Summer 7-29-2024

Essays on Experimental and Environmental Economics

Sharad K C

ABSTRACT

ESSAYS ON EXPERIMENTAL AND ENVIRONMENTAL ECONOMICS

By

SHARAD KC

AUGUST, 2024

Committee Chair: Dr. James C. Cox

Major Department: Economics

This dissertation has three chapters on experimental and environmental economics. Chapter one examines the economic impact of fake reviews on market outcomes. The second chapter, co-authored with my colleague Xiangyu Meng, investigates China's location choice of air quality monitors. The third chapter analyzes the corporate reporting behavior of firms for Toxic Release Inventory (TRI) program of Environmental Protection Agency (EPA).

The first chapter uses lab experiment to measure the impact of fake reviews on buyer seller payoffs, trust and trustworthiness in an online marketplace setting. As consumers increasingly rely on online reviews to make decisions, firms have an incentive to use fake reviews to build up their reputation and, consequently, sales revenue. Fake reviews distort the online reputation market, resulting in buyers making sub-optimal choices and truthful sellers facing unfair competition. However, the extent of the impact of fake reviews is hardly known because researchers cannot easily identify fake reviews. This problem is further exacerbated by sellers who quickly learn to mimic truthful reviews or provide monetary incentives for buyers to leave positive feedback. To solve this problem, the paper uses a laboratory experiment to accurately identify fake reviews and measure its impact on buyer

and seller payoff, trust, and trustworthiness in the marketplace. The paper also differentiates between two types of goods, experience and credence goods. We find that there is trade off between prices and fake reviews, sellers post lower price to offset the occurrence of fake reviews. The paper also finds that fake reviews lowers trust by 4.3% in the experience goods market.

The second chapter studies the location choice of air quality monitors in China. Since 2013, China has added more than four thousand air quality monitoring stations that provide the public with real-time information on six airborne pollutants, i.e., particular matter (P.M.) 2.5, P.M.10, sulfur dioxide (SO2), nitrogen dioxide (NO2), ozone (O3), and carbon monoxide (CO). Authorities manage these monitors at four levels of the government: state, provincial, municipal, and county. Typically, pollution monitors are located where they could be deemed useful, for example, within more air-polluted areas or near schools, hospitals, road traffic, and industries. While the real-time information has helped individuals, firms, and governments make decisions; it is unclear how a government body makes siting decisions. This chapter aims to answer three questions: Where are the pollution monitors located? What drives the decision to add a new monitor in a particular location? Is there a difference in location choice behavior between central and local governments in China? We find that, in 2021, central monitors were located in cleaner areas than local monitors and monitors were located around public buildings like schools and hospitals. However, when it comes to placing a monitor, central monitors are installed in polluted areas and local monitors are placed in cleaner areas. We also find that both, central and local monitors, are clustered around each other.

The third chapter analyzes the corporate reporting behavior for TRI program of EPA. Without the monitoring and enforcement activities by the Environmental Protection Agency

(EPA), environmental laws are primarily non-binding. Toxic Release Inventory is one such policy that requires constant monitoring by the EPA to ensure that the firms comply. The self-reported nature of firms' toxic releases makes it crucial for the EPA to inspect and punish violations of misreporting or non-reporting regularly. This paper examines the impact of past regulatory actions by the EPA on firms' compliance behavior regarding the Toxic Release Inventory (TRI) reporting. Using a dataset spanning more than 30 years, our analysis reveals several key findings. Firstly, there exists a significant correlation between prior inspections and subsequent enforcement actions, with firms having a notable 0.9% likelihood of facing enforcement following a previous inspection. Additionally, we observe a nuanced relationship between enforcement history and inspection likelihood, suggesting potential resource prioritization by regulatory bodies. Furthermore, our study unveils a "neighborhood effect" in regulatory outcomes, wherein both Neighbor Inspection and Neighbor Enforcement significantly influence a firm's probability of facing enforcement actions. These findings provide valuable insights into the complexities of EPA monitoring and enforcement strategies and their implications for promoting compliance with environmental regulations within the TRI reporting framework.

ESSAYS ON EXPERIMENTAL AND ENVIRONMENTAL ECONOMICS

BY

SHARAD KC

A Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree
of
Doctor of Philosophy
in the
Andrew Young School of Policy Studies
of
Georgia State University

GEORGIA STATE UNIVERSITY
2024

ACCEPTANCE

This dissertation was prepared under the direction of the candidate's Dissertation Committee. It has been approved and accepted by all members of that committee, and it has been accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Economics in the Andrew Young School of Policy Studies of Georgia State University.

Dissertation Chair: Dr. James C. Cox

Committee: Dr. H. Spencer Banzhaf
Dr. Vjollca Sadiraj
Dr. Mark Isaac

Electronic Version Approved:

Dr. Thomas J. Vicino, Dean
Andrew Young School of Policy Studies
Georgia State University
August, 2024

DEDICATION

For my parents, sisters, and my partner.

Without your support,

I wouldn't be able to become a Ph.D.

ACKNOWLEDGEMENTS

I am profoundly grateful to all those who have supported me throughout my journey at Georgia State University. This dissertation would not have been possible without the guidance, encouragement, and support of many individuals.

Firstly, I would like to express my deepest gratitude to my advisor, Dr. James Cox, for his unwavering support, insightful guidance, and valuable feedback throughout the research process. I vividly remember the first class I took with Dr. Cox on Experimental Economics. It was an eye opening and engaging class that encouraged me to pursue a lab experiment for one of my dissertation chapters. Dr. Cox has been more than an advisor; he has been a mentor, a source of inspiration, and a constant pillar of support. Whether through countless meetings, detailed feedback, or simply providing encouragement during challenging times, his contributions have been invaluable. I would also like to extend my sincere thanks to my professors and committee members, whose knowledge and expertise have greatly contributed to my understanding and growth in this field. Special thanks to Dr. Vjollca Sadiraj, Dr. Spencer Banzhaf, and Dr. Mark Isaac for their invaluable advice and mentorship. It was their constructive criticisms that pushed me to improve this dissertation.

I am grateful to my Ph.D. cohort and fellow students for their camaraderie, collaboration, and the stimulating discussions that have enriched my research experience. Their friendship and support have been a source of strength and inspiration throughout this journey. I specially want to thank Dr. Xiangyu Meng for being a great Co-author in helping shape the second chapter of my dissertation.

Furthermore, I am deeply appreciative of the support and encouragement from my family and friends. To my family, thank you for your unwavering belief in me and for providing the emotional and moral support necessary to persevere through the challenges of this endeavor. Your love and encouragement have been my foundation. To my friends, thank you for your understanding, patience, and for always being there to lift my spirits. Your support has been invaluable in helping me stay motivated and focused.

Finally, I want to thank the professors who commented and made suggestions on this dissertation at Andrew Young School of Policy Studies: Dr. Garth Heutel, Dr. Daniel Kreisman, Dr. Carlianne Patrick, and the list go on.

Table of Contents

List of Tables

List of Figures

Introduction

This dissertation has three chapters on experimental and environmental economics. Chapter one examines the economic impact of fake reviews on market outcomes. The second chapter, co-authored with my colleague Xiangyu Meng, investigates China's location choice of air quality monitors. The third chapter analyzes the corporate reporting behavior of firms for Toxic Release Inventory (TRI) program of Environmental Protection Agency (EPA).

The first chapter uses lab experiment to measure the impact of fake reviews on buyer seller payoffs, trust and trustworthiness in an online marketplace setting. As consumers increasingly rely on online reviews to make decisions, firms have an incentive to use fake reviews to build up their reputation and, consequently, sales revenue. Fake reviews distort the online reputation market, resulting in buyers making sub-optimal choices and truthful sellers facing unfair competition. However, the extent of the impact of fake reviews is hardly known because researchers cannot easily identify fake reviews. This problem is further exacerbated by sellers who quickly learn to mimic truthful reviews or provide monetary incentives for buyers to leave positive feedback. To solve this problem, the paper uses a laboratory experiment to accurately identify fake reviews and measure its impact on buyer and seller payoff, trust, and trustworthiness in the marketplace. The paper also differentiates between two types of goods, experience and credence goods. We find that there is trade off between prices and fake reviews, sellers post lower price to offset the occurrence of fake reviews. The paper also finds that fake reviews lowers trust by 4.3% in the experience goods market.

The second chapter studies the location choice of air quality monitors in China. Since

2013, China has added more than four thousand air quality monitoring stations that provide the public with real-time information on six airborne pollutants, i.e., particulate matter (P.M.) 2.5, P.M.10, sulfur dioxide (SO2), nitrogen dioxide (NO2), ozone (O3), and carbon monoxide (CO). Authorities manage these monitors at four levels of the government: state, provincial, municipal, and county. Typically, pollution monitors are located where they could be deemed useful, for example, within more air-polluted areas or near schools, hospitals, road traffic, and industries. While the real-time information has helped individuals, firms, and governments make decisions; it is unclear how a government body makes siting decisions. This chapter aims to answer three questions: Where are the pollution monitors located? What drives the decision to add a new monitor in a particular location? Is there a difference in location choice behavior between central and local governments in China? We find that, in 2021, central monitors were located in cleaner areas than local monitors and monitors were located around public buildings like schools and hospitals. However, when it comes to placing a monitor, central monitors are installed in polluted areas and local monitors are placed in cleaner areas. We also find that both, central and local monitors, are clustered around each other.

The third chapter analyzes the corporate reporting behavior for TRI program of EPA. Without the monitoring and enforcement activities by the Environmental Protection Agency (EPA), environmental laws are primarily non-binding. Toxic Release Inventory is one such policy that requires constant monitoring by the EPA to ensure that the firms comply. The self-reported nature of firms' toxic releases makes it crucial for the EPA to inspect and punish violations of misreporting or non-reporting regularly. This paper examines the impact of past regulatory actions by the EPA on firms' compliance behavior regarding the Toxic

Release Inventory (TRI) reporting. Using a dataset spanning more than 30 years, our analysis reveals several key findings. Firstly, there exists a significant correlation between prior inspections and subsequent enforcement actions, with firms having a notable 93% likelihood of facing enforcement following a previous inspection. Additionally, we observe a nuanced relationship between enforcement history and inspection likelihood, suggesting potential resource prioritization by regulatory bodies. Furthermore, our study unveils a ”neighborhood effect” in regulatory outcomes, wherein both Neighbor Inspection and Neighbor Enforcement significantly influence a firm's probability of facing enforcement actions. These findings provide valuable insights into the complexities of EPA monitoring and enforcement strategies and their implications for promoting compliance with environmental regulations within the TRI reporting framework.

Chapter I The Economic Impact of Fake Reviews

1.1 Introduction

A recent study by Fakespot, a free website that analyzes online product reviews to filter out fake reviews, claims one-third of Amazon reviews are not trustworthy. As consumers increasingly rely on online reviews to make decisions, firms have an incentive to use fake reviews to build up their reputation and, consequently, sales revenue. Fake reviews distort the online reputation market, resulting in buyers making suboptimal choices and truthful sellers facing unfair competition. This paper studies the effect of fake reviews on market outcomes such as buyer payoff, seller payoff and trust.

Fake reviews erode trust in markets since consumers make sub-optimal choices. Luca (2016) finds that a one-star increase in Yelp reviews leads to a five to nine percent increase in restaurants' revenue. Sellers have a strong economic incentive to game the reputation system using fake reviews. Recent research into the field of fake reviews is limited to establishing and estimating the magnitude of the problem (Luca 2016, Mayzlin et al. 2014). However, these studies often face the challenge of properly identifying fake reviews. Furthermore, the economic impact of fake reviews on buyers and sellers is unknown. This paper eliminates the problem of identifying fake reviews through a lab experiment and explore the economic impact of fake reviews on buyers and sellers.

Identifying fake reviews is particularly challenging in online markets as fake reviewers continuously adapt to the changes made by platform. To study the economic impact of fake reviews, the paper uses a lab experiment to accurately identify fake reviews and use that identification to assess if buyers make sub-optimal choices. The paper compares average buyer and seller payoff, trust, and trustworthiness between markets with fake reviews and markets without fake reviews to know the effect on buyers/sellers. The hypotheses of the paper is that sellers have a higher incentive to provide fake reviews and bear a higher economic burden than buyers.

The specific aims of the paper are:

1. To estimate the economic impact of fake reviews on buyers and sellers by correctly identifying fake reviews from real reviews.

2. To measure the effect of fake reviews on different types of goods categorized as experience, and credence goods.

The paper is one of the few to explore the economic impact of fake reviews on buyers and sellers. The paper finds that sellers post lower prices in fake review treatments to compensate for bad rating system, in some cases by up to 6.8 percent. The paper also finds that there is lower trust in fake review by 4.3 percent. An interesting result from the paper is that there is lower trust in sessions with any kind of review system, fake or verified. One explanation for this result is that buyers upon seeing the very low rating of a seller in treatment session choose not to enter into the transaction. We discuss more about this result later in the paper.

The remainder of the paper proceeds as follows. Section 1.2 provides background and literature review. We then present our experimental design in Section 1.3, hypothesis in Section 1.4, and empirical results in Section 1.5. Finally, Section 1.6 concludes.

1.2 Literature Review

Reputation systems play a vital role in our daily lives. We use ratings and reviews to make purchasing decisions, decide where to eat or even when deciding to buy a house. There is a vast literature confirming the economic importance of reputation systems (Luca, 2011; Mayzlin et al., 2014; Resnick et al., 2000; Solimine & Isaac, 2023). Reputation systems partly solve problems arising from marketplace information asymmetry. Reputation systems also help foster competition between sellers and thus raise the overall quality of the goods and services. However, since most of the reputation mechanisms are voluntary and incur some cost, accurate ratings and reviews are hard to get.

The effect of online feedback system on fostering trust in marketplace has been well documented in the literature. Bolton et al. (2004) use experimental method to find the effect of online feedback on transaction efficiency. While they find that feedback improves efficiency, these mechanisms also act as a public good as in the benefits or trust, and all agents reap trustworthy behavior in the platform (Bolton et al., 2004). Ert et al. (2016) use controlled experiment using simulated choice on Airbnb to find the effect of photos on host's trustworthiness. They report that those perceived trustworthy through host's photo (visual information) can demand higher listing price as well as a higher probability of matching. Solimine and Isaac (2023) examines how the inclusion of a reputation aggregation system in online marketplaces affects market dynamics, demonstrating that it reduces false advertising and enhances trust among participants, though it falls short of achieving full market efficiency. Through the analysis of transaction data using bipartite networks, the study quantifies how ratings contribute to the development of diverse, trustworthy, and high-quality market environments.

Reputation systems also help sellers to increase their revenue. Luca (2011) finds that a one-star increase in Yelp rating lead to a five to nine percent increase in revenue. Similarly, Chevalier and Mayzlin (2006) also find that an improvement in a book's reviews leads to a rise in relative sales. It provides a considerable incentive to sellers to improve their ratings and reviews. Often this could mean that seller resorts to illegal means such as fake reviews. The Federal Trade Commission recently alleged two companies of engaging in posting deceptive or inaccurate information to mislead consumers.[1]

Online marketplaces such as Amazon, eBay, Uber, Airbnb, and others have a critical task of eliciting actual reputation of sellers to foster greater trust in their platforms. However, problems such as reciprocal reviewing, reputation inflation, and J shaped distribution of inflation plague online reputation systems. In recent years, the problem of fake reviews has been a significant concern for online marketplaces. Research by consumer association *Which?* found that Facebook groups that traded fake reviews were still operating and posting large number of posts on the social media platform.[2] A recent study by He et al. (2022) find that soliciting fake reviews on Facebook leads to a substantial increase in average rating of the seller in the short-term. They find that this increase, however, disappeared a month after a seller stopped buying reviews followed by a significant increase in one-star reviews.

In the online markets, fake reviews can be of various types. Sellers can give positive reviews to themselves, or they can give negative reviews to their competitors. Sellers can also buy fake reviews from buyers after a sale in return for a gift card. There have been

[1] https://www.ftc.gov/news-events/press-releases/2019/10/devumi-owner-ceo-settle-ftc-charges-they-sold-fake-indicators

[2] https://techcrunch.com/2019/08/06/facebook-still-full-of-groups-trading-fake-reviews-says-consumer-group/

numerous efforts from online marketplaces to tackle the problem of fake reviews. Usually, these platforms have algorithms that look for clues, such as the length of the review, the total number of reviews left by the reviewer and many more. However, fake reviewers have also learned to adapt to these algorithms. So, it is difficult to identify a fake review from a true one.

The problem of identifying a fake review from a true one has also made it difficult for economist to study the economic effects of fake reviews. There are a few approaches used in the literature to study the impact of fake reviews. Mayzlin et al. (2014) use differences in the reviews posted at TripAdvisor (anyone can post review) and Expedia (only a customer can post a review) for different types of hotels. Rather than directly identifying fake reviews, they rely on the assumption that ratings on TripAdvisor are higher than on Expedia since the cost of posting a fake review is less than in Expedia. They find that hotels with next-door neighbors have more negative reviews on TripAdvisor. The study also finds that independent hotels with small owners and small management companies have more positive reviews on TripAdvisor.

Luca and Zervas (2016) use a different approach. They use data from Yelp to estimate the incidence of review fraud and to understand the conditions under which it is most prevalent. Yelp uses an algorithm that flags reviews that it believes are unreliable. Luca and Zervas (2016) use this difference in properties between filtered and flagged reviews to identify fake reviews. They find that low ratings increase incentives for positive review fraud, and having more reviews reduces incentives for positive review fraud. They also find that competition encourages negative review fraud. Another strategy used in the literature by He et al. (2022) is by following products (Sellers) from Amazon.com that are active on secret online (Facebook) groups that solicit fake reviews. The authors then compare the before and after average ratings of these products. They find that the number of ratings per week increased by seven, and the average ratings increase by 0.08 in the weeks following the purchase of fake reviews.

The effect of reviews and fake reviews can also affect market outcomes based on the type of product, experience and credence goods. The exploration of experience and credence goods markets began with Akerlof's seminal paper on the "lemons problem," which highlighted how quality uncertainty affects market mechanisms. Akerlof demonstrated how markets with asymmetric information (e.g., used cars) lead to adverse selection and thus high-quality goods are driven out because buyers cannot differentiate them from low-quality goods ("lemons"), resulting in a market failure where only poor-quality goods are sold (Akerlof, 1978). Nelson expanded on this by examining how information and advertising influence consumer behavior in markets for experience goods. Nelson found that consumers rely on advertising as a proxy for quality in experience goods. Furthermore, high advertising expenditures signal higher quality, influencing purchasing decisions and the availability of information through advertising helped mitigate some information asymmetries (Nelson, 1970, 1974). Darby and Karni introduced the concept of credence goods and discussed the optimal level of fraud in such markets. They highlighting how information asymmetry leads to fraud. They showed that in the absence of regulation, sellers might overcharge or provide unnecessary services. They suggested that regulation and certification can reduce fraud and improve market efficiency (Darby & Karni, 1973). Plott and Wilde's experiment showed that consumers heavily depend on professional diagnosis in credence goods markets, emphasizing the potential for professionals to exploit their informational advantage. They found that expert services can reduce information asymmetry, but also highlighted the potential for professionals to exploit their informational advantage. (Plott & Wilde, 1980). Lynch, Miller, Plott, and Porter found that in posted offer markets, sellers adjust prices based on perceived product quality. High-quality products command higher prices and attract more consumers, demonstrating the importance of quality signaling in experience goods markets. (Lynch et al., 1991). Emons analyzed how experts might exploit information asymmetry in credence goods markets and suggested mechanisms like liability and warranties to mitigate fraud (Emons, 1997). Wolinsky found that competition among experts can improve service quality and pricing, though too much competition might lead to an oversupply of low-quality services, suggesting a balance between competition and regulation is necessary (Wolinsky, 1993). Dulleck, Kerschbamer, and Sutter's experimental study found that liability, verifiability, and reputation significantly reduce fraudulent behavior by experts, and that competition, when combined with these factors, enhances market efficiency by reducing information asymmetry and encouraging honest behavior (Dulleck et al., 2011). These differences in experience and credence goods market provide motivation for the paper to look at the differential effects of reviews and fake review based on the type of product.

1.3 The Experiment

1.3.1 Experimental Game Design

The experiment design for the credence goods market follows Dulleck et al. (2011) experiment that studies the determinants for efficiency in the credence goods market. The design employed in this paper is a modified version of the Dulleck et al. (2011) experiment. It is based on a transaction in an credence goods marketplace. There are two types of players, buyers and sellers. Furthermore, buyers can be of two types, type A and type B, and there are two types of goods in the market, good A and good B. Type A buyers are high type of buyers that prefer a high type of good and type B buyers prefer type B good or low type good. Type B buyers are satisfied with type A good, but incur a higher cost and for this reason they prefer type B good although they are satisfied with the quality of good in both cases.

The design for experience goods closely matches the design of the credence goods, except that the buyer knows their type. Furthermore, unlike in the credence goods market, the price that buyer pays depends on the buyer type and not on the quality of good seller provides. This is consistent with an experience goods like a hotel room. As a buyer, you can select the room and pay for it when booking it. The seller can then choose to provide the high or low quality service but still receive the price already paid for by the buyer.

In the case of experience good, as shown in Figure 1, the game proceeds in the following sequence of events:

◦ First seller posts prices for the two types of goods, (p_a, p_b) which costs the seller (c_a, c_b) to provide such that $p_a > p_b$ and $c_a > c_b$.

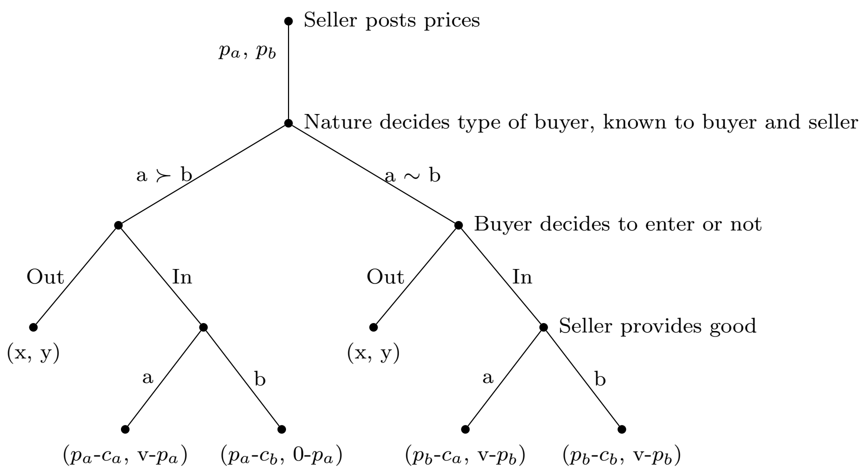
◦ Then nature decides the quality of the buyer (a or b), which is revealed to both the buyer and the seller.

◦ After nature decides the type, buyer will choose between buying from the seller versus opting out. If buyer chooses to buy from the seller, the game proceeds to the next step else, the game ends here.

◦ Seller chooses the quality of the good and charges for it. Seller charges p_a for high type of buyer and p_b for low type of buyer. For the baseline game (No review), the game ends here.

○ In the baseline game, there is not rating. For Verified Review treatment, only the matched buyer can provide a rating between one to five. For Fake Review treatment, all sellers can provide a rating to one other seller. Starting round 2 of the game, the buyer will know the average (arithmetic mean) rating of the seller he/she is matched with.

Figure 1 Experience Good



The final node in Figure 1 shows the payoffs for buyer and sellers in the following format: (Seller payoff, Buyer payoff). Seller payoff is equal to the price charged minus the cost of good provided. Buyer payoff is dependent on whether the buyer receives sufficient quality. The buyer receives a value v if the quality of the good meets their type. The low type of buyer always receives value v whereas the high-quality buyer only receives a value v if the seller provides high-quality good.

For credence goods games, the sequence of the game is the same as in experience good expect for the buyer must decide whether to transact with the seller before nature decides the type of the buyer. Buyer will not know his/her type throughout the game. Seller, on the other hand will know the type of the buyer and subsequently choose to provide the good and charge for it[3].

1.3.2 Experimental Treatments

The experimental design incorporates a 2 X 3 treatment design framework. The first criterion is on the type of good, experience and credence. The second is on whether sellers have a review system and if they can manipulate it, No review, Verified review, and Fake review.

Type of Good. Type of good, experience and credence, is based on the following criteria:

◦ If the buyer knows the quality of the good they want i.e. buyers know their type.

◦ If the buyer knows the quality of good they get from the seller i.e. buyer is notified of the type of good they get.

Experience goods are those where the buyer knows the quality of the good they desire but remains uncertain about the quality of the product they will receive from the seller. The buyer has a general idea of what they want, but the specific attributes of the product are uncertain until experienced. A classic example of an experience good is booking a hotel room. Before making a reservation, a guest may have an idea of the type of room they prefer based on descriptions, reviews, and photographs. However, until they physically stay in the room, they cannot be certain about its cleanliness, comfort, or other qualitative aspects.

Credence goods, on the other hand, present a higher level of uncertainty for the buyer. In this case, the buyer not only lacks knowledge about the quality of the good they want but also remains uninformed about the quality of the product they receive from the seller. These goods often require specialized knowledge or expertise to evaluate, making it challenging

[3]See Appendix figure A.1 for credence good game tree.

for consumers to make informed decisions. Common examples of credence goods include visits to professionals such as doctors or car mechanics. When seeking medical advice or car repairs, consumers rely heavily on the expertise and integrity of the service provider since they lack the knowledge to assess the quality of the service themselves.

In the experiment, I distinguish experience and credence good by letting the buyers know their type in experience goods, whereas in credence goods buyers do not know their type. I also inform the subjects in the subject instructions on whether the buyer knows their type or not. In all treatments sellers always know the type of the matched buyer. In addition, sellers also know whether buyers know their type. In addition, I also let the subjects know whether buyers or sellers know the type of buyer, In Decision 4 of Subject Instructions for Player B in Appendix 3.7, I change the wording from, "You and your matched player know your type." for experience goods to, "You will NOT know your type." Similarly for Player A in Appendix 3.7, I change the wording from, "You and your matched player know your type." to "Only you will know the type of Player B."

Understanding the distinction between experience and credence goods is essential for both, consumers and producers. For consumers, it highlights the importance of research, reviews, and reputation when making purchasing decisions, especially for goods and services with inherent uncertainties. For producers, it underscores the significance of transparency, reputation management, and quality assurance to build trust and loyalty among consumers in markets characterized by informational asymmetry.

Type of Rating System. In Verified Review, buyers have the authority to provide ratings only after finalizing a transaction with a matched seller. These ratings, known as "verified reviews," act as evidence of the transaction's validity, promoting transparency and trust within the marketplace. Before each game round, the average seller rating, calculated from verified reviews, is shared with buyers, allowing them to make well-informed decisions based on the reputation of potential trading partners.

Unlike the verified reviews, the fake review setup introduces a more complex scenario where all sellers, regardless of transactional participation, can post reviews for each other. This broad review system goes beyond genuine transactions, possibly paving the way for the spread of inaccurate or deceptive feedback. As with the verified reviews system, the average seller rating, incorporating both authentic and deceptive reviews, is revealed to buyers before each round, offering a nuanced view of seller credibility.

Table 1 Treatment Table

Type of good/ Type of review	No review	Verified Review	Fake Review
Experience good	38	40	40
Credence good	38	38	40
Total	76	78	80

Table 1 provides the treatment table with number of subjects in each treatment [4]. Number of subjects across treatment are fairly even. No review serves as the baseline scenario, devoid of any review mechanisms or reputation systems. In each round of the game, buyers and sellers possess only private information regarding their previous transactions, with no external feedback or ratings influencing their decisions. In verified review treatment, buyers can provide one review to the paired seller. They are also provided with the average

[4] See Table A.2 in Appendix for incidence of fake review for each round of the game.

rating of the seller from previous rounds of the game, starting from round 2. Finally in the Fake review treatment, in addition to buyers providing the rating to their paired seller, sellers can choose to provide themselves and one other seller a "fake" review. Buyers are aware of the possibility that sellers can provide reviews in the fake review treatment.

1.3.3 Matching Between Subjects

The matching process between buyers and sellers is characterized by a randomized mechanism, wherein participants are paired anew following the conclusion of each round of the game. This random matching protocol ensures that neither buyers nor sellers retain any knowledge or identification of their previous counterparts in the marketplace interactions. It fosters an environment of impartiality and unpredictability, thereby minimizing any potential biases or strategic maneuvers stemming from past interactions.

Within this framework, buyers are afforded no insight into the identities or characteristics of sellers they have previously engaged with, and conversely, sellers remain oblivious to the historical interactions with specific buyers. This anonymity and lack of persistent identification not only uphold the integrity of the experimental design but also emulate real-world market conditions where transactions often occur between unfamiliar parties.

The experimental protocol unfolds across a total of eighteen rounds, divided into two distinct phases: initial practice rounds and subsequent paid rounds. During the preliminary phase, participants partake in two practice rounds, providing them with an opportunity to acquaint themselves with the mechanics of the game and familiarize themselves with the interface. Importantly, this introductory phase also affords subjects the liberty to ask questions and seek clarifications on any aspects of the game mechanics or rules that may be unclear, thus ensuring a comprehensive understanding before transitioning to the paid rounds.

Following the completion of the practice rounds, participants proceed to engage in sixteen paid rounds, wherein their decisions and actions carry tangible consequences and potential rewards. This delineation between practice and paid rounds not only serves as a pedagogical tool for participants but also enables researchers to discern between exploratory behavior during the initial phase and strategic decision-making in the subsequent rounds.

1.3.4 Experimental Procedure

The experiment was conducted, using oTree version 2.1.28(Chen et al., 2016), in the Experimental Economics Center (ExCEN) center at Georgia State University between May 1, 2022 and May 31, 2022. A total of 234 undergraduate students participated in the experiments with 1,216 observations over the course of 16 rounds of the game.

Upon arrival at the lab, participants were briefed on the study's details and provided informed consent, after which they were randomly assigned to computer terminals. Once all participants were seated, they received printed instructions [5]. To help participants understand the experimental task, there were two practice rounds of the game. The experimenter then answered any clarification questions participants had. Information regarding the distribution across treatment conditions was not disclosed beforehand, and participants were prohibited from communicating with each other during the experiment. After the end of experiment, all participants answered a questionnaire asking for their demographic information. Participants did not receive monetary incentive for the questionnaire. All participants received a show up fee of 5 US dollars. On average, the subjects earned a total of 20 US

[5] Subject instructions for treatments are included in the Appendix 3.7.

dollars.

Table 2 provides summary statistics of the experiment subjects. It presents key de-

Table 2 Summary Statistics

	No review			Verified review			Fake review		
	N	Mean	SD	N	Mean	SD	N	Mean	SD
Female	76	0.70	0.46	78	0.65	0.48	80	0.63	0.48
Age	76	20.20	2.97	78	20.54	2.27	80	20.7	2.88
Asian	76	0.28	0.45	78	0.21	0.40	80	0.30	0.46
Black	76	0.57	0.50	78	0.58	0.49	80	0.53	0.50
Hispanic	76	0.03	0.16	78	0.03	0.16	80	0.06	0.24
Multiracial	76	0.04	0.19	78	0.08	0.27	80	0.04	0.19
White	76	0.08	0.27	78	0.09	0.29	80	0.08	0.26
N/A	76	0.01	0.11	78	0.03	0.16	80	0.00	0.00

mographic statistics across three experimental groups: no review, verified review, and fake review. In terms of gender distribution, the proportion of females is slightly higher in the no review group (70%) compared to both the verified review (65%) and fake review (63%) groups. Regarding age, the mean age increases marginally from the no review group (20.20) to the verified review group (20.54) and further to the fake review group (20.7), with increasing variability within age observed in the fake review group. Analysis of ethnicity reveals nuanced patterns: while the proportion of Asian participants decreases from the no review (28%) to verified review (21%) before rising again in the fake review group (30%), other ethnic categories exhibit relatively stable or minor fluctuations across the groups. Notably, the multiracial category experiences an increase in the verified review group (8%) compared to the other groups. Overall, there is no substantial difference of subject pool between the treatment sections.

1.4 Hypothesis

1.4.1 Nash Equilibrium Predictions

Using backward induction, we can find the Nash Equilibrium predictions for both players under either types in both experience and credence goods.

For experience goods, the seller in the final node will choose b for high and low type of buyer. Since the buyers know their type, high type buyer will choose Out since the buyer will incur a negative payoff whereas, low type buyer will choose In as the payoff for the low type buyer is indifferent whether seller chooses a or b. So, the subgame perfect equilibria will be high type buyer choosing Out option throughout the game and Low type buyer choosing In and the Seller providing type b goods to both types of players.

In the credence goods case, since the induced value for buyer is 18 if he/she gets the right type of product. If Seller offers prices 15 - ϵ for both types, it will prompt buyers to choose "In", because the outside option payoff is 3. ϵ is a positive perturbation. The nature decides buyer types with a 50% probability of being high type and 50% of being low type. The seller gains 15-10-ϵ = 5 - ϵ if the buyer is high type, where 10 is the cost of providing the high type product. The seller gains 15 - 4 - ϵ = 11 - ϵ if the buyer is low type, where 4 is the cost for low type product. If seller provides the right type of product, the expected profit is 8 - ϵ for seller and 3 + ϵ for buyers. Therefore, despite the type, buyer is better off selecting "In".

1.4.2 Payoff, Trust, Trustworthiness Hypothesis

The paper hypothesizes the following about payoffs, trust, and trustworthiness:

◦ **Ambiguity in Buyer and Seller Payoffs in Verified Review Treatment:** In the verified review treatment, where reviews are authenticated, it's hypothesized that the changes in payoffs for both buyers and sellers might be ambiguous. Since sellers can offset market trust with lower prices, payoffs for buyers and sellers may be higher or lower depending on the prices they post. Furthermore, buyers can be dissuaded in the markets where there is a review system if the overall rating is low signalling bad sellers in the market and hence lower trust.

◦ **Lower Trust in Fake Review Treatment Compared to Verified Review Treatment:** It's anticipated that trust levels would be lower in the fake review treatment compared to the verified review treatment. In the fake review treatment, where reviews lack authenticity, buyers may perceive them as less reliable indicators of product or service quality. This skepticism towards reviews could diminish trust in the sellers and the overall marketplace. Additionally, trust levels might also be influenced by the prevalence of fake reviews in the online ecosystem, further undermining confidence in the platform.

◦ **Lower Trust in Credence Goods Treatment Compared to Experience Goods Treatment:** Trust levels are expected to be lower in the credence goods treatment compared to the experience goods treatment. In the case of credence goods, where quality attributes are difficult for buyers to assess even after purchase, reliance on reviews becomes crucial. Since buyers are uncertain about the true quality of the product or service, they may exhibit lower trust levels overall. Conversely, in the experience goods treatment, where quality can be evaluated through direct experience, buyers may rely less on reviews and thus exhibit higher trust levels.

- **Lower Trustworthiness in Fake Review Treatment:** Trustworthiness, particularly among sellers, is predicted to be lowest in the fake review treatment. With fake reviews, sellers have the opportunity to manipulate ratings to their advantage, potentially deceiving buyers about the true quality of their offerings. This manipulation undermines the credibility of the rating system and erodes trust between buyers and sellers. Consequently, sellers' trustworthiness is expected to be compromised the most in the context of fake reviews, as they engage in deceptive practices to bolster their reputation. These hypotheses outline the expected relationships between ratings, trust levels, and trustworthiness across different treatment conditions, shedding light on the complexities of online market dynamics influenced by reviews and ratings.

1.5 Results

1.5.1 Fake Reviews

Table 3 provides a breakdown of the occurrence of fake reviews categorized by type, indicating whether they were present ("Yes") or absent ("No"), along with the average rating associated with each category. It presents the prevalence of fake reviews among sellers, revealing a notable reliance on such deceptive practices to influence both their own ratings and those of their peers. Sellers were found to utilize fake reviews extensively, with 83 percent of reviews being self-assigned and 71 percent targeting other sellers. As anticipated, sellers tend to assign significantly higher ratings to themselves compared to others, averaging 4.54 out of 5 for self-reviews and 3.48 for reviews of others. Subsequent sections will delve into the implications of these fake reviews on buyer/seller payoffs, trust, trustworthiness, and market efficiency.

Table 3 Incidence of Fake Review

Fake review type	Yes	No	Avg. Rating
Self	530	110	4.54
Other	451	189	3.48

1.5.2 Market Efficiency

Market efficiency in all treatments is low as shown in Table 4. Market efficiency is calculated as sum of realized payoffs for buyer and seller in each round of the game as a percent of total possible payoff aggregated over treatment types. The total surplus gained by buyers and sellers for high type of buyers is 8 (18-10), where 18 is the maximum value gained by buyers and 10 is the cost of providing high type of good to sellers. Similarly, total surplus in low type buyer case is 14 (18-4), where 18 is the maximum value gained by the buyer and 4 is the cost of providing low type good for the seller.

In Table 4, for Credence goods, reputation systems provide value in capturing consumer and producer surplus. However, for experience goods case, it is unclear that having a reputation system provides any value in terms of market efficiency. In both types of goods, Fake review has a higher market efficiency than verified review. The results highlight the importance of having a review system in place. However, it cannot be determined if fake reviews are detrimental to the market as a whole. In further sections, we will look at the effect on buyers and sellers separately.

Table 4 Market Efficiency in Treatments

Treatment type	All cases	Buyer chooses to enter
CreBaseline	27.42%	27.65%
CreFake	29.52%	30.28%
CreVerified	28.45%	29.21%
ExpBaseline	33.08%	35.14%
ExpFake	33.35%	35.71%
ExpVerified	30.20%	30.82%

1.5.3 Buyer and Seller Payoffs

Figure 2 below shows the mean payoff in each round for Experience and Credence goods for the baseline, verified review, and fake review treatments. In the Experience goods case, having review system leads to a higher payoff. However, it is not so clear for Credence goods. It is interesting to see that fake review had higher mean payoff in both types of goods. The result holds even after analyzing the data for only when buyer chooses to enter the transaction with the seller [6].

We need to delve into buyer and seller payoff. Figure 3 below shows the average payoffs in each round of the game for buyers and sellers. Sellers have higher payoffs than buyers. If we compare between treatments, seller payoff increases in verified review treatment and decreases in the fake review treatment. Similarly, buyer payoff decreases in the verified review treatment and increases in the fake review treatment. This suggests that sellers enjoy a higher bargaining power in the verified review sessions.

[6] See Figure A.3 and figure A.2 for payoff graph when Buyer decides to enter the transaction with the Seller.

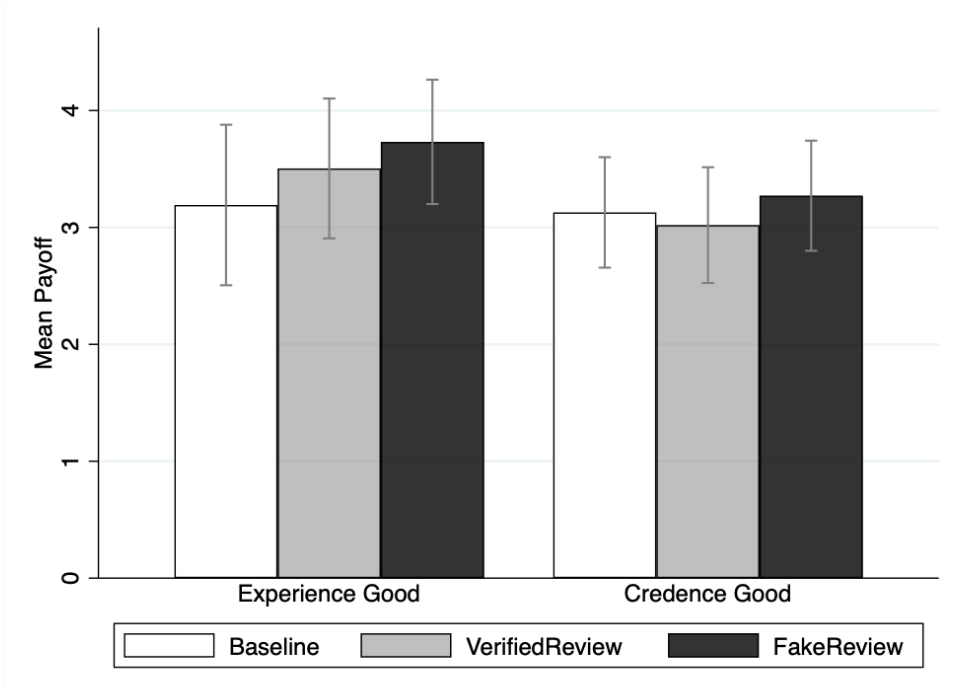


Figure 2 Mean Player Payoffs by Type of Good for Each Round

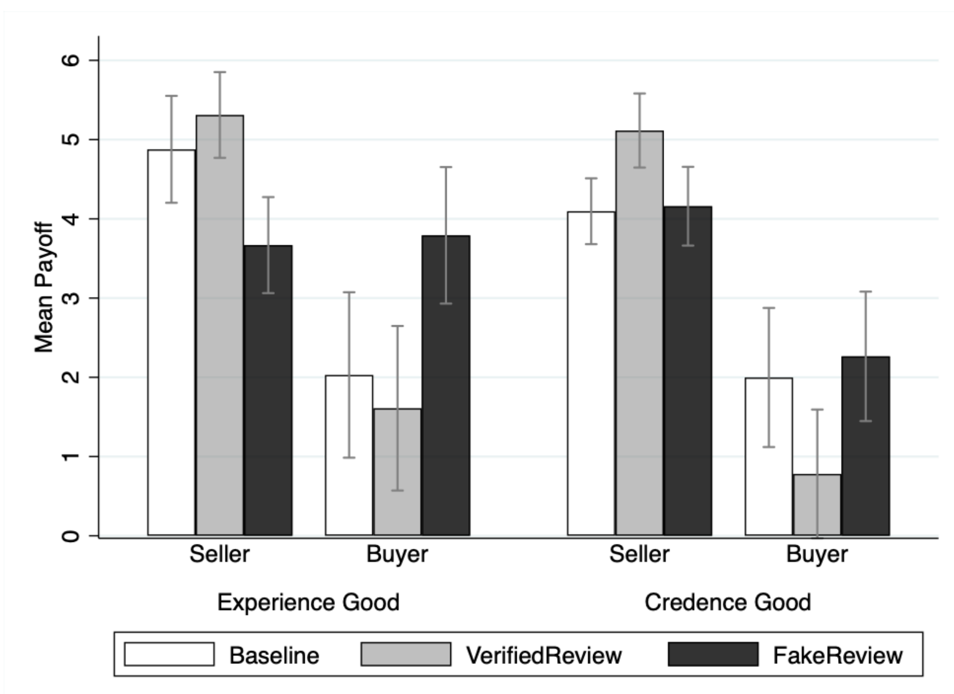


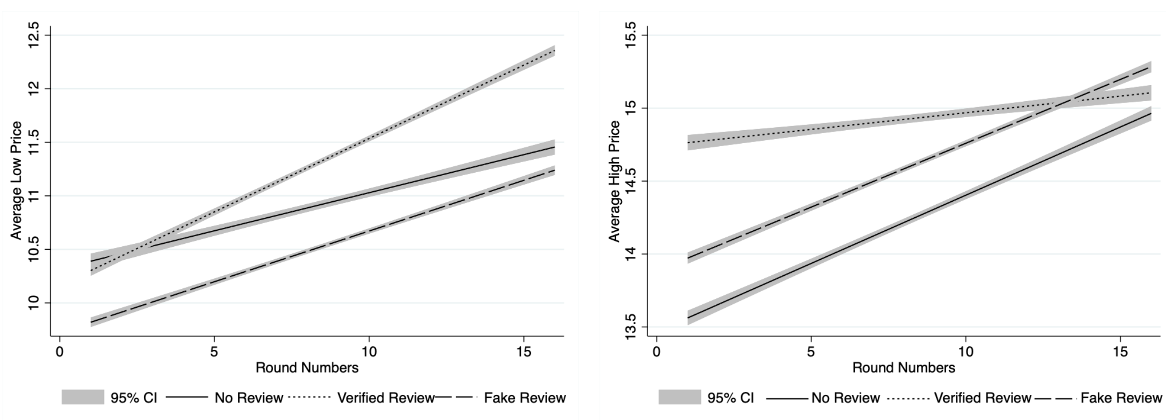
Figure 3 Mean Player Payoffs by Type of Good, Type of Player and Treatments

Sellers posted significantly higher prices in the verified review treatments than in fake review treatments. One possible explanation is that sellers in fake review treatments charge lower prices to compensate for the bad review system where fake reviews are prevalent. In both figures, sellers posted higher prices than in the baseline and fake review treatments. Figure 4 shows average high price posted by the seller over rounds of the game. Figure 4 shows the average low price posted by the seller. In the verified review treatment, buyers paid a higher price than in the baseline and fake review treatments. Sellers appeared to be posting lower prices in the fake review treatment to compensate for the bad review system where fake reviews were prevalent. Figure 5 provides a clear picture of the price paid by buyers in all three review treatments.

Figure 4 Mean Low and High Price Posted by Seller Over Review Treatments

(a) Low Price

(b) High Price



Using a t-test, we found that buyers paid significantly more in the verified review treatment than in the no review treatment. The average price paid was 0.49 ECU higher in the verified review treatment. Similarly, prices paid were 0.86 ECU (6.8%) higher in the verified review treatment than in the fake review treatment. Relative to the baseline, buyers in the fake review treatment paid 0.37 ECU (3.0%) more.

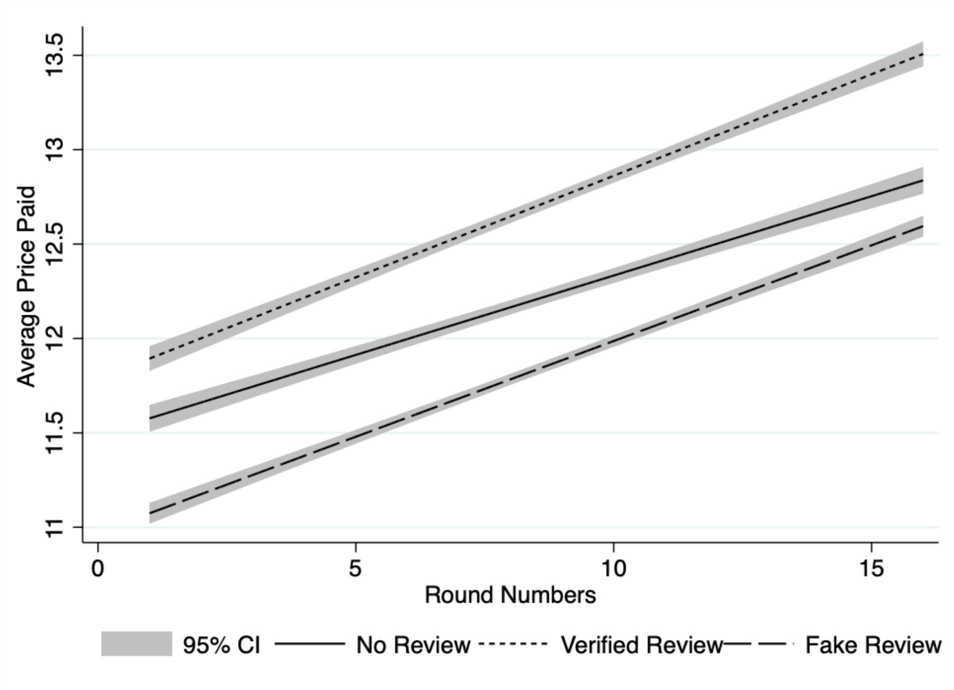


Figure 5 Mean Price Paid by Buyer Over Review Treatments

If a larger number of buyers enter the market with verified reviews, then a higher number of high-type buyers may experience negative payoffs from untrustworthy sellers. This could negatively affect the average buyer payoff. The figure shows that the mean buyer decision to engage in a transaction with the seller is much higher for both types of goods than for fake review treatments. A simple t-test of means shows that buyers chose to enter the market 4.3 percent more in the verified review treatment than in the fake review treatment. There is no statistical difference in means between the baseline and verified review treatment.

Table 5 shows the linear regression results for payoff. Columns (1), (2), and (3) provide payoff results for sellers and columns (4), (5), and (6) show payoff results for buyers. Columns (2), (3), (5), and (6) include treatments with review system only. In these specifications, we control for average rating of seller at the start of the round. Furthermore, columns (3) and

(6) control for High Type Buyer. Standard errors are clustered at session level.

Linear regression results for buyer and seller payoffs shows sellers have significantly higher payoffs and buyers have lower payoffs in verified review treatment. The results also show that fake review negatively impacts sellers, whereas positively impacts buyers in all specifications. The table also shows that rating provides benefits for sellers. An increase of one star rating improves seller payoffs by 0.37 ECU in each round. As hypothesized, payoffs are lower for buyers in credence goods than in experience goods. Since buyers do not know their type, the result is driven by high type buyers choosing not to transact with the seller and hence losing out on higher gains from engaging in the market trade.

Result 1 : There is a trade off between prices and fake reviews. Sellers gain a higher payoff from a market with verified reviews. This result seems to be driven mainly by sellers posting lower prices in the fake review treatment to compensate for the market with fake reviews. Sellers charged 6.8% lower price in fake review than in the verified review treatment.

1.5.4 Trust

The paper measures trust as the number of times buyer chooses to enter the market and transact with the matched seller. As shown in Figure 6 trust is the highest in baseline. This is surprising since the baseline has no reputation mechanism. However, if we look at the average rating for verified review treatment and fake review treatment, ratings are very low and thus contributed to a lower trust by buyers in a market with review systems. If we compare between the two review systems, we find that fake review has a significantly lower trust for experience good. For credence goods, there is not any significant difference and this can be explained by the fact that in credence goods market buyers do not know their type.

Table 5 Linear Regression Results for Buyer/Seller Payoff

	(1)	(2)	(3)	(4)	(5)	(6)
	seller	seller	seller	buyer	buyer	buyer
Credence Good	-0.158	0.195	0.191	-0.630*	-0.749*	-0.556
	(0.201)	(0.248)	(0.248)	(0.349)	(0.434)	(0.380)
Fake Review	-0.444*	-1.210***	-1.207***	0.779*	1.824***	1.737***
	(0.252)	(0.265)	(0.265)	(0.426)	(0.446)	(0.390)
Verified Review	0.760***			-1.072**		
	(0.244)			(0.435)		
Round	-0.0138	-0.00110	-0.00178	0.0465	0.0609	0.0970**
	(0.0213)	(0.0285)	(0.0285)	(0.0369)	(0.0489)	(0.0428)
Rating		0.364***	0.365***		-0.443**	-0.450**
		(0.118)	(0.118)		(0.200)	(0.175)
High Type Buyer			0.135			-6.928***
			(0.243)			(0.365)
N	1872	1177	1177	1872	1177	1177

Standard errors in parentheses

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

Consequently, trust is lower in all kinds of market.

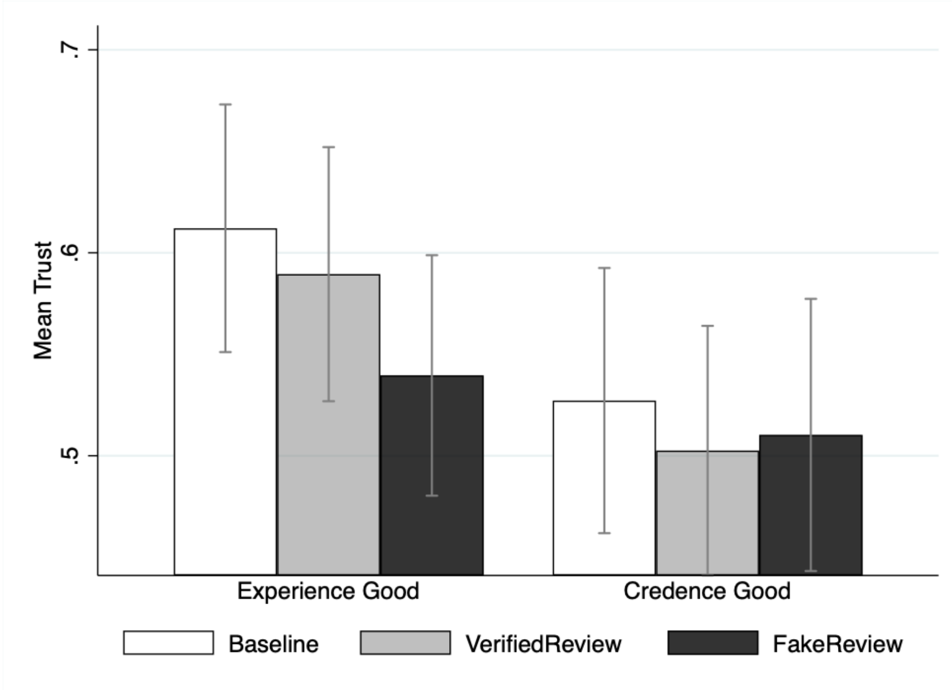


Figure 6 Mean Trust by Buyer

Table 6 shows the logit regression results of trust on treatment types. Column (1) and (3) show results for all treatments, whereas Column (2) and (4) only use treatments with review system. We also control for rating in Column (2) and (4). Fake review significantly decreases trust amongst high type of buyers but not for low type buyer. This is as expected since low type buyers payoffs are not affected by design. It is also more likely that trust decreases with round. Buyers are less trusting of sellers in later rounds as they face negative consequences of bad sellers in the earlier rounds of the game. Rating itself has a positive effect on buyers, implying buyers use the rating of sellers as signaling method to asses their quality. Credence goods treatment increases trust among buyers. One surprising result is that verified review treatment does not have a significant effect on trust.

Table 6 Subsection Logit Regressions for High/Low Type Buyer

	(1) High	(2) High	(3) Low	(4) Low
Credence Goods	0.0597*	0.115***	-0.155***	-0.144***
	(0.0316)	(0.0392)	(0.0303)	(0.0394)
Fake Review	-0.164***	-0.148***	0.0308	0.0405
	(0.0381)	(0.0406)	(0.0385)	(0.0422)
Verified Review	-0.0204		-0.0143	
	(0.0389)		(0.0398)	
Round	-0.0252***	-0.0188***	-0.0116***	-0.00609
	(0.00302)	(0.00425)	(0.00336)	(0.00473)
Rating		0.0535***		0.00896
		(0.0190)		(0.0195)
N	933	590	939	587

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

Result 2: Ratings negatively affect trust in the system if they are low on average. Fake reviews lowers trust further, but only in experience goods markets where buyers know what they want.

1.5.5 Trustworthiness

The paper measures trustworthiness as the number of times seller chooses to provide the right type of good to the buyer, i.e. good type matches buyer type. Table 7 shows the subsection logit regressions for high and low type of buyers. Column (1) and (3) show results for all treatments, whereas Column (2) and (4) only show results for treatments with review system. We also control for rating in Column (2) and (4). Similar to the results on trust above, fake review significantly decreases trustworthiness towards high type of buyers. An interesting result is that trustworthiness increases with credence good market for high type of buyers but decreases for low type of buyers, although statistically not significant. Rating does not have any effect on trustworthiness for seller. Sellers are more trustworthy in later rounds for high type of buyers.

Table 7 Subsection Logit Regressions of Trustworthiness for High/Low Type Buyer

	(1) Low	(2) Low	(3) High	(4) High
Credence goods	-0.0330	-0.0185	0.0611**	0.131***
	(0.0231)	(0.0286)	(0.0240)	(0.0301)
Fake Review	-0.0411	-0.0320	-0.0747**	-0.0612**
	(0.0282)	(0.0307)	(0.0311)	(0.0288)
Verified Review	-0.00954		-0.0210	
	(0.0297)		(0.0266)	
Round	0.00807***	0.00754**	-0.00776***	-0.000885
	(0.00255)	(0.00349)	(0.00248)	(0.00303)
Rating		0.0141		0.0171
		(0.0136)		(0.0133)
N	917	581	933	581

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

1.6 Conclusion

In this paper, we have examined the economic impact of fake reviews on online market outcomes, focusing on buyer and seller behavior as well as trust in the review system. Our study contributes to the literature by providing empirical evidence on how fake reviews influence market dynamics and consumer perceptions.

Our findings reveal several key insights. Firstly, we observe a trade-off between prices and the prevalence of fake reviews. Sellers in markets with fake reviews tend to offer lower prices, potentially as a strategy to mitigate the negative effects of fake reviews on consumer trust and purchasing decisions. This highlights the dynamic nature of pricing strategies in response to changing market conditions.

Secondly, we find that the overall ratings of products or sellers significantly affect trust in the review system. Low average ratings decrease trust, and the presence of fake reviews further diminishes trust, particularly in markets where consumers have a clear understanding of their preferences (experience goods markets). This underscores the importance of maintaining the integrity of online review systems to preserve consumer trust and facilitate efficient market outcomes.

Our study has important implications for both policymakers and market participants. Policymakers need to develop effective strategies to detect and mitigate the proliferation of fake reviews, thereby safeguarding the credibility of online review systems. Market participants, including sellers and platform operators, should prioritize transparency and authenticity to foster consumer trust and enhance market efficiency.

While our study sheds light on the economic consequences of fake reviews, several avenues for future research remain. Further exploration into the mechanisms driving seller behavior in response to fake reviews, as well as the long-term effects of fake reviews on market dynamics, would enrich our understanding of this phenomenon.

In conclusion, our research underscores the need for concerted efforts to combat fake reviews and uphold the integrity of online marketplaces. By addressing the challenges posed by fake reviews, we can promote trust, transparency, and efficiency in e-commerce, ultimately benefiting both consumers and sellers alike.

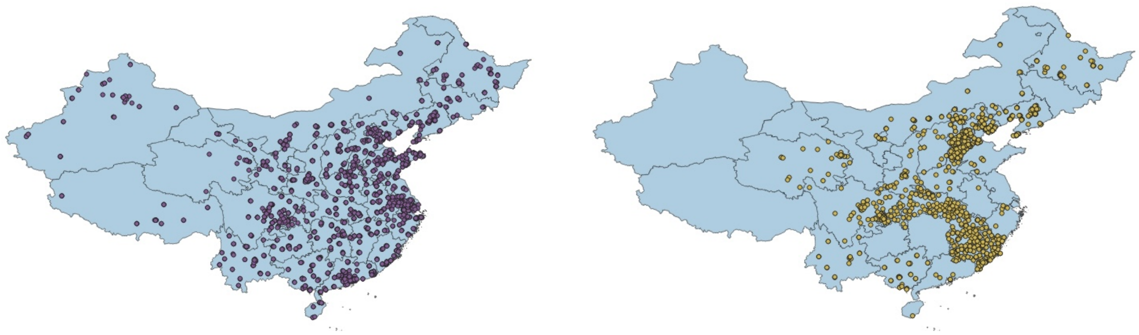
Chapter II Location Choice of Air Quality Monitors in China

2.1 Introduction

According to the World Health Organization, air pollution caused the deaths of 7 million people in 2016, with over 4 million of those deaths being attributed to ambient air pollution. This public health crisis is also responsible for millions of people being diagnosed with respiratory diseases each year. One strategy that countries have implemented to address this issue is through the use of air quality monitoring stations, also known as pollution monitors. These monitors provide valuable information to the public and government agencies to help combat ambient air pollution.

The information collected by the monitors can be used by individuals to alter individual behavior in an effort to reduce exposure to pollution, such as going out during times of lower pollution levels, using air purifiers, spending less time outdoors when pollution levels are high, or purchasing high-quality masks. Barwick et al. (2019) found that increasing public access to air pollution data in China led to an increase in people searching for pollution-related topics online, changing their consumption patterns to avoid pollution exposure, and being willing to pay more for homes in cleaner areas. In addition, individuals can use this information to pressure local and central governments to address the sources of air pollution. Governments can also use this data to guide the public and implement policies to reduce air pollution.

The United States, Western Europe, and East Asia have the highest concentration of air quality monitoring stations. Among the countries in East Asia, China has significantly increased its pollution monitoring efforts. Figure **??** from 2021 demonstrates the scope and



(a) Central Monitors (b) Local Monitors

Figure 7 Air Quality Monitors in 2021

scale of air quality monitors added by China. Since 2013, China has added more than 5000 air quality monitoring stations that provide real-time information on six types of air pollutants (P.M.2.5, P.M.10, sulfur dioxide, nitrogen dioxide, ozone, and carbon monoxide). These stations are managed by state, provincial, municipal, and county level governments. In this paper, we call the monitors placed by the state central (controlled) monitors. The monitors placed and controlled by provincial, municipal, and county governments are called local (controlled) monitors. Table 8 in section 2.3.2 shows the changes in the number of central and local air quality monitoring stations from 2013 to 2021. While the central government initially added a large number of monitors, local governments have more recently been responsible for the majority of new monitor placements. However, there is a lack of research on the specific factors that influence the government's decision on where to locate these monitors.

There are several reasons why governments might choose to place air quality monitors in specific locations. One reason is to place monitors in urban area aiming to provide air quality information for most people. The second is to provide accurate information to protect vulnerable population such as children and hospital patients by placing monitors near schools and hospitals. Additionally, governments may place monitors near sources of pollution, such as road traffic or polluting industries, to accurately assess the impacts of these sources on air quality. Local governments may also have incentives to strategically locate monitors to avoid punishment from higher levels of government for poor air quality[7], or to respond to public pressure to improve air quality.[8]. Previous research has shown that local governments may have incentives to misreport air quality data by placing monitors in cleaner areas, to appear to meet air quality standards or promote economic growth. Grainger et al. (2016) find that counties that are marginal to the non-attainment threshold for National Ambient Air Quality Standards (NAAQS) placed newly sited monitors in cleaner areas relative to counties non-marginal in the US. Furthermore, local governments can encourage production in polluting firms to collect tax revenue and promote economic growth (Qi & Zhang, 2014).

In this paper, we aim to identify the factors that influence the government's decision on where to locate air quality monitors, with a focus on the case of China. This project will be one of the first to examine the behavior of governments in placing pollution monitors in developing countries. It is important to study these decision-making processes because the location of the monitors can significantly impact the pollution readings they report.

2.2 Literature Review

There is a lack of research on the factors that influence the location of air quality monitors, with most existing studies coming from fields outside of economics, such as geology. Yu et al. (2018) found that air quality monitoring stations in the Beijing-Tianjin-Hebei region

[7] Previous research on environmental policy (Kahn et al., 2015) in China suggests that central government often uses policy achievement as a tool to evaluate local government officials.

[8] See Xu et al. (2019) to learn about the progress of environmental activism in China and how Chinese NGOs involved in the air quality measurement activities use social media and other communication methods to fulfill their organizational objectives and connect fragmented supportive forces.

of China were concentrated in areas with high levels of pollution. Muller and Ruud (2018) examined the factors that influenced the addition or removal of air quality monitors in the Netherlands and found that peak ozone readings in the previous period significantly affected these decisions. Yang et al. (2020) studied the roll out of central monitors in China and found that they had a significant impact on local air quality near the monitors, but not in other areas.

Other research has focused on the optimization of monitoring networks, considering factors such as population density, land use, and city size (Haas, 1992; Lu et al., 2011; Maji et al., 2017; Pires et al., 2008). However, these studies are not comprehensive and do not fully address the economic, demographic, and infrastructure factors that may influence the location of air quality monitors.

Our paper aims to fill this gap in the literature by examining the location choice of air quality monitors in China. We consider both current monitors and those that have been added in recent years, and consider a range of economic, demographic, and infrastructure factors in order to provide a comprehensive understanding of the air quality monitoring system in China. We also use machine learning methods to supplement our analysis in answering the specific questions of the paper.

2.3 Data

2.3.1 Monitor location

We obtained data on the locations of air quality monitoring stations in China from the websites of individual provinces and the central government, where this data is made available to the public. The data was provided by Shanghaiqingyue (http://data.epmap.org/), an organization that promotes the transparency of environmental data and supports scientific research. The data includes 1,998 central government-controlled monitors and 3,486 local government-controlled monitors from 2013 to 2021. Due to limited data availability, the local controlled monitor data only covers 25 provinces out of 31 provinces in mainland China. Figure **??** show the distribution of central and local monitors in China in 2021. Figure B.1 illustrates the 25 provinces we have local monitor location. Table 8 displays the changes in the number of central and local monitors from 2013 to 2021. While the initial increase in the number of monitors was primarily driven by the central government, in recent years there has been a significant increase in the number of local monitors being added to the monitoring network. This growth highlights the importance of local monitors in providing air quality information to the public.

2.3.2 Pollution Data

We use satellite pollution data from Xu et al. (2019) to measure P.M.2.5 levels. This estimate is based on an empirical model that includes only satellite-derived Aerosol Optical Depth (AOD)[9] measurements at a resolution of $0.05° \times 0.05°$[10]. The raw AOD data comes from NASA's Moderate Resolution Imaging Spectroradiometer(MODIS). Figure 8 displays the levels of P.M. 2.5 in 2018 across China. Except for the desert region in Xinjiang autonomous region, the northeast region of the country had the highest levels of pollution, particularly in the Hebei, Henan, Shandong, Beijing, and Tianjin provinces.

[9] Aerosol optical depth is a measure of the extinction of the solar beam by dust and haze. In other words, particles in the atmosphere (dust, smoke, pollution) can block sunlight by absorbing or by scattering light. AOD tells us how much direct sunlight is prevented from reaching the ground by these aerosol particles. It is a dimensionless number that is related to the amount of aerosol in the vertical column of atmosphere over the observation location.

[10] The $0.05° \times 0.05°$ resolution can be understood roughly as a 5km-by-5km spatial grid.

Table 8 Entry and Exit of Air Quality Monitors from 2013 to 2021

	Central Monitors		Local Monitors	
Year	Entry	Exit	Entry	Exit
2013	829	0	0	0
2014	225	1	26	0
2015	463	27	456	0
2016	37	41	350	14
2017	36	35	586	19
2018	24	24	183	52
2019	13	15	453	31
2020	39	87	1,129	525
2021	332	-*	303	-*
Total	1,998	230	3,486	641

Note: * When we collected the data in 2021, there was no monitor exit. But we cannot claim there is no exit for the whole year.

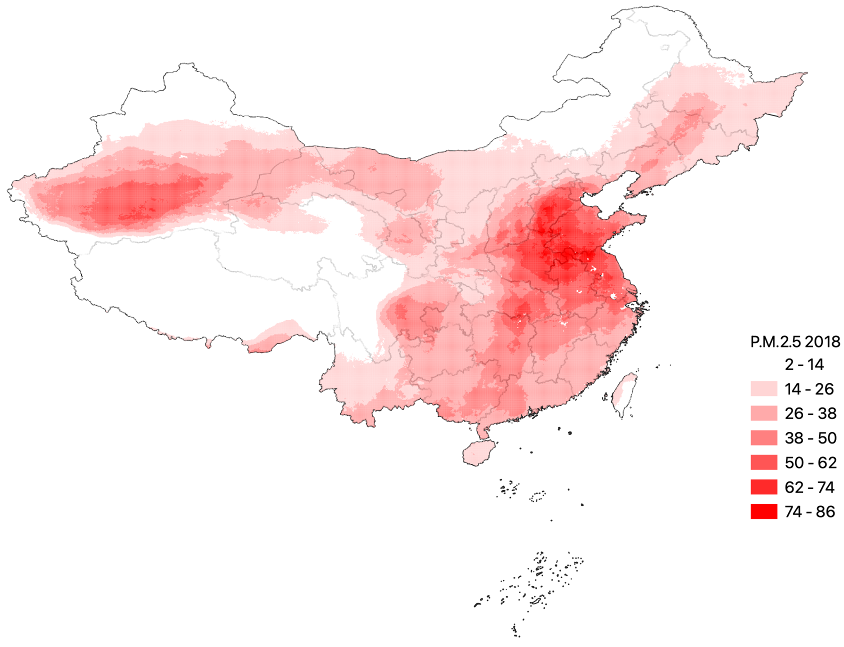


Figure 8 Satellite-derived P.M. 2.5

2.3.3 Other Variables

For our control variables, we obtained data on gross domestic product (GDP), GDP for primary industries, government revenue, the number of large companies, and the number of high school students at the county level from China Statistical Yearbooks. We also collected data on the locations of schools and hospitals from the Gaode map, a Chinese version of Google maps, and data on the length of primary highways from OpenStreetMaps (OpenStreetMap contributors, 2017). Population density data is from census estimates at a 1km resolution in 2010, 2015, and 2020, with estimates for other years obtained through linear interpolation.

Summary statistics for 2021 are presented in Tables 9 and **??**. Table 9 provides the summary statistics for central monitor analysis, and Table ?? provides summary statistics for local monitor analysis. The summary statistics show that local monitors are placed in polluted areas more than central monitors on average. One explanation for this discrepancy can be that the authorities clean up areas around the monitors once the monitor is placed there. Since most of the central monitors were placed in earlier years, from 2013 to 2015, areas around the monitors are already cleaned, hence the lower average PM 2.5. A recent study (Yang et al., 2020) found that local governments in China targeted pollution reduction efforts in areas closer to the monitors after their installation.

Another possible explanation for the difference in pollution levels between central and local monitors is that central monitors are installed in areas with fewer people and less economic activity, leading to lower pollution levels. The summary statistics support this theory, as local monitors are typically located in grids with higher population density and greater economic activity. This pattern may also be partly explained since we only have local monitor data for 25 provinces in China, most of which are highly populated.

2.4 Methodology

To conduct our study, we first create a grid of approximately 5km x 5km for all of China. As shown in Figure 9 of Chongqing City, some of these grids contain pollution monitors while others do not. The central monitors are more clustered than the local monitors. We then perform two sets of analyses. The first is a cross-sectional analysis of the current pollution monitors (in 2021). The second examines the factors that influence the location choice for air quality monitors, with separate analyses for central and local monitors.

Table 9 Summary Statistics for Central Monitor Data in 2021

Variable	N	Mean	S.D.
Central Monitor Dummy	385,567	0.004	0.062
Average P.M. 2.5	385,567	29.41	21.02
Length of Highway	385,567	964.04	3,348.48
Population Density	385,567	145.15	726.20
Number of Schools	385,567	0.54	3.11
Number of Hospitals	385,567	0.15	1.49
Distance to the Nearest Central Monitor	385,567	1.31	1.23
Distance to the Nearest Local Monitor	385,567	3.61	4.77
GDP	381,378	43,250.39	8.28E+05
GDP of Primary Industry	381,355	57,191.49	1.63E+06
Government Budget Revenue	381,378	2,462.84	1.24E+05
Number of Large Companies	338,090	1.46	19.72
Number of High School Students	381,378	233.68	2564.66

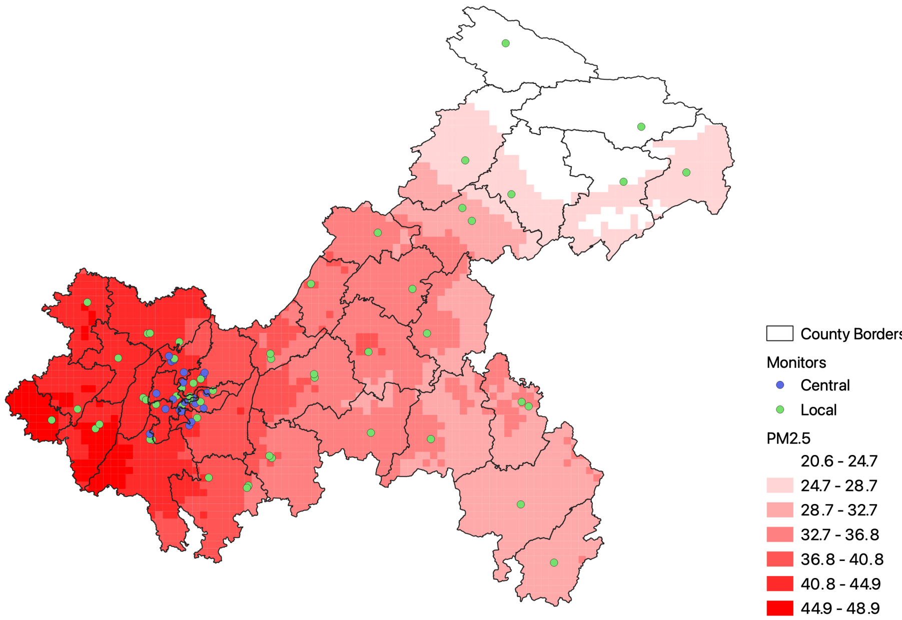


Figure 9 Chongqing City PM2.5 (2018) and Monitor Location (2021)

2.4.1 Where are the Monitors Located?

To study where the monitors are currently located, we estimate a cross-sectional logit model for both central and local monitors. Meanwhile, to test whether we have found the determining factor for the location choice of the monitors, we predict the location choice of the central and local monitors using machine learning techniques using logit regression.

Correlation Tests. To determine the main factors in whether a grid has at least one monitor, we conducted correlation analysis of the outcome variables, central monitor dummy, and local monitor dummy, with all the variables we collected. We differentiate between local and central monitors as we have limited data for the local monitors.

Regression Analysis. We use logit regression analysis to predict the determining factors to the monitor location. The outcome in the models is a binary variable indicating whether there is a pollution monitor within a particular grid (1) or not (0).

$$Pr(Monitor_i = 1|X_i) = \phi(\alpha + \beta_1 PM2.5_i + \beta_2 Controls_i + u_i) \tag{1}$$

The explanatory variable $Pr(Monitor_i = 1|X_i)$ is the probability of whether grid i has at least one monitor. $\phi()$ indicates the functional form of the logistic transformation of the linear estimate. $PM2.5_i$ is the average P.M. 2.5 at grid i from 2011 to 2018. $Controls_i$ are the control variables for grid i such as population density, length of highways within the grid, number of schools and hospitals, distance to the nearest central monitor, distance to the closest local monitor, GDP, GDP of primary industries, government revenue, number of large companies, and number of high school students. Finally, u_i is the error term. We cluster standard errors at the county level for all analysis and control for provincial-level fixed effects in some analysis.

Machine Learning Predictive Analysis. The identification strategy and functional form of the machine learning model is the same as the regression analysis. Since only 0.4% of the grids have central monitors and 1.0% of the grids have local monitors, the positive class is relatively rare to the negative class. Therefore, in the machine learning model prediction, we adjust the weights to balance the two classes before fitting the regression models to gain a fair prediction. We use a 80/20 train/test split of our data to conduct the analysis.

To provide a ranking of different factors, we generate a correlation graph for all relevant variables. We start the prediction with the same variables as the regression analysis and then slowly reduce the factors to the most relevant. In the end, we can find the most important factors in predicting the location choice of the monitors.

To show the accuracy of monitor location predictions, we introduce various evaluation metrics such as the true positive rate (TPR), true negative rate (TNR), and balanced accuracy (BA). TPR focuses on the success rate of predicting the positive class, i.e. the grid with at least one air quality monitor. TNR focuses on the success rate of predicting the negative class. BA is the mean of TPR and TNR, which indicates the accuracy of the model in general.

2.4.2 What Influences the Location Choice?

In this section, we study the factors that influence the location choice of monitors, we estimate a panel logit model for both central and local monitors. Similar to the cross-sectional analysis to find the determining factor for monitor locations, we also conduct regression analysis with the entire dataset as well as machine learning predictive analysis. In contrast from the cross-sectional machine learning prediction in 2.4.1, we predict whether the grid has a new monitor in a given year.

Panel Regression Analysis. We use the panel regression logit model below to find the leading factors in placing a new monitor.

$$Pr(New\ Monitor_{i,t} = 1|X_{i,t-1}) = \phi(\beta_0 + \beta_1 P.M.2.5_{i,t-1} + \beta_2 Controls_{i,t-1} + u_i + \sigma_t + \epsilon_{i,t}) \quad (2)$$

In our model, the explanatory variable $Pr(NewMonitor_{i,t} = 1|X_{i,t-1})$ is the probability of whether grid i at time t-1 has at least one monitor. $\phi()$ indicates the function form of the logistic transformation of the linear estimate. $P.M.2.5_{i,t-1}$ is the P.M. 2.5 level at grid i and time t-1. $Controls_{i,t-1}$ includes control variables for grid i at time t-1 such as population density, length of highways within the grid, number of schools and hospitals, distance to the nearest central monitor, distance to the closest local monitor, GDP, GDP of primary industries, government revenue, number of large companies, and number of high school students. We control for time-invariant location-specific effects through u_i and time-specific effects through σ_t. Finally, $\epsilon_{i,t}$ is the error term. We cluster standard errors at the county level for all analyses and control for provincial-level fixed effects in some analyses. The coefficient β_1 is of particular interest as it tells us how likely the government is to place a monitor in a more polluted grid.

Machine Learning Predictive Analysis. We use the same machine learning model in 2.4.1 to predict the location choice of the pollution monitors. The only difference in this section is we use a panel data instead of a cross-section to conduct the analysis. We then discuss the features that are most relevant in giving us the best prediction results as measured by balanced accuracy (BA).

2.5 Results

We present two sets of main results. The first is on the cross-section analysis of where the monitors are located in 2021 and then on the location choice of monitors placed over time. For each section, we start with the correlation analysis, followed by the logit regression results and finally, machine learning model results.

2.5.1 Where are the Monitors Located?

Correlation and analysis tests. Figure 10 and 11 shows the correlation between variables in our cross-sectional data from 2021. The figures highlight the fact that monitors are around densely populated areas. Central and Local Monitors are highly correlated with demographic variables such as Number of Hospitals, and Length of Highway. However, they are less correlated with economic indicators such as GDP of Primary Industries, Number of large companies, and Government Revenue. The results are as expected since the primary aim of the monitors was to provide pollution information to the public rather than measure pollution around industrial economic activities.

Comparing the P.M.2.5 correlation numbers between central and local monitors, we observe that correlation for Local Monitor is higher than Central Monitor. This indicates Local Monitors are located in grids with higher pollution than Central Monitors. A simple explanation of the difference between Central and Local Monitors is that central monitors are being placed in cleaner areas relative to local monitors. However, as we discuss in subsequent sections, it can also result from various other reasons such as strategic behavior from local governments in cleaning up areas around central monitors. A key incentive feature associated with pollution reduction effort in China is that local government performance is measured by Central Monitors. As a result, officials have greater incentive in reducing pollution around Central Monitors.

Regression Analysis Results. In Table 10, columns (1) and (2) show the marginal effects from logit regressions for central monitors only, while Columns (3) and (4) show the marginal effects from logit regressions for local monitors only. Columns (1) and (3) do not

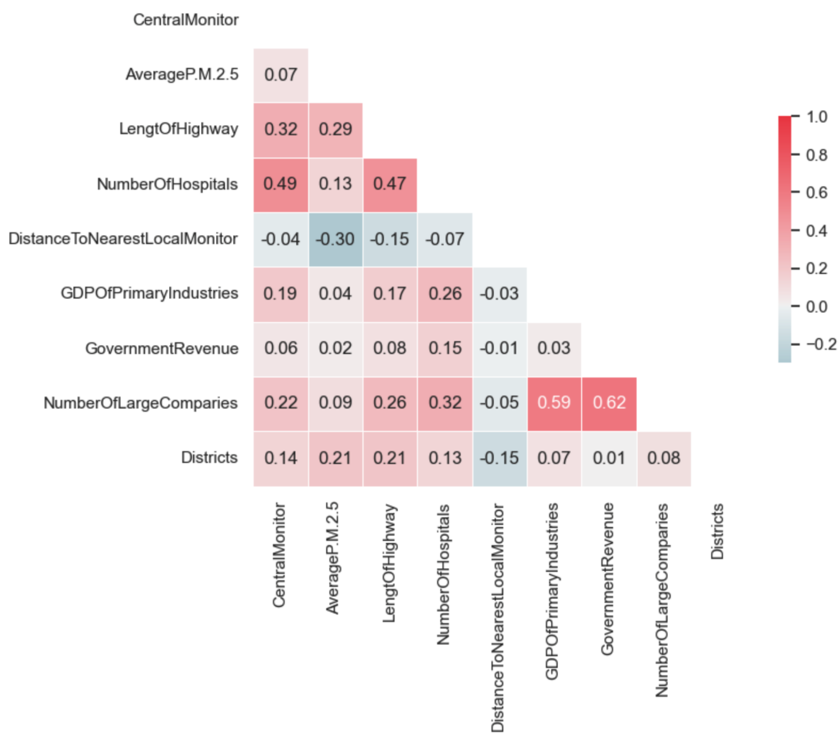


Figure 10 Determining Factors Correlation Graph in Central Monitor Location Choices

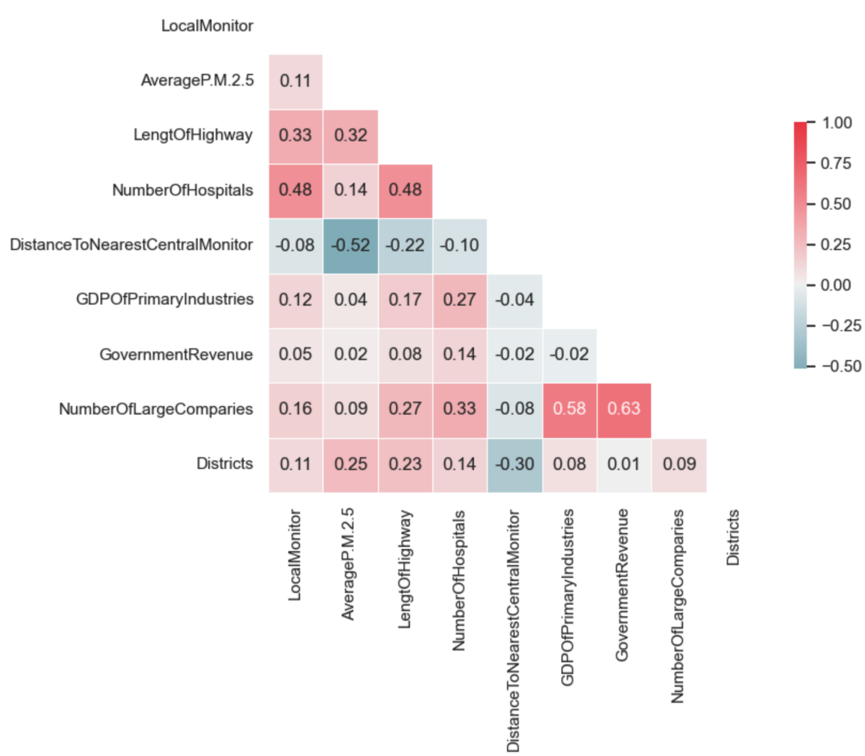


Figure 11 Determining Factors Correlation Graph in Local Monitor Location Choices

include province fixed effects, and columns (2) and (4) include province fixed effects.

The marginal effect results from the logit models for monitor location in 2021 provide detailed insights into the factors affecting the location of central and local monitors. Rregression results suggest that Local and Central Monitors are positively related to P.M.2.5 and statistically significant. For average P.M.2.5 levels, the effect on central monitors is 0.0004, while for local monitors, it is 0.0020, indicating a stronger impact on local monitors. A major difference between correlation analysis and regression analysis is the coefficient for Length of Highway and Number of Hospitals. The length of highways shows a significant positive effect for central monitors at 0.0007 and for local monitors at 0.0020, indicating a strong effect of local monitors. The number of hospitals also has a significant positive effect, with values of 0.0009 for central and 0.0035 for local monitors. This difference shows that Local Monitors are likely placed closer to highway and Hospitals relative to the Central Monitors.

Economic factors like GDP of primary industries and government revenue show mixed effects, while the number of large companies have positive impacts for central and mixed impacts for local monitors. Factors such as GDP of primary industries show a positive effect for central monitors but a negative effect for local monitors. Government revenue has a negative effect for both types of monitors. Number of large companies shows a positive effect on central monitors but a negative effect on local monitors.[11]

Machine learning predictive results. We present results from the machine learning predictive model in Table 11 and 12. Our model achieved a high balanced accuracy of 93.43 percent for Central Monitors and 91.77 percent for Local Monitors including all fea-

[11]The appendix includes Table B.3, which shows the linear probability results for central and local monitor models. It indicates that central monitors tend to be in cleaner grids than local monitors in 2021. The linear regressions are all controlled for provincial fixed effects.

Table 10 Marginal Effects for Monitor Location in 2021 Using Logit Models

	(1)	(2)	(3)	(4)
	Central Monitor		Local Monitor	
Average P.M. 2.5	0.0002**	0.0004**	0.0020***	0.0012***
	(9.6e-05)	(1.8e-04)	(2.1e-04)	(2.8e-4)
Length of Highway	0.0007***	0.0007***	0.0019***	0.0020***
	(3.5e-05)	(3.6e-05)	(9.0e-05)	(9.3e-05)
Number of Hospitals	0.0009***	0.0009***	0.0033***	0.0035***
	(3.1e-05)	(3.2e-05)	(9.4e-05)	(9.8e-05)
Distance to the Nearest Central Monitor			-0.0107***	-0.0101***
			(7.6e-04)	(7.9e-4)
Distance to the Nearest Local Monitor	-0.0014***	-0.0010**		
	(2.4e-04)	(4.7e-04)		
GDP of Primary Industries	0.0001***	0.0001***	-0.0002	-0.0003***
	(3.7e-05)	(3.7e-05)	(1.3e-4)	(1.1e-4)
Government Revenue	-9.58e-05**	-8.49e-05**	-0.0002	-0.0003**
	(3.9e-05)	(3.5e-05)	(1.4e-04)	(1.3e-4)
Number of Large Companies	9.535e-05**	8.69e-05**	-0.0004***	-0.0002*
	(4.2e-05)	(4.2e-05)	(1.1e-04)	(1.2e-04)
District Dummy	0.0022***	0.0022***	0.0003*	0.0004**
	(8.8e-05)	(9.2e-05)	(1.6e-04)	(1.7e-04)
Provincial Dummies	No	Yes	No	Yes
N	338,067	338,067	212,123	212,123

Notes: 1. All explanatory variables are standardized.
2. Standard errors in parentheses.
3. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

tures[12].

We want to find out the most important feature among all included. Hence, we ran the model with different combinations of our features. Out of all the combinations included

[12]Balance Accuracy is the arithmetic average between TPR and TNR.

in the model, Number of Hospitals is the most important feature in predicting the location of the air quality monitor for both Central and Local Monitors. We were able to get a balanced accuracy of 88.28 for Central and 88.52 for Local monitors using only the Number of Hospital feature, only a 5 percent decrease when using all available features. This is a strong indication that governments are placing air quality monitors closer to hospitals.

Table 11 Logit Model Prediction Accuracy for Central Monitor Location Choice (2021)

	(1)	(2)	(3)	(4)	(5)
Predictors	All	All but Provincial Dummies	P.M. 2.5, Number of Hospitals, and District Dummy	Number of Hospitals and District Dummy	Number of Hospitals
TPR	0.9113	0.9048	0.9144	0.9306	0.7823
	(0.0161)	(0.0154)	(0.0160)	(0.0134)	(0.0209)
TNR	0.9573	0.9608	0.9286	0.8930	0.9832
	(0.0013)	(0.0008)	(0.0025)	(0.0011)	(0.0004)
BA	0.9343	0.9328	0.9215	0.9119	0.8828
	(0.0078)	(0.0075)	(0.0078)	(0.0067)	(0.0105)
Train Size	270,453	270,453	308,453	308,453	308,453
Test Size	67,614	67,614	77,114	77,114	77,114

Notes: 1. All predictors include Average P.M.2.5, Length of Highway, Number of Hospitals, Distance to the Nearest Local Monitor, GDP of Primary Industries, Government Revenue, Number of Large Companies, District Dummy, and Provincial Fixed Effects.
2. Standard errors in parentheses.

2.5.2 What Influences the Location Choice?

Panel Regression Analysis Results. Table 13 presents the results of our panel data analysis on the entry of central or local monitors in grid cells. The analysis is conducted separately for central and local monitors, with the results appearing in columns (1) and (2) and columns (3) and (4), respectively. We included lagged explanatory variables in the analysis, as we assumed that the government would consider past air pollution levels and

Table 12 Logit Model Prediction Accuracy for Local Monitor Location Choice (2021)

	(1)	(2)	(3)	(4)	(5)
Predictors	All	All but Provincial Dummies	P.M. 2.5, Number of Hospitals, and District Dummy	Number of Hospitals and District	Number of Hospitals
TPR	0.8879	0.8729	0.8596	0.8325	0.7922
	(0.0148)	(0.0150)	(0.0159)	(0.0158)	(0.0170)
TNR	0.9475	0.9516	0.9427	0.9686	0.9784
	(0.0016)	(0.0013)	(0.0020)	(0.0007)	(0.0006)
BA	0.9177	0.9123	0.9012	0.9006	0.8852
	(0.0071)	(0.0072)	(0.0075)	(0.0079)	(0.0085)
Train Size	169,698	169,698	182,855	182,855	182,855
Test Size	42,425	42,425	45,714	45,714	45,714

Notes: 1.All predictors include Average P.M.2.5, Length of Highway, Number of Hospitals, Distance to the Nearest Central Monitor, GDP of Primary Industries, Government Revenue, Number of Large Companies, District Dummy, and Provincial Fixed Effects.

2. Standard errors in parentheses.

other factors when deciding where to install a monitor.

The results suggest that central monitors are typically installed in areas with higher levels of pollution, while local monitors are installed in cleaner areas. Central and Local governments have different priorities when installing an air quality monitor. Since the entire air quality monitoring program was brought about by the central government to reduce pollution, it explains why the monitors are placed in areas with higher PM 2.5. Furthermore, the priority to clean up the most polluted areas are more important to the central government due to the centralized government structure of China.

On the contrary, local monitors are placed in cleaner areas. This could be due to the crowding out effect from the placement of central monitors. Local governments had to place the new monitors in places where there was no pollution monitoring, away from the polluted areas where central monitors were already present. In addition, we find that central monitors are installed near each other and local monitors are installed near each other, as indicated by the positive marginal effects with distance to the nearest central monitor and nearest local monitor. The result suggests that local governments are covering air quality monitoring for areas where central government does not have presence.

Table 13 Panel Logit Marginal Effects for Location Choice of Monitor Entries

	(1)	(2)
	Central Monitor	Local Monitor
L.P.M. 2.5	0.0064***	-0.0317
	(0.0020)	(0.0398)
L.Length of Highway	-1.5e-4	0.0101**
	(1.5e-4)	(0.0039)
L.Population Density	0.0031	-0.0075
	(0.0021)	(0.0490)
L.Distance to the Nearest Central Monitor	0.1073***	-0.0741
	(0.0053)	(0.0719)
L.Distance to the Nearest Local Monitor	0.0011	0.0758***
	(0.0014)	(0.0161)
L.GDP of Primary Industries	1.1e-4	0.0166
	(2.7e-4)	(0.0124)
L.Government Revenue	9.2e-4***	0.0119
	(3.2e-4)	(0.0073)
N	2,355	7,835

Notes: 1. All explanatory variables are standardized.
2. All four regressions control for the individual grid and year fixed effects.
3. Standard errors in parentheses.
4. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Machine Learning Predictive Analysis Results. In this section we present the machine learning model results for new monitor placement using logit model. Table 14 and 15 show the results of the model. As in the cross-section analysis we are able to achieve a high level of balance accuracy for both central and local monitors. We achieve a 93.09 percent balanced accuracy for central monitors and 88.04 percent for local monitors. For both type of monitors, population density is the primary predictor for placing an air quality monitor. The higher precision for central monitor mainly comes from population density feature. As discussed in previous sections, central governments tend to place monitors in population dense areas. Furthermore, central monitors were placed ahead of local monitors, resulting in densely populated areas already covered by central monitors.

Table 14 Logit Model Prediction Accuracy for Central Monitor Entry Location Choice

	(1)	(2)	(3)	(4)	(5)
Predictors	All	All but Year Dummies	P.M. 2.5, Population Density, and District and County Dummies	Population Density and District and County Dummies	Population Density
TPR	0.9165	0.8827	0.9031	0.9101	0.8240
	(0.0268)	(0.0311)	(0.0171)	(0.0148)	(0.0196)
TNR	0.9453	0.9290	0.9349	0.9285	0.9612
	(0.0013)	(0.0027)	(0.0019)	(0.0032)	(6.4e-4)
BA	0.9309	0.9059	0.9190	0.9193	0.8926
	(0.0130)	(0.0149)	(0.0082)	(0.0071)	(0.0096)
Train Size	1,527,668	1,527,668	2,159,175	2,776,082	2,776,082
Test Size	381,917	381,917	539,794	694,021	694,021

Notes: 1. All predictors include L.P.M.2.5, L.Length of Highway, L.Population Density, L.Distance to the Nearest Central Monitor, L.Distance to the Nearest Local Monitor, L.GDP of Primary Industries, L.Government Revenue, and Yearly Dummies.

2. Standard errors in parentheses.

Table 15 Logit Model Prediction Accuracy for Local Monitor Entry Location Choice

	(1)	(2)	(3)	(4)	(5)
Predictors	All	All but Year Dummies	P.M. 2.5, Population Density, and District and County Dummies	Population Density and District and County Dummies	Population Density
TPR	0.8566	0.8432	0.8043	0.7859	0.7736
	(0.0184)	(0.0190)	(0.0171)	(0.0165)	(0.0172)
TNR	0.9043	0.9046	0.9049	0.9258	0.9302
	(0.0016)	(0.0015)	(0.0017)	(9.2e-4)	(7.9e-4)
BA	0.8804	0.8739	0.8546	0.8558	0.8519
	(0.0088)	(0.0091)	(0.0094)	(0.0080)	(0.0084)
Train Size	904,748	904,748	1,097,131	1,462,841	1,462,841
Test Size	226,187	226,187	274,283	365,711	365,711

Notes: 1. All predictors include L.P.M.2.5, L.Length of Highway, L.Population Density, L.Distance to the Nearest Central Monitor, L.Distance to the Nearest Local Monitor, L.GDP of Primary Industries, L.Government Revenue, and Yearly Dummies.

2. Standard errors in parentheses.

2.5.3 Discussions

Our analysis of cross-sectional and panel data has provided insight into the relationship between pollution and the placement of air quality monitors. Pollution level is crucial in determining where a monitor is installed. Initially, central monitors were placed in areas with high levels of pollution. In contrast, the local monitors were placed in cleaner areas.

We find interesting results regarding pollution and the placement of air quality monitors. While initially, central monitors were placed in polluted areas, these areas seem to have been cleaned up over time. However, it is difficult to know whether the reduction is actual reduction due to abatement technologies/change in behavior or just a shift of pollution from monitored areas to unmonitored areas. For local monitors, we find the opposite result. We find that although monitors were placed in cleaner areas, pollution in the grids were much higher in 2021, the last year for our dataset. Since the central government only uses central monitors for environmental performance evaluation, it will be interesting to see if local governments also take steps to reduce pollution in areas near local monitors.

We also find that factors such as number of hospitals, length of highway, economic indicators such as GDP of primary industries, and proximity to other air quality monitors, are crucial to identify where the air quality monitors are located. However, our results show that those factors are less relevant when determining the entry of a new air quality monitor.

Another critical area of interest is if local governments move local monitors to show a more favorable pollution level in the province. According to table 8, we see many local pollution monitors entering and exiting the grid. It is very well possible that these monitors are relocated from the current location to a favorable location nearby. Due to the limitation of our data, we cannot observe this behavior and hence cannot account for this behavior.

Chapter III Monitoring and Compliance of the Toxics Release Inventory: Does EPA Enforcement Action Change Corporate Reporting Behavior?

3.1 Introduction

The Union Carbide disaster in Bhopal, India, in 1984, and the incident in West Virginia in 1985 served as stark reminders of the potential catastrophic consequences of chemical emergencies. In Bhopal, a gas leak from a pesticide plant owned by Union Carbide resulted in thousands of deaths and injuries, while the West Virginia disaster involved a chemical leak from a storage tank, threatening local communities and ecosystems. These incidents underscored the need for better preparedness for chemical emergencies and increased transparency regarding the presence and handling of hazardous substances.

In response to these concerns, the U.S. Congress passed the Emergency Planning and Right to Know Act (EPCRA) in 1986. EPCRA aimed to address the lack of information available to the public about potentially harmful substances being used and stored in their communities. One of the key provisions of EPCRA was the creation of the Toxics Release Inventory (TRI). The TRI mandates that certain industries report their use, storage, and release of toxic chemicals to the environment.

The TRI serves as a valuable tool for both government agencies and the public. By requiring firms to disclose information about their use of toxic chemicals, the TRI enables state and local governments to better assess potential risks to public health and the environment. Armed with this information, authorities can develop emergency response plans and take proactive measures to protect communities from chemical hazards.

Moreover, the public availability of TRI data empowers citizens to hold companies accountable for their environmental practices. Transparency regarding toxic chemical use creates social pressure on firms to adopt more responsible manufacturing processes and reduce pollution. By incentivizing companies to minimize their environmental impact, the TRI contributes to the overall goal of promoting sustainability and safeguarding public health.

The implementation of the EPCRA and the creation of the Toxics Release Inventory (TRI) represent significant steps toward enhancing chemical safety and environmental protection. By requiring companies to disclose information about their use of toxic chemicals, the TRI enables better preparedness for emergencies, fosters community awareness, and encourages pollution reduction efforts through social pressure and accountability.

The Toxics Release Inventory (TRI) program, while instrumental in promoting transparency regarding the use and release of toxic chemicals, relies heavily on self-reporting by companies. As such, ensuring the accuracy and reliability of the data requires robust monitoring and enforcement mechanisms. Without adequate oversight by the Environmental Protection Agency (EPA), compliance with TRI reporting requirements becomes non-binding, potentially undermining the program's effectiveness.

Given the significant reliance on TRI data for various economic and public health analyses, the accuracy of the information is paramount. However, several factors can impact the reliability of the reported data. For instance, firms may have incentives to manipulate reporting to present a more favorable image of their environmental performance. This could include overreporting pollution reduction efforts or underreporting actual chemical releases. Additionally, reporting thresholds set by the EPA may introduce biases, as companies may strategically adjust their processes to fall below these thresholds and avoid reporting requirements.

To address these challenges and ensure compliance with TRI regulations, the EPA employs a combination of monitoring, enforcement, and regulatory oversight measures. The agency regularly conducts inspections and audits of regulated facilities to verify the accuracy of reported data and assess compliance with TRI requirements. EPA inspectors, along with attorneys stationed in regional offices and headquarters, play a crucial role in enforcing TRI regulations.

In instances of non-compliance, the EPA has the authority to issue civil penalties, which may include monetary fines and corrective actions to address violations. These penalties serve as deterrents and incentivize companies to adhere to reporting obligations and maintain the integrity of TRI data. Moreover, the threat of enforcement actions helps reinforce the importance of accurate reporting among regulated entities.

Despite these efforts, ensuring the accuracy of TRI data remains an ongoing challenge. Continuous monitoring and oversight by the EPA are necessary to identify and address potential discrepancies or instances of non-compliance effectively. Additionally, stakeholders, including policymakers, researchers, and the public, play a crucial role in scrutinizing TRI data and holding companies accountable for their environmental responsibilities. By maintaining vigilance and strengthening enforcement measures, the EPA can enhance the reliability and usefulness of the TRI program in informing environmental policy and protecting public health.

This paper examines how EPA inspections influence the reporting behavior of firms under the TRI program. The repercussions of these inspections, particularly if they result in civil penalties, impact firms in two significant ways. Firstly, they can act as a deterrent against dishonest reporting, thereby potentially reducing the rate of violations. Secondly, these inspections may also impact the actual quantity of chemicals released. Furthermore, given that EPA inspections and violations are publicized, the penalties may trigger spillover effects on neighboring facilities, both in terms of their geographic proximity and industry type. Awareness of penalties among nearby firms could prompt changes in their reporting practices regarding chemical releases. Understanding the deterrence effect of EPA inspections is crucial due to the widespread use of this information by various individuals and institutions, making it pertinent for policymaking.

Existing research on compliance behavior has primarily centered on legislation like the Clean Air Act (CAA), Clean Water Act (CWA), and Resource Conservation and Recovery Act (RCRA). However, there is a scarcity of studies examining the accuracy of information reported under the Toxics Release Inventory (TRI) program and the effectiveness of the Environmental Protection Agency (EPA) in maintaining data quality through its monitoring efforts. This empirical study aims to address this gap by investigating the compliance behavior of TRI reporting firms concerning both monetary and non-monetary sanctions. Specifically, the research poses the question: Do firms exhibit a greater or lesser tendency to misreport pollution data to the TRI database following an EPA inspection or enforcement action?

The structure of the paper is as follows: Section 2 offers a concise review of relevant literature. Section 3 outlines the data sources used in the study, along with their associated limitations. Section 4 elaborates on the empirical methodology employed. Section 5 presents the findings of the analysis. Finally, Section 6 concludes the paper.

3.2 Literature Review

The Toxics Release Inventory (TRI) dataset has been a valuable resource for economists examining the impact of pollution on various outcomes. Studies utilizing TRI information have investigated its effects on diverse areas such as housing prices (Banzhaf & Walsh, 2008; Mastromonaco, 2015), health risks (Currie et al., 2015), worker chemical exposure (Finger & Gamper-Rabindran, 2013), and firm behavior (Gibson, 2019). Moreover, the release of TRI information has been associated with significant reductions in pollution emissions, with some studies attributing decreases of up to 46% (Graham & Miller, 2001).

Several explanations have been proposed to account for the substantial decline in pollution emissions observed following the dissemination of TRI data. One explanation suggests that pressure from stakeholders such as the stock market and journalists played a role (Hamilton, 1995). Environmental advocacy groups and public awareness campaigns have also been credited with influencing firms to reduce their pollution emissions (Lynn & Kartez, 1994). Additionally, the anticipation of future regulatory actions may have incentivized firms to proactively decrease their environmental impact (Grant, 1997).

Despite its usefulness, the accuracy of TRI information remains a significant concern, primarily due to its self-reported nature. Firms are tasked with providing estimates of their pollution releases, as stipulated on the EPA website, which allows for a degree of subjectivity in reporting. De Marchi and Hamilton (2006) shed light on this issue by demonstrating that the substantial decreases in air emissions reported by firms do not align with corresponding reductions in measured concentrations recorded by EPA monitors (Marchi & Hamilton, 2006). This discrepancy raises questions about the reliability of the reported data and underscores the need for closer scrutiny of TRI information.

Further evidence of potential underreporting comes from Koehler and Spengler (2007), who uncover instances of aluminum facilities failing to fully disclose their toxic chemical releases (Koehler & Spengler, 2007). Similarly, Bennear (2008) examines the reporting thresholds set by TRI and reveals that firms may engage in strategic behavior to manipulate their reporting practices. Bennear's findings suggest that up to 40% of the observed reduction in toxic releases in Massachusetts can be attributed to firms' strategic decisions rather than genuine pollution abatement efforts (Bennear, 2008).

Addressing these inconsistencies in TRI data is crucial for ensuring that self-reported data, such as the TRI, remain valuable tools for both researchers and regulators in efforts to reduce pollution levels. One approach taken by the Environmental Protection Agency (EPA) involves auditing select firms and imposing penalties on those found to be misreporting their data. Oestreich (2015) investigates the theoretical incentives for firms' emissions and self-reporting behavior under two different audit mechanisms: random audit and competitive audit. Oestreich's findings suggest that a competitive audit mechanism, where more audit resources are allocated to firms with lower reported emissions relative to their peers, leads to more truthful reporting (Oestreich, 2015).

The effectiveness of EPA monitoring in influencing the compliance behavior of firms has been extensively studied in the literature, with the majority of studies indicating a positive impact. Lui (2012) explores the compliance behavior of firms subject to multiple environmental regulations and finds evidence of negative cross-program effects, with compliance with the Resource Conservation and Recovery Act (RCRA) negatively affecting compliance with the Clean Air Act (CAA) (Liu, 2012). Similarly, Hanna and Oliva (2010) analyze plant-level data on inspections, fines, and emissions, revealing that plants, on average, reduce air emissions by fifteen percent following enforcement actions (Hanna & Oliva, 2010). Shimshack and Ward (2005) observe similar trends in compliance behavior under the Clean Water Act (CWA), with fines for water pollutant violations leading to a two-thirds reduction in statewide violation rates the following year (Shimshack & Ward, 2005). Gray and Shadbegian (2005) study paper mills to find the determinants of compliance with air pollution regulations. Although enforcement activity incresed compliance, they find that plants that include the pulping process, older plants, and larger plants were less likely to be in compliance (Gray & Shadbegian, 2005). In another study by Gray and Shadbegian (2007), the authors use spatial analysis to study the spatial factors affecting environmental performance. They find that compliance is positively spatially correlated (Gray & Shadbegian, 2007).

However, when it comes to the Toxics Release Inventory (TRI) reporting, the question of how EPA monitoring influences compliance behavior remains less explored. Zou (2017) addresses a related question that examines the impact of intermittent monitoring of environmental standards on polluting activities. The paper suggests that under the federal Clean Air Act, there are strategic responses by local entities to the once-every-six-day air quality monitoring schedule. The study utilizes satellite data to demonstrate that air quality tends to be worse on unmonitored days, indicating short-term suppression of pollution on monitored days, particularly during high-pollution periods when non-compliance risk is elevated. Additionally, the paper suggests that cities increase their use of air quality warnings on monitored days, implying a role for local governments in coordinating emission reductions (Zou, 2021).

These studies collectively offer two key insights. Firstly, EPA monitoring and enforcement activities have proven effective in improving firms' compliance behavior with environmental regulations. However, secondly, firms may engage in strategic behavior to circumvent regulations, suggesting that continuous and vigilant monitoring by the EPA may be necessary to ensure sustained compliance. Yet, the limited budget available for EPA inspections poses a challenge, potentially resulting in insufficient monitoring frequency. In such cases, firms may perceive limited consequences for non-compliance, undermining the effectiveness of regulatory enforcement efforts. Addressing this imbalance between monitoring capacity and regulatory compliance incentives is crucial for maintaining environmental standards and achieving desired pollution reduction outcomes.

Building on the insights from the literature review, the next section delves into the empirical analysis using available data to explore the relationship between EPA monitoring, firm compliance behavior, and the accuracy of TRI reporting.

3.3 Data

The compliance data in this study is from the Enforcement and Compliance History Online (ECHO) database. Spanning the years 1988 to 2018, the ECHO database has comprehensive information regarding each EPA inspection and enforcement action. Each entry in the database includes details such as the commencement and conclusion dates of the case, its outcome, and notably, a distinct identifier for each firm. This unique identifier enables the integration of the compliance data with information extracted from the TRI database. In conjunction with the ECHO data, yearly pollution statistics and firm location data spanning the same period are sourced from the TRI database. Table 16 provides a summary table for the total number of EPA inspection cases, enforcement cases, cases penalized, and the average penalty per case for all EPA programs and only for TRI. The average penalty per case assessed for TRI is less than one-third of all EPA programs.

Table 16 Summary Stats about Inspection and Enforcement Cases.

	Total	TRI
No. of Inspection cases	188,437	3,241
No. of Enforcement cases	112,203	3,030
No. of cases penalized	25,763	1,458
Average penalty per case assessed	$60,445	$18,545

An inherent limitation of relying on ECHO data is the potential underestimation of a firm's misbehavior, as it is contingent upon detection by the EPA. In other words, instances of firms misreporting TRI data may go unnoticed by regulatory authorities, thereby resulting in an incomplete picture of non-compliance. Consequently, the proxy for misreporting derived from ECHO data may overstate the efficacy of EPA monitoring and enforcement activities in deterring firms from engaging in non-compliant behavior. This limitation underscores the need for a cautious interpretation of the results, recognizing the possibility of underestimating the true extent of misreporting.

Table 17 presents a detailed analysis of EPA enforcement and inspection cases categorized by industry sector, providing valuable insights into the enforcement landscape across various economic domains. Among the sectors, Fabricated Metals emerge as noteworthy, showcasing a particularly high enforcement-to-inspection ratio of 36.59%. This implies that out of the 1,260 inspections conducted within this sector, a substantial portion resulted in 461 enforcement cases, reflecting heightened scrutiny and regulatory action. Similarly, the Computers and Electronic Products industry exhibits a significant enforcement ratio of 35.03%, indicating robust regulatory oversight and enforcement efforts. Conversely, sectors such as Hazardous Waste, Petroleum Bulk Terminals, and Other industries report lower enforcement ratios, ranging between 1.57% and 2.23%, suggesting comparatively fewer instances of regulatory non-compliance or stricter adherence to environmental standards. Additionally, the Chemicals sector, with 3,779 inspections and 364 enforcement cases, demonstrates a ratio of 9.63%, underscoring the nuanced regulatory landscape within this industry. These findings shed light on the varying degrees of regulatory enforcement and compliance challenges across different industrial sectors, thereby informing policy decisions and resource allocations aimed at enhancing environmental protection and enforcement effectiveness.

3.4 Benford's Law

De Marchi and Hamilton (2006) use Benford's law to test whether the first digits of the TRI data follow a monotonically decreasing distribution (Marchi & Hamilton, 2006). I use a similar method to examine potential changes in reported numbers before and after an inspection or enforcement action, I employ Benford's law as a diagnostic tool. Benford's Law, also known as the "first-digit law" or the "law of anomalous numbers," is a statistical phenomenon that describes the frequency distribution of leading digits in many naturally occurring datasets. The law states that in many sets of numerical data, the leading digit (i.e., the first digit in a number) is more likely to be small, such as 1, 2, or 3, rather than large digits like 8 or 9. This counter-intuitive phenomenon has been observed in diverse datasets, including financial records, population statistics, scientific data, and even physical constants (Fewster, 2009; Hill, 1995; Nigrini, 1996). The results of this analysis offer insights

Table 17 EPA Enforcement and Inspection Cases by Industry Sector

Industry Sector	Inspection	Enforcement	Enforcement Ratio
Chemicals	3,779	364	9.63 %
Fabricated Metals	1,260	461	36.59 %
Primary Metals	1,111	237	21.33 %
Food	916	132	14.41 %
Petroleum	797	69	8.66 %
Other	752	12	1.60 %
Transportation Equipment	666	136	20.42 %
Electric Utilities	630	13	2.06 %
Hazardous Waste	583	13	2.23 %
Petroleum Bulk Terminals	574	9	1.57 %
Plastics and Rubber	503	130	25.84 %
Computers and Electronic Products	431	151	35.03 %
Chemical Wholesalers	429	25	5.83 %
Nonmetallic Mineral Product	427	70	16.39 %
Machinery	330	93	28.18 %
Paper	320	36	11.25 %
Wood Products	282	36	12.77 %
Electrical Equipment	249	68	27.31 %
Miscellaneous Manufacturing	182	50	27.47 %
Printing	88	31	35.23 %

into whether firms engage in misreporting behavior.

Figure 12 and Figure 13 present the outcomes of applying Benford's law to firms subjected to inspection or enforcement actions. Comparing the distribution of first digits in reported numbers one year before and after an inspection reveals no substantial changes, suggesting limited evidence of manipulation in reporting behavior by the firms. One interesting result is that we see an increased use of the digit 5 for before and after inspection.

Figure 12 Application of Benford's Law for Inspection Cases

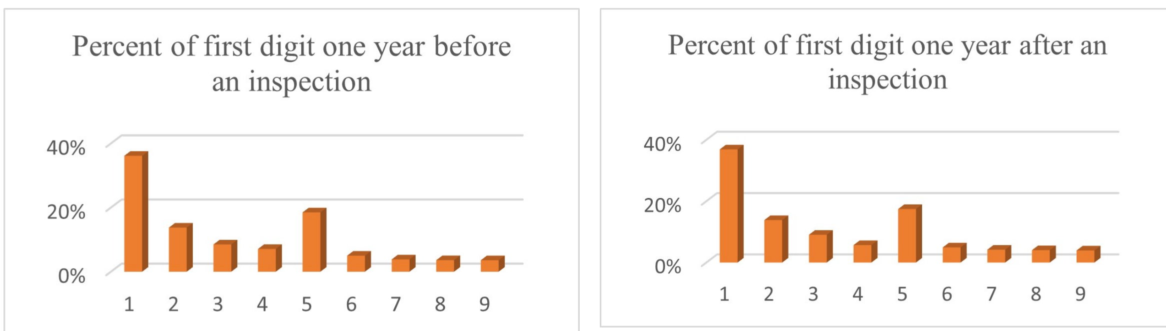
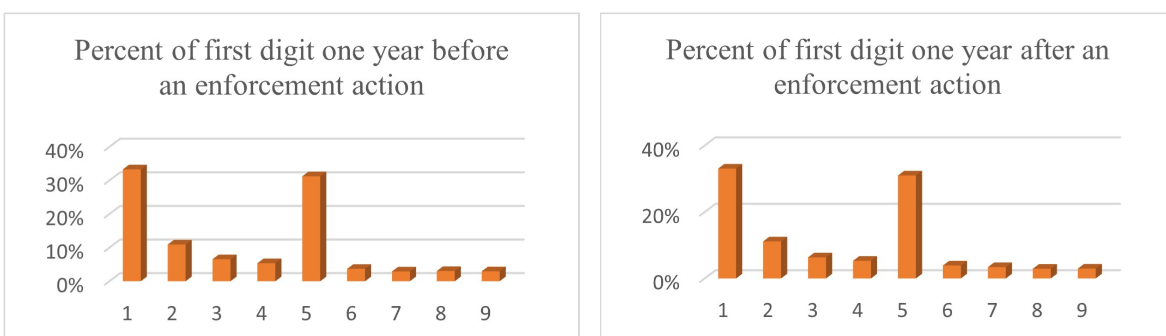


Figure 13 Application of Benford's Law for Enforcement Actions



3.5 Methodology

To analyze how past EPA inspections and enforcement actions influence a firm's future compliance behavior, I employ a Probit model to explain the probability change of compliance and inspection. EPA first inspects a facility, followed by an enforcement action if the firm is found to be in non-compliance. As a result, enforcement actions are a joint outcome of firms action and the EPA action. Furthermore, a significant number of the violations in regards to TRI are for clerical or paperwork error and not pollution violations. In the following equation, $Enforcement_{i,t}$ and $Inspection_{i,t}$ is the outcome variable, which equals 1 if the firm is inspected (enforcement action) at time t. $X_{i,t-1}$ is the set of explanatory variables such as industry sector, neighbor inspection history, neighbor compliance history, and individual compliance history in the previous period. ind_j and v_t are the industry fixed effect and year fixed effect and $\epsilon_{i,t}$ is the error term.

$$\begin{aligned}
Pr(Enforcement_{i,t}) = 1|X_{i,t-1}) &= \phi(\alpha +_{i,t-1} + ind_j + v_t + \epsilon_{i,t}) \\
&= \phi(\alpha + \beta_1 Enforcement_{t-1} + \beta_2 Inspection_{t-1} + \beta_3 Neighbor\ Inspection_{t-1} \\
&\qquad + \beta_4 Neighbor\ Enforcement_{t-1} + ind_j + v_t + \epsilon_{i,t})
\end{aligned}$$

If the enforcement and inspection actions of the EPA are effective, they should deter firms from making the same mistakes in subsequent periods, which is explained by β_1 and β_2 values. Furthermore, the action can also affect nearby firms as the EPA makes enforcement and inspection information public through its website. Neighbor Inspection and Neighbor Enforcement are calculated as the fraction of Inspection/Enforcement cases (not including the firm itself) out of the total TRI reporting in the same zip code in the previous year. I hypothesize that $Enforcement_{i,t}$, and $Inspection_{i,t}$ variables should have negative impacts on the probability that a firm faces an enforcement action or an inspection. Prior EPA inspection or enforcement should decrease the likelihood that the firm misreports. The following section presents the results of the two Probit regressions.

3.6 Results

Table 18 presents the marginal effects for Enforcement and Inspection, shedding light on the interplay between past regulatory actions and their impact on current outcomes. The findings indicate a strong relationship between prior inspection and subsequent enforcement for a firm. Specifically, there is a notable 0.91% likelihood that a firm will face enforcement activity following a prior inspection. This high probability can be attributed to the time lag between inspection and the regulatory body's assessment of any misreporting, which may lead to enforcement actions in the current period as a result of findings from the previous inspection.

Conversely, if a firm has undergone enforcement action in the previous period, it becomes less likely to face an inspection in the current year, although this effect is not statistically significant. This suggests a potential prioritization by regulatory bodies to allocate inspection resources to firms without recent enforcement history. Additionally, prior inspection significantly increases the likelihood of the firm being inspected in the current period, with a probability of 2.4%. However, enforcement actions in the prior period resulted in a decreased likelihood of inspection in the current year, indicating a potential focus shift towards firms with no recent enforcement activity.

Moreover, the analysis highlights the neighborhood effect of regulatory actions, wherein both Neighbor Inspection and Neighbor Enforcement in the last period amplify a firm's probability of facing an enforcement action. Specifically, if a neighboring firm undergoes enforcement, there is a substantial 1.9% likelihood that the focal firm will be inspected. Similarly,

Table 18 Probit Results for Enforcement and Inspection

Variables	Enforcement	Inspection
$Enforcement_{t-1}$	-0.00025	-0.00139
	(0.0011)	(0.00376)
$Inspection_{t-1}$	0.00910***	0.02406***
	(0.00031)	(0.00085)
$Neighbor\ Inspection_{t-1}$	0.00033	0.03459***
	(0.00123)	(0.00250)
$Neighbor\ Enforcement_{t-1}$	0.00713***	-0.01954**
	(0.00271)	(0.01108)
Year Fixed Effects	Yes	Yes
Industry Fixed Effects	Yes	Yes
Observations	646,056	536,271

Notes: 1. Standard errors in parentheses
2. *** p<0.01, ** p<0.05, * p<0.1

Neighbor Inspection significantly increases the likelihood of a firm undergoing inspection, with a probability of 3.5%. Furthermore, Neighbor Enforcement also influences the probability of a firm facing enforcement, with a probability of 0.7%. These findings underscore the interconnectedness of regulatory outcomes within neighboring firms and their implications for enforcement activities.

3.7 Conclusion

The findings from this study suggest that EPA inspections and enforcement actions may not necessarily lead to improved compliance behavior regarding TRI reporting among firms. While prior enforcement actions may decrease the probability of subsequent enforcement from the EPA, firms with a history of inspection are more likely to undergo further inspection, indicating a complex relationship between regulatory actions and compliance outcomes.

A notable observation is the potential strategic behavior by the EPA in allocating inspection resources and learning from past monitoring efforts. The significant likelihood of a firm facing enforcement following a prior inspection suggests targeted enforcement efforts, possibly aimed at addressing recurring offenders. This challenges the assumption of random inspections and raises concerns about the potential bias.

Moreover, the disparities in penalties for TRI violations compared to other environmental regulations raise questions about the effectiveness of enforcement mechanisms and the perceived importance of TRI compliance within the EPA's enforcement priorities. The relatively lower penalties for TRI violations, coupled with potential concerns about community health impacts from other environmental violations, may influence the EPA's enforcement strategies and resource allocation.

In conclusion, while EPA inspections and enforcement actions play a crucial role in environmental regulation, the findings suggest a need for further examination of regulatory practices and their implications for compliance behavior. Addressing issues such as strategic enforcement, resource allocation, and penalty adequacy is essential for enhancing the effectiveness of environmental regulations and promoting sustainable compliance practices among firms reporting TRI data.

Appendices

Appendix A. Chapter I Supplementary Tables, Figures, and Instructions

Appendix A.1 Supplementary Figures

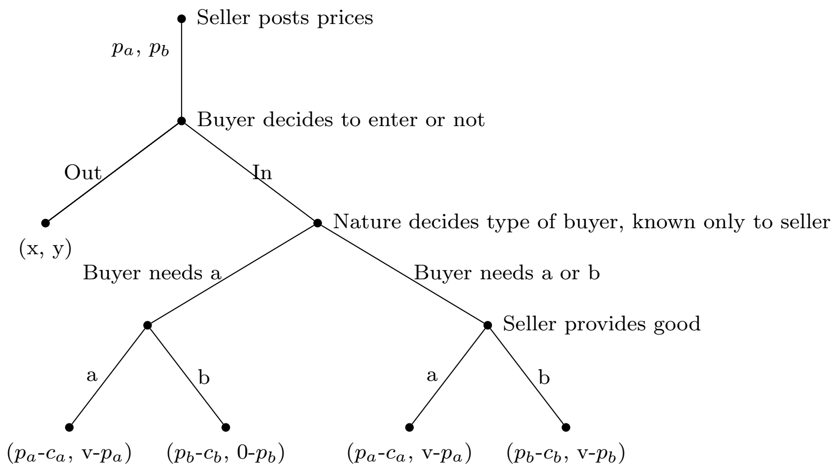


Figure A.1 Credence Good Game Tree

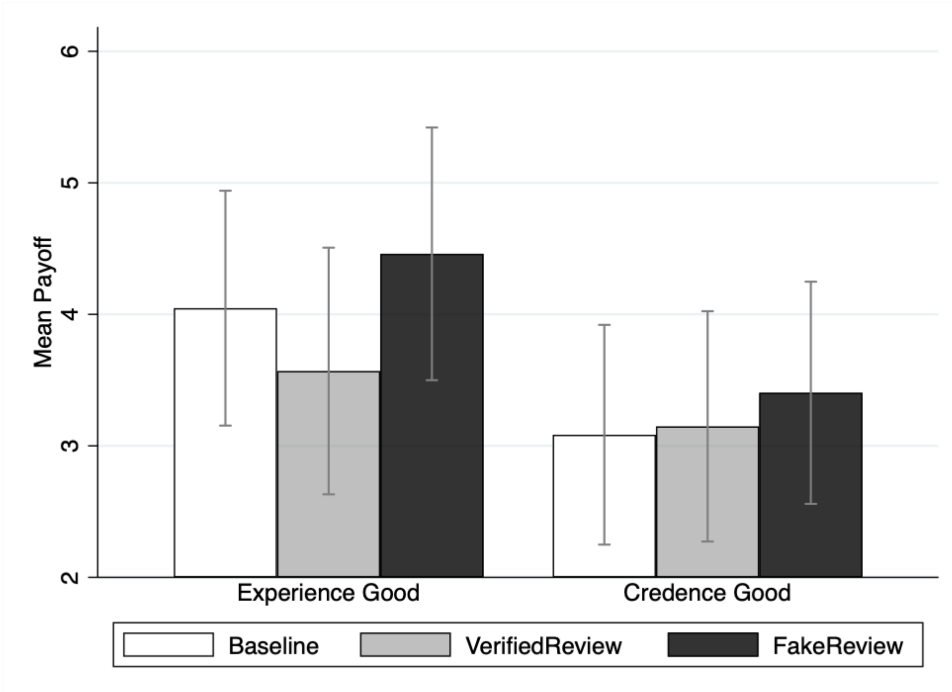


Figure A.2 Mean Payoffs by Type of Good Each Round When Buyers ”Enter”

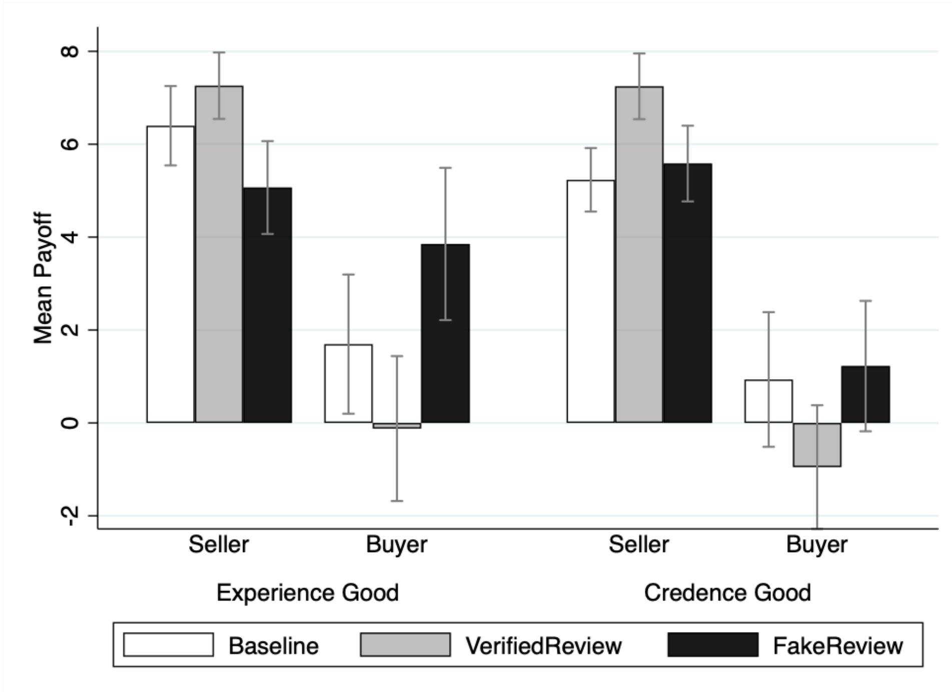


Figure A.3 Mean Payoffs by Type of Good & Player Each Round When Buyers "Enter"

Post experiment survey

1) What is your Gender?
    a. Male
    b. Female

2) What is your Ethnicity?
    a. Asian
    b. Black or African American
    c. Hispanic
    d. Multiracial
    e. Prefer Not to Answer
    f. White or Caucasian

3) What is your age?
    a.

Figure A.4 Post Experiment Survey Form

Appendix A.2 Supplementary Tables

Table A.1 Linear Regression Results for Buyer/Seller Payoff Controlled for Real Rating

	(1)	(2)	(3)	(4)	(5)	(6)
	seller	seller	seller	buyer	buyer	buyer
Credence Goods	0.416	0.0686	0.0621	-1.467**	-0.816	-0.533
	(0.355)	(0.382)	(0.382)	(0.630)	(0.684)	(0.594)
Round	-0.0283	0.00485	0.00258	0.0714	0.0397	0.113*
	(0.0374)	(0.0422)	(0.0423)	(0.0606)	(0.0692)	(0.0604)
Rating		0.597***	0.600***		-0.452*	-0.505**
		(0.156)	(0.156)		(0.251)	(0.218)
Real rating		-1.120***	-1.116***		1.248***	1.126***
		(0.217)	(0.217)		(0.344)	(0.299)
High Type Buyer			0.220			-6.736***
			(0.348)			(0.496)
N	640	573	573	640	573	573

Notes: 1. Standard errors in parentheses
2. * p<0.1, ** p<0.05, *** p<0.01

Table A.2 Fake Review Incidence by Round Number

	Self Fake Rate		Other Fake Rate		
Round Number	No	Yes	No	Yes	Total
1	7	33	9	31	40
2	6	34	9	31	40
3	8	32	14	26	40
4	10	30	19	21	40
5	7	33	13	27	40
6	5	35	9	31	40
7	8	32	14	26	40
8	5	35	12	28	40
9	8	32	12	28	40
10	4	36	10	30	40
11	6	34	14	26	40
12	4	36	11	29	40
13	6	34	11	29	40
14	9	31	12	28	40
15	9	31	12	28	40
16	8	32	8	32	40
Total	110	530	189	451	640

Appendix A.3 Instructions - Credence, Verified Review, Buyer

Welcome! You are Player B.

Thank you for taking part in the experiment. **Please do not talk.** If you have a question after you finish reading the instructions, please raise your hand, and the experimenter will approach you and answer your question in private. **The use of mobile devices is PROHIBITED.** Violation of this rule may result in the experimenter removing you from subject pool to participate in future experiments. Your decisions and earnings in the experiment will remain anonymous.

Your Decision

In the experiment today, you will be making decisions over multiple rounds of the game. You will be randomly matched with a player before each round of the game. You may or may not be matched with the same player more than once. You will play 2 practice rounds and **16** paid rounds. In each round, you will earn experimental points. Your points will be converted to US dollars at the end of the experiment. The rate of exchange is 1 point = $ 0.25. You will also receive 10 points once before the start of the game as an endowment. You will make **Decision 2 and 4.**

Decision 1,

You will wait for Player A to make a decision. Player A chooses price for Action 1 and Action 2.

Decision 2,

Before the start of the decision, the computer will randomly assign your type. You can be one of the two types: type 1 or type 2. This type is determined for you in each new round. With a probability of 50% you are of type 1, and with a probability of 50% you are of type 2. **Only** your matched player will know your type. You will NOT know your type. You can either choose 'In' or 'Out.' If you choose 'In,' the game will move to Decision 3. If you choose 'Out,' the game ends here for both players A and B.

Decision 3,

You do not have to make any decision in this round. Please wait for Player A to make a decision. Player A can choose between Action 1 and Action 2.

Decision 4,

You will provide a rating between 1 and 5 to Player A with 1 being the lowest rating and 5 being the highest rating. Starting from round 2, you will know the average rating of Player A from previous rounds of the game before the start of each round. The average rating is the arithmetic average of ratings provided by previous Player B's that were matched with Player A.

Payoff for Player A

◦ If you choose 'Out' in Decision 2, Player A gets 3 points for this round of the game.

◦ If you are type 1 and if you choose 'In' during Decision 1, Player A payoff is dependent on Decision 3, and Player A payoff will be the following:

- If Player A choose action 1, Player A payoff is (price for Action 1 – 10 points).
- If Player A choose action 2, Player A payoff is (price for Action 2 – 4 points).

◦ If you are type 2 and if you choose 'In' during Decision 1, Player A payoff is dependent on Decision 3, and Player A payoff will be the following:

- If Player A choose action 1, Player A payoff is (price for Action 1 – 10 points).
- If Player A choose action 2, Player A payoff is (price for Action 2 – 4 points).

Payoff for Player B

∘ If you choose 'Out' in decision 2, you get 3 points for this round of the game.

∘ If you are type 1 and if you choose 'In' during Decision 1, your payoff is dependent on Decision 3, and your payoff will be the following:

- If Player A chooses action 1, your payoff is (18 points – price for Action 1).

- If player A chooses action 2, your payoff is (0 – price for Action 2).

∘ If you are type 2 and if you choose 'In' during Decision 1, your payoff is dependent on Decision 3, and your payoff will be the following:

- If Player A chooses action 1, your payoff is (18 points – price for Action 1).

- If player A chooses action 2, your payoff is (18 points – price for Action 2).

Your total earnings in this task will be the sum of your earnings in each round plus the endowment (1 point = 0.25 US dollars).

Appendix A.4 Instructions - Credence, Verified Review, Seller

Welcome! You are Player A.

Thank you for taking part in the experiment. **Please do not talk.** If you have a question after you finish reading the instructions, please raise your hand, and the experimenter will approach you and answer your question in private. The **use of mobile devices is PROHIBITED.** Violation of this rule may result in the experimenter removing you from subject pool to participate in future experiments. Your decisions and earnings in the experiment will remain anonymous.

Your Decision

In the experiment today, you will be making decisions over multiple rounds of the game. In each round, you will be making multiple decisions. You will be randomly matched with a player before each round of the game. You may be matched with the same player more than once. You will play **2** practice rounds and **16** paid rounds. In each round, you will earn experimental points. Your points will be converted to US dollars at the end of the experiment. The rate of exchange is 1 point = $ 0.25. You will also receive 10 points once before the start of the game as an endowment. You will make **Decision 1, 3, and 5.**

Decision 1,

You will choose a price for Action 1 and Action 2. You can choose the price between 1 and 18 points.

Decision 2,

Please wait for Player B to make a decision. Before the start of the decision, the computer will randomly assign Player B a type. Player B can be one of the two types: type

1 or type 2. Type for Player B is determined before each new round. With a probability of 50% Player B is of type 1, and with a probability of 50% Player B is of type 2. Only you know the type of Player B. You do not have to make any decision in this round. If Player B chooses ‘In,’ the game will move to Decision 3. If Player B chooses ‘Out,’ the game ends here for both players A and B.

Decision 3,

You will choose between Action 1 and Action 2. Action 1 will cost you 10 points and action 2 will cost you 4 points.

Decision 4,

You do not have to make any decision in this round. Player B will provide you with a rating between 1 and 5 with 1 being the lowest rating and 5 being the highest rating. Starting from round 2, Player B will know your average rating before the start of each round. The average rating is the arithmetic average of ratings provided by previous Player B’s that you were matched with.

Decision 5,

You can provide a rating between 1 and 5 with 1 being the lowest rating and 5 being the highest rating to yourself and provide rating to one other Player A. The average rating is the arithmetic average of ratings provided by previous Player B’s that you were matched with, the possible rating you rated yourself, and the possible rating provided by another Player A.

Payoff for Player A

○ If Player B chooses ‘Out’ in Decision 1, you get 3 points for this round of the game. ○ If Player B is type 1 and if Player B chooses ‘In’ during Decision 1, your payoff is dependent on Decision

3, and your payoff will be the following:

- If you choose action 1, your payoff is (price for Action 1 – 10 points).

- If you choose action 2, your payoff is (price for Action 2 – 4 points).

◦ If Player B is type 2 and if Player B chooses 'In' during Decision 1, your payoff is dependent on Decision 3, and your payoff will be the following:

- If you choose action 1, your payoff is (price for Action 1 – 10 points).

- If you choose action 2, your payoff is (price for Action 2 – 4 points).

Payoff for Player B

◦ If Player B chose 'Out' in decision 2, Player B gets 3 points for this round of the game.

◦ If Player B is type 1 and if Player B chooses 'In' during Decision 1, Player B payoff is dependent on Decision 3, and Player B payoff will be the following:

- If you choose action 1, Player B payoff is (18 points – price for Action 1).

- If you choose action 2, Player B payoff is (0 – price for Action 2).

◦ If Player B is type 2 and if Player B chooses 'In' during Decision 1, Player B payoff is dependent on Decision 3, and Player B payoff will be the following:

- If you choose action 1, Player B payoff is (18 points – price for Action 1).

- If you choose action 2, Player B payoff is (18 points – price for Action 2).

Your total earnings in this task will be the sum of your earnings in each round plus the endowment (1 point = 0.25 US dollars).

Appendix A.5 Instructions - Experience, Verified Review, Buyer

Welcome! You are Player B.

Thank you for taking part in the experiment. **Please do not talk.** If you have a question after you finish reading the instructions, please raise your hand, and the experimenter will approach you and answer your question in private. **The use of mobile devices is PROHIBITED.** Violation of this rule may result in the experimenter removing you from subject pool to participate in future experiments. Your decisions and earnings in the experiment will remain anonymous.

Your Decision

In the experiment today, you will be making decisions over multiple rounds of the game. You will be randomly matched with a player before each round of the game. You may or may not be matched with the same player more than once. You will play 2 practice rounds and **16** paid rounds. In each round, you will earn experimental points. Your points will be converted to US dollars at the end of the experiment. The rate of exchange is 1 point = $ 0.25. You will also receive 10 points once before the start of the game as an endowment. You will make **Decision 2 and 4.**

Decision 1,

You will wait for Player A to make a decision. Player A chooses price for Action 1 and Action 2.

Decision 2,

Before the start of the decision, the computer will randomly assign your type. You can be one of the two types: type 1 or type 2. This type is determined for you in each new round. With a probability of 50% you are of type 1, and with a probability of 50% you are of type 2. You and your matched player know your type. You will NOT know your type. You can either choose 'In' or 'Out.' If you choose 'In,' the game will move to Decision 3. If you choose 'Out,' the game ends here for both players A and B.

Decision 3,

You do not have to make any decision in this round. Please wait for Player A to make a decision. Player A can choose between Action 1 and Action 2.

Decision 4,

You will provide a rating between 1 and 5 to Player A with 1 being the lowest rating and 5 being the highest rating. Starting from round 2, you will know the average rating of Player A from previous rounds of the game before the start of each round. The average rating is the arithmetic average of ratings provided by previous Player B's that were matched with Player A.

Payoff for Player A

◦ If you choose 'Out' in Decision 2, Player A gets 3 points for this round of the game.

◦ If you are type 1 and if you choose 'In' during Decision 1, Player A payoff is dependent on Decision 3, and Player A payoff will be the following:

- If Player A choose action 1, Player A payoff is (price for Action 1 – 10 points).

- If Player A choose action 2, Player A payoff is (price for Action 1 – 4 points).

◦ If you are type 2 and if you choose 'In' during Decision 1, Player A payoff is dependent on Decision 3, and Player A payoff will be the following:

- If Player A choose action 1, Player A payoff is (price for Action 2 – 10 points).

- If Player A choose action 2, Player A payoff is (price for Action 2 – 4 points).

Payoff for Player B

◦ If you choose 'Out' in decision 2, you get 3 points for this round of the game.

◦ If you are type 1 and if you choose 'In' during Decision 1, your payoff is dependent on Decision 3, and your payoff will be the following:

- If Player A chooses action 1, your payoff is (18 points – price for Action 1).

- If player A chooses action 2, your payoff is (0 – price for Action 1).

◦ If you are type 2 and if you choose 'In' during Decision 1, your payoff is dependent on Decision 3, and your payoff will be the following:

- If Player A chooses action 1, your payoff is (18 points – price for Action 2).

- If player A chooses action 2, your payoff is (18 points – price for Action 2).

Your total earnings in this task will be the sum of your earnings in each round plus the endowment (1 point = 0.25 US dollars).

Appendix A.6 Instructions - Experience, Verified Review, Seller

Welcome! You are Player A.

Thank you for taking part in the experiment. **Please do not talk.** If you have a question after you finish reading the instructions, please raise your hand, and the experimenter will approach you and answer your question in private. The **use of mobile devices is PROHIBITED.** Violation of this rule may result in the experimenter removing you from subject pool to participate in future experiments. Your decisions and earnings in the experiment will remain anonymous.

Your Decision

In the experiment today, you will be making decisions over multiple rounds of the game. In each round, you will be making multiple decisions. You will be randomly matched with a player before each round of the game. You may be matched with the same player more than once. You will play **2** practice rounds and **16** paid rounds. In each round, you will earn experimental points. Your points will be converted to US dollars at the end of the experiment. The rate of exchange is 1 point = \$ 0.25. You will also receive 10 points once before the start of the game as an endowment. You will make **Decision 1, 3, and 5.**

Decision 1,

You will choose a price for Action 1 and Action 2. You can choose the price between 1 and 18 points.

Decision 2,

Please wait for Player B to make a decision. Before the start of the decision, the computer will randomly assign Player B a type. Player B can be one of the two types: type

1 or type 2. Type for Player B is determined before each new round. With a probability of 50% Player B is of type 1, and with a probability of 50% Player B is of type 2. You and your matched player knows the type of Player B. You do not have to make any decision in this round. If Player B chooses 'In,' the game will move to Decision 3. If Player B chooses 'Out,' the game ends here for both players A and B.

Decision 3,

You will choose between Action 1 and Action 2. Action 1 will cost you 10 points and action 2 will cost you 4 points.

Decision 4,

You do not have to make any decision in this round. Player B will provide you with a rating between 1 and 5 with 1 being the lowest rating and 5 being the highest rating. Starting from round 2, Player B will know your average rating before the start of each round. The average rating is the arithmetic average of ratings provided by previous Player B's that you were matched with.

Decision 5,

You can provide a rating between 1 and 5 with 1 being the lowest rating and 5 being the highest rating to yourself and provide rating to one other Player A. The average rating is the arithmetic average of ratings provided by previous Player B's that you were matched with, the possible rating you rated yourself, and the possible rating provided by another Player A.

Payoff for Player A

◦ If Player B chooses 'Out' in Decision 1, you get 3 points for this round of the game.

◦ If Player B is type 1 and if Player B chooses 'In' during Decision 1, your payoff is dependent on Decision 3, and your payoff will be the following:

- If you choose action 1, your payoff is (price for Action 1 – 10 points).

- If you choose action 2, your payoff is (price for Action 1 – 4 points).

◦ If Player B is type 2 and if Player B chooses ‘In’ during Decision 1, your payoff is dependent on Decision 3, and your payoff will be the following:

- If you choose action 1, your payoff is (price for Action 2 – 10 points).

- If you choose action 2, your payoff is (price for Action 2 – 4 points).

Payoff for Player B

◦ If Player B chose ’Out’ in decision 2, Player B gets 3 points for this round of the game.

◦ If Player B is type 1 and if Player B chooses ‘In’ during Decision 1, Player B payoff is dependent on Decision 3, and Player B payoff will be the following:

- If you choose action 1, Player B payoff is (18 points – price for Action 1).

- If you choose action 2, Player B payoff is (0 – price for Action 1).

◦ If Player B is type 2 and if Player B chooses ‘In’ during Decision 1, Player B payoff is dependent on Decision 3, and Player B payoff will be the following:

- If you choose action 1, Player B payoff is (18 points – price for Action 2).

- If you choose action 2, Player B payoff is (18 points – price for Action 2).

Your total earnings in this task will be the sum of your earnings in each round plus the endowment (1 point = 0.25 US dollars).

Appendix B. Chapter II Supplementary Tables and Figures

Figure B.1 Local Monitor Available Provinces

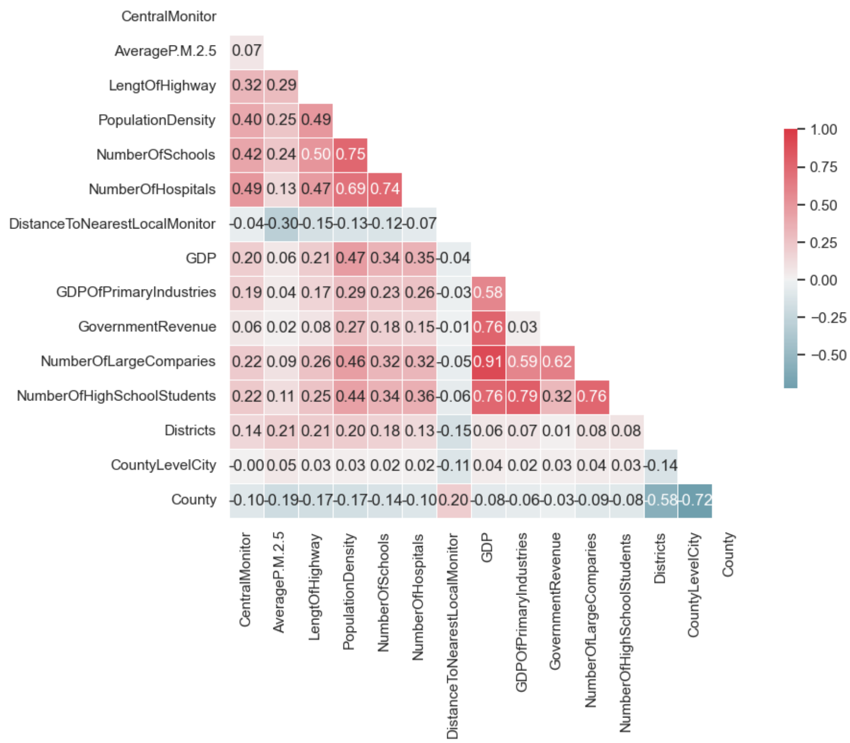


Figure B.2 Correlation Graph of Central Monitor Location with Determining Factors

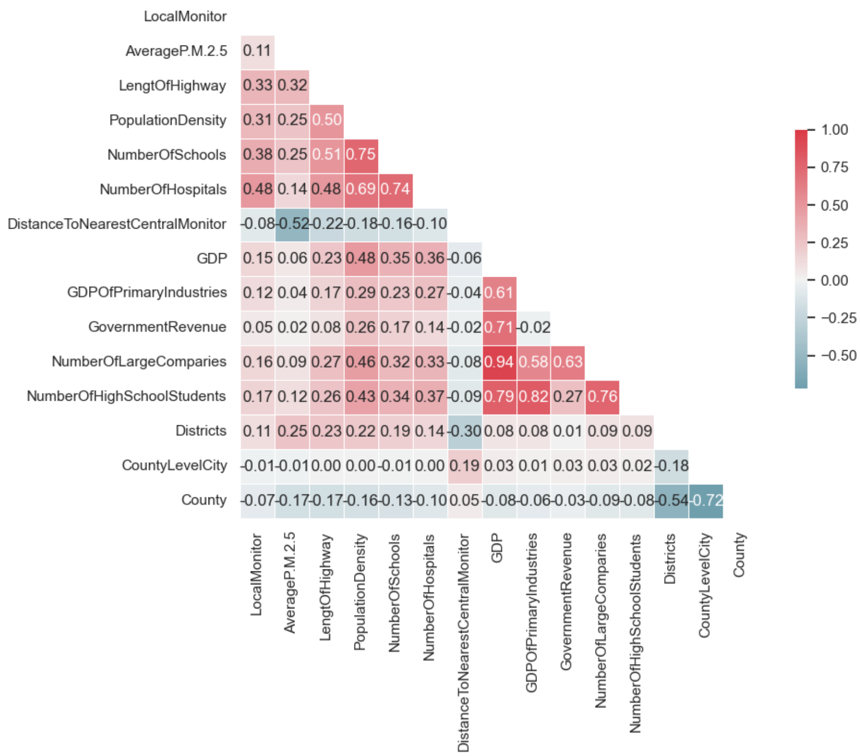


Figure B.3 Correlation Graph of Local Monitor Location with Determining Factors

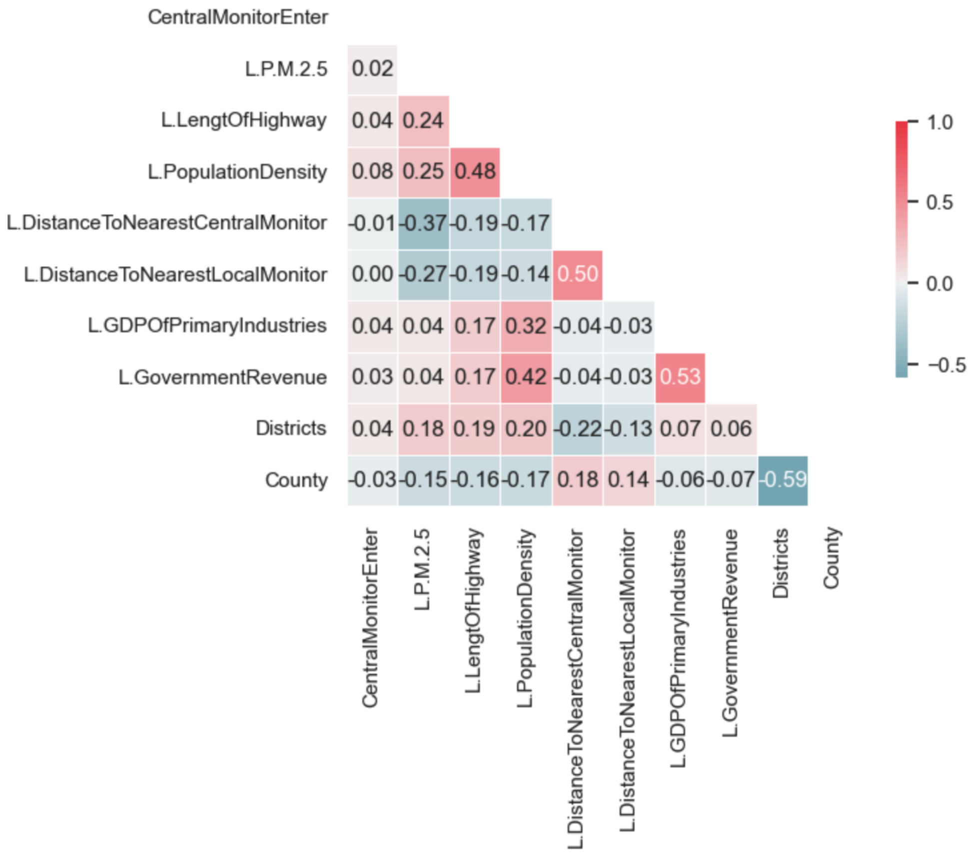


Figure B.4 Correlation Graph of Central Monitor Entries with Determining Factors

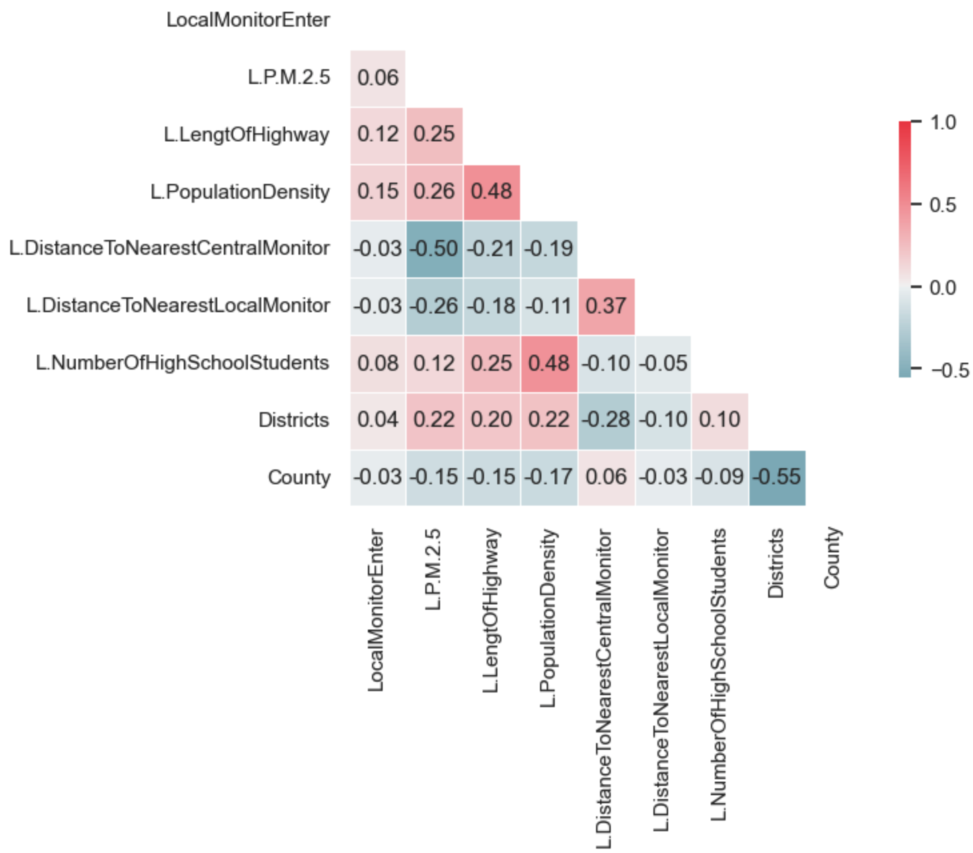
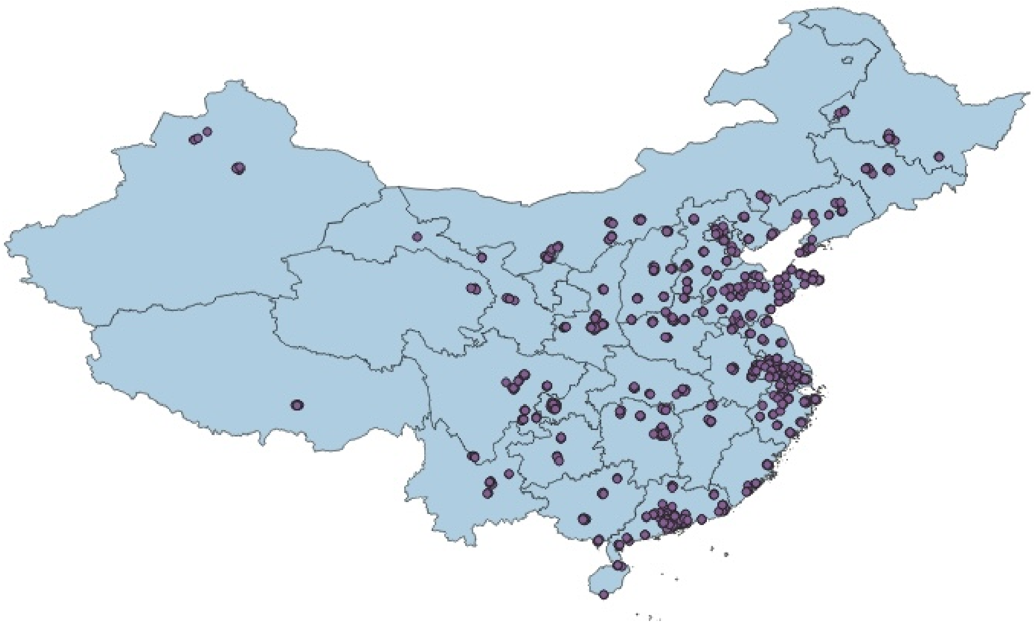
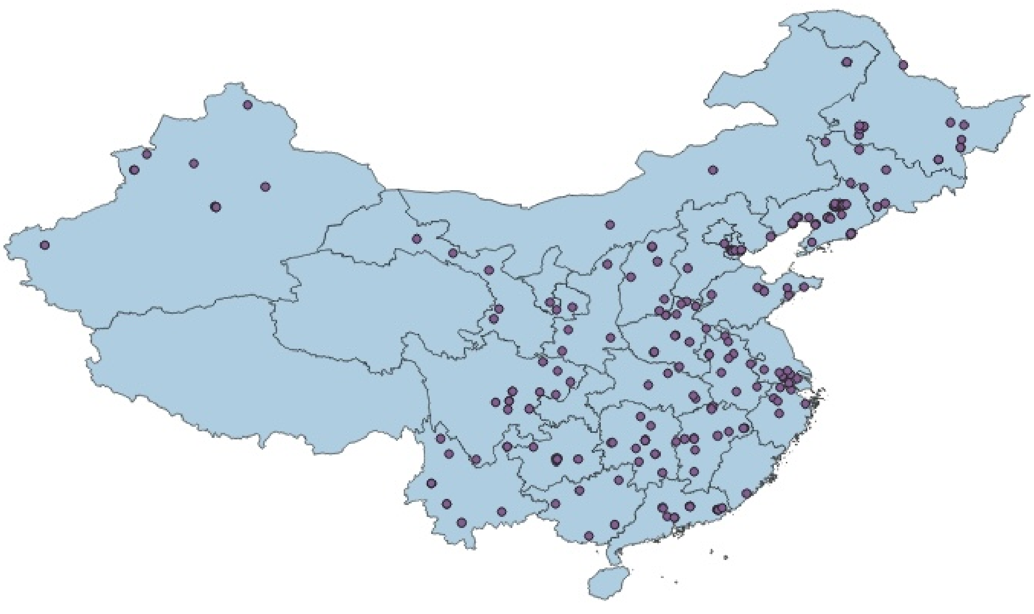


Figure B.5 Correlation Graph of Local Monitor Entries with Determining Factors

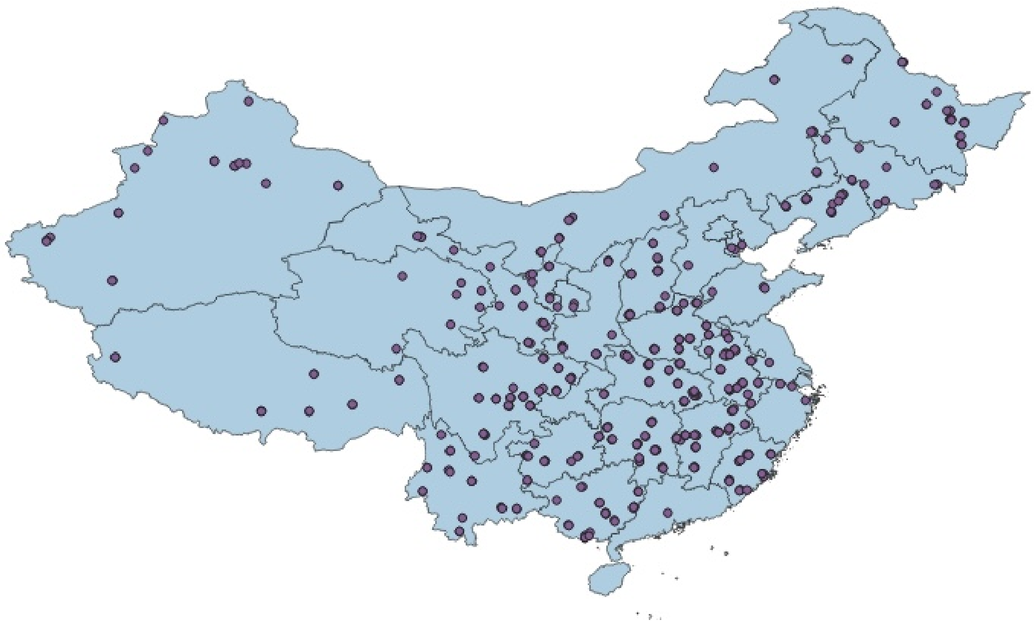


(a) 2013

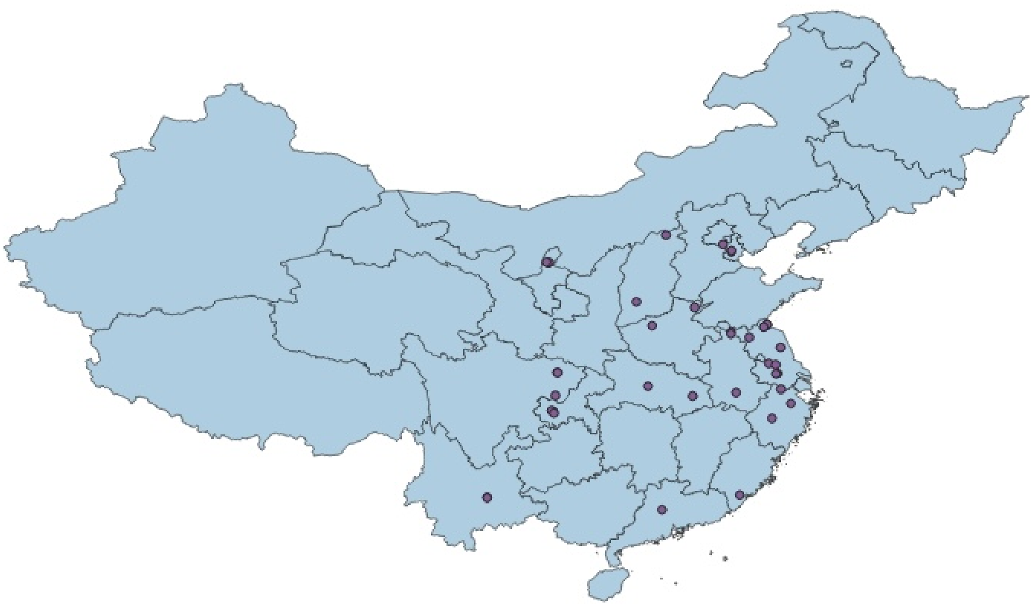


(b) 2014

Figure B.6 Central Monitor Entry 2013-2014



(a) 2015



(a) 2016

Figure B.7 Central Monitor Entry 2015-2016

(a) 2017

(a) 2018

Figure B.8 Central Monitor Entry 2017-2018

Table B.1 Linear Probability Model Marginal Effects for Monitor Location (2021)

	(1)	(2)	(3)	(4)
	Central Monitor		Local Monitor	
Average P.M. 2.5	-0.0022***	-0.0015***	3.8e-4	-0.0034***
	(1.1e-4)	(1.7e-4)	(2.3e-4)	(2.8e-4)
Length of Highway	0.0066***	0.0069***	0.0122***	0.0083***
	(1.1e-4)	(1.1e-4)	(2.2e-4)	(1.7e-4)
Number of Hospitals	0.0257***	0.0256***	0.0414***	0.0299***
	(1.1e-4)	(1.1e-4)	(2.2e-4)	(1.6e-4)
Distance to the Nearest Central Monitor			-2.1e-4	-0.0016***
			(2.7e-4)	(2.2e-4)
Distance to the Nearest Local Monitor	5.3e-4***	7.5e-4***		
	(1.0e-4)	(2.2e-4)		
GDP of Primary industries	7.1e-4***	6.3e-4***	-0.0020***	7.5e-4***
	(1.3e-4)	(1.3e-4)	(2.9e-4)	(2.2e-4)
Government Revenue	-0.0045***	-0.0046***	-0.0029***	-0.0049***
	(1.4e-4)	(1.4e-4)	(3.0e-4)	(2.3e-4)
Number of Large Companies	0.0060***	0.0062***	0.0010**	0.0077***
	(1.8e-4)	(1.9e-4)	(3.9e-4)	(3.0e-4)
District Dummy	0.0139***	0.0151***	0.0055***	0.0182***
	(3.2e-4)	(3.3e-4)	(6.0e-4)	(4.8e-4)
Provincial Dummies	No	Yes	No	Yes
N	338,067	338,067	212,123	212,123

Notes: 1. All explanatory variables are standardized.
2. Standard errors in parentheses.
3. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.2 Pooled Logit Model Marginal Effects for Monitor Entry Location Choice

	(1)	(2)
	Central Monitor	Local Monitor
L.P.M. 2.5	3.0e-06**	2.9e-5***
	(1.4e-06)	(7.1e-06)
L.Length of Highway	1.4e-06***	7.9e-06***
	(3.1e-07)	(1.8e-06)
L.Population Density	1.1e-06***	2.7e-5***
	(2.1e-07)	(5.2e-06)
L.Distance to the Nearest Central Monitor	-1.8e-5***	-2.4e-4***
	(5.6e-06)	(3.9e-5)
L.Distance to the Nearest Local Monitor	1.0e-5***	1.8e-5**
	(3.0e-06)	(8.3e-06)
L.GDP of Primary Industries	3.7e-07***	-6.8e-06***
	(1.2e-07)	(2.6e-06)
L.Government Revenue	-4.6e-07	1.6e-06*
	(3.0e-07)	(9.6e-07)
District Dummy	1.9e-5***	4.9e-5***
	(3.8e-06)	(1.4e-5)
County Dummy	-3.1e-5***	1.6e-06
	(4.6e-06)	(9.1e-06)
Yearly and Provincial Dummies	Yes	Yes
N	1,898,380	809,630

Notes: 1. All explanatory variables are standardized for each year.
3. Standard errors in parentheses.
4. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.3 Marginal Effects for Monitor Entry Using Panel Linear Probability Models

	(1) Central Monitor	(2) Local Monitor
L.P.M. 2.5	2.0e-4***	1.6e-5
	(7.4e-5)	(2.9e-4)
L.Length of Highway	-4.1e-4***	7.6e-4***
	(2.7e-5)	(8.4e-5)
L.Population Density	0.0545***	-0.0072***
	(5.8e-4)	(0.0018)
L.Distance to the Nearest Central Monitor	6.6e-4***	-1.5e-4
	(4.2e-5)	(1.6e-4)
L.Distance to the Nearest Local Monitor	2.8e-4***	3.2e-4***
	(5.1e-5)	(7.9e-5)
L.GDP of Primary industries	0.0019***	0.0023***
	(7.9e-5)	(2.5e-4)
L.Government Revenue	0.0028***	0.0023***
	(1.2e-4)	(3.2e-4)
N	1,909,585	1,130,935

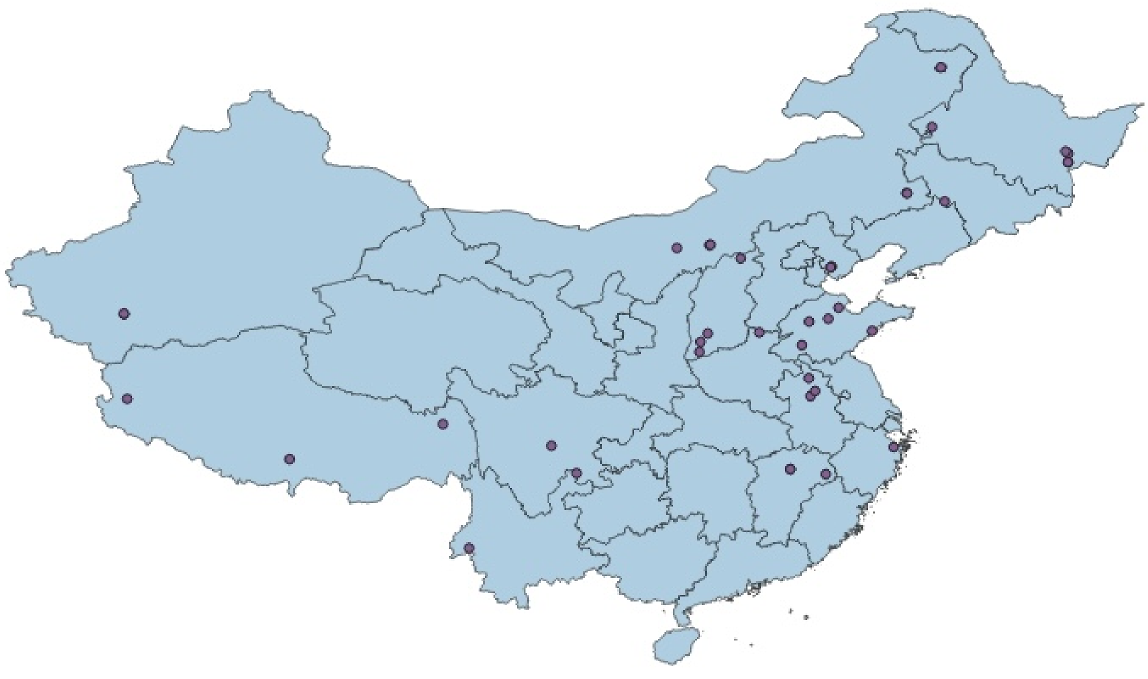
Notes: 1. All explanatory variables are standardized.
2. Standard errors in parentheses.
3. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.4 Marginal Effects for Monitor Entry Using Pooled Linear Probability Models

	(1) Central Monitor	(2) Local Monitor
L.P.M. 2.5	-3.5e-5*	5.6e-4***
	(1.9e-5)	(6.8e-5)
L.Length of Highway	2.6e-4***	0.0026***
	(1.4e-5)	(4.3e-5)
L.Population Density	0.0011***	0.0041***
	(1.5e-5)	(4.6e-5)
L.Distance to the Nearest Central Monitor	2.0e-5	5.7e-05
	(1.5e-5)	(5.5e-5)
L.Distance to the Nearest Local Monitor	1.5e-4***	2.7e-4***
	(2.6e-5)	(6.1e-5)
L.GDP of Primary industries	3.3e-4***	4.2e-4***
	(1.4e-5)	(4.3e-5)
L.Government Revenue	-2.1e-4***	-3.3e-4***
	(1.4e-5)	(4.5e-5)
District Dummy	0.0012***	4.0e-4***
	(5.2e-05)	(1.5e-4)
County Dummy	-1.8e-4***	3.0e-04***
	(5.2e-05)	(1.0e-04)
Yearly and Provincial Dummies	Yes	Yes
N	1,909,585	1,130,935

Notes: 1. All explanatory variables are standardized.
2. Standard errors in parentheses.
3. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

(a) 2019



(a) 2020

Figure B.9 Central Monitor Entry 2019-2020

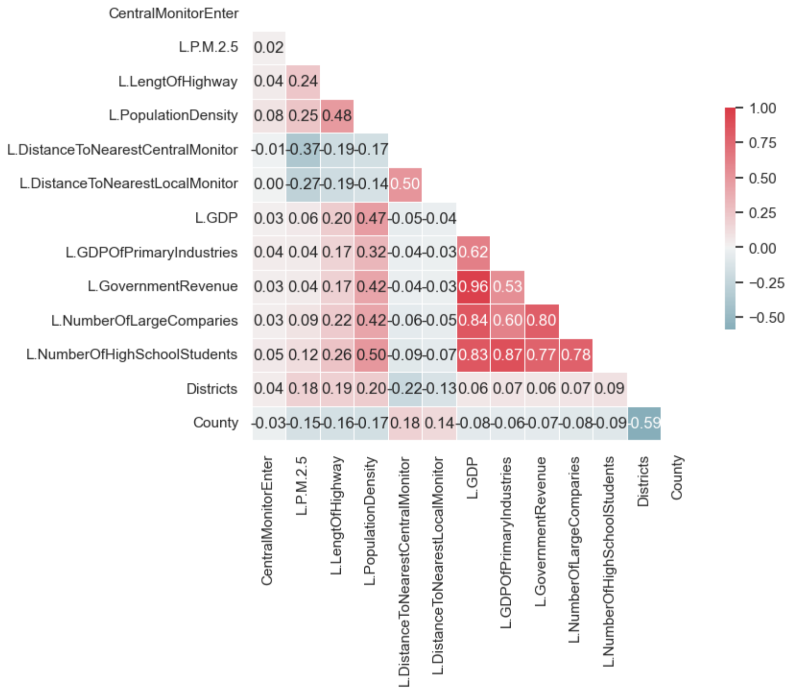


Figure B.10 Correlation Graph of Central Monitor Entries with Determining Factors

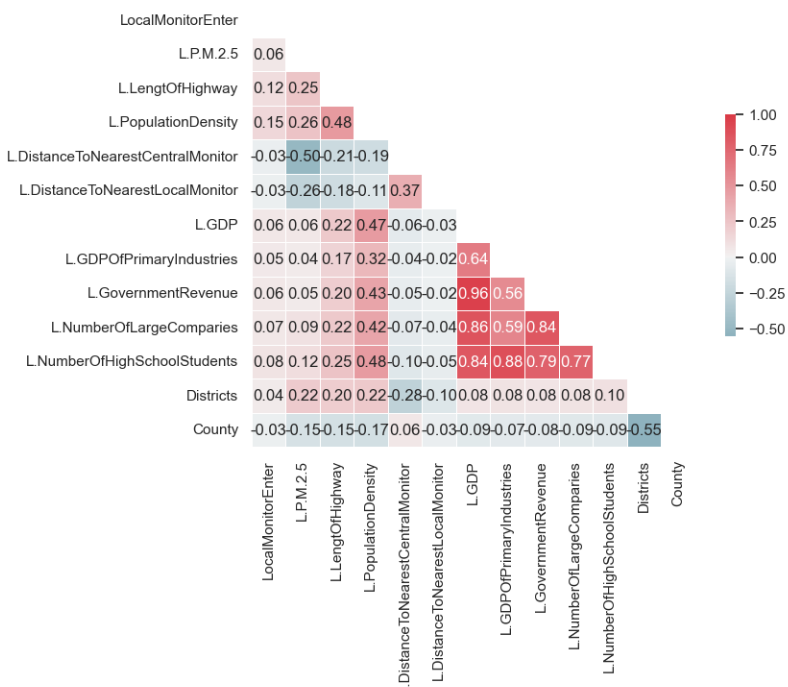


Figure B.11 Correlation Graph of Local Monitor Entries with Determining Factors

References

Akerlof, G. A. (1978). The market for “lemons”: Quality uncertainty and the market mechanism. In *Uncertainty in economics* (pp. 235–251). Elsevier.

Banzhaf, H. S., & Walsh, R. P. (2008). Do people vote with their feet? an empirical test of tiebout’s mechanism. *American economic review*, *98*(3), 843–863.

Barwick, P. J., Li, S., Lin, L., & Zou, E. (2019). *From fog to smog: The value of pollution information* (tech. rep.). National Bureau of Economic Research.

Bennear, L. S. (2008). What do we really know? the effect of reporting thresholds on inferences using environmental right-to-know data. *Regulation & Governance*, *2*(3), 293–315.

Bolton, G. E., Katok, E., & Ockenfels, A. (2004). How effective are electronic reputation mechanisms? an experimental investigation. *Management science*, *50*(11), 1587–1602.

Chen, D. L., Schonger, M., & Wickens, C. (2016). Otree—an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, *9*, 88–97.

Chevalier, J. A., & Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of marketing research*, *43*(3), 345–354.

Currie, J., Davis, L., Greenstone, M., & Walker, R. (2015). Environmental health risks and housing values: Evidence from 1,600 toxic plant openings and closings. *American Economic Review*, *105*(2), 678–709.

Darby, M. R., & Karni, E. (1973). Free competition and the optimal amount of fraud. *The Journal of law and economics*, *16*(1), 67–88.

Dulleck, U., Kerschbamer, R., & Sutter, M. (2011). The economics of credence goods: An experiment on the role of liability, verifiability, reputation, and competition. *American Economic Review*, *101*(2), 526–555.

Emons, W. (1997). Credence goods and fraudulent experts. *The RAND Journal of Economics*, 107–119.

Ert, E., Fleischer, A., & Magen, N. (2016). Trust and reputation in the sharing economy: The role of personal photos in airbnb. *Tourism management*, *55*, 62–73.

Fewster, R. M. (2009). A simple explanation of benford's law. *The American Statistician*, *63*(1), 26–32.

Finger, S. R., & Gamper-Rabindran, S. (2013). Mandatory disclosure of plant emissions into the environment and worker chemical exposure inside plants. *Ecological economics*, *87*, 124–136.

Gibson, M. (2019). Regulation-induced pollution substitution. *Review of Economics and Statistics*, *101*(5), 827–840.

Graham, M., & Miller, C. (2001). Disclosure of toxic releases in the united states. *Environment: Science and Policy for Sustainable Development*, *43*(8), 8–20.

Grainger, C., Schreiber, A., & Chang, W. (2016). *How states comply with federal regulations: Strategic ambient pollution monitoring* (tech. rep.). Working paper, University of Wisconsin-Madison.

Grant, D. S. (1997). Allowing citizen participation in environmental regulation: An empirical analysis of the effects of right-to-sue and right-to-know provisions on industry's toxic emissions. *Social science quarterly*, 859–873.

Gray, W. B., & Shadbegian, R. J. (2005). When and why do plants comply? paper mills in the 1980s. *Law & Policy*, *27*(2), 238–261.

Gray, W. B., & Shadbegian, R. J. (2007). The environmental performance of polluting plants: A spatial analysis. *Journal of Regional Science*, *47*(1), 63–84.

Haas, T. C. (1992). Redesigning continental-scale monitoring networks. *Atmospheric Environment. Part A. General Topics*, *26*(18), 3323–3333.

Hamilton, J. T. (1995). Pollution as news: Media and stock market reactions to the toxics release inventory data. *Journal of environmental economics and management*, *28*(1), 98–113.

Hanna, R. N., & Oliva, P. (2010). The impact of inspections on plant-level air emissions. *The BE Journal of Economic Analysis & Policy*, *10*(1).

He, S., Hollenbeck, B., & Proserpio, D. (2022). The market for fake reviews. *Marketing Science*.

Hill, T. P. (1995). A statistical derivation of the significant-digit law. *Statistical science*, 354–363.

Kahn, M. E., Li, P., & Zhao, D. (2015). Water pollution progress at borders: The role of changes in china's political promotion incentives. *American Economic Journal: Economic Policy*, *7*(4), 223–42.

Koehler, D. A., & Spengler, J. D. (2007). The toxic release inventory: Fact or fiction? a case study of the primary aluminum industry. *Journal of Environmental Management*, *85*(2), 296–307.

Liu, L. (2012). Analysis of firm compliance with multiple environmental regulations. *Available at SSRN 2127370*.

Lu, W.-Z., He, H.-D., & Dong, L.-y. (2011). Performance assessment of air quality monitoring networks using principal component analysis and cluster analysis. *Building and Environment*, *46*(3), 577–583.

Luca, M. (2011). Reviews, reputation, and revenue: The case of yelp. com (no. 12-016). *Harvard Business School*.

Luca, M., & Zervas, G. (2016). Fake it till you make it: Reputation, competition, and yelp review fraud. *Management Science*, *62*(12), 3412–3427.

Lynch, M., Miller, R. M., Plott, C. R., & Porter, R. (1991). Product quality, informational efficiency, and regulations in experimental markets. *Research in experimental economics*, *4*, 269–318.

Lynn, F. M., & Kartez, J. D. (1994). Environmental democracy in action: The toxics release inventory. *Environmental Management*, *18*, 511–521.

Maji, K. J., Dikshit, A. K., & Deshpande, A. (2017). Can fuzzy set theory bring complex issues in sizing air quality monitoring network into focus? *International Journal of System Assurance Engineering and Management*, *8*(4), 2118–2128.

Marchi, S. d., & Hamilton, J. T. (2006). Assessing the accuracy of self-reported data: An evaluation of the toxics release inventory. *Journal of Risk and uncertainty*, *32*, 57–76.

Mastromonaco, R. (2015). Do environmental right-to-know laws affect markets? capitalization of information in the toxic release inventory. *Journal of Environmental Economics and Management*, *71*, 54–70.

Mayzlin, D., Dover, Y., & Chevalier, J. (2014). Promotional reviews: An empirical investigation of online review manipulation. *American Economic Review*, *104*(8), 2421–55.

Muller, N. Z., & Ruud, P. A. (2018). What forces dictate the design of pollution monitoring networks? *Environmental Modeling & Assessment*, *23*, 1–14.

Nelson, P. (1970). Information and consumer behavior. *Journal of political economy*, *78*(2), 311–329.

Nelson, P. (1974). Advertising as information. *Journal of political economy*, *82*(4), 729–754.

Nigrini, M. J. (1996). A taxpayer compliance application of benford's law. *The Journal of the American Taxation Association*, *18*(1), 72.

Oestreich, A. M. (2015). Firms' emissions and self-reporting under competitive audit mechanisms. *Environmental and Resource Economics*, *62*, 949–978.

OpenStreetMap contributors. (2017). Planet dump retrieved from https://planet.osm.org.

Pires, J., Sousa, S., Pereira, M., Alvim-Ferraz, M., & Martins, F. (2008). Management of air quality monitoring using principal component and cluster analysis—part i: So2 and pm10. *Atmospheric Environment*, *42*(6), 1249–1260.

Plott, C. R., & Wilde, L. L. (1980). *Professional diagnosis vs. self-diagnosis: An experimental examination of some special features of markets with uncertainty*. California Institute of Technology, Division of the Humanities; Social Sciences.

Qi, Y., & Zhang, L. (2014). Local environmental enforcement constrained by central–local relations in china. *Environmental Policy and Governance*, *24*(3), 216–232.

Resnick, P., Kuwabara, K., Zeckhauser, R., & Friedman, E. (2000). Reputation systems. *Communications of the ACM*, *43*(12), 45–48.

Shimshack, J. P., & Ward, M. B. (2005). Regulator reputation, enforcement, and environmental compliance. *Journal of Environmental Economics and Management*, *50*(3), 519–540.

Solimine, P., & Isaac, R. M. (2023). Reputation and market structure in experimental platforms. *Journal of Economic Behavior & Organization*, *205*, 528–559.

Wolinsky, A. (1993). Competition in a market for informed experts' services. *The RAND Journal of Economics*, 380–398.

Xu, H., Bechle, M. J., Wang, M., Szpiro, A. A., Vedal, S., Bai, Y., & Marshall, J. D. (2019). National pm2. 5 and no2 exposure models for china based on land use regression, satellite measurements, and universal kriging. *Science of the Total Environment*, *655*, 423–433.

Yang, L., Lin, Y., Wang, J., & Peng, F. (2020). *Pollution monitoring, strategic behavior, and dynamic representativeness* (tech. rep.). Working paper.

Yu, T., Wang, W., Ciren, P., & Sun, R. (2018). An assessment of air-quality monitoring station locations based on satellite observations. *International Journal of Remote Sensing*, *39*(20), 6463–6478.

Zou, E. Y. (2021). Unwatched pollution: The effect of intermittent monitoring on air quality. *American Economic Review*, *111*(7), 2101–2126.

Vita

Sharad KC is a Ph.D. Candidate in the Economics Department Andrew Young School of Policy Studies at Georgia State University. He is expected to graduate in the Summer of 2024. He holds an Masters degree in Economics from Georgia State University, Atlanta, GA. His research interests are applied microeconomics, experimental economics, environmental economics, and data science. His dissertation focuses on individual, corporate and government decision making in information as a public good. In his dissertation in the doctorate program, he conducted research using lab experiment and causal inference methods to find the effect of fake reviews on market outcomes, air quality monitor placement, and corporate reporting behavior. He is also interested in different research methods, such as lab in the field methods, and is looking forward to more research using these methods.