and the bullwhip effect during the 2007–2009 recession: evidence from the manufacturing sector. *Journal of Supply Chain Management*, 46(1), 12-18.
- Duflo, E., Greenstone, M., Pande, R., & Ryan, N. (2013). Truth-telling by third-party auditors and the response of polluting firms: Experimental evidence from India. *Quarterly Journal of Economics*, 128(4), 1499-1545.
- Easley, D., & O'Hara, M. (2009). Ambiguity and nonparticipation: The role of regulation. *The Review of Financial Studies*, 22(5), 1817-1843.
- Eisenhardt, K. M., & Brown, S. L. (1998). Time pacing: Competing in markets that won't stand still. *Harvard Business Review*, 76(2), 59-70.
- Elfenbein, D. W., Fisman, R., & McManus, B. (2023). The impact of socioeconomic and cultural differences on online trade. *Management Science*, 69(10), 6181-6201.
- Elking, I., Paraskevas, J. P., Grimm, C., Corsi, T., & Steven, A. (2017). Financial dependence, lean inventory strategy, and firm performance. *Journal of Supply Chain Management*, 53(2), 22-38.
- Ellis, S. C., Shockley, J., & Henry, R. M. (2011). Making sense of supply disruption risk research: A conceptual framework grounded in enactment theory. *Journal of Supply Chain Management*, 47, 65–96.
- Eroglu, C., & Hofer, C. (2011). Lean, leaner, too lean? The inventory-performance link revisited. *Journal of Operations Management*, 29(4), 356-369.
- Fainmesser, I. P., Olié Lauga, D., & Ofek, E. (2021). Ratings, reviews, and the marketing of new products. *Management Science*, 67(11), 7023-7045.
- Fan, D., Zhou, Y., Yeung, A. C., Lo, C. K., & Tang, C. (2022). Impact of the US–China trade war on the operating performance of US firms: The role of outsourcing and supply base complexity. *Journal of Operations Management*, 68(8), 928-962.

- FDA. (2022). Food Safety Modernization Act (FSMA). Retrieved from <https://www.fda.gov/food/guidance-regulation-food-and-dietary-supplements/food-safety-modernizationact-fsma>
- FDA. (2023). FSMA Final Rule for Preventive Controls for Human Food. Retrieved from <https://www.fda.gov/food/food-safety-modernization-act-fsma/fsma-final-rule-preventive-controls-human-food>
- Federalpay.org. (2024). FederalPay.org - The Civil Employee's Resource. Retrieved from <https://www.federalpay.org/>
- Field, T. F., & Dunn, P. L. (2013). Transferability of skills: Historical foundations and development. *Foundations of Forensic Vocational Rehabilitation*, 133.
- FINRA. (2015). The Final Frontier: Investing in Frontier Markets. Retrieved from <https://www.finra.org/investors/insights/final-frontier-investing-frontier-markets>
- Folger, R., & Konovsky, M. A. (1989). Effects of procedural and distributive justice on reactions to pay raise decisions. *Academy of Management journal*, 32(1), 115-130.
- Folkman, S. (1984). Personal control and stress and coping processes: a theoretical analysis. *Journal of Personality and Social Psychology*, 46(4), 839.
- Fowler, J. H. (2006). Elections and markets: The effect of partisanship, policy risk, and electoral margins on the economy. *The Journal of Politics*, 68(1), 89-103.
- Fritz, C., & Sonnentag, S. (2006). Recovery, well-being, and performance-related outcomes: The role of workload and vacation experiences. *Journal of Applied Psychology*, 91(4), 936-945.
- Fritz, C., Sonnentag, S., Spector, P. E., & McInroe, J. A. (2010). The weekend matters: Relationships between stress recovery and affective experiences. *Journal of Organizational Behavior*, 31(8), 1137-1162.
- Froehle, C. M., & White, D. L. (2014). Interruption and forgetting in knowledge-intensive service environments. *Production and Operations Management*, 23(4), 704-722.
- Fugate, B., Pagell, M., & Flynn, B. (2019). From the editors: introduction to the emerging discourse incubator on the topic of research at the intersection of supply chain management and public policy and government regulation. *Journal of Supply Chain Management*, 55(2), 3-5.
- Gilbert, C. G. (2005). Unbundling the structure of inertia: Resource versus routine rigidity. *Academy of Management Journal*, 48(5), 741-763.
- Gishkayeva, L. L., Aziyeva, R. K., & Abubakarov, M. V. (2019). Employment And Salary As Indicators Of Social Quality Of Life. *European Proceedings of Social and Behavioural Sciences*.
- Goes, P. B., Ilk, N., Lin, M., & Zhao, J. L. (2018). When more is less: Field evidence on unintended consequences of multitasking. *Management Science*, 64(7), 3033-3054.
- Golini, R., & Kalchschmidt, M. (2011). Moderating the impact of global sourcing on inventories through supply chain management. *International Journal of Production Economics*, 133(1), 86-94.
- Gray, J. V., Anand, G., & Roth, A. V. (2015). The influence of ISO 9000 certification on process compliance. *Production and Operations Management*, 24(3), 369-382.
- Griffis, S. E., Rao, S., Goldsby, T. J., Voorhees, C. M., & Iyengar, D. (2012). Linking order fulfillment performance to referrals in online retailing: an empirical analysis. *Journal of Business Logistics*, 33(4), 279-294.

- Grover, A. K., & Dresner, M. (2022). A theoretical model on how firms can leverage political resources to align with supply chain strategy for competitive advantage. *Journal of Supply Chain Management*, 58(2), 48-65.
- Grover, A. K., Chopra, S., & Mosher, G. A. (2016). Food safety modernization act: A quality management approach to identify and prioritize factors affecting adoption of preventive controls among small food facilities. *Food Control*, 66, 241-249.
- Gu, T., Sanders, N. R., & Venkateswaran, A. (2017). CEO incentives and customer-supplier relations. *Production and Operations Management*, 26(9), 1705-1727.
- Gualandris, J., Klassen, R. D., Vachon, S., & Kalchschmidt, M. (2015). Sustainable evaluation and verification in supply chains: Aligning and leveraging accountability to stakeholders. *Journal of Operations Management*, 38, 1-13.
- Gundlach, G. T., Frankel, R., & Krotz, R. T. (2019). Competition policy and antitrust law: implications of developments in supply chain management. *Journal of Supply Chain Management*, 55(2), 47-67.
- Gurvich, I., O'Leary, K. J., Wang, L., & Van Mieghem, J. A. (2020). Collaboration, interruptions, and changeover times: Workflow model and empirical study of hospitalist charting. *Manufacturing & Service Operations Management*, 22(4), 754-774.
- Haans, R. F., Pieters, C., & He, Z. L. (2016). Thinking about U: Theorizing and testing U-and inverted U-shaped relationships in strategy research. *Strategic Management Journal*, 37(7), 1177-1195.
- Han, C., Dong, Y., & Dresner, M. (2013). Emerging market penetration, inventory supply, and financial performance. *Production and Operations management*, 22(2), 335-347.
- Handley, S. M., & Gray, J. V. (2013). Inter-organizational quality management: The use of contractual incentives and monitoring mechanisms with outsourced manufacturing. *Production and Operations Management*, 22(6), 1540-1556.
- Hankins, W., Stone, A. L., Cheng, C. H. J., & Chiu, C. W. J. (2020). Corporate decision-making in the presence of political uncertainty: The case of corporate cash holdings. *Financial Review*, 55, 307-337.
- Harris, M. M., Anseel, F., & Lievens, F. (2008). Keeping up with the Joneses: A field study of the relationships among upward, lateral, and downward comparisons and pay level satisfaction. *Journal of Applied Psychology*, 93(3), 665-673.
<https://doi.org/10.1037/0021-9010.93.3.665>
- Hassan, T. A., Hollander, S., van Lent, L., & Tahoun, A. (2019). Firm-level political risk: Measurement and effects. *The Quarterly Journal of Economics*, 134, 2135-2202.
- Hayuth, Y. (1994). The overweight container problem and international intermodal transportation. *Transportation Journal*, 18-28.
- Heckman, J. J., Ichimura, H., & Todd, P. (1998). Matching as an econometric evaluation estimator. *The Review of Economic Studies*, 65(2), 261-294.
- Helper, S., Gray, J. V., Hughes, M. M., & Roman, A. V. (2021). Public policy and operations management. *Journal of Operations Management*, 67(7), 780-802.
- Helpman, E., & Itskhoki, O. (2010). Labour market rigidities, trade and unemployment. *The Review of Economic Studies*, 77(3), 1100-1137.
- Hendey GW, Barth BE, Soliz T (2005) Overnight and Postcall Errors in Medication Orders. *Acad. Emerg. Med.* 12(7):629-634.
- Herman, J. L., & Tetrick, L. E. (2009). Problem-focused versus emotion-focused coping strategies and repatriation adjustment. *Human Resource Management: Published in*

- Cooperation with the School of Business Administration, The University of Michigan and in Alliance with the Society of Human Resources Management*, 48(1), 69-88.
- Hilbe, J. M. (2011). *Negative binomial regression*. Cambridge University Press.
- Hill, A. D., Aime, F., & Ridge, J. W. (2017). The performance implications of resource and pay dispersion: The case of Major League Baseball. *Strategic Management Journal*, 38(9), 1935-1947. <https://doi.org/10.1002/smj.2616>
- Hitt, M. A., Dacin, M. T., Levitas, E., Arregle, J. L., & Borza, A. (2000). Partner selection in emerging and developed market contexts: Resource-based and organizational learning perspectives. *Academy of Management Journal*, 43(3), 449-467.
- Hofer, C., Eroglu, C., & Hofer, A. R. (2012). The effect of lean production on financial performance: The mediating role of inventory leanness. *International Journal of Production Economics*, 138(2), 242-253.
- Hoffmann, S., & Ahn, J. W. (2021). Updating economic burden of foodborne diseases estimates for inflation and income growth.
- Hollander, S., Pronk, M., & Roelofsen, E. (2010). Does silence speak? An empirical analysis of disclosure choices during conference calls. *Journal of Accounting Research*, 48, 531–563.
- Hook, J. G., & Cook, T. D. (1979). Equity theory and the cognitive ability of children. *Psychological Bulletin*, 86(3), 429.
- Hoopes, J. L., Merkley, K. J., Pacelli, J., & Schroeder, J. H. (2018). Audit personnel salaries and audit quality. *Review of Accounting Studies*, 23, 1096-1136.
- Hsee, C. K., Yang, A. X., & Wang, L. (2010). Idleness aversion and the need for justifiable busyness. *Psychological Science*, 21(7), 926-930.
- Hunter, E. M., & Wu, C. (2016). Give me a better break: Choosing workday break activities to maximize resource recovery. *Journal of Applied Psychology*, 101(2), 302.
- Huq, F. A., Chowdhury, I. N., & Klassen, R. D. (2016). Social management capabilities of multinational buying firms and their emerging market suppliers: An exploratory study of the clothing industry. *Journal of Operations Management*, 46, 19-37.
- Hutchison-Krupat, J., & Chao, R. O. (2014). Tolerance for failure and incentives for collaborative innovation. *Production and Operations Management*, 23(8), 1265-1285.
- Iacus, S. M., King, G., & Porro, G. (2012). Causal inference without balance checking: Coarsened exact matching. *Political Analysis*, 20(1), 1–24.
- Ibanez, M. R., & Toffel, M. W. (2020). How scheduling can bias quality assessment: Evidence from food-safety inspections. *Management Science*, 66(6), 2396-2416.
- Ibanez, M., Palmarozzo, A., Short, J. L., & Toffel, M. W. (2023). Second-versus Third-party Audit Quality: Evidence from Global Supply Chain Monitoring. *Available at SSRN 4496823*.
- Imbens, G. W., & Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, 47(1), 5-86.
- Jones, F. (2020, May, 17). McDonald's and Yum Brands face new food safety issues in China. Retrieved from <https://fooddigital.com/franchising/mcdonalds-and-yum-brands-face-new-food-safety-issues-china>
- Julio, B., & Yook, Y. (2012). Political uncertainty and corporate investment cycles. *Journal of Finance*, 67, 45–83.

- Kacperczyk, A., & Balachandran, C. (2018). Vertical and horizontal wage dispersion and mobility outcomes: Evidence from the Swedish microdata. *Organization Science*, 29(1), 17-38. <https://doi.org/10.1287/orsc.2017.1169>
- Karimi, A., Mishra, A., Natarajan, K. V., Sinha, K. K., Balsara, O., & Lamphere, B. (November 2023). Empowering Frontline Health Workers to Manage Stock-outs in LMICs: Evaluating Training Programs in Indonesia. Working Draft.
- Kauffmann, C. (2013). International Regulatory Co-operation. Retrieved from <https://www.oecd.org/gov/regulatory-policy/international-regulatory-co-operation-9789264200463-en.htm>
- Kawasaki, T., & Matsuda, T. (2015). Containerization of bulk trades: a case study of US–Asia wood pulp transport. *Maritime Economics & Logistics*, 17, 179-197.
- Kc, D. S. (2014). Does multitasking improve performance? Evidence from the emergency department. *Manufacturing & Service Operations Management*, 16(2), 168-183.
- Kell, G. (2005). The global compact selected experiences and reflections. *Journal of Business Ethics*, 59, 69-79.
- Kesavan, S., Lambert, S. J., Williams, J. C., & Pendem, P. K. (2022). Doing Well by Doing Good: Improving Retail Store Performance with Responsible Scheduling Practices at the Gap, Inc. *Management Science*, 68(11), 7818-7836.
- Kim, C. F., Pantzalis, C., & Park, J. C. (2012). Political geography and stock returns: The value and risk implications of proximity to political power. *Journal of Financial Economics*, 106(1), 196-228.
- Kim, D. Y., & Zhu, P. (2018). Supplier dependence and R&D intensity: The moderating role of network centrality and interconnectedness. *Journal of Operations Management*, 64, 7–18.
- Kistler, J. T., Janakiraman, R., Kumar, S., & Tiwari, V. (2021). The effect of operational process changes on preoperative patient flow: Evidence from field research. *Production and Operations Management*, 30(6), 1647-1667.
- Kleindorfer, P. R., & Saad, G. H. (2005). Managing disruption risks in supply chains. *Production and Operations Management*, 14(1), 53-68.
- Ko, K., Mendeloff, J., & Gray, W. (2010). The role of inspection sequence in compliance with the US Occupational Safety and Health Administration's (OSHA) standards: Interpretations and implications. *Regulation & Governance*, 4(1), 48-70.
- Koh, K., Rajgopal, S., & Srinivasan, S. (2013). Non-audit services and financial reporting quality: Evidence from 1978 to 1980. *Review of Accounting Studies*, 18(1), 1-33.
- Kök, A. G., & Shang, K. H. (2007). Inspection and replenishment policies for systems with inventory record inaccuracy. *Manufacturing & Service Operations Management*, 9(2), 185-205.
- Kouvelis, P., & Tang, S. Y. (2012). On optimal expediting policy for supply systems with uncertain lead-times. *Production and Operations Management*, 21(2), 309-330.
- Krasna, H., Venkataraman, M., & Patino, I. (2024). Salary Disparities in Public Health Occupations: Analysis of Federal Data, 2021–2022. *American Journal of Public Health*, 114(3), 329-339.
- Krasna, H., Venkataraman, M., Robins, M., Patino, I., & Leider, J. P. (2024). Standard Occupational Classification codes: gaps in federal data on the public health workforce. *American Journal of Public Health*, 114(1), 48-56.
- Kube, R., von Graevenitz, K., Löschel, A., & Massier, P. (2019). Do voluntary environmental programs reduce emissions? EMAS in the German manufacturing sector. *Energy*

- Economics*, 84, 104558.
- Kull, T., & Closs, D. (2008). The risk of second-tier supplier failures in serial supply chains: implications for order policies and distributor autonomy. *European Journal of Operational Research*, 186(3), 1158-1174.
- Kuntz, L., Mennicken, R., & Scholtes, S. (2015). Stress on the ward: Evidence of safety tipping points in hospitals. *Management Science*, 61(4), 754-771.
- Kusaba, K., Moser, R., & Rodrigues, A. M. (2011). Low-cost country sourcing competence: A conceptual framework and empirical analysis. *Journal of Supply Chain Management*, 47(4), 73-93.
- Laack, R. V. (2017). OIG Report on FDA's Inspections of Domestic Food Facilities: Challenges Remain. Retrieved from <https://www.thefdalawblog.com/2017/10/oig-report-on-fdas-inspections-of-domestic-food-facilities-challenges-remain/>
- Lapr , M. A., & Scudder, G. D. (2004). Performance improvement paths in the US airline industry: linking trade-offs to asset frontiers. *Production and Operations Management*, 13(2), 123-134.
- Lapr , M. A., & Tsikriktsis, N. (2006). Organizational learning curves for customer dissatisfaction: Heterogeneity across airlines. *Management Science*, 52(3), 352-366.
- Lapr , M. A., Mukherjee, A. S., & Van Wassenhove, L. N. (2000). Behind the learning curve: Linking learning activities to waste reduction. *Management Science*, 46(5), 597-611.
- Lawler, E. E. (1968). Equity theory as a predictor of productivity and work quality. *Psychological Bulletin*, 70(6p1), 596.
- Lazarus, R. S. (1998). The stress and coping paradigm. *Fifty years of the research and theory of RS Lazarus: An analysis of historical and perennial Issues*, 182-220.
- Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. Springer publishing company.
- Lee, H. H., & Li, C. (2018). Supplier quality management: Investment, inspection, and incentives. *Production and Operations Management*, 27(2), 304-322.
- Lee, H. L., Padmanabhan, V., & Whang, S. (1997). Information distortion in a supply chain: The bullwhip effect. *Management Science*, 43(4), 546-558.
- Lee, Y. S., Ribbink, D., & Eckerd, S. (2018). Effectiveness of bonus and penalty incentive contracts in supply chain exchanges: Does national culture matter?. *Journal of Operations Management*, 62, 59-74.
- Leroy, S., Schmidt, A. M., & Madjar, N. (2020). Interruptions and task transitions: Understanding their characteristics, processes, and consequences. *Academy of Management Annals*, 14(2), 661-694.
- Levie, J., & Autio, E. (2011). Regulatory burden, rule of law, and entry of strategic entrepreneurs: An international panel study. *Journal of Management Studies*, 48(6), 1392-1419.
- Levine, D. I., & Toffel, M. W. (2010). Quality management and job quality: How the ISO 9001 standard for quality management systems affects employees and employers. *Management Science*, 56(6), 978-996.
- Levinson, D.R. (2017). Challenges remain in FDA's inspections of domestic food facilities. Retrieved from <https://oig.hhs.gov/oei/reports/oei-02-14-00420.asp>
- Li, Y., Zhang, Y. A., & Shi, W. (2020). Navigating geographic and cultural distances in international expansion: The paradoxical roles of firm size, age, and ownership. *Strategic Management Journal*, 41(5), 921-949.

- Li-Ping Tang, T., Kim, J. K., & Shin-Hsiung Tang, D. (2000). Does attitude toward money moderate the relationship between intrinsic job satisfaction and voluntary turnover?. *Human Relations*, 53(2), 213-245.
- Lu, G., Du, R. Y., & Peng, X. (2022). The impact of schedule consistency on shift worker productivity: An empirical investigation. *Manufacturing & Service Operations Management*, 24(5), 2780-2796.
- Lu, R., Shen, Q., Wang, T., & Zhang, X. (2022). Frenemies: Corporate advertising under common ownership. *Management Science*, 68(6), 4645-4669.
- Lucca, D., Seru, A., & Trebbi, F. (2014). The revolving door and worker flows in banking regulation. *Journal of Monetary Economics*, 65, 17-32.
- Makadok, R., Burton, R., & Barney, J. (2018). A practical guide for making theory contributions in strategic management. *Strategic Management Journal*, 39(6), 1530-1545.
- Mani, V., & Muthulingam, S. (2019). Does learning from inspections affect environmental performance? Evidence from unconventional well development in Pennsylvania. *Manufacturing & Service Operations Management*, 21(1), 177-197.
- Manuj, I., & Mentzer, J. T. (2008). Global supply chain risk management. *Journal of business Logistics*, 29(1), 133-155.
- Marshall, B. R., Nguyen, N. H., & Visaltanachoti, N. (2015). Frontier market transaction costs and diversification. *Journal of Financial Markets*, 24, 1-24.
- Martinez, C., Steven, A. B., & Dresner, M. (2016). East Coast vs. West Coast: The impact of the Panama Canal's expansion on the routing of Asian imports into the United States. *Transportation Research Part E: Logistics and Transportation Review*, 91, 274-289.
- Masorgo, N., Mir, S., & Hofer, A. R. (2023). You're driving me crazy! How emotions elicited by negative driver behaviors impact customer outcomes in last mile delivery. *Journal of Business Logistics*, 44(4), 666-692.
- Mayo, K., Webb, E., Ball, G., & Bretthauer, K. (2024). Scheduling smarter: Scheduling decision impact on nurse-aide turnover. *Manufacturing & Service Operations Management*, 26(1), 182-196.
- McFadden, K. L., & Hosmane, B. S. (2001). Operations safety: an assessment of a commercial aviation safety program. *Journal of Operations Management*, 19(5), 579-591.
- McKinsey & Company. (2022). Taking the pulse of shifting supply chains. Retrieved from <https://www.mckinsey.com/capabilities/operations/our-insights/taking-the-pulse-of-shifting-supply-chains>
- Miller, J. W., Bolumole, Y., & Schwieterman, M. A. (2020). Electronic logging device compliance of small and medium size motor carriers prior to the December 18, 2017, mandate. *Journal of Business Logistics*, 41(1), 67-85.
- Miller, J. W., Saldanha, J. P., Rungtusanatham, M., & Knemeyer, M. (2017). How does driver turnover affect motor carrier safety performance and what can managers do about it?. *Journal of Business Logistics*, 38(3), 197-216.
- Morrison, B. (2015). The problem with workarounds is that they work: The persistence of resource shortages. *Journal of Operations Management*, 39, 79-91.
- Morvan, C., & O'Connor, A. (2017). *An analysis of Leon Festinger's a theory of cognitive dissonance*. Macat Library.
- MSCI (2024). MSCI Market Classification Framework. Retrieved from <https://www.msci.com/our-solutions/indexes/market->

- classification#:~:text=MSCI%20evaluates%20equity%20markets%20around,emerging%2C%20frontier%20or%20standalone%20market.
- Mulki, J. P., & Jaramillo, F. (2011). Workplace isolation: Salespeople and supervisors in USA. *The International Journal of Human Resource Management*, 22(04), 902-923.
- Muraven, M., & Baumeister, R. F. (2000). Self-regulation and depletion of limited resources: Does self-control resemble a muscle?. *Psychological Bulletin*, 126(2), 247.
- Nagar, V., Schoenfeld, J., & Wellman, L. (2018). The effect of economic policy uncertainty on investor information asymmetry and management disclosures. *Journal of Accounting and Economics*.
- Oliver, C. (1991). Strategic responses to institutional processes. *Academy of Management Review*, 16(1), 145-179.
- Pagell, M., & Krause, D. R. (2004). Re-exploring the relationship between flexibility and the external environment. *Journal of Operations Management*, 21, 629–649.
- Pagell, M., Fugate, B., & Flynn, B. (2018). Research at the intersection of supply chain management and public policy and government regulation. *Journal of Supply Chain Management*.
- Palit, S., Hora, M., & Ghosh, S. (2022). Global buyer–supplier networks and innovation: The role of technological distance and technological breadth. *Journal of Operations Management*, 68(6-7), 755-774.
- Palmarozzo, A., Short, J., & Toffel, M. (2021). Auditor Independence and Outsourcing: Aligning Incentives to Mitigate Shilling and Shirking. In *Academy of Management Proceedings* (Vol. 2021, No. 1, p. 11847). Briarcliff Manor, NY 10510: Academy of Management.
- Parmigiani, A., Klassen, R. D., & Russo, M. V. (2011). Efficiency meets accountability: Performance implications of supply chain configuration, control, and capabilities. *Journal of Operations Management*, 29(3), 212-223.
- Pendem, P. K., Green Jr, P. I., & Staats, B. R. (2022). The microstructure of work: Understanding productivity benefits and costs of interruptions. *Manufacturing & Service Operations Management*, 24(4), 2202-2220.
- Pettengill, G. N. (1993). An experimental study of the “blue-Monday” hypothesis. *The Journal of Socio-Economics*, 22(3), 241-257.
- Pierce, L., & Toffel, M. W. (2013). The role of organizational scope and governance in strengthening private monitoring. *Organization Science*, 24(5), 1558-1584.
- Podsakoff, N. P., LePine, J. A., & LePine, M. A. (2007). Differential challenge stressor-hindrance stressor relationships with job attitudes, turnover intentions, turnover, and withdrawal behavior: A meta-analysis. *Journal of Applied Psychology*, 92(2), 438-454. <https://doi.org/10.1037/0021-9010.92.2.438>
- Powell, D. A., Erdozain, S., Dodd, C., Costa, R., Morley, K., & Chapman, B. J. (2013). Audits and inspections are never enough: A critique to enhance food safety. *Food Control*, 30(2), 686-691.
- Power, D., & Terziovski, M. (2007). Quality audit roles and skills: Perceptions of non-financial auditors and their clients. *Journal of Operations Management*, 25(1), 126-147.
- Prentice, D. A., & Miller, D. T. (1992). When small effects are impressive. *Psychological Bulletin*, 112(1), 160-164.
- Rafferty, A. E., & Griffin, M. A. (2006). Perceptions of organizational change: a stress and coping perspective. *Journal of Applied Psychology*, 91(5), 1154.

- Repenning, N. P. (2000). Drive out fear (unless you can drive it in): The role of agency and job security in process improvement. *Management Science*, 46(11), 1385-1396.
- Rodrik, D. (2008). Second-best institutions. *American Economic Review*, 98(2), 100-104.
- Ronen, J. (2010). Corporate audits and how to fix them. *Journal of Economic Perspectives*, 24(2), 189-210.
- Rubel, O., & Prasad, A. (2016). Dynamic incentives in sales force compensation. *Marketing Science*, 35(4), 676-689.
- Salancik, G. R., & Pfeffer, J. (1978). A social information processing approach to job attitudes and task design. *Administrative Science quarterly*, 224-253.
- Schultz, K. L., McClain, J. O., & Thomas, L. J. (2003). Overcoming the dark side of worker flexibility. *Journal of Operations Management*, 21(1), 81-92.
- Schutze, H., Manning, C. D., & Raghavan, P. (2008). *Introduction to information retrieval* (Vol. 39). Cambridge University Press.
- Scott, A., Balthrop, A., & Miller, J. W. (2021). Unintended responses to IT-enabled monitoring: The case of the electronic logging device mandate. *Journal of Operations Management*, 67(2), 152-181.
- Scott, A., Li, M., Cantor, D. E., & Corsi, T. M. (2023). Do voluntary environmental programs matter? Evidence from the EPA SmartWay program. *Journal of Operations Management*, 69(2), 284-304.
- Scott-Young, C., & Samson, D. (2008). Project success and project team management: Evidence from capital projects in the process industries. *Journal of Operations Management*, 26(6), 749-766.
- Sebo, P. (2023). How well does NamSor perform in predicting the country of origin and ethnicity of individuals based on their first and last names?. *Plos One*, 18(11), e0294562.
- Shah, R., Ball, G. P., & Netessine, S. (2016). Plant operations and product recalls in the automotive industry: An empirical investigation. *Management Science*, 63(8), 2439-2459.
- Shen, L., Zhou, K. Z., Wang, K., & Zhang, C. (2023). Do political ties facilitate operational efficiency? A contingent political embeddedness perspective. *Journal of Operations Management*, 69(1), 159-184.
- Shi, W., Sun, S. L., Yan, D., & Zhu, Z. (2017). Institutional fragility and outward foreign direct investment from China. *Journal of International Business Studies*, 48, 452-476.
- Short, J. L., & Toffel, M. W. (2016). The integrity of private third-party compliance monitoring. *Administrative and Regulatory Law News*, 42(1), 22-25.
- Short, J. L., Toffel, M. W., & Hugill, A. R. (2016). Monitoring global supply chains. *Strategic Management Journal*, 37(9), 1878-1897.
- Short, J. L., Toffel, M. W., & Hugill, A. R. (2019). Code contingencies: Designing monitoring regimes to promote improvement in supply chain working conditions. Working Paper No. 17-001. Harvard Business School, Boston.
- Shou, Z., Zheng, X. V., & Zhu, W. (2016). Contract ineffectiveness in emerging markets: An institutional theory perspective. *Journal of Operations Management*, 46, 38-54.
- Shrestha, N. (2020). Detecting multicollinearity in regression analysis. *American Journal of Applied Mathematics and Statistics*, 8(2), 39-42.
- Sieweke, J., Köllner, B., & Süß, S. (2017). The relationship between employees' objective internal and external pay standing and their job performance: A within-person analysis. *Journal of Business and Psychology*, 32(5), 533-546. <https://doi.org/10.1007/s10869-016-9457-6>

- Skitka, L. J. (2009). Exploring the “lost and found” of justice theory and research. *Social Justice Research*, 22, 98-116.
- Somerfield, M. R., & McCrae, R. R. (2000). Stress and coping research: methodological challenges, theoretical advances, and clinical applications. *American Psychologist*, 55(6), 620.
- Song, M., & Wu, Y. (2009). *Handbook of research on text and web mining technologies*. Hershey, PA: IGI Global.
- Song, P., Wang, Q., Liu, H., & Li, Q. (2020). The value of buy online-and-pickup-in-store in omni-channel: Evidence from customer usage data. *Production and Operations Management*, 29(4), 995–1010.
- Staats, B. R., Dai, H., Hofmann, D., & Milkman, K. L. (2017). Motivating process compliance through individual electronic monitoring: An empirical examination of hand hygiene in healthcare. *Management Science*, 63(5), 1563-1585.
- Stahl, G. K., & Caligiuri, P. (2005). The effectiveness of expatriate coping strategies: The moderating role of cultural distance, position level, and time on the international assignment. *Journal of Applied Psychology*, 90(4), 603.
- Stanczyk, A., Cataldo, Z., Blome, C., & Busse, C. (2017). The dark side of global sourcing: a systematic literature review and research agenda. *International Journal of Physical Distribution & Logistics Management*, 47(1), 41-67.
- Starbird, S. A. (2005). Moral hazard, inspection policy, and food safety. *American Journal of Agricultural Economics*, 87(1), 15-27.
- Steven, A. B., & Britto, R. A. (2016). Emerging market presence, inventory, and product recall linkages. *Journal of Operations Management*, 46, 55-68.
- Steven, A. B., & Corsi, T. M. (2012). Choosing a port: An analysis of containerized imports into the US. *Transportation Research Part E: Logistics and Transportation Review*, 48(4), 881-895.
- Steven, A. B., Dong, Y., & Corsi, T. (2014). Global sourcing and quality recalls: An empirical study of outsourcing-supplier concentration-product recalls linkages. *Journal of Operations Management*, 32(5), 241-253.
- Sweeney, P. D., & McFarlin, D. B. (2005). Wage comparisons with similar and dissimilar others. *Journal of Occupational and Organizational Psychology*, 78(1), 113-131.
- Swift, C., Guide Jr, V. D. R., & Muthulingam, S. (2019). Does supply chain visibility affect operating performance? Evidence from conflict minerals disclosures. *Journal of Operations Management*, 65(5), 406-429.
- Ta, H., Esper, T. L., Rossiter Hofer, A., & Sodero, A. (2023). Crowdsourced delivery and customer assessments of e-Logistics Service Quality: An appraisal theory perspective. *Journal of Business Logistics*, 44(3), 345-368.
- Talluri, S., Kull, T. J., Yildiz, H., & Yoon, J. (2013). Assessing the efficiency of risk mitigation strategies in supply chains. *Journal of Business logistics*, 34(4), 253-269.
- Tan, T. F., & Netessine, S. (2020). At your service on the table: Impact of tabletop technology on restaurant performance. *Management Science*, 66(10), 4496-4515.
- Tang, C., & Tomlin, B. (2008). The power of flexibility for mitigating supply chain risks. *International Journal of Production Economics*, 116(1), 12-27.
- Thomas, R. W., Fugate, B. S., & Koukova, N. T. (2011). Coping with time pressure and knowledge sharing in buyer–supplier relationships. *Journal of Supply Chain Management*, 47(3), 22-42.

- Thomas, S. P., Murfield, M. L. U., & Eastman, J. K. (2021). I wasn't expecting that! The relational impact of negotiation strategy expectation violations. *Journal of Supply Chain Management*, 57(4), 3-25.
- Toffel, M. W., Short, J. L., & Ouellet, M. (2015). Codes in context: How states, markets, and civil society shape adherence to global labor standards. *Regulation & Governance*, 9(3), 205-223.
- Tokar, T., & Swink, M. (2019). Public policy and supply chain management: Using shared foundational principles to improve formulation, implementation, and evaluation. *Journal of Supply Chain Management*, 55(2), 68-79.
- Tong, X., Lai, K. H., Lo, C. K., & Cheng, T. C. E. (2022). Supply chain security certification and operational performance: The role of upstream complexity. *International Journal of Production Economics*, 247, 108433.
- Torre, E. D., Pelagatti, M., & Solari, L. (2015). Internal and external equity in compensation systems, organizational absenteeism and the role of explained inequalities. *Human Relations*, 68(3), 409-440.
- Tremblay, M., St-Onge, S., & Toulouse, J. M. (1997). Determinants of salary referents relevance: A field study of managers. *Journal of Business and Psychology*, 463-484.
- Trevor, C. O., & Wazeter, D. L. (2006). A contingent view of reactions to objective pay conditions: Interdependence among pay structure characteristics and pay relative to internal and external referents. *Journal of Applied Psychology*, 91(6), 1260.
- Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review*, 117(2), 440.
- Tröster, C., Van Quaquebeke, N., & Aquino, K. (2018). Worse than others but better than before: Integrating social and temporal comparison perspectives to explain executive turnover via pay standing and pay growth. *Human Resource Management*, 57(2), 471-481.
- Trougakos, J. P., Chawla, N., & McCarthy, J. M. (2020). Working in a pandemic: Exploring the impact of COVID-19 health anxiety on work, family, and health outcomes. *Journal of Applied Psychology*, 105(11), 1234.
- Tucker, A. L. (2007). An empirical study of system improvement by frontline employees in hospital units. *Manufacturing & Service Operations Management*, 9(4), 492-505.
- Tymon Jr, W. G., Stumpf, S. A., & Doh, J. P. (2010). Exploring talent management in India: The neglected role of intrinsic rewards. *Journal of World Business*, 45(2), 109-121.
- U.S. FDA. (2024, February 5). Food Safety Modernization Act (FSMA). Retrieved from <https://www.fda.gov/food/guidance-regulation-food-and-dietary-supplements/food-safety-modernization-act-fsma>
- Ulfa, M., Azuma, M., & Steiner, A. (2022). Burnout status of healthcare workers in the world during the peak period of the COVID-19 pandemic. *Frontiers in Psychology*, 13, 952783.
- Upton, D. M. (1995). Flexibility as process mobility: the management of plant capabilities for quick response manufacturing. *Journal of Operations Management*, 12(3-4), 205-224.
- Urda, J., & Loch, C. H. (2013). Social preferences and emotions as regulators of behavior in processes. *Journal of Operations Management*, 31(1-2), 6-23.
- Van Dyne, L. Ve Ellis, J. B. (2004). Job creep: a reactance theory perspective on organizational citizenship behavior as over-fulfillment of obligations. (Eds.: J.C. Shapiro, L.Shore, S. Taylor Ve L. Tetrick). *The employment relationship: 79 examining psychological and contextual perspectives*, 181-205. Oxford, NY: Oxford University Press.

- Varanasi, L. (2023). Billionaire businessman says young Indians should work 70 hours a week. Retrieved from <https://www.businessinsider.com/indian-billionaire-infosys-founder-calls-70-hour-workweek-2023-10>
- Villena, V. H., & Dhanorkar, S. (2020). How institutional pressures and managerial incentives elicit carbon transparency in global supply chains. *Journal of Operations Management*, 66(6), 697-734.
- Wagner, S. M., & Bode, C. (2008). An empirical examination of supply chain performance along several dimensions of risk. *Journal of Business Logistics*, 29(1), 307-325.
- Wang, B., Pan, X., Dresner, M. E., & Yao, Y. O. (2023). Eying Sales Gains with Virtual Try-On Technology: A Natural Experiment in Webrooming. Available at SSRN 4566923.
- Wang, Q., Jiang, S., Ngai, E. W., & Huo, B. (2024). Vendor selection in the wake of data breaches: A longitudinal study. *Journal of Operations Management*.
- Westman, M., & Eden, D. (1997). Effects of a respite from work on burnout: vacation relief and fade-out. *Journal of Applied Psychology*, 82(4), 516.
- Wiedmer, R., Rogers, Z. S., Polyviou, M., Mena, C., & Chae, S. (2021). The dark and bright sides of complexity: a dual perspective on supply network resilience. *Journal of Business Logistics*, 42(3), 336-359.
- Wiengarten, F., Fan, D., Lo, C. K., & Pagell, M. (2017). The differing impacts of operational and financial slack on occupational safety in varying market conditions. *Journal of Operations Management*, 52, 30-45.
- Williams, M. L., McDaniel, M. A., & Nguyen, N. T. (2006). A meta-analysis of the antecedents and consequences of pay level satisfaction. *Journal of Applied Psychology*, 91(2), 392-413. <https://doi.org/10.1037/0021-9010.91.2.392>
- Wirtz, D., Kruger, J., Scollon, C. N., & Diener, E. (2003). What to do on spring break? The role of predicted, on-line, and remembered experience in future choice. *Psychological Science*, 14(5), 520-524.
- Witt, M. A., & Lewin, A. Y. (2007). Outward foreign direct investment as escape response to home country institutional constraints. *Journal of International Business Studies*, 38(4), 579-594.
- Wooldridge, J. M. (2009). *Introductory econometrics: A modern approach* (4th ed.). Mason, OH: SouthWestern Cengage Learning.
- Wooldridge, J. M. (2015). Control function methods in applied econometrics. *Journal of Human Resources*, 50(2), 420-445.
- Wright, R. A., Junious, T. R., Neal, C., Avello, A., Graham, C., Herrmann, L., ... & Walton, N. (2007). Mental fatigue influence on effort-related cardiovascular response: Difficulty effects and extension across cognitive performance domains. *Motivation and Emotion*, 31, 219-231.
- Wu, Z., & Salomon, R. (2016). Does imitation reduce the liability of foreignness? Linking distance, isomorphism, and performance. *Strategic Management Journal*, 37(12), 2441-2462.
- Xu, X., Cropanzano, R., McWha-Hermann, I., & Lu, C. (2024). Multiple salary comparisons, distributive justice and employee withdrawal. *Journal of Applied Psychology*.
- Yao, Y., Dresner, M., & Zhu, K. X. (2019). "Monday effect" on performance variations in supply chain fulfillment: How information technology-enabled procurement may help. *Information Systems Research*, 30(4), 1402-1423.
- Yee, R. W., Yeung, A. C., & Cheng, T. E. (2008). The impact of employee satisfaction on

- quality and profitability in high-contact service industries. *Journal of operations management*, 26(5), 651-668.
- Yestiana, Y., Kurniati, T., & Hidayat, A. A. A. (2019). Predictors of burnout in nurses working in inpatient rooms at a public hospital in Indonesia. *Pan African Medical Journal*, 33(1).
- Yu, Q., Mankad, S., & Shunko, M. (2023). Evidence of the unintended labor scheduling implications of the minimum wage. *Manufacturing & Service Operations Management*.
- Zantedeschi, D., Feit, E. M., & Bradlow, E. T. (2017). Measuring multichannel advertising response. *Management Science*, 63(8), 2706-2728.
- Zhang, C., Henke Jr, J. W., & Griffith, D. A. (2009). Do buyer cooperative actions matter under relational stress? Evidence from Japanese and US assemblers in the US automotive industry. *Journal of Operations Management*, 27(6), 479-494.
- Zhang, R., Hora, M., John, S., & Wier, H. A. (2022). Competition and slack: The role of tariffs on cost stickiness. *Journal of Operations Management*, 68(8), 855-880.
- Zourrig, H., Chebat, J. C., & Toffoli, R. (2009). Consumer revenge behavior: a cross-cultural perspective. *Journal of Business Research*, 62(10), 995-1001.