Essays on the Economics of Digital Markets

Kyung Sun Rhee

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2019

Reading Committee:

Yong Tan, Chair

Jinyang Zheng

Hyeunjung (Elina) Hwang

Program Authorized to Offer Degree:

Business Administration

University of Washington

Abstract

Essays on the Economics of Digital Markets

Kyung Sun Rhee

Chair of the Supervisory Committee:
Michael G. Foster Endowed Professor Yong Tan
Information Systems and Operations Management at Foster School of Business

This dissertation provides a comprehensive investigation of the IT economy in the digital market by examining user behaviors and institutions' strategies. The essays include research questions from two different dynamic settings and answer questions by applying corresponding empirical setups. The first essay examines the impact of technology restriction and demand shifts in transportation dynamics. A transportation network company (TNC) is an app-based, two-sided platform that matches passengers with vehicles through information sharing. In 2015, the Shanghai government introduced a policy to restrict taxi drivers' use of TNC apps during certain hours. I leverage this exogenous policy shock to investigate the economic impact of information sharing on the use of taxis and the entire transportation dynamics. I identify (1) a significant

decrease in the total number of taxi transactions during the enforcement time but a significant increase during non-enforcement times, (2) simultaneous traffic increase on public transportation during policy enforcement and non-enforcement times (3) increase in congestion level on surface streets, indicating a spillover to private vehicle use for short-distance rides. Interestingly, my mechanism analysis reveals an increased average travel distance after restricting information sharing suggesting the dominant mechanism of information sharing. The second essay investigates the right online connection that helps users in the job market. Recognizing recruiters' increasing reliance on professional networking sites, job seekers strive to create as many professional connections as possible. In particular, a common tactic is to make LinkedIn connections with professionals who work for the target companies in the job seekers' target fields, as they can offer timely information about new job openings and become potential referrers. This paper empirically investigates whether such a networking tactic is actually instrumental in obtaining referrals. Much to my surprise, the analysis reveals that this common tactic is not a promising way to get referrals: I find that job seekers are less likely to be referred by employees who are in the target company and the target field due to the peer competition. I further find that this adverse effect of job similarity on referrals weakens as the hierarchical level of the referring employee gets higher than that of the job candidate, because they are less likely to compete directly. Although one might expect that gender homophily may weaken this competition effect, I find that it does not.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

Foremost, I would like to express my deepest gratitude to my advisor, Yong Tan. I am grateful for his unconditional support, guidance, patience, encouragement, and immense knowledge. Throughout the process, he was always there for me. In hard times, he has not given up on me even though I gave up myself. For my achievements, he was happier than me and celebrated with me. I truly enjoyed every encounter with him and felt privileged having him as my advisor. I will always remember and cherish the support he has given me, without which I might never have done this far of doctoral study.

My thanks also go to the members of my dissertation committee: Ming Fan, Jinyang Zheng, Hyeunjung (Elina) Hwang, and Dong-Jae Eun. I was very fortunate to have a chance to closely work with Jinyang Zheng, whom I express my sincere thanks not only for his tremendous guidance on the research but also for his caring as an academic brother. My appreciation also goes to Hyeunjung (Elina) Hwang, as her insights and passion for research have certainly inspired me. I am thankful to Ming Fan for his knowledgeable and far-sighted suggestions to improve this dissertation. I am also thankful to Dong-Jae Eun who has offered me helpful and insightful comments on my dissertation.

A very special gratitude goes out to Jaime Banaag, Shawna Reimers, Jessica Aceves, Beau Kirkeby and other staff at the Foster School of Business for their ceaseless help during the five-years study. I thank so much to my Mackenzie crew for the intensive discussions on the research and their unlimited emotional support that helped me to survive the challenging process of study:

Sharyhar Doosti, Omid Rafieian, Maggie Meng, Eugene Pavlov, Mohammad Arbabian, Keehae Moon, Aravinda Garimella, Amirezza Fazli, and David Shin. Thanks for staying with me when I am struggling and sharing the happiness of my achievements. I am also grateful to Mr. MacCabe, Mr.Martinex, Mrs. O'Connell-Cahill, Mr. Cahill for helping and providing the funding for the work. Your support has increased the chances of achievements and success in my research career.

DEDICATION

I dedicate my dissertation work to my dear family. A very special feeling of gratitude to my loving parents, Yeong Hyun Rhee and Sun Hee Rhee, whose words of encouragement, caring and push for tenacity helped me to what I have accomplished today. I will never finish thank you for everything you do every day for me. With a special mention to my brother Jae Sang and sister Myun Sun who never left me feeling alone and are very special.

Last, not the least, a very special thanks to my grandma, Yoon Hwan Kim, in memoriam. I believe you are always with me wherever I go.

Chapter 1. INTRODUCTION

Improvements in technology and the introduction of two-sided platforms redirected ways of consumption, running businesses and enacting regulations. Accordingly, features of digital technology created enormous social and economic values from the perspective of users as well as the organizations and the public society. However, there is a concern of growing inequality between the people who can use such technologies to advantage and those who don't. Further, there is an equality of the benefits among the users who are involved in the digital markets. In this dissertation, I examine how the economic value is created from newly rising digital markets, and whether unfairness problems can be potentially addressed. This dissertation provides a comprehensive investigation of the IT economy in the digital market by examining user behaviors and institution strategies. The essays include research questions from two different dynamics and answer those questions by applying corresponding empirical setups.

The first essay examines the impact of technology restriction and demand shifts in transportation dynamics. A transportation network company (TNC) is an app-based, two-sided platform that matches passengers with vehicles through information sharing. In 2015, the Shanghai government introduced a policy to restrict taxi drivers' use of TNC apps during certain hours. I leverage this exogenous policy shock and collect comprehensive data on various transportation usage to investigate the economic impact of information sharing on the use of taxis and the entire transportation dynamics. I identify a significant decrease in the total number of taxi transactions during the enforcement time but a significant increase during non-enforcement times. Further, the restriction simultaneously increases traffic on public transportation and during non-enforcement times. Moreover, the restriction leads to increasingly congested surface streets during both policy enforcement and non-enforcement times, indicating a spillover to private vehicle use for short-distance rides. Interestingly, my mechanism analysis reveals an increased average distance after restricting information sharing, including sharing destinations, suggesting the dominant mechanism of information sharing as expanding drivers' demand choice set, alleviating their uncertainty over idle time and, thus, encouraging them to take shorter orders due to their higher marginal profit. From an academic perspective, I contribute to the information systems literature by isolating the economic value of information sharing for a TNC and the transportation systems and by providing information-sharing designs to improve TNC and transportation efficiency and manage drivers' selectivity behavior. Practically, I highlight how information sharing should be considered in the platform design and regulation of a TNC.

The second essay looks into the right online connections that are instrumental for users in the job market. According to a recent survey, LinkedIn comprises over 500 million professionals, and the majority of recruiters utilize LinkedIn to identify qualified job candidates. Recognizing recruiters' increasing reliance on professional networking sites, job seekers strive to create as many professional connections as possible. In particular, a common tactic is to make LinkedIn connections with professionals who work for the target companies in the job seekers' target fields, as they can offer timely information about new job openings and become potential referrers. This paper empirically investigates whether such a networking tactic is actually instrumental in obtaining referrals. For empirical analysis, I use unique data from a global consulting company that utilizes LinkedIn for its hiring. Much to my surprise, the analysis reveals that this common tactic is not a promising way to get referrals: I find that job seekers are less likely to be referred by employees who are in the target company and the target field. I attribute this result to peer competition: Employees tend not to refer qualified candidates with similar expertise in order to protect themselves from future competition. I further find that this adverse effect of job similarity on referrals weakens as the hierarchical level of the referring employee gets higher than that of the job candidate, because they are less likely to compete directly. Although one might expect that gender homophily may weaken this competition effect, I find that it does not.

Chapter 2. LITERATURE REVIEW

I investigate the economic value of digital markets in two different contexts. First, I overlay the related research on the effect of transportation network companies (TNC)'s an information sharing on the urban transportation system. Next, I review the theoretical and empirical literature on referral-based job searches. In particular, it is related to the streams of research that examine (1) the benefit of referral-based hiring, (2) the factors that influence the decision to refer, and (3) the role of an online social network in a job search.

2.1 Transportation Network Company

2.1.1 *Economic Value of Information Sharing and TNCs*

With the surge of interest in TNCs, studies in various fields have focused on the impact of mobile-based, on-demand TNC services on individuals' behaviors, industry sectors, and public society. One stream of literature attempts to investigate particular features, typically those more related to a certain field, embedded in TNC services, e.g., dynamic pricing and online promotion at the level of individual users. Zheng et al. (2016) explore the impact of a two-sided sales promotion on drivers' willingness to be involved in a TNC platform and how the platform provides optimal promotions accordingly. The authors construct a structural model of drivers' decisions in the presence of a sales promotion and adapt Bayesian learning processes to cover decisions under uncertainty. Cohen et al. (2016) focus on the effects of dynamic pricing on passengers and examine the variation in surging prices. They estimate the elasticity of demand for Uber X and find that passenger demand is inelastic to the price change and results in a large consumer surplus. Hall et al. (2017) further extend the investigation on dynamic pricing to the driver side and find that, when the fare increases, drivers' hourly earning rates rise at the initial stage but decline somewhat to become constant. They interpret this as due to the fall in drivers' utilization; that is, drivers work less when the fare is high. Lam and Liu (2019) show that the dynamic pricing in ridesharing might help to mitigate geographical disparity. Chen et al. (2017) find that the real-time flexibility of Uber's working environment generates a surplus for drivers. Although information sharing plays a esssential role in the value creation of TNCs, to the best of my knowledge, very limited research has investigated the effects and the economic value of information sharing through information technology in TNCs.

Another research stream examines the TNC platform's impact and, thus, provides an understanding of the societal impact of TNCs as a whole. Rogers (2015) quantifies the social costs of Uber, e.g., political regulation, labor standards, while Wallsten (2015) examines the competitive impact of Uber entry on the taxi industry and finds that Uber's increasing popularity is associated with a decline in consumer complaints about taxis in New York City and a reduction in a particular type of complaints about taxis in Chicago. Rayle et al. (2016) also examine the substitution effect of TNCs on taxis. They focus on who and why certain people use ride-sharing services as compared with other transportation modes. They find that, to a certain extent, TNCs replace taxi trips and transportation modes, as they decrease passengers' waiting time and enable point-to-point travel. Li et al. (2016) investigate Uber entry's impact on traffic congestion and show that ride-sharing services significantly reduce traffic congestion in an urban area. Gong et al. (2017) investigate Uber entry's impact on vehicle purchase and find a considerable increase in new vehicle ownership. Greenwood and Wattal (2017) and Park et al. (2017) extend the scope of research on Uber's impact to public safety and find that the use of Uber decreases the rate of alcohol-impaired crashes and rape occurrences. The impact found in this stream of research, however, cannot be attributed to information sharing because the findings concern Uber as a whole and could be attributed to other factors. For example, the Uber platform uses dynamic pricing and brings additional supply to transportation systems, which might explain its substitution and competitive effect more than does information sharing.

My research contributes to this literature stream by studying the impact of information sharing in a TNC on the urban transportation systems. The Tech Restrict Policy enforced by the government provides a valuable opportunity to clearly identify the impact of information sharing. A comprehensive analysis of multiple forms of spillover effects from the policy on transportation also illuminates the economic value of information sharing in a TNC. My mechanism analysis further reveals the value-creation process of information sharing. Those findings not only contribute to the research stream on the economic value of TNCs by enhancing the understanding of how a TNC creates values for transportation systems but also provides a direct contribution to the growing literature stream about the impacts of information sharing (Bapna et al. 2012; Burtch et al. 2013; Chan and Ghose 2014; Greenwood and Agarwal 2016) through extending the scope of research on information sharing in TNCs.

2.1.2 *Designing a Transportation Market*

Efficiency has long been a major objective for the market design of transportation systems; thus, earlier transportation literature focused largely on various perspectives of market design to improve efficiency. For example, Gan et al. (2013) propose a pricing scheme to optimize the efficiency of a taxi system. Zhan et al. (2014) investigate the efficiency of taxis from a recommendation system perspective and show a significant gain (20% to 90% reduction of vacant time) if a centralized system-wide recommendation scheme is applied. Further, solutions for efficient routing/scheduling of mass transit vehicles that resolve problems of traffic congestion and demand management have been studied (Arnott 2001; Balakrishna et al. 2008; Farahani et al. 2013; Cheng et al. 2019).

As the traditional transportation market was not based on transparent platforms, previous studies were less focused on the solutions that can be derived from the platform's attributes. Considering the main advantage of utilizing online platforms is information sharing due to its nature of having two-sided markets, the recent advent of TNCs shifts the focus of market design as related to information sharing. Such work suggests the value of information sharing for market efficiency based on a theoretical framework as well as on empirical evidence from earlier counterparts of two-sided online platforms (Bakos 1997; Bailey and Bakos 1997; Bakos and Katsamakas 2008). For instance, Bakos and Bailey (1997) argue that platforms improve market efficiency by ensuring the integrity of the market and facilitating the buyer-seller match. It is also shown that increasing information transparency and reducing information asymmetries improve market efficiency (Brynjolfsson and Smith 2000; Brynjolfsson et al. 2003; Rochet and Tirole 2004). Hagiu and Halaburda (2014) study the effect of different levels of information on two-sided platform profits. Wang et al. (2018) and Ozkan and Ward (2017) showed empirically that, by providing additional information on passengers, e.g., passenger arrival rates, willingness to wait, drivers could reduce mismatches between drivers and passengers. Wang et al. (2019) also showed empirically that providing the destination of passengers helps drivers' forward-looking planning and, thus, improves their utility maximization efficiency.

Another objective of the market design for TNC platforms is fairness. Although the literature on improving the fairness of TNCs through information sharing is extremely limited, there is a related literature stream in regard to digital inequality that concerns an inequality in the access and use of information and communication technologies (DiMaggio et al. 2004; Kvasny 2002; Van

Dijk and Hacker 2003). Bertsimas et al. (2011, 2012) examine the trade-off between efficiency and fairness for market designers and have developed several managerial prescriptions developed from a selected objective. The most relevant work on the perspective of information sharing is that of Wang et al. (2019), who show that the improved efficiency of drivers, by revealing passengers' destination, sacrifices the fulfillment rate of passengers.

My work contributes to this literature on the following dimensions: First, I show the value of information sharing for efficiency for the market design of taxis; Second, for the market design of TNCs, I extend the scope of efficiency to externality effects. Last, I consider information sharing design that simultaneously maintains the efficiency and allows the platform manager to moderate the selectivity behavior of drivers and, thus, to incorporate the concern on fairness. The findings further generate practically valuable insights for TNC platform designers and transportation policymakers regarding how to manage the selectivity behavior of drivers.

2.2 Social Hiring

This essay contributes to the theoretical and empirical literature that studies referral-based job searches. In particular, my study is related to the streams of research that examine (1) the benefit of referral-based hiring, (2) the factors that influence the decision to refer, and (3) the role of an online social network in a job search.

2.2.1 *The Benefit of Employee Referral-based Recruiting*

A growing body of literature has explored the economic outcomes of employee referral-based hiring (Fernandez et al. 2000; Mouw 2003; Castilla 2005; Pallais and Sands 2016, Burks et al. 2015, Dustmann et al. 2016). Burks et al. (2015) quantified the benefits of employee referrals in terms of overall profits of a firm. By comparing referred candidates to non-referred candidates across three different industries, the authors showed that referred candidates are more likely to take an offer, conditional on receiving offers. Interestingly, this pattern persists even when referred and non-referred candidates are similar on observable and unobservable characteristics. The authors also found that new hires through employee referrals exhibit higher productivity than non-referred hires. For instance, referred hires experienced lower accident rates in trucking firms and higher patent inventions in high-tech firms. They also found that referred hires are less likely to quit.

Several economists have used theoretical models to explain why firms benefit from employee referrals. Simon and Warner (1992) used Jovanovic (1979, 1984)'s job matching model and showed that employees' "old boy" networks reduce companies' uncertainty about employees' productivity. They demonstrated that referred employees earn higher initial rates due to a good match and stay on a job longer compared to other employees hired outside the network. Dustmann et al. (2016) built on Simon and Warner (1992) and a learning model by Jovanovic (1979, 1984) to show that utilizing employees' networks in a job search helps companies reduce uncertainty and information deficiencies about the labor market. Consequently, employee referrals ensure that new employees are better matched to their firms than those hired through external markets. The authors demonstrated that, as a result, employee referrals result in an approximately 2% lower probability of a new hire leaving the firm. Brown et al. (2016) provided corroborating empirical evidence for the above theoretical results.

Another line of research has empirically examined how employee referrals benefit job seekers. Using geographically detailed employer-employee level data, Schmutte (2014) showed that local referral networks within narrowly defined neighbors have a significant positive effect on the earnings outcomes from job searches. The author explained that this pattern occurs because referral networks facilitate the exchange of information about particularly attractive job opportunities (i.e., jobs with higher wage premiums).

My paper complements this line of research studying post-referral-stage outcomes (i.e., how firms and job seekers benefit from employee referrals) by investigating the pre-referral-stage factors that affect employee referrals. This investigation is important to understanding the environments, or the conditions, that foster employee referral. Knowing such factors helps us understand the underlying mechanisms of why employee referrals bring superior economic outcomes to firms as well as to job seekers.

2.2.2 *The Factors influencing the Decision to Refer*

Given the benefit of referral-based job search, scholars have examined the factors that influence employees' decisions to refer. First, network scholars examined how network tie strength between a referrer and a potential referral affects referral outcomes (Liu and Duff 1972; Granovetter 1973; Yakubovich 2005; Marin 2012; Gee et al. 2017). Tie strength is a relative measure of relationship intensity based on the amount of time or the emotional intensity people share. For instance, tie strength with a close friend whom I meet every day is stronger than tie strength with acquaintances. Studies have documented that, although strong ties are more helpful in gaining job offers (Obukhova 2012), weak ties enable the discovery of unknown job leads (Yakubovich 2005, Montgomery 1992).

Second, homophily is documented as one of the main factors that influence employees' referral decisions (e.g., Rees 1966, Verbrugge 1977). Homophily is the principle that describes the preference of people to favor similar others more than dissimilar others (McPherson et al. 2001). The homophily principle explains why I tend to form friendships with others of similar age, education, ethnic background, etc. In the referral setting, studies have found a strong influence of gender homophily: The same-gender referrals are made significantly more frequently than different-gender referrals (Berger 1995, Fernandez and Sosa 2005). In some cases, such preference for the same gender referrals is so strong that it even leads to segregation of jobs by gender (Bielby and Baron 1986, Chan and Wang 2018, Kalleberg et al. 1996).

In addition to the above-mentioned dyadic characteristics of a referrer and a job seeker, a few individual characteristics of a referrer or a job seeker are found to influence a referrer's decision to refer. For instance, a job seeker's qualification is found to affect employees' referral decisions. That is, employees tend not to refer candidates with low qualifications, to protect their own reputations. As referrals represent connections between referrers and referred candidates, referrers fear that their referrals might prove unreliable and might compromise their own reputations with employers (Burke et al. 2015, Pieper et al. 2018, Smith 2005). Because of this potential risk, employers prescreen the quality and fit of job seekers before making referrals (Beaman and Magruder 2012, Rees 1966, Vecchio 1995).

My work extends this research stream by investigating how job similarity between a referrer and a job seeker influences an employee's referral decision and how its effect is moderated by a hierarchical difference and gender homophily. By simultaneously considering various influencing factors (i.e., job similarity, hierarchical difference, gender homophily, tie strength, individual characteristics), my study offers a more nuanced understanding of the factors that influence referral decisions.

2.2.3 *The Role of Online Network Ties on Job Search*

The increasing use of online social networks by job seekers and recruiters has sparked the interest of scholars in understanding the role of online social networks in job search outcomes. One group of scholars explored how recruiters utilize online social networks in their hiring process. Studies found that recruiters attract and screen candidates using the published information in online social networks (Ollington et al. 2013). Conducting a field experiment, Manant et al. (2014) provided evidence that online Facebook profiles are used in the early stage of the job hiring process in the French labor market. The authors further found that the search for extra information on candidates through the social network platform is not conducted in-depth: It is more likely that recruiters limit their searches to the front page of profiles.

Even with such limited use, it is found that information disseminated through online networking platforms influences job search outcomes. For instance, Acquisti and Fong (2015) examined how personal information published in online networking platforms, especially information that is hard to obtain in a formal application or interview process, affects the job search outcome. From randomized experiments, the authors found that the online disclosure of personal features, which are illegal for companies to request (e.g., religion, sexual orientation), influences the hiring decisions of U.S. firms: There are 13% fewer callbacks for Muslim candidates than for Christian candidates, nationwide. Based on this finding, the authors claim that shared information in online networking sites creates unexpected hiring discrimination.

Another group of scholars focused on how the strength of online network ties affects job search outcomes. As discussed in §2.2, prior empirical studies conducted in an offline setting (e.g., Granovetter 1973, Montgomery 1992) found that weak ties are more useful in finding job leads than strong ties. In contrast, in the context of online network ties (e.g., Facebook), Burke and Kraut (2013) found that strong ties are more helpful for finding a job than weak ties. Similarly, Garg and Telang (2017) found that strong LinkedIn connections are more beneficial than weak ties for obtaining job leads. This discrepancy in findings across offline and online settings suggests that prior findings based upon an offline context may not be safely extrapolated to the online context.

Given the growing reliance of job searches on online professional networking sites, this study aims to advance my understanding of its impact on job referrals.

Chapter 3. TECHNOLOGY RESTRICTION IN DIGITAL MARKETS – TRANSPORTATION NETWORK COMPANY

3.1 INTRODUCTION

A transportation network company (TNC), also known as a ride-hailing service, matches passengers with vehicles via a website or mobile app (California Public Utilities Commission 2013). Known through established companies such as Uber, Lyft, Hailo, and DiDi, a TNC creates a two-sided market through information technology to connect two distinct user groups, drivers and passengers, and provides each with network values (Hagiu and Wright 2014). The key benefit of TNC services for the drivers is the information sharing that reduces the information asymmetry (DePillis 2016, Langfitt 2014). More specifically, a TNC provides information on potential passengers, i.e., pick-up points and destinations, which not only reduces drivers' searching cost but also enables drivers to better plan future rides. Similarly, on the passenger side, a TNC provides driver information, i.e., waiting times, estimated prices, and profiles (White 2015), which enables passengers to make an appointment with simplicity and flexibility.

As one of the many features of a TNC, the value of information sharing is easily disguised or overshadowed by the compound effect of various features. Some TNCs in the United States (U.S.), such as Uber and Lyft, depend on peer-to-peer drivers or introduce a competitive pricing scheme, which challenges the incumbent industry, e.g., taxi (Dickenson 2018). This results in concerns about passenger safety (LaFrance and Eveleth 2015), increased traffic congestion (Wolfe 2018), and reduced use of public transportation (Badger 2017), among others. Platform designers and regulators overlook the various attributes of TNCs and launch policies to completely restrict TNC operations. For instance, in 2016, Austin, Texas, ruled out the business of the two biggest ride-sharing companies, Uber and Lyft, after being faced with regulations that required drivers' fingerprints as a means to address public safety issues (McPhate 2016). Through addressing the negative effects of TNCs, such policies simultaneously erased the potential values generated from TNCs. To best extract the economic value of TNCs, while eliminating their adverse effects, practitioners should disentangle the economic value from their other components so that a precisely guided policy in regard to platform design can be established.

From the standpoint of information systems research, I aim to examine the economic value of the most essential component of TNC apps: information sharing. Functioning as a two-sided market, a TNC leverages information technology to share the information of the users of one side, e.g., passengers, to the users of the other side, e.g., drivers. The information sharing alleviates the information asymmetry between the two sides, generating better matching for the passengers with drivers/vehicles and, thus, is the core of a TNC two-sided market. I leverage a unique policy shock launched in a well-controlled TNC market in Shanghai, where TNCs work only with licensed taxi drivers, and no peer-to-peer drivers are allowed. Taxi drivers have access to both online (TNC apps) and offline (road hailing) channels, and, thus, this policy restricts taxi drivers' use of TNC apps for a certain period of the day to protect the passengers with less access to technology (apps). During my observation period, the TNC market structure rules out additional TNC supplies from peer-to-peer drivers with private vehicles and controls over the pricing scheme. This results in restricting taxi drivers' apps usage as equivalent to restricting their information sharing, hence, successfully isolating and extracting the true economic value of information sharing for a TNC.

To comprehensively gauge the economic value of information sharing in the TNC market, I build an econometric model that examines how restricting drivers' apps usage affects the TNC (taxi) operations during the policy enforcement periods and investigate its externality effects on non-policy enforcement periods; alternative transportation modes, i.e., public transportation systems (metro, bus, ferry, or parking associated with metro) and private vehicles; and the combination of non-policy enforcement times and alternative transportation modes. I further uncover how restricting information sharing in apps usage affects drivers' behaviors by proposing two major underlying mechanisms of plausible opposite forces on drivers' average travel distances and empirically identifying the predominant mechanism.

My empirical analyses reveal findings that are robust to multiple specifications and falsification tests. First, restricting taxi drivers' apps usage significantly decreased the number of taxi transactions during the time period of policy but simultaneously increased the use of public transportation systems. The restriction also increased the congestion level on surface streets, which implies that there is an increased use of private vehicles for short-distance travels. Second, the restriction on apps usage significantly increased the number of taxi transactions during non-policy enforcement times. Finally, I observed an increased usage of public transportation systems and the congestion level during non-policy enforcement times. In sum, I suggest that restricting information sharing discourages the efficiency of taxi operations and causes traffic spillovers to alternative time or transportation modes, which generates a burden on the use of alternative transportation modes. Although passengers with less technology access may have benefited from the increased offline taxi rides after restricting drivers' apps usage, I conclude that there are significant side effects and costs due to reduced taxi efficiency and unexpected negative externalities, e.g., traffic on alternative transportation modes.

Although policymakers and, perhaps, the public believe that concealing certain information about a taxi order, e.g., destination, might discourage drivers' selectivity in orders, my mechanism analysis reveals that the average travel distance per transaction increased after restricting information sharing, including sharing destinations, for drivers. Considering that the taxi fare system in the city maximizes drivers' profit when they take more orders that are shorter in travel distance, my results suggest that the predominant mechanism of information sharing is to expand drivers' demand choice set, thus alleviating their uncertainty over future demand/idle time. The reduced risk of future idle time influences drivers not to insist on long-distance orders as a "safer" choice.

Theoretically, I advance the understanding of the economic value of information sharing in TNCs. Previous studies have documented the compound economic value of Uber from various perspectives (Gong et al. 2017, Greenwood and Wattal 2017, Li et al. 2016, Park et al. 2017, Rayle et al. 2016, Rogers 2015, Wallsten 2015). My study disentangles this compound value to extract the finer-grained value of information sharing. This not only enhances the understanding of TNCs' value creation but also widens the scope of the literature stream on information sharing to include the TNC context (Bapna et al. 2016; Burtch et al. 2013; Chan and Ghose 2014; Greenwood and Agarwal 2016). In addition, my study contributes to the literature stream on the TNC platform design in multiple ways: I show the extent to which information sharing improves the efficiency of a TNC platform and its impact on the overall transportation system. Further, I suggest designing information sharing in the TNC platform to moderate drivers' selective behavior, thus improving fairness for passengers. In terms of practical implementation, this study evaluates the side effects of restricting users' behavior via a technology restriction policy (Tech Restrict Policy) and its post-revision enactment to generalize how policy regulations should manage the users in two-sided market platforms. I also provide valuable managerial insight on a broader spectrum, i.e., how policymakers should manage information sharing in an aspect of the overall transportation ecosystem and how platform designers should design TNC platforms for efficient and optimal information sharing between the users.

The remainder of the paper is organized as follows. Section 3.2 presents the related literature and the contribution of this paper. Section 3.3 provides contextual background and data. Section 3.4 includes the identification strategy and model specification. Section 3.5 presents the results and a check of their robustness. Section 3.6 provides an extension of the scope of analysis to the underlying mechanism, and Section 3.7 concludes.

3.2 CONTEXT AND DATA

3.2.1 *TNCs and the Taxi Industry in Shanghai*

The general setup of the major TNCs in Shanghai, China, are similar to those in other countries. During my observation, two large TNCs, Didi and Kuaidi, had been dominating the TNC market for roughly one year. The TNCs establish a two-sided market to connect passengers with ride-providers and, accordingly, two app versions are provided to passengers and ride-providers, separately. The passenger-side app helps passengers to file requests for rides. In each request, passengers input their pick-up location and destination. The platform centrally pushes those requests, along with the passengers' pick-up location and destination information, to the active drivers nearby (typically within 3 km). To receive ride requests, ride-providers need to install and log into the driver-side apps on their smartphone. Figure 3.1 displays the information that is given to the ride-providers from the app.[1] Ride-providers can select/accept a request among incoming ones from their app and, after acceptance, fulfill that request by driving to the pick-up location to pick up the passengers and drive them to the destination. Both passengers and drivers can cancel a request during the process, but a penalty by the platform will be imposed. Once a request is completed, passengers pay the rider-provider the fare through an electronic payment system, e.g., WeChat Pay, or cash, and the ride-provider receives the fare immediately.

[1] Due to a data non-disclosure agreement, I cannot present the actual screenshot of the mobile application.

Distance between Pick-up point and the Driver
• Pick-up location • Destination location

Figure 3.1. Information that taxi drivers receive from the apps

During my observation, TNCs in China differed from Uber in the U.S. due to heavy regulation and the early stage of the TNC industry. Unlike other companies such as Uber, no TNC in China operated its own vehicles and crew during my observation period. Instead, TNCs operate through the existing taxis by licensed drivers.[2] TNCs, in my context, thus, do not provide additional supply to the urban transportation system. In addition, taxi drivers do not work exclusively for TNCs. Instead, they have two different channels by which they find their customers: on the road (offline) and via TNC apps (online).

Other than the regulations for TNCs, the Shanghai government also regulates taxis. There are seven licensed companies that are authorized to run approximately 50,000 taxis in Shanghai. The taxi companies lease their taxis to the taxi drivers, who pay the costs for leasing, gas, insurance, and so forth before collecting the residuals as their profit. The pricing is regulated by the city government; hence, even though the TNC system is embedded in the taxi, the fare is constant, regardless of whether a driver takes an order online or offline. Any trip shorter than 3km will incur a 14 Yuan (RMB) fare. The distance over the first 3 km additionally incurs a 2.40 Yuan fare per km. Notably, neither the taxi companies nor taxi drivers share revenue with the TNC platform. In fact, just like Uber and Lyft, the TNCs in China, during their early stage, focus on increasing their user base rather than generating revenue. The taxi companies also do not intervene into the drivers' app use at the very beginning.

[2] The entry date for a private car to operate as a ride-sharing vehicle in Shanghai was October 8, 2015, when the Shanghai Municipal Transportation Commission (SMTC), the official department in charge of the transportation policy enactment and enforcement in Shanghai, issued the first certificate for operating a TNC through private vehicles to Didi, the dominant TNC in China.

The intention of allowing TNC systems into the taxi industry, from the perspective of the city regulator, is to improve the efficiency of taxi operations.[3] The platform might more effectively match passengers with taxi drivers, which improves the profitability of taxi drivers through the reduction of idle time and enhances passengers' convenience in finding a taxi. Given that the total number of taxis is fixed, the introduction of TNC services is expected to help the taxi market to absorb the increased demand, which is a major challenge for the transportation system of a growing city, such as Shanghai. Providing an additional online channel with the introduction of TNCs, however, resulted in unexpected side effects. Note that the online channel provides taxi drivers with unique information that cannot be obtained from traditional offline orders, e.g., destination. With the fare as constant, taxi drivers became selective about passengers, based on the exclusive information provided by the apps (Tuss 2016) and prefer to take passengers from the online channel.

3.2.2 *Tech Restrict and Rush Hour Policies*

The transportation regulator in Shanghai, the Shanghai Municipal Transportation Commission (SMTC), witnessed substantially increased difficulty for roadside passengers to find an available taxi after the entry of TNC apps. To protect offline passengers from losing the opportunity for a taxi ride, the Shanghai government intend to partially reset the market structure to how it was before the entry of TNCs and introduced a policy to restrict taxi drivers' using TNC apps, beginning on April 13, 2015 (Yue 2015). According to Jianping Sun, the director of SMTC,[4] the

[3]As discussed by Jianping Sun, the director of SMTC, in April 2015. Reported by eastday.com at http://sh.eastday.com/m/20150414/u1ai8667710.html

[4] The discussion was reported, along with the policy announcement, on China Central Television (CCTV), Channel 13.

government not only aimed to support citizens who have less, or no, access to the online platform but also to develop the taxi industry by supporting passengers' accessibility via apps.[5]

Under the policy, 40,000 taxis in Shanghai's largest five taxi companies, which account for four-fifths of the taxi capacity in Shanghai, are divided into five groups based on the last digit of the taxi drivers' license plate number. Every working day (weekdays minus holidays), taxis in one group, roughly 6,000 cabs (12% of all the taxis in Shanghai), are restricted from using any TNC apps for a period of time, and signs are placed on the dashboard of participating vehicles so that passengers can easily identify them (Yue 2015). As Figure 3.2 shows, the badge indicates the day and time of the restriction. Throughout my observation window, the time periods are 07:30–10:30 and 16:30–19:30. Because this is a government policy, it is enforced through the transportation police. When the policy is in effect, when police officers sample drivers and administer an alcohol test, taxi drivers can be sampled and checked. The companies monitor the taxis through their GPS tracking system. Passengers also report cases in which they spot taxi drivers who are not following the regulation, and offending drivers will be dealt with by their taxi company (Yue 2015). The Chinese media reports that the policy has been well executed.[6] Drivers, however, can still operate their taxi and pick up their passengers from offline channels, e.g., roadside. To this end, the policy resets the demand information for drivers, and the affected drivers are forced to take passengers from non-TNC channels.

Throughout my observation period, another policy, termed the Rush Hour Policy, is applied to reduce traffic congestions levels on the expressway during rush hours, when the transportation volume is high. Residents who live in the city rely heavily on the expressways for daily travel

[5] Discussed during an interview with eastday.com; reported at http://sh.eastday.com/m/20150414/u1ai8667710.html
[6] Reported by large media outlets, e.g., people.com, sina.com, jiefang daily, eastday.com; http://sh.people.com.cn/n/2015/0414/c357908-24490696.html

between their workplace and home. To ensure smooth trips in the city, this policy regulates private vehicles registered outside of Shanghai city, temporarily registered vehicles, vehicles by new drivers, and vacant taxis, which are restricted in terms of driving on the expressways during rush hours. In my observation period, that time is 07:30–09:30 and 16:30–18:30 every working day (weekdays minus holidays) before April 15 and is extended to 07:00–10:00 and 16:00–19:00 starting on April 15 (Yan 2015). There were no transportation-related policies in my observed time window besides the Tech Restrict Policy and Rush Hour Policy.

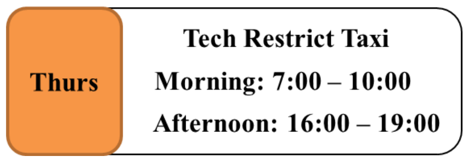

Figure 3.2. Badge for taxi drivers with app-usage restriction

3.2.3 *Data*

To estimate empirically the effects of the Tech Restrict Policy on macro-level transportation systems, I construct a unique data set that contains information from various official sources. Data collection from companies and institutions was coordinated by a Shanghai government agency, the Shanghai Municipal Commission of Economy & Informatization. Table 3.1 shows four different organizations that host the raw data. The taxi/GPS data are generated by taxi drivers who work for Shanghai Qiangsheng Holding Co., Ltd., the largest taxi company in Shanghai, with approximately 23,000 full-time employees and $600 million in annual revenue.[7] This company is also the major participant in the Tech Restrict Policy, and all of its drivers are affected. With these data, I construct a dependent variable that refers to the number of transactions made by the taxi for each half-hour session.

[7] Information on Shanghai Qiangsheng Holding Co., Ltd. can be found at the *Financial Times* site: https://markets.ft.com/data/equities/tearsheet/profile?s=600662:SHH

Table 3.1 Summary of Data Sources

Information	**DataSource**
Taxi-related data	Shanghai Qiangsheng Holding Co., Ltd.
Transactions via SPTC	Shanghai Public Transportation Card, Ltd.
Weather	Shanghai Environmental Protection Bureau
Congestion levels	Shanghai Construction and Transportation Research Institute

Note. Data collection was coordinated by the Shanghai Municipal Commission of Economy & Informatization.

Dependent variables that represent alternative transportations usage, i.e., metro, buses, ferries, and parking, are created based on transportation card data, provided by Shanghai Public Transportation Card (SPTC) Co., Ltd. Its raw data provide records of transactions made on each transportation mode by each cardholder, which my data aggregates. The company that operates the card is state owned and the only issuer of transportation cards in Shanghai. Like other transportation cards, such as the MetroCard in New York City and the Oyster Card in London, SPTC is a prepaid card; it can be purchased at the service booths in every metro station or at banks and convenience stores. Users could take any metro, bus, or ferry in Shanghai as long as they have a positive balance on the card, and they also can use the card to pay, for example, for a taxi or parking space in a lot affiliated with a metro station. They can recharge the card easily through service booths and multiple channels of e-payment. They swipe the card on the site of the public transportation, such as taxi meter or the gateway of the metro station. The fare, which depends on the distance of the user’s journey, is deducted from the balance of the user’s card account.

SPTC was the dominant public transportation payment method in Shanghai during my observation period. It is largely preferred by citizens and tourists in Shanghai due to its ease of use on public transportation. Alternatively, passenger who purchase a one-time ticket for a metro/ferry will have to wait in a line to process the ticket purchase for every ride, and passengers need to have the exact fare amount in cash for bus rides. Therefore, the majority of citizens and tourists prefer to use SPTC when taking buses and the metro. In addition, it should be noted that it was only after 2017 that other channels of e-payment i.e., Alipay and WeChat, were introduced into the bus and metro systems in Shanghai. Therefore, during my observation period, SPTC captured the major share of daily public transportation usage for those who commute via public transportation. Its overall usage is quite stable and, thus, is representative of the overall public transportation user population. In addition to public transportation, SPTC is also used for taxis, gasoline, and parking lots affiliated with a metro station.

To account for possible factors that influence the impact of the Tech Restrict Policy on transportation dynamics, I consider weather conditions that change passengers' decisions about travel activities and transportation modes (Saneinejad et al. 2012). The Shanghai Environmental Protection Bureau is in charge of collecting and maintaining the data related to weather. Weather data include temperature and precipitation levels for Shanghai, which are recorded at 3-hour intervals. The values are captured as the mean of precipitation and temperature from 10 weather monitors located across the city, managed by the bureau.

I further collect data on the congestion levels of roads from the Shanghai Construction & Transportation Research Institute, a government institution. Similarly, the traffic values are calculated as the mean of traffic density from 42 traffic monitors on the freeway and that from 68 traffic monitors on the surface streets. The Shanghai Construction & Transportation Research Institute distributes and manages those monitors and collects data from them. To better sample the traffic and weather conditions of the city, both institutions distribute their monitors evenly across the city. Overall, this rich data set provides us with information on the number of transactions conducted within taxi and alternative transportation systems, allowing us to capture plausible demand shifts in transportation dynamics.

3.2.4 *Description of Variables*

I construct a half-hour index i by dividing 24 hours into 30-minute intervals. For instance, the half-hour index for 01:00–01:30 is 0, 01:30–02:00 is 1, and so forth. The day index t is each day in April 2015, ranging from 1 to 30. As a result, I have 48 half-hour indices for 30 days, resulting in 30 x 48 = 1,440 observations.

3.2.4.1 Dependent Variables

The dependent variable for the main analysis is the total number, i.e., count, of transactions conducted during half-hour session i on day t for each transportation mode. For instance, *Taxi* is the total number of transactions conducted for the taxi during the half-hour i on day t. In particular, I calculate the number of transactions of taxis from the taxi-related data. I also measure the numbers of *Metro*, *Bus*, *Ferry*, and *Parking* affiliated with the subway station transactions, accordingly, from the data of SPTC. Finally, for the extended analysis, congestion-related dependent variables are measured from the data by the Shanghai Construction & Transportation Research Institute.

3.2.4.2 Independent Variables

The main independent variable of interest is the dichotomous indicator T_{it}, which indicates whether the technology is restricted at half-hour index i and day t. R_{it} is also a binary indicator that is 1 if the Rush Hour Policy is applied at half-hour index i and day t, and zero otherwise.

Additional control variables $Precipitation_{it}$ and $Temperature_{it}$ come from the weather data. They are the citywide mean of precipitation and temperature at half-hour index i and day t. As the government institution recorded both precipitation and temperature in 3-hour intervals, I assume that the weather is consistent and stable during the three hours. A description of the variables and the summary statistics are found in Tables 3.2 and 3.3.

Table 3.2. Description of the Variables

Variable	**Description**
Day	Days 1–30 in April 2015 Day index *t* ranges from 1 to 30
Half-Hour	30-minute time intervals within a day. Half-Hour index *i* ranges from 0 to 47
Dependent Variables	
Taxi	Number of transactions conducted on taxi at half-hour *i* and day *t*
TaxiSub	Number of transactions conducted via Card on taxi at half-hour *i* and day *t*
Metro	Number of transactions conducted via SPTC on metro at half-hour *i* and day *t*
MetroSub	Number of transactions conducted via SPTC on metro at half-hour i and day t, during metro operation hours
Bus	Number of transactions conducted via SPTC on bus by half-hour *i* and day *t*
Ferry	Number of transactions conducted via SPTC on ferry by half-hour *i* and day *t*
Parking	Number of transactions conducted via SPTC for metro parking usage by half-hour *i* at day *t*
PublicTrans	Number of transactions conducted via SPTC on public transportations at half-hour *i* at day *t*
ln(totalTravelDist)	Total distance taxis traveled at/during half-hour *i* and day *t*
ln(avgTravelDist)	Average distance taxi traveled in each transaction during half-hour *i* and day *t*
ln(congFree)	Mean of congestion level of freeway at half-hour *i* and day *t* over 42 different districts
ln(congSurface)	Mean of congestion level of surface at half-hour *i* and day *t* over 68 different districts
Independent Variables	
Tech Restrict Policy (*T*)	Tech Restrict Policy
Rush Hour Policy (*R*)	Rush Hour Policy
Temporal Distance of Tech Restrict Policy (*D*)	Temporal distance between half-hour *i* and the closest *i* when Tech Restrict Policy is applied
Temporal Distance of Rush Hour Policy (*d*)	Temporal distance between half-hour *i* and the closest *i* when Rush Hour Policy is applied
Control Variables	
Precipitation	Mean precipitation over 10 districts in town at half-hour *i* at day *t*
Temperature	Mean temperature over 10 districts in town at half-hour *i* at day *t*
Holiday	Day when it was official holiday in Beijing, China in April 2015

3.2.5 *Model-Free Evidence*

I develop numerical baseline estimates to explore the impact of technology restrictions on transportation dynamics. Taking into account that the Rush Hour Policy is always applied in my time window, I suggest that the difference between the trend of the mixture of the two policies and the trend of the Rush Hour Policy is generated by the technology restriction of the Tech Restrict Policy. The model-free evidence is illustrated by the graphs provided in Figure 3.

3. I plot the normalized number of transactions for each transportation mode: taxis, public transportation (including metro, bus, and ferry), and parking lots affiliated with metro stations. I compare the average number of transactions during policy enforcement time across three conditions: when only the Rush Hour Policy exists, when there is a mixture of the Rush Hour Policy and the Tech Restrict Policy, and when there is a mixture of two policies with the extended Rush Hour Policy. The shaded area in graphs represents the last case; hence, the policy implemented time is illustrated at 07:00–10:00 ($12 \leq i \leq 17$) and 16:00–19:00 ($30 \leq i \leq 35$).

Table 3.3. Summary Statistics

Variable	Mean	Std. Dev	Min	Max
Taxi	8,090.06	3,500.40	510	14,583
TaxiSub	5,452.56	2,948.14	186	9,449
Metro	172,237.30	166,314.00	0	735,069
MetroSub	216,924.60	159,190.10	0	735,069
Bus	108,124.30	90,904.50	293	357,715
Ferry	1,311.41	1,222.41	0	5,439
Parking	70.34	86.88	0	580
PublicTrans	281,673.00	252,674.00	304	1,087,610
ln(totalTravelDist)	10.89	0.75	8.24	12.56
ln(avgTravelDist)	2.21	0.15	1.81	3.19
ln(congFree)	3.23	0.26	2.42	3.83
ln(congSurface)	3.25	0.45	2.16	3.87
Precipitation	0.58	0.56	0.1	5.57
Temperature	14.7	4.68	4.23	27.71
Holiday	0.1	0.3001042	0	1

Figure 3.3(a) shows the average number of taxi transactions made during the day under three different conditions. The number of taxi transactions does not show a clear trend with a mixture of policies; restricting drivers' taxi apps usage causes both an increase and a decrease in the average number of taxi transactions. During the morning policy time, there is a higher number of taxi transactions with the mixture of the Rush Hour Policy and the Tech Restrict Policy than with only the Rush Hour Policy. Even though the number of taxi transactions is higher at the initial stage of the policy time, it eventually falls and becomes similar to the number of taxi transactions

with the Rush Hour Policy only. In the afternoon policy time, the number of taxi transactions is significantly higher when there is no technology restriction. As a result, the average number of taxi transactions with a mixture of two policies does not show consistent results; thus, further analysis with econometric models is needed.

Figure 3.3 (b) shows the average number of public transportation transactions during the day. Considering that the Rush Hour Policy is always in existence, the technology restriction by the Tech Restrict Policy causes a significant increase in the average number of public transportation transactions. As seen in the graph, the number of transactions is higher with a mixture of policies than with only the Rush Hour Policy for both the morning and afternoon policy time periods. The results are consistent with and without the extended Rush Hour Policy. The average uses of parking lots affiliated with metro stations, as seen in Figure 3.3 (c), is similar to that of public transportation transactions.

As a result, the Tech Restrict Policy introduction generates a different trend in the proportion of taxi transactions to public transportation. Despite my preliminary insights, concern remains because I cannot establish statistically whether the results are due specifically to the technology restriction. Therefore, I construct an econometric model to supplement my model-free evidence and find significant effects of technology restriction. I particularly consider the Rush Hour Policy to ensure that a change in transportation usage is due to the technology restriction only and not to the extended policy hours. Overall, the preliminary analysis yields result that reinforce my main analysis.

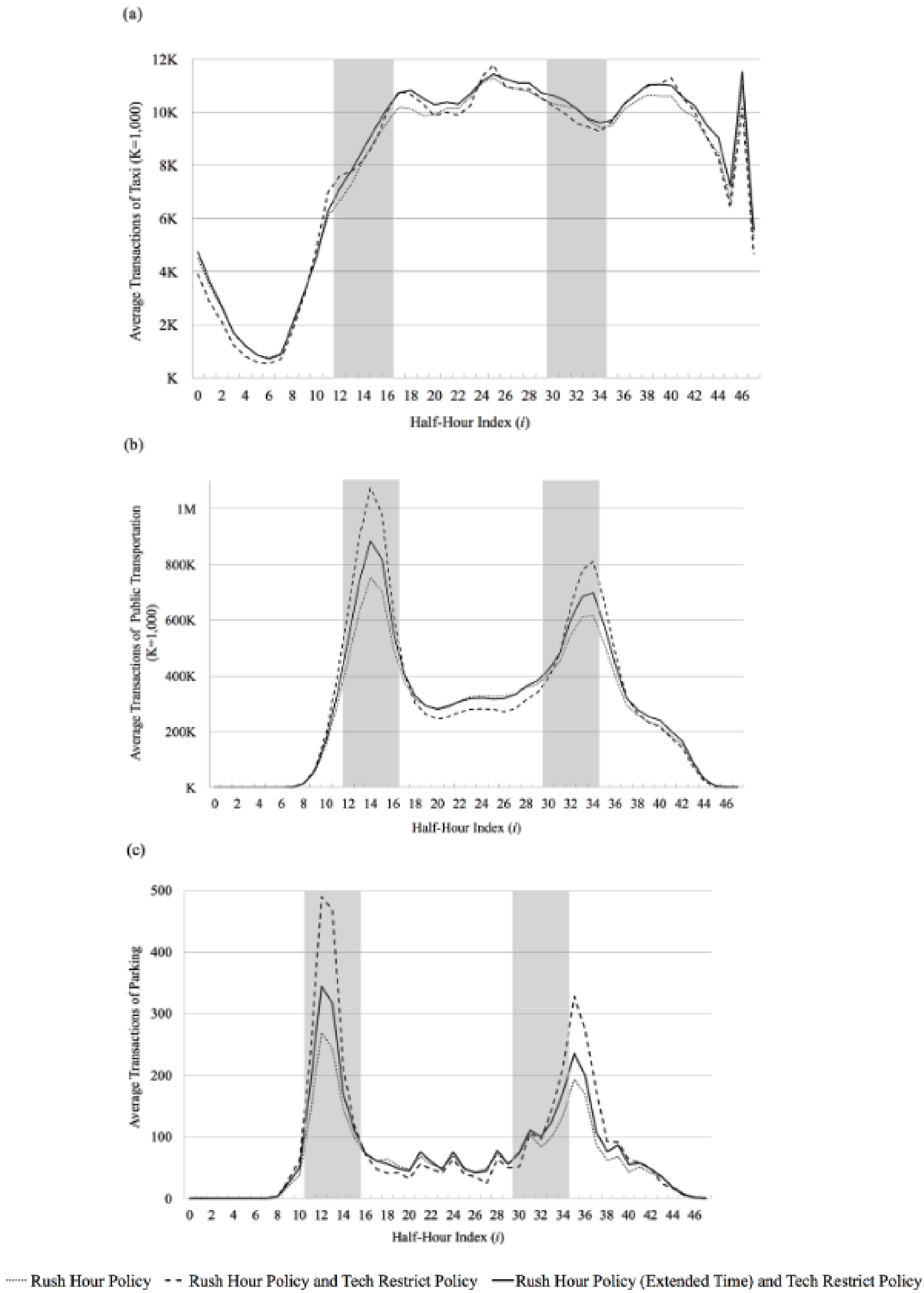

Figure 3.3. Model-free evidence: Transportation trends during the day

3.3 METHODS

My core interest is to examine the effects of information sharing on transportation dynamics. Note that the regulations of my context control the overall taxi driver supply and the potential effects of a pricing strategy. Thus, I can largely attribute the effects of restricting drivers' ride-sharing app usage (policy effects) to the effects of information sharing. In particular, I capture four dimensions of the policy effects: *main*, *horizontal*, *vertical*, and the combination of horizontal and vertical (*mixed*) effects. I define *main effects* as the direct impact of the policy on taxis when enforced and *horizontal effects* as policy effects that are generated by rides' shifting to alternative transportation modes when enforced. *Vertical effects*, or temporal spillover effects, occur when taxi drivers or passengers shift to times with no policy enforcement. Finally, *mixed effects*, the combination of horizontal and vertical effects, are generated when taxi passengers shift to alternative transportation modes at a time without policy enforcement.

Although the vertical effects and horizontal effects are clear, the mixed effects could be missed. Note that the alternative transportation included in my context is public transportation, which is typically heavily utilized and nearly full during the rush hours. An increased usage of public transportation from taxi passenger spillover, if observed, might generate negative externality, e.g., every metro rider's utility might deteriorate because the metro is more crowded. If the ridership of the same transportation mode but at non-policy-enforcement time remains the same and, thus, the utility at that time stays constant, passengers who use that transportation mode at the enforcement times might further spill over into the non-enforcement times until the non-enforcement time's utility also deteriorates by the spillover's negative externality, and there are no additional passengers. From another perspective, the passengers who spill over to taxi use at a non-enforcement time face the constraints of the taxi's capacity. The additional demand shifted from the enforcement time periods makes taxis busier and, thus, increases the cost of getting a taxi. If the utility associated with alternative transportation modes at the same time stays constant, the increased cost of a taxi will drive some potential taxi riders to an alternative transportation mode until a new equilibrium is reached. In sum, from both perspectives, I would hypothetically expect positive mixed effects. In the following, I discuss how I specify and identify these effects.

3.3.1 *Main and Horizontal Effects of the Tech Restrict Policy*

In this section, I introduce my identification strategy for the *main* and *horizontal* effects of the Tech Restrict Policy. Note that these two types of effects use the same identification strategy but for different transportation modes, i.e., taxi and alternative transportation modes, accordingly. I combine the discussions on their identification strategies, and, in the following, I term those effects as the effects of the Tech Restrict Policy for a succinct presentation.

The major challenge comes from the introduction of another policy, the Rush Hour Policy, in the city before the Tech Restrict Policy to mitigate the traffic congestions at rush hours. It has been applied every working day (weekdays except holidays) during a similar duration of time[8] as the Tech Restrict Policy. This results in a large overlap of two different policy effects, which makes it difficult to capture whether the change in the number of transportation transactions is due to the Tech Restrict Policy or the Rush Hour Policy.

To address this challenge, I first extract the combined effect of the Rush Hour Policy and Tech Restrict Policy and the effect of Rush Hour Policy alone. Then I take the difference to partial out the net effect of the Tech Restrict Policy. More specifically, using a within-day half-hour index i and a daily index t to represent the observation period in a 2-dimensional space, I visualize my

[8] The overlap occurred on Days 13 and 14, then was extended to three hours: 7:00–10:00 and 16:00–19:00.

identification strategy, as seen in Figure 3.4. The *x*-axis indicates day t, which ranges from 1 to 30, and the *y*-axis indicates the half-hour index i, covering 24 hours. β_1 represents the effect of the Tech Restrict Policy, β_3 represents that of the Rush Hour Policy, and the shaded areas represent the treated time scopes, accordingly.[9]

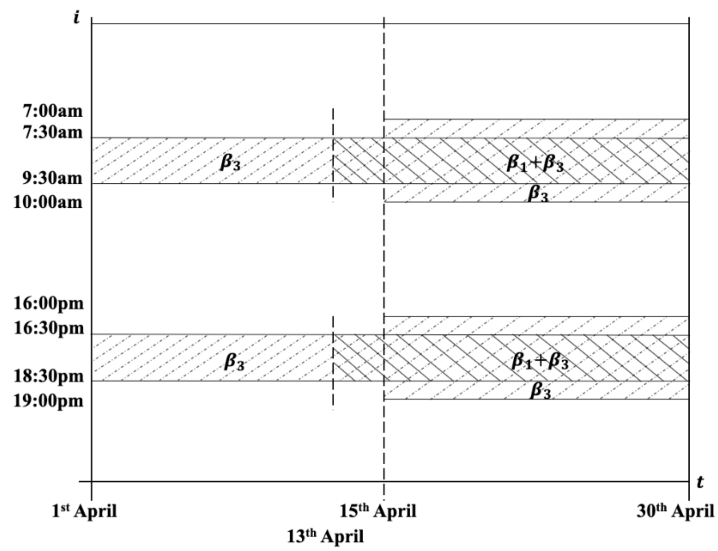

Figure 3.4. Identification strategy

As is seen, the Tech Restrict Policy-treated areas (shaded with β_1) are within the Rush Hour Policy-treated areas (shaded with β_3); thus, the Tech Restrict Policy-treated areas are double-shaded. Therefore, by taking the difference between the double-shaded areas and the non-treated (non-shaded) areas, I can identify only the combined effect of both policies, $\beta_1 + \beta_3$. I can, however, separately identify the net effect of β_3 by taking the difference between the Rush Hour Policy-treated areas (shaded with β_3) and the non-treated (non-shaded) areas. I can thus identify the net effect of Tech Restrict Policy, β_1, by taking the difference between $\beta_1 + \beta_3$ and β_3.

[9] For a clearer presentation, I ignore the non-working days in Figure 3.4.

I want to clarify that, due to the “within” nature of the Tech Restrict Policy, an interactive effect between these two policies is not identifiable, and β_3 incorporates the interactive effects of these two policies. Note that my objective is to accurately identify the net effect of the Tech Restrict Policy, which should encompass its synergy with any other factors, and the Rush Hour Policy variable is included to control for confounding bias, as it overlaps with the Tech Restrict Policy. This identification strategy satisfactorily addresses my needs, and, thus, I do not attempt to tease out the interactive effect between the two policies.

Further, I note that day-specific and time-specific fixed effects could be incorporated in the model so that both $\beta_1 + \beta_3$ and β_3 are not confounded by day-specific unobserved factors and time-specific unobserved factors. This is because I can always find within-half-hour variations and within-day variations for $\beta_1 + \beta_3$ and β_3 (i.e., 07:30 to 09:30 and 16:30 to 18:30 for $\beta_1 + \beta_3$, any working day later than April 13 for $\beta_1 + \beta_3$, 07:00 to 07:30, 09:30 to 10:00, 16:00 to 16:30, and 18:30 to 19:00 for β_3, and any day for β_3), and, thus, both $\beta_1 + \beta_3$ and β_3 are identifiable after minus day-specific average across half-hour periods and half-hour specific average across days. I can thus eliminate confounders such as time trend, seasonality, and rush hour.

Finally, the policy of interest had been discussed, designed, and planned to be launched before the first day of its execution and was designed as a permanent policy change unless a further policy change was made. That is, the policy has no variation in my observation after its launching. This ensures that the policy effect identification is free from simultaneity bias. Along with the half-hour and day fixed effects as controlling for time-specific and day-specific confounding bias, the policy variable should be free from endogeneity concerns.

3.3.2 *Vertical and Mixed Effects of the Tech Restrict Policy*

Both vertical and mixed effects are policy effects on alternative time periods but differ on the targeted transportation modes. Therefore, similar to the identification strategies for *main effects* and *horizontal effects*, I can apply a similar specification of *vertical effects* on alternative transportation modes to identify the *mixed effects*. Thus, in the following, I discuss only the *vertical effects* of the Tech Restrict Policy.

For *vertical effects*, I create measurements to gauge policy effects across time within a day. As time is continuous, I extend the analysis to estimate the spillover effects of policy enforcement at times when no policy is applied. As the Tech Restrict Policy is enforced at two different policy periods, 07:30–09:30 and 16:30–18:30, two independent vertical effects are derived from each policy enforcement period. Therefore, I construct a new measurement, $\mathit{Vertical\ Effects}_i$, that captures the possible overlap of two independent vertical effects over a continuous time when the Tech Restrict Policy is not enforced.

I take into account three considerations to construct $\mathit{Vertical\ Effects}_i$ accurately. First, by my definition, vertical effects occur only before or after the time ranges of policy enforcement. Thus, no vertical effects are present during the periods of policy enforcement. Hence, I define D_{1i} as a minimum temporal distance between half-hour i and the time with policy enforcement between 07:30 and 09:30. Similarly, D_{2i} is a minimum temporal distance between half-hour i and the time when the policy is applied between 16:30 and 18:30. D_{1i} and D_{2i} are positive integers, as they are temporal distances that indicate absolute differences between half-hour indices.

Second, I assume that each vertical effect is monotonically decreasing.[10] The nature of *vertical effects* is the substitution of time to travel, and I expect passengers to prefer alternative time periods that are closer to their planned time periods. That is, a passenger who originally planned to travel at 09:00 might be more likely to travel at 10:00 than 18:00 when affected by the policy. Therefore, I hypothetically expect diminishing vertical effects as time becomes further away from the initial policy enforcement time. I apply multiple absolutely monotonically decreasing functions on temporal distances, $f(D_{1i})$ and $f(D_{2i})$ e.g., $\exp(-D_{1i})$ and $\exp(-D_{2i})$, and D_{1i}^{-2} and D_{2i}^{-2}. These functions increase when time period i approaches the policy enforcement time period but decreases as i becomes further away. Because the two independent vertical effects from different policy enforcement time periods can overlap at each half-hour i, I model the final vertical effect as follows:

$$Vertical\ Effects_i = \beta_2(f(D_{1i}) + f(D_{2i})) \tag{3.1}$$

where i is a within-day half-hour, and β_2 is the parameter (to be identified) that modifies the magnitude of the vertical effects.

This specification reflects the symmetry characteristic of a time. I shift the start of the new day from noon to the time period 01:00–01:30 and indicate it as half-hour index i = 0 (or 48). The time period 01:00–01:30 is at the halfway point between i = 12 and i = 35, which is the beginning of the morning Tech Restrict Policy time and the end of the afternoon policy time, respectively. This setup confirms that the Tech Restrict Policy has the least impact on the midpoint of the non-policy enforced time, as it is at the point of furthest temporal distance from both the morning and afternoon policy enforcement time ranges.

[10] This assumption is later relaxed and tested in the model specification test (Subsection 5.1).

In Figure 3.5, I present the $Vertical\ Effects_i$ across time within a day. The time periods for the Tech Restrict Policy enforcement are stated as "Tech Restrict" in blank spaces, indicating that there are no $Vertical\ Effects_i$. The overlap of vertical effects between two policy periods is presented as the shaded area around the clock. The wider the shaded area, the larger the $Vertical\ Effects_i$, reflecting that $Vertical\ Effects_i$ decrease as they get further away from the policy implementation time. The shaded area at i = 24 is thicker than the area at i = 0 because the temporal distances between the policy enforcement time and i = 24 are shorter than the distances between that and i = 0.

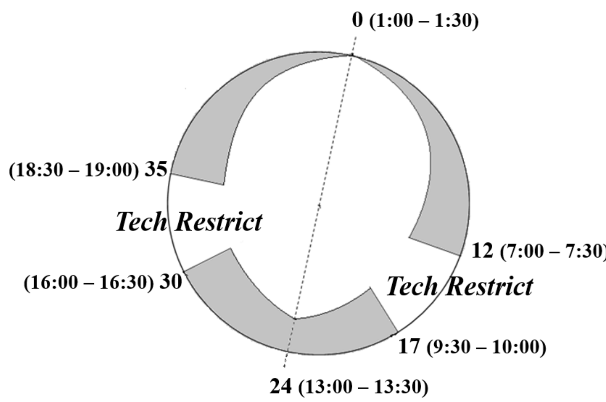

Figure 3.5. Conceptualization of time symmetry and temporal spillover effects

To control for the temporal spillover effects of the Rush Hour Policy, a term for the vertical effects of the Rush Hour Policy is included. I apply a similar specification with a different magnitude parameter, β_4, and the minimum temporal distances between half-hour i and the times of the Rush Hour Policy, d_{1it} and d_{2it}. Note that the times for this policy change across days (t); thus, d_{1it} and d_{2it} depend on t. The monotonically decreasing function also could be different from that for the vertical effects of the Tech Restrict Policy; thus, I denote this as $\beta_4(f_r(d_{1i}) + f_r(d_{2i}))$.

Note that the *vertical effects* (and the *mixed effects*) are fundamentally an interaction term between the policy treatment and a function of temporal distance between a specific half-hour and the enforcement time. Therefore, its identification comes from the variation provided by the temporal distance, in addition to that of the policy effect, as previously discussed. Also note that the enforcement time of the Rush Hour Policy differs from that of the Tech Restrict Policy. This indicates that, for a particular half-hour, both policy treatments and functions of temporal distance vary across these two policies, and this additional variation further ensures a separate identification of the *vertical effects* of the Rush Hour Policy.

3.3.3 *Model Specification*

To estimate the *main* and *vertical effects* of the Tech Restrict Policy, I specify *Taxi*, the count of taxi order arrivals in 30-minute intervals, as the dependent variable for Model 1a. Considering the nature of the primary dependent variable, I use a Poisson regression model. Negative binomial regressions do not apply to my data, as there is no overdispersion in the transaction distributions that are conditional on individuals. In addition, distributions do not show the typical long right tail of the transaction distributions. There is no heterogeneity in the city, as the policy affects only one city, Shanghai. By controlling half-hour-specific factors (i) and day-specific factors (t), the equation is as follows:

$$P(y_{it} \mid \lambda_{it}) = \frac{\lambda_{it}^{\,y_{it}}}{y_{it}!} e^{-\lambda_{it}} \tag{3.2}$$

where $\lambda_{it} = \beta_1 \cdot T_{it} + \beta_2 \cdot (f(D_{1i}) + f(D_{2i}))_{it} + \beta_3 \cdot R_{it} + \beta_4 \cdot (f_r(d_{1i}) + f_r(d_{2i})) + \beta_5 \cdot Controls_{it} + \mu + \gamma_i + \tau_t$.

Both T_{it} and R_{it} are dummy variables that equal 1 if the half-hour i on day t has policy enforcement, and 0 otherwise. γ_i is half hour-specific fixed effects that control the day-invariant factors across half-hours, and τ_t is day-specific fixed effects that controls unobserved, common across all i, daily varying shocks. μ represents the constant, and the variable $Controls_{it}$ includes

(a) $Holiday_t$ to control for public holidays within the time window and (b) $Precipitation_{it}$ and $Temperature_{it}$ because weather conditions are considered to have an impact on passengers' travel behaviors and emotions and affect choices of transportation modes (Böcker et al. 2016). Finally, parameters $\beta_{it} = (\beta_1, \beta_2, \beta_3, \beta_4, \beta_5)$ are coefficients, where β_1 and β_2 are the estimates for my interests.

To estimate the *horizontal* and *mixed effects* of the Tech Restrict Policy, similar to Model 1a, I apply the Poisson regression to all other count dependent variables: *Metro* (Model 1b), *Bus* (Model 1c), *Ferry* (Model 1d), and *Parking* affiliated with metro station transactions (Model 1e). Note that I allow the error terms of various transportation modes to be correlated for a more robust estimate. As the dependent variables are the number of transactions conducted for each transportation mode, the results allow us to examine the impact of the technology restriction on macro-level transportation dynamics.

3.4 RESULTS

3.4.1 *Model Specification Test*

As discussed in Section 3.3.2, I compare the fitness performance of alternative functions to model the adopt-alternative specification of vertical effects, shown in Table 3.4. Note that all of the models have the same degree of freedom, and, thus, a log-likelihood comparison will be equivalent to AIC and BIC comparisons. I show the log-likelihoods alternative models in Table 3.4 and find that the e^{-x} function for the Tech Restrict Policy's vertical/mixed effect f and $x^{-1/4}$ for the Rush Hour Policy's vertical/mixed effect f_r best fit my data. Therefore, I choose those two functions for my model.

Table 3.4. Model Specification Test for Vertical Effects

Tech Restrict f	$x^{-1/4}$	$x^{-1/2}$	x^{-2}	x^{-4}	e^{-x}
Rush Hour f_r					
$x^{-1/4}$	-39041.906	-39032.483	-39012.291	-39029.639	**-39004.925**
$x^{-1/2}$	-39062.001	-39045.881	-39026.755	-39055.524	-39006.254
x^{-2}	-39154.785	-39146.615	-39088.731	-39108.329	-39062.017
x^{-4}	-39169.314	-39166.913	-39107.478	-39117.883	-39089.416
e^{-x}	-39154.104	-39145.083	-39088.631	-39108.988	-39061.198

Note. **Bold** signifies the maximum

To further justify that the vertical effects are monotonically decreasing, I compare the best model above with alternative models that do not assume monotonically decreasing vertical effects, i.e., constant vertical effects and monotonically increasing ones with different rates. As shown in Table 3.5, the log-likelihood of alternative models with uniform or increasing vertical effects are lower than hours with monotonically decreasing vertical effects. This result justifies the monotonically decreasing functions.

Table 3.5. Model Specification Test for Vertical Effects

Vertical Effects	Best decreasing	Constant	x^2	x^4	e^x
Log-likelihood	**-39004.925**	-39513.427	-39320.558	-39224.401	-39164.169

Note. **Bold** signifies the maximum

3.4.2 *Estimation Results*

Table 3.6 presents the impact of technology restriction on the number of transportation transactions. My empirical analysis reveals that the number of fulfilled taxi transactions declines when taxi drivers' ride-sharing app usage is restricted. The significant negative result (β_{1a} = -0.021, p < 0.001) for taxi transaction counts indicates that approximately 2.08% fewer taxi transactions are likely to be made during a month if the usage of the app is restricted. Therefore, I suggest that the Tech Restrict Policy negatively affects the efficiency of taxi operations. At the same time, the Tech Restrict Policy causes taxi passengers to shift to alternative transportation systems or to take a taxi at other times. I analyze such policy effects in three different dimensions: *horizontal*, *vertical*, and *mixed effects*. First, the *horizontal effects* are significantly positive (β_{1b} = 0.364, $p < 0.001$; β_{1c} = 0.245, $p < 0.001$; β_{1d} = 0.190, $p < 0.001$; β_{1e} = 0.312, $p < 0.001$) when the policy is enforced. This suggests that taxi passengers shift to alternative transportation systems with the introduction of the Tech Restrict Policy.

Table 3.6. Estimation Results

	Model 1a	Model 1b	Model 1c	Model 1d	Model 1e
Method			**Poisson**		
Dependent Variable	**Taxi**	**Metro**	**Bus**	**Ferry**	**Parking**
Main/Horizontal Effects	-0.021 ***	0.364 ***	0.245 ***	0.190 ***	0.312 ***
	(0.002)	(0.017)	(0.008)	(0.017)	(0.018)
Vertical/Mixed Effects	0.126 ***	0.575 ***	0.430 ***	0.293 ***	0.571 ***
	(0.008)	(0.054)	(0.014)	(0.044)	(0.020)
Rush Hour Policy	-0.057 ***	0.089 ***	0.051 ***	0.052 ***	0.018
	(0.005)	(0.008)	(0.003)	(0.001)	(0.011)
Vertical/Mixed Effects of Rush Hour	-0.018 ***	0.155 ***	0.113 ***	0.111 ***	0.119 ***
	(0.002)	(0.013)	(0.006)	(0.012)	(0.012)
Holiday	-0.145 ***	-0.352 ***	-0.313 ***	-0.598 ***	-0.348 **
	(0.006)	(0.014)	(0.001)	(0.019)	(0.116)
Precipitation	-0.031 **	-0.173 ***	-0.124 ***	-0.128 ***	-0.147 ***
	(0.011)	(0.002)	(0.005)	(0.010)	(0.007)
Temperature	-0.004 ***	0.000	-0.002 ***	0.003 **	-0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.011)
Constant	8.520 ***	0.936	6.861 ***	3.097 ***	-0.448 *
	(0.015)	(1.178)	(0.031)	(0.005)	(0.208)
Half-Hour Fixed Effects	Yes	Yes	Yes	Yes	Yes
Day Fixed Effects	Yes	Yes	Yes	Yes	Yes
No. of Observations	1,410	1,410	1,410	1,410	1,410
McFadden's R-squared	0.973	0.991	0.993	0.989	0.907

Note. Robust standard errors are in parentheses.
†$p < 0.1$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Note that the usage of private vehicles indicates use of metro station-affiliated parking lots. Thus, I interpret that there is an increase of passengers who utilize a combination of private vehicles and metro service. Second, the *vertical effect* is significantly positive (β_{2a} = 0.126, $p <$ 0.001), reflecting an increase in taxi transactions at times when the policy is not applied. In particular, the time periods next to the policy enforcement time receive a 4.93% increment, and those 1 hour away from the enforcement receive a 1.81% increment. This indicates that taxi passengers shift to times without technology restriction. Finally, the *mixed effects* are significantly positive, (β_{2b} = 0.575, $p < 0.001$; β_{2c} = 0.430, $p < 0.001$; β_{2d} = 0.293, $p < 0.001$; β_{2e} = 0.571, $p < 0.001$), revealing that passengers shift to alternative public transportation at a time when the policy is not applied. Altogether, my findings suggest that restricting information sharing deteriorates the overall efficiency of taxi operations and leads to positive spillovers to alternative times, publication transportation, and a combination thereof. These findings provide empirical evidence for the economic value of information sharing in TNCs, as it improves the efficiency of taxi operations and further alleviates the burden on public transportation, shown as a negative externality. Figure 3.6 presents the estimated correlation between the error terms. ρ_{mn} is the correlation between the error term of Model 1*m* and that of Model 1*n*, and *n* and *m* are model designations that range from *a* to *e*.

$$\rho_{mn} = \begin{pmatrix} \square & \begin{array}{c}\text{1.65E-05}\\(0.00001)\end{array} & \begin{array}{c}\text{2.08E-05}\\(0.00002)\end{array} & \begin{array}{c}\text{-1.15E-05}\\(0.00001)\end{array} & \begin{array}{c}\text{3.74E-04}^{***}\\(0.00000)\end{array} \\ & \square & \begin{array}{c}\text{1.10E-05}^{***}\\(0.00000)\end{array} & \begin{array}{c}\text{-3.73E-04}^{***}\\(0.00000)\end{array} & \begin{array}{c}\text{9.43E-05}\\(0.00007)\end{array} \\ & & \square & \begin{array}{c}\text{3.35E-04}^{***}\\(0.00001)\end{array} & \begin{array}{c}\text{7.54E-04}^{***}\\(0.00000)\end{array} \\ & & & \square & \begin{array}{c}\text{-6.64E-04}^{***}\\(0.00003)\end{array} \\ & & & & \square \end{pmatrix}$$

Figure 3.6. Correlation Matrix of Error Terms.

3.4.3 *Robustness Checks*

My model and data have several limitations that are subject to further robustness checks. These checks are presented below.

3.4.3.1 Alternative Subsamples

I consider alternative subsamples of the taxi and metro population to check the robustness of my findings. Instead of using *Taxi*, which reflects taxi transactions conducted regardless of the type of transaction method, I use *TaxiSub*, which is the number of taxi transactions conducted via the SPTC. Table 7 shows the comparison between estimates of *Taxi* and *TaxiSub*, in which results are consistent across the models. I also conduct an analysis of the smaller but stricter subsample of the metro population for estimating passengers' shift to the metro. The metro in Shanghai operates from 05:30; thus, there is little metro service when the half-hour index ranges from 0 to 9. I remove data points that have a half-hour index less than or equal to 9 in the metro population. The variable *MetroSub* indicates the number of metro transactions conducted when the half-hour index i is above 9. Subsampling is not suggested for the remainder of the dependent variables, as there are no strict operation constraints to be considered. Table 3.7 presents the comparison between estimates of *Metro* and *MetroSub*, in which results are consistent across the models and, thus, support the robustness of my findings.

Table 3.7. Subsample Analysis

	Model 1a	Model 2a	Model 1b	Model 2b
Method	**Poisson**			
Dependent Variable	**Taxi**	**TaxiSub**	**Metro**	**MetroSub**
Main/Horizontal Effects	-0.021 ***	-0.014 ***	0.364 ***	0.370 ***
	(0.002)	(0.001)	(0.017)	(0.016)
Vertical/Mixed Effects	0.126 ***	0.133 ***	0.575 ***	0.564 ***
	(0.008)	(0.010)	(0.054)	(0.057)
Rush Hour Policy	-0.057 ***	-0.051 ***	0.089 ***	0.098 ***
	(0.005)	(0.007)	(0.008)	(0.010)
Vertical/Mixed Effects	-0.018 ***	-0.007 **	0.155 ***	0.160 ***
of Rush Hour	(0.002)	(0.002)	(0.013)	(0.012)
Holiday	-0.145 ***	-0.076 ***	-0.352 ***	-0.345 ***
	(0.006)	(0.003)	(0.014)	(0.014)
Precipitation	-0.031 **	-0.053 ***	-0.173 ***	-0.179 ***
	(0.011)	(0.014)	(0.002)	(0.001)
Temperature	-0.004 ***	-0.005 **	0.000	-0.001
	(0.001)	(0.002)	(0.001)	(0.001)
Constant	8.520 ***	7.326 ***	0.936	11.160 ***
	(0.015)	(0.040)	(1.178)	(0.007)
Half-Hour Fixed Effects	Yes	Yes	Yes	Yes
Day Fixed Effects	Yes	Yes	Yes	Yes
No. of Observations	1,410	1,410	1,410	1,116
McFadden's R-squared	0.973	0.977	0.991	0.984

Note. Robust standard errors are in parentheses.
†p < 0.1, *p < 0.05, **p < 0.01, ***p < 0.001.

3.4.3.2 Alternative Time Span

One additional concern with my data construction is the interval of the time span that I use for half-hour index i. In the main analysis, because both the Tech Restrict and Rush Hour Policies were applied for two hours every day, with the Rush Hour Policy extension to three hours starting on April 15, 2015, by adding 30 minutes at the start and end of the time ranges of the original policy enforcement time, I believe that 30 minutes is a critical time range when implementing transportation-related policies. Thus, I split time into 30-minute intervals. The 30-minute interval, however, can be a large time span in terms of reflecting drivers' or passengers' travel behaviors. Therefore, I split time into the granular level of 15-minute intervals and repeat the regression with the new data set of 2,880 observations. Table 3.8 presents consistent results from my main analysis.

Table 3.8. Analysis in 15-minute Intervals

	Model 3a	Model 3b	Model 3c	Model 3d	Model 3e
Method			**Poisson**		
Dependent Variable	**Taxi**	**Metro**	**Bus**	**Ferry**	**Parking**
Main/Horizontal Effects	-0.022 ***	0.365 ***	0.242 ***	0.184 ***	0.330 ***
	(0.001)	(0.013)	(0.000)	(0.014)	(0.006)
Vertical/Mixed Effects	0.101 **	0.569 ***	0.407 ***	0.276 ***	0.596 ***
	(0.036)	(0.027)	(0.027)	(0.028)	(0.056)
Rush Hour Policy	-0.037 **	0.147 ***	0.099 ***	0.081 ***	0.096 **
	(0.013)	(0.006)	(0.010)	(0.007)	(0.029)
Vertical/Mixed Effects	-0.020 ***	0.187 ***	0.136 ***	0.128 ***	0.155 ***
of Rush Hour	(0.003)	(0.014)	(0.000)	(0.014)	(0.011)
Holiday	-0.141 ***	-0.470 ***	-0.399 ***	-0.681 ***	-0.534 ***
	(0.027)	(0.047)	(0.066)	(0.005)	(0.141)
Precipitation	-0.018 **	-0.087 ***	-0.084 ***	-0.082 ***	-0.101 ***
	(0.006)	(0.010)	(0.004)	(0.005)	(0.020)
Temperature	-0.003	-0.005	-0.007	-0.002 ***	-0.017
	(0.002)	(0.004)	(0.006)	(0.000)	(0.011)
Constant	7.865 ***	0.184	6.256 ***	2.482 ***	-0.727 ***
	(0.030)	(1.104)	(0.096)	(0.035)	(0.205)
Half-Hour Fixed Effects	Yes	Yes	Yes	Yes	Yes
Day Fixed Effects	Yes	Yes	Yes	Yes	Yes
No. of Observations	2,820	2,820	2,820	2,820	2,820
McFadden's R-squared	0.968	0.990	0.993	0.982	0.861

Note. Robust standard errors are in parentheses.
†p < 0.1, *p < 0.05, **p < 0.01, ***p < 0.001.

3.4.3.3 Alternative Specification of Vertical Effects

I further validate that the results are not particular to my specification of vertical/mixed effects. I repeat the regressions, using the alternative specification for vertical/mixed effects. I also repeat the analysis, using the alternative measure with a slower diminishing rate to reflect the vertical/mixed effects as approaching zero with an increase of temporal distances. Therefore, in equation (6), I replace specification e^{-x} with x^{-2}. Table 3.9 presents consistent results with the main measurements.

Table 3.9. Alternative Specification of Vertical Effects

	Model 4a	Model 4b	Model 4c	Model 4d	Model 4e
Method			**Poisson**		
Dependent Variable	**Taxi**	**Metro**	**Bus**	**Ferry**	**Parking**
Main/Horizontal Effects	-0.011 ***	0.415 ***	0.286 ***	0.221 ***	0.367 ***
	(0.001)	(0.009)	(0.006)	(0.009)	(0.014)
Vertical/Mixed Effects	0.049 ***	0.276 ***	0.205 ***	0.146 ***	0.268 ***
	(0.002)	(0.013)	(0.005)	(0.013)	(0.012)
Rush Hour Policy	-0.049 ***	0.084 ***	0.052 ***	0.050 ***	0.023 ***
	(0.003)	(0.007)	(0.002)	(0.004)	(0.002)
Vertical/Mixed Effects of Rush Hour	-0.007 ***	0.158 ***	0.117 ***	0.109 ***	0.131 ***
	(0.001)	(0.005)	(0.004)	(-0.005)	(0.007)
Holiday	-0.145 ***	-0.354 ***	-0.315 ***	-0.599 ***	-0.351 **
	(0.006)	(0.014)	(0.001)	(0.019)	(0.115)
Precipitation	-0.031 **	-0.173 ***	-0.124 ***	-0.128 ***	-0.147 ***
	(0.011)	(0.002)	(0.005)	(0.010)	(0.007)
Temperature	-0.005 ***	-0.001	-0.002 ***	0.003 *	-0.003
	(0.001)	(0.001)	(0.001)	(0.001)	(0.011)
Constant	8.524 ***	0.933	6.858 ***	3.092 ***	-0.447 *
	(0.013)	(1.186)	(0.026)	(0.012)	(0.213)
Half-Hour Fixed Effects	Yes	Yes	Yes	Yes	Yes
Day Fixed Effects	Yes	Yes	Yes	Yes	Yes
No. of Observations	1,410	1,410	1,410	1,410	1,410
McFadden’s R-squared	0.973	0.991	0.993	0.989	0.907

Note. Robust standard errors are in parentheses.
†p < 0.1, *p < 0.05, **p < 0.01, ***p < 0.001.

3.4.3.4 Hypothetical Control Group Test

One limitation of my setting is the lack of a control group, which hinders us from interpreting the findings as if they were based on a natural experiment. For the purpose of ensuring the robustness of my results, I hypothetically create a “control group” of time that does not receive any effects of the policy. I choose 12:30 to 2:00 as the control group because it is the furthest time from the treated time period and, thus, theoretically, receives the fewest vertical/mixed effects, following my diminishing effects assumption. In fact, I believe that an individual who originally plans to travel during the rush hour but is forced change his or her plan due to the policy is unlikely to pick 12:30 to 2:00 as the substitute because it does not encompass working hours. I therefore assume and normalize this time period as receiving zero vertical/mixed effects. Since these periods receive neither main/horizontal nor vertical/mixed effects, they play the role of the control group. I also consider two cases of the control group: In the first case, I assume that the control group does not receive any effects from only the Tech Restrict Policy; in the second case, I assume that the control group does not receive any effects from either the Tech Restrict Policy or the Rush Hour Policy. The results are reported in Table 3.10 and 3.11, respectively. Both results reveal consistency with the main results.

3.5 EXTENDED ANALYSIS AND MECHANISM ANALYSIS

3.5.1 *Extended Analysis*

I further attempt to extend the scope of the policy impact to include private vehicle usage, which could be another major substitution for taxis other than public transportation. Given that fine-grained private vehicle usage data are difficult to collect, I use congestion-level data to approximate. Although the congestion level may reflect the use of private vehicles, taxis, and buses, it is largely dominated by private vehicles. Buses and taxis' contribution to the congestion level is constant before and after the policy, as they are in constant operation. From a policymaker's perspective, a high level of traffic congestion is an issue to be resolved. Specifically, I examine the policy's impact on the overall congestion level on the expressway and ground, *ln(congFree)* (Model 8g) and *ln(congSurface)* (Model 8h), respectively. Note that the congestion levels are continuous variables that follow a normal distribution after I take a log-transformation. To ensure the normality assumption of the error terms, I conduct the log-normal regression as below:

$$\log(y_{it}) = \beta_1 T_{it} + \beta_2 (f(D_{1i}) + f(D_{2i}))_{it} + \beta_3 R_{it} + \beta_4 (f_r(d_{1i}) + f_r(d_{2i})) + \beta_5 Controls_{it} + \mu + \gamma_i + \tau_t + \varepsilon_{it} \quad (3.\ 3)$$

where the error term is ε_{it}, and other parameters are consistent with equation (2). To better present the results, I also include an analysis (Model 8f), as a comparison. This aggregates four forms of public transportation to reflect the overall impact of apps usage restriction on public transportation systems.

Table 3.10. Control Group Test 1

	Model 5a	Model 5b	Model 5c	Model 5d	Model 5e
Method			**Poisson**		
Dependent Variable	**Taxi**	**Metro**	**Bus**	**Ferry**	**Parking**
Main/Horizontal Effects	-0.021 ***	0.364 ***	0.245 ***	0.190 ***	0.312 ***
	(0.002)	(0.017)	(0.008)	(0.017)	(0.018)
Vertical/Mixed Effects	0.126 ***	0.575 ***	0.430 ***	0.293 ***	0.571 ***
	(0.008)	(0.054)	(0.014)	(0.044)	(0.020)
Rush Hour Policy	-0.057 ***	0.089 ***	0.051 ***	0.052 ***	0.018
	(0.005)	(0.008)	(0.003)	(0.001)	(0.011)
Vertical/Mixed Effects of Rush Hour	-0.018 ***	0.155 ***	0.113 ***	0.111 ***	0.119 ***
	(0.002)	(0.013)	(0.006)	(0.012)	(0.012)
Holiday	-0.145 ***	-0.352 ***	-0.313 ***	-0.598 ***	-0.348 **
	(0.006)	(0.014)	(0.001)	(0.019)	(0.116)
Precipitation	-0.031 **	-0.173 ***	-0.124 ***	-0.128 ***	-0.147 ***
	(0.011)	(0.002)	(0.005)	(0.010)	(0.007)
Temperature	-0.004 ***	0.000	-0.002 ***	0.003 **	-0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.011)
Constant	8.520 ***	0.936	6.861 ***	3.097 ***	-0.448 *
	(0.015)	(1.178)	(0.031)	(0.005)	(0.208)
Half-Hour Fixed Effects	Yes	Yes	Yes	Yes	Yes
Day Fixed Effects	Yes	Yes	Yes	Yes	Yes
No. of Observations	1,410	1,410	1,410	1,410	1,410
McFadden's R-squared	0.973	0.991	0.993	0.989	0.907

Note. Robust standard errors are in parentheses.
†$p < 0.1$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Table 3.11. Control Group Test 2

	Model 6a	Model 6b	Model 6c	Model 6d	Model 6e
Method			**Poisson**		
Dependent Variable	**Taxi**	**Metro**	**Bus**	**Ferry**	**Parking**
Main/Horizontal Effects	-0.025 **	0.362 ***	0.242 ***	0.186 ***	0.311 ***
	(0.009)	(0.017)	(0.008)	(0.018)	(0.015)
Vertical/Mixed Effects	0.134 ***	0.576 ***	0.430 ***	0.290 ***	0.575 ***
	(0.008)	(0.053)	(0.015)	(0.043)	(0.020)
Rush Hour Policy	-0.064 ***	0.087 ***	0.048 ***	0.049 ***	0.015 +
	(0.013)	(0.007)	(0.003)	(0.001)	(0.009)
Vertical/Mixed Effects of Rush Hour	-0.022 *	0.153 ***	0.110 ***	0.108 ***	0.117 ***
	(0.009)	(0.013)	(0.006)	(0.014)	(0.010)
Holiday	-0.143 ***	-0.352 ***	-0.313 ***	-0.598 ***	-0.347 **
	(0.005)	(0.014)	(0.001)	(0.019)	(0.116)
Precipitation	-0.031 **	-0.173 ***	-0.124 ***	-0.128 ***	-0.147 ***
	(0.011)	(0.002)	(0.005)	(0.010)	(0.007)
Temperature	-0.004 ***	0.000	-0.002 ***	0.003 **	-0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.011)
Constant	8.506 ***	1.009	6.914 ***	3.152 ***	-0.390 *
	(0.032)	(1.107)	(0.079)	(0.045)	(0.154)
Half-Hour Fixed Effects	Yes	Yes	Yes	Yes	Yes
Day Fixed Effects	Yes	Yes	Yes	Yes	Yes
No. of Observations	1,410	1,410	1,410	1,410	1,410
McFadden's R-squared	0.973	0.991	0.993	0.989	0.907

Note. Robust standard errors are in parentheses.
†$p < 0.1$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

As presented in Table 3.12, I find consistent results as implied by my main results: an increase in public transportation usage at times with and without technology restriction ($\beta_{1f} = 0.144$, $p < 0.001$; $\beta_{2f} = 0.095$, $p < 0.001$), indicating that passengers are likely to shift to public transportation if taxi drivers' TNC app usage is restricted. I further find a significant increase in the congestion level on surface streets after the introduction of the Tech Restrict Policy ($\beta_{1g} = 0.072$, $p < 0.001$) and a marginally significant increase in the congestion level for other-than-enforcement time ($\beta_{2g} = 0.053$, $p < 0.001$).

Table 3.12. Impact of Technology Restriction on Taxi Demand and Congestion

	Model 7f	**Model 7g**	**Model 7h**
Method	**Poisson**	**Log-Normal**	
Dependent Variable	**PublicTrans**	**ln(congSurface)**	**ln(congFree)**
Main/Horizontal Effects	0.292 ***	0.146 ***	0.065
	(0.019)	(0.011)	(0.091)
Vertical/Mixed Effects	0.528 ***	0.176 +	-0.079
	(0.044)	(0.072)	(0.240)
Rush Hour Policy	0.059 ***	0.066 *	0.089 **
	(0.004)	(0.023)	(0.011)
Vertical/Mixed Effects of Rush Hour	0.135 ***	0.074 *	0.051
	(0.013)	(0.017)	(0.045)
Holiday	-0.358 ***	-0.026	-0.067
	(0.045)	(0.017)	(0.046)
Precipitation	-0.166 ***	-0.014	0.017
	(0.009)	(0.016)	(0.013)
Temperature	0.000	0.005	0.008
	(0.004)	(0.003)	(0.006)
Constant	6.910 ***	2.123 ***	2.875 ***
	(0.011)	(0.057)	(0.103)
Half-Hour Fixed Effect	Yes	Yes	Yes
Day Fixed Effects	Yes	Yes	Yes
No. of Observations	1,410	1,410	1,410
McFadden's R-	0.977	0.980	0.830

Note. Robust standard errors are in parentheses.
†$p < 0.1$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

There is, however, no significant change in congestion on the expressway for all of the time periods. Note that surface streets are more for short-distance trips, whereas the expressway is for longer trips. My findings indicate that passengers shift to private vehicles for short trips when the technology is restricted, which further implies the spillover of taxi rides due to technology restriction and that restricting technology decreases the efficiency of taxi operations.

3.5.2 *Mechanism Analysis*

Although I have shown the overall impact of technology restriction on transportation systems, my findings do not explain the underlying mechanism for how information sharing allows taxis to absorb more orders and, from the perspective of policymakers, why the policy results in fewer taxi transactions, which makes it even more difficult to find a taxi overall. I thus investigate the travel distance of taxis to uncover the underlying mechanism and to generate practical insights regarding policymaking. I provide two possible underlying mechanisms. Note that the policy directly affects drivers only and that the major change for the affected drivers is their access to the information of demand. Such information, however, may have varying effects on the driver and, consequently, their taxi operations.

The first mechanism takes into consideration the demand information that reflects order distance/duration: the *destination* information. Note that the major cost of taxi operations lies in idle time. When the uncertainty of the future demand/idle time is high, risk-averse drivers prefer long travel distance/duration orders from the apps to minimize their overall idle time. Such selectivity is widely observed in a taxi industry that is not well regulated and is reflected in the policymaking in my context. The destination disclosure allows risk-averse drivers to select long-distance orders from online channel. Because the objective of the policy is to protect the offline passengers with less access to the apps, removing the demand information from the apps is likely to restrict drivers from taking advantage of online orders with long travel distances. As a consequence, fairness between online and offline channels is rebuilt. Following this mechanism, I expect a decreased average distance after the policy shock, as drivers' selectivity will be dampened.

Another mechanism is the demand information of *availability of orders.* Because the apps significantly extend the choice set of passengers and reduce the search cost to find passengers, drivers' uncertainty over demand is significantly reduced. As drivers could always easily find passengers, they are less concerned about idle time. Note that the regulated pricing scheme for the taxi industry in Shanghai encourages drivers to benefit from shorter orders. The marginal revenue for taxi rides less than 3km is more than 4.66 RMB/km, whereas that for taxi rides more than 3km is only 2.40 RMB/km. Given the fixed capacity of drivers' travel distances and working hours, taxi drivers who are less concerned with idle time would maximize their profits by frequently taking a number of orders with short travel distances. Following this mechanism, I expect increased average travel distances because the policy brings back the uncertainty of future demand/idle time, which decreases drivers' preference for shorter distance orders.

To show which mechanism dominates, I investigate the impact of the policy that restricts drivers' apps usage in terms of average travel distance. I also consider the total distance to build a connection between the analysis of the total number of taxi transactions and that for average travel distance. Note that the distance is also a continuous variable that exhibits a normal shape after a log-transformation; thus, I repeatedly apply the model specified in Equation 3.3.

Table 3.13 presents the estimation results. The average travel distance per transaction is positive and statistically significant ($\beta_{1g} = 0.047$, $p < 0.05$), whereas total distance does not change significantly ($\beta_{1f} = 0.057$, $p > 0.1$). This indicates that the impact of the availability of orders is the predominant mechanism as compared to the revealed destination information. I note that this finding is consistent with my analysis of the congestion level, as I show that short-distance trips use private vehicles when drivers' apps usage is restricted. In addition, this finding is consistent with my main results: a decline in the total number of taxi transactions, possibly due to the lower availability of a vacant taxi, which results in potential taxi passengers' shifting to alternative transportation systems or taking a taxi at times when there is no policy enforcement.

Table 3.13. Impact of Technology Restriction on Travel Distance

	Model 8j	**Model 8k**
Method	**Log-Normal**	
Dependent Variable	**ln(totalTravelDist)**	**ln(avgTravelDist)**
Main/Horizontal Effects	0.057	0.047 *
	(0.063)	(0.022)
Vertical/Mixed Effects	0.376 +	0.119 *
	(0.187)	(0.045)
Rush Hour Policy	-0.081 *	-0.003
	(0.032)	(0.016)
Vertical/Mixed Effects	0.007	0.013
of Rush Hour	(0.038)	(0.014)
Holiday	-0.178 **	-0.027
	(0.052)	(0.029)
Precipitation	-0.025 *	0.002
	(0.012)	(0.008)
Temperature	0.000	0.006 *
	(0.006)	(0.003)
Constant	10.470 ***	1.947 ***
	(0.110)	(0.047)
Half-Hour Fixed Effects	Yes	Yes
Day Fixed Effects	Yes	Yes
No. of Observations	1,410	1,410
McFadden's R-squared	0.970	0.870

Note. Robust standard errors are in parentheses.
†$p < 0.1$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

The results of my mechanism analysis provide managerial insight into policymaking. Based on the disclosed motivation for the policy (Hu 2017), policymakers attempt to reset fairness between online and offline passengers by erasing the information advantage of drivers so that they cannot select trips with longer travel distances. This, however, overlooks the reduced uncertainty of future demand or its interactive effects with the taxi fare scheme in Shanghai. Information on the availability of orders provided from TNC apps enables drivers to be efficient and risk taking, as they were able to find subsequent taxi orders after each trip and earn more from many short trips than from fewer but longer trips. The results also provide practical insight into the problem of efficiency and fairness that is related to whether the platform should reveal the destination information of passengers. Although information about the destination is shown to improve the efficiency of TNC operations (Wang et al. 2019), some TNCs, e.g., Uber, Lyft, conceal it before drivers take the passengers. Considering my context, my results suggests that selectivity might depend more on the pricing scheme than on the revealed destination. A TNC could use pricing schemes to direct drivers' preferences, simultaneously gaining improved efficiency and directing the selectivity to a favored outcome.

3.6 Discussion and Conclusion

Employing information technology for information sharing in TNC services fosters transportation efficiency and supports passengers' convenience and flexibility. Nevertheless, the economic value of information sharing for TNC services and for the transportation system is not clearly understood. Differing from studies that examine the overall TNC impact, I investigate the impact of information sharing through information technology on transportation systems by leveraging a policy shock of technology restriction enforced by the Shanghai government. The policy allows us to clearly gauge the economic value of information sharing apart from other attributes of TNCs, e.g., pricing, additional supply, and to further investigate the underlying mechanism. I find concrete evidence that information sharing significantly improves the efficiency of taxis: A fixed number of taxis could undertake more orders for a fixed time and alleviate the traffic burden in alternative times; alternative public transportation modes, e.g., metro, bus, and ferry; and a combination thereof. I extend the scope to private vehicle usage and find results similar to those for public transportation. Taken together, these findings comprehensively demonstrate the externality value of information sharing in a TNC, along with its intrinsic economic value for TNC operations. My study thus contributes to both the extant literature stream on the impact of information sharing and the growing stream of the economic value of information sharing for TNCs.

I also contribute to the market design literature for taxis/TNCs by considering the mechanism of information sharing. I show the effectiveness of information sharing for improving TNC platform efficiency, thus alleviating the burden on alternative forms of transportation or alternative times. More interestingly, opposite to the perspective of policymakers and, perhaps, the public, I provide evidence that revealing the destination does not necessarily result in drivers' selectivity into long-distance orders. Considering drivers' rationality, I suggest that the information on demand, which extends the choice set of drivers, plays a dominant role and alleviates the uncertainty of drivers about idle time. The alleviation fundamentally corrects drivers' preference for long-distance orders and resets their selectivity to be determined by the pricing scheme. I thus provide a solution for the TNC platform to not only achieve the objective of maintaining the efficiency gain in operations derived from information sharing on the destination (Wang et al. 2019) but also moderate the "fairness" structure, using accommodating pricing schemes.

My study also provides managerial insight into practical concerns. First, I supply insight into policy evaluation. I show that restricting TNC app usage generates significant side effects, as it decreases the capacity of taxi rides. The cost of the policy was enormous when the initial objective of the policy was to protect offline passengers: Reduced transactions indicates that there is even lower chance of taking a taxi for the entire taxi market. Based on the mechanism analysis of the increasing average travel distance, I found a gap in policy design between the intended and realized outcomes. A later removal of the policy in December 2017 beyond my observation window confirms that that results were unexpected side effects and a miscalculation. To this end, encouraging and supporting offline passengers to be involved in utilizing TNC services is a better option than restricting drivers' use of TNCs. As TNCs in many countries share similarities with my context,[11] my evaluation could be helpful for their future policymaking in regard to the TNC market. Further, my results can be applied to other IT services with two-sided markets when investigating the satisfactory equilibrium point for information sharing of two sides of users.

Beyond evaluating the focal Tech Restrict Policy and identifying its side effects, I offer managerial implications for TNC platform designers and policymakers. First, by showing the isolated impact of information sharing, I justify the economic value of TNCs. Thus, I suggest that policymakers and regulators not restrict the information sharing of TNCs, as it is not only valuable for TNC users but also efficient for the overall transportation ecosystem. With the restriction on drivers' TNC usage, passengers significantly shift to alternative forms of transportation, such as public transit. Although one can say that this is a positive effect of policy, as it encourages the use of public transit, it can be a serious negative externality to the incumbent public transit users if there is sudden rush of users when the infrastructure is constant. This also applies to daily travelers or residents, whereby there is an increase in the congestion level on the surface roads with a policy introduction; this can be interpreted as another externality, as it results in an increase in private vehicle use on surface roads.

Second, for platform designers, I suggest customizing information sharing as an effective solution to improve the platform's operation efficiency. For instance, experiments are required to

[11] The TNC in many countries (i.e., parts of Australia, Belgium, Bulgaria, Canada, Czech, Denmark, Hungary, Germany, Greece, Israel, Italy, Japan, Netherland, South Korea, and Turkey) are based on taxis rather than on private vehicles. Most other countries (i.e.., European Union-member countries, the United Kingdom, South Africa, and Saudi Arabia, and parts of India, Brazil, Australia, and Canada) have required or are in the process of requiring TNC drivers to be licensed, as required by taxi services, with limits on the licenses so that the supply is regulated. In the United States, Uber recently included taxi drivers and provided UberTaxis so that taxi drivers also could leverage the value of information technology from TNC.s

test the type and extent of information to reveal as well as how, when, and to whom to reveal certain information. In my research context, the availability of orders should be revealed for drivers' efficiency, whereas destination information should be concealed to protect the passengers. I suggest that platform designers consider the net value of information sharing by disentangling environmental structures, e.g., pricing schemes, when designing the platform. Although concealing the destination information establishes fairness for the passengers, alternative approaches can control the drivers' selective behaviors and achieve better efficiency by satisfying both platform designers and policymakers' specific objectives. In sum, the TNC platform and related policies could be designed more precisely for specific objectives.

My study has a few limitations that suggest avenues for future research. First, in the current study, I focused on the aggregate level of the Shanghai municipality due to data limitations. If finer-grained, e.g., district-level, data are available, researchers might generate further insight by considering spatial factors. Second, in my setting, the company did not provide the identified channel (TNC-based or roadside) for each transaction made for a taxi. It would be intriguing to examine the differing impact of the Tech Restrict Policy between offline and online channels. Third, my context is specific to TNC technology. I hope that future studies investigate various industries to generate deeper insight into technology restriction.

Chapter 4. THE RIGHT ONLINE CONNECTION THAT HELPS YOU LAND A JOB – SOCIAL HIRING VIA LINKEDIN

4.1 INTRODUCTION

Social hiring, also known as social recruiting, describes a job hiring process wherein employers and job seekers rely on social media channels to engage and inform each other (Jacobs 2009). The key benefit of professional networking sites (e.g., LinkedIn.com) is access to a large-scale professional social network. According to *Fortune* magazine (Darrow 2017), LinkedIn comprises over 500 million professionals who are either actively or passively searching for jobs. Among these professionals, more than 50% were reported to be active users as of March 2017. Leveraging this large-scale network, recruiters can reach out to a broad pool of talented job seekers. A study documented that 94% of recruiters utilize LinkedIn to search for appropriate job seekers (Stadd 2013). For instance, when a position is open, many employers start their hiring process by identifying qualified job candidates from professional networking sites. The benefit of professional networking sites to job seekers is the opportunity to expand their social networks beyond their family, friends, or alumni networks, which is expected to increase their chance of securing referrals (Bersin 2012, Jacobs 2009).

Recognizing recruiters' increasing reliance on professional networking sites, job seekers make efforts to create connections that are helpful to their job searches. In particular, a common tactic is to make connections with a target company's employees, because these connections can not only offer timely and accurate information about job openings but also increase exposure to the company's recruiters. Hence, many MBA programs advise students to build connections with employees who are currently working at the company to which they plan to apply, and preferably those working at the target function (Gilchrist 2018). For instance, if students are looking for a data analyst position at Google, they are advised to make connections with the company's data analysts or software engineers rather than with corporate finance personnel. Through these connections, job seekers aim to increase their chance of obtaining referrals.

The goal of this paper is to empirically investigate whether this common networking tactic is actually beneficial in getting referrals. To answer this question, I examine how job similarity between a job seeker and a referring employee affects the likelihood of referral. Although people in similar fields seem to have a lot to offer to job seekers, I propose that they might not be such a promising source of referrals. This is because, from the perspective of an employee, qualified job candidates in similar fields are potential competitors who can threaten their status in the workplace. I further explore the factors that can weaken this proposed adverse effect of job similarity on referral outcome.

My empirical analyses reveal interesting findings that are robust to multiple falsification tests. First, much to my surprise, I find that this common networking tactic actually delivers an unfavorable referral outcome: The likelihood of referral decreases as job similarity between an employee and a job candidate increases. This indicates that candidates are less likely to be referred by employees who are doing similar jobs. This finding would be disappointing to many job seekers as they strive to make LinkedIn connections with employees who have similar expertise in their target companies. I attribute this result to peer competition. That is, employees tend to protect themselves from potential competitors who can vie with them for tasks, promotions, or rewards in the future. Further empirical investigation reveals that this adverse effect of job similarity on referral weakens as the hierarchical level of the referring employee becomes greater than that of the job candidate, because they are less likely to compete directly. Although I expect that gender homophily between a job seeker and an employee might help dilute the competitive force, I find that it does not necessarily overcome the force in my context. In sum, a hierarchical difference between a job candidate and an employee helps to surmount the negative job similarity effect on referrals, whereas gender homophily is not powerful enough to overcome it.

This study advances knowledge in several ways. First of all, it advances my understanding of the antecedents of referral-based hiring. Previous studies have documented several dyad-level as well as individual-level characteristics that influence referral decision. For instance, tie strength (Granovetter 1973), gender homophily (Berger 1995), and qualification (Smith 2005) are documented as antecedents of referrals. My study extends this stream of literature by introducing job similarity as an inhibitor of job referrals. Also, by simultaneously examining various influencing factors (i.e., job similarity, hierarchical difference, gender homophily, tie strength, individual characteristics), my study offers a more nuanced understanding of the factors that influence referral decisions. In addition, this study joins a small number of recent studies investigating the impact of online network ties on job search outcomes. Despite the growing reliance on online network sites for job search, only a few studies have examined its implications on job search outcomes. I provide more in-depth discussions about my contribution to the prior literature in §2. Practically, my findings reveal that the common strategy of job seekers making connections with employees in the target company and job function may not be instrumental in obtaining referrals. Further, my results inform companies about the potential drawback of utilizing employee referrals.

The remainder of this paper is organized as follows. In §2, I review related literature and discuss the contribution of this study. In §3, I develop my hypotheses about how job similarity between a job candidate and a referrer may affect employees' decision to refer. In §4, I describe my empirical context, data, and estimation model. In §5 and §6, I present my results and check their robustness. In §7, I conclude.

4.2 Theory and Hypotheses

4.2.1 *Job Similarity as an Inhibitor of Job Referral*

Recognizing recruiters' increasing reliance on professional networking sites, job seekers strive to create many professional connections. In particular, job seekers aim to make connections with professionals who work for the target companies in their target fields. This is because people in similar fields can offer accurate and timely information about new job openings and can introduce job seekers to useful contacts (Makra 2015). In addition, such connections may also enhance job seekers' exposure to target recruiters because recruiters actively utilize their own employees' social connections to search for qualified candidates (Petrone 2016). Additionally, job seekers can directly ask for employees' referrals (Perez 2018).

Although employees in similar fields seem to have a lot to offer to job seekers, when it comes to referrals, they may choose not to. According to Bendersky and Hays (2012), individuals are concerned with their status compared to that of fellow group members and, consequently, tend to compete with each other to maintain status. Research has shown, for instance, that an excessive similarity in status among group members tends to reduce collaboration and the quality of information exchange (Hambrick 1994). Moreover, individuals make efforts to prevent other members from gaining status (Overbeck et al. 2005). Previous research also found that people tend to compare themselves with others who are most similar to them. This means that resembling rivals exhibit greater social comparison or competition concerns than those who are less similar in terms of performance or characteristics (Garcia et al. 2013, Goethals and Darley 1977, Kilduff et al.

2010). Given this tendency of individuals to compete with similar group members, it is plausible to think that employees are likely to be reluctant to refer job seekers with similar expertise.

In brief, job seekers may be overlooking a potential adverse effect of job similarity between themselves and employees in obtaining referrals: Employees may not want to refer candidates with similar expertise to protect themselves from future competition. I therefore hypothesize the following.

H1: As job similarity between a candidate and an employee increases, candidates are less likely to be referred to a job.

4.2.2 *Moderating Effect of Hierarchical Difference*

I further propose that the adverse effect of job similarity on referral outcome is likely to weaken as the hierarchical level of an employee gets higher relative to that of a job candidate, because the hierarchical difference lessens direct competition. Even within the same field, different job levels require different skillsets (Pavett and Lau 1983). For instance, software developers are evaluated mainly by their technical skills (e.g., creating an algorithm). In contrast, soft skills like communication and project management skills are also crucial for project leads because they need to manage software development projects and their team members, on top of software development tasks. Due to this difference in required skillsets, individuals in different hierarchical levels usually do not directly compete with each other (Leicht and Marx 1997). This diminished direct competition is likely to lead individuals to be less concerned about maintaining their status, and in turn, to be more willing to refer others in different hierarchical levels. The referring employees' power is likely to alleviate the competition threat that is driven by job similarity on referrals. According to Salancik and Pfeffer (1977), power is the ability to get things done the way one wants them to be done. It has been shown that the power of a referrer can play a positive role on the likelihood of referrals (Schlachter 2018). If employees' power is stronger than that of the candidate, which is the case when employees are in a higher job level than job candidates, it is likely that employees feel less threat of competition from job candidates.

In sum, the hierarchical difference between job candidates and employees, through which they can avoid direct competition, is likely to weaken the negative effect of job similarity on referral outcome. I therefore hypothesize the following.

H2: The negative influence of job similarity on a referral is likely to be reduced as the hierarchical level of an employee gets higher relative to that of a job candidate.

4.2.3 *Moderating Effect of Gender Similarity*

Gender homophily is known to play an important role in seeking as well as making referrals. It has been found that job candidates prefer the same gender contacts when utilizing their professional network connections for job searches: Men prefer male contacts and women prefer female contacts (Berger 1995). Likewise, in making referral decisions, referrers are known to favor same-gender job seekers. For instance, in the setting of the call center of a large bank, Fernandez and Sosa (2005) found that in the employee referral environment, both men and women showed a significant level of bias towards the same gender in their referrals. This tendency is so strong that it even contributes to segregation in working environments, as one gender tends to dominate certain types of occupations (Bielby and Baron 1986, Kalleberg et al. 1996). Preference toward the same gender is also found in an online labor market. In an online freelancer market, Chan and Wang (2018) found that the presence of a gender hiring bias resulted in unfair outcomes, as it restricted job access to a particular group of workers based on their gender, even though workers of the opposite gender were equipped with relevant abilities and credentials.

Considering this strong preference for the same gender referrals, I expect that gender homophily between a job candidate and an employee can relieve the negative competition effect generated from job similarity. That is, if a job candidate is the same gender as the employee, the employee's strong preference for the same gender may overcome the competition threat driven by job similarity. Hence, targeting employees who are of the same gender might help job candidates procure referrals in the employee referral process. I therefore hypothesize the following.

H3: The negative influence of job similarity on referral is likely to be reduced if a candidate is the same gender as an employee.

4.3 EMPIRICAL METHOD

In this section, I describe my research context and data, followed by the measure descriptions and my model specifications.

4.3.1 *Research Context*

The data for my study comes from a global management consulting company (which I abbreviate to "the company") that provides various services including strategy consulting, digital technology, and operations services to its client companies. The company is a Fortune 500 company with net revenues of $34.9 billion at the end of 2017. In 2014, the company introduced a social hiring initiative that leverages its own employees' LinkedIn connections in the hiring process. Specifically, the company incorporates its employees' LinkedIn connections in identifying qualified job candidates in the following way: For each job opening, the company searches through the LinkedIn profiles of people who are directly connected to its employees[12] and sorts them based on their fit with the open position. This screening process is conducted based on the degree of

[12] The LinkedIn connections of employees are automatically updated in the system.

suitability of a connected individual's LinkedIn profile information (e.g., job title, job level, industry, location, etc.) to the job description of an open position.

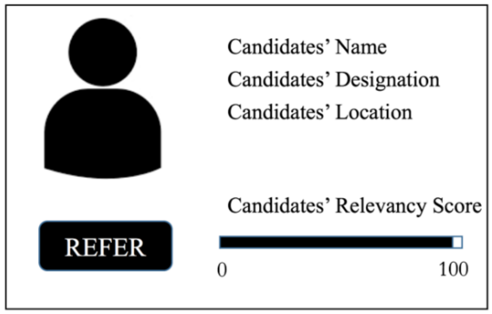

Figure 4.1. Candidate Profile Shown to Employee

Based on the fit determined based on the objective information, a handful of individuals are identified as "qualified job candidates" (referred to as "candidates" from now on) for the open job. The company then sends a notification to employees who are connected to those qualified job candidates and asks whether they are willing to refer them for the open position. This notification is sent through a mobile application that provides information about a job candidate's level of fitness for a job (i.e., relevancy score), a picture, name, and the current job title. Figure 4.1 illustrates the reproduced screenshot of such a notification.[13] Upon receipt, employees can refer a potential job candidate by clicking the refer button on the app. If a candidate is referred, the candidate is notified directly via personal or LinkedIn email. Figure 4.2 illustrates the process of how employee referrals take place based on LinkedIn connections.

4.3.2 *Data and Job Referral Network*

My data span 23 months from January 2014 to November 2015. During my empirical period, the company had a total of 8,026 job openings. The company provided us with a random sample of 8,609 qualified job candidates who were identified through the LinkedIn connections of 109 of

[13] Due to the non-disclosure agreement, I cannot present the actual screenshot of the mobile application.

the company's employees. Based on the data, I constructed the referral network between employees (j) and qualified job candidates (i) for each job k. In the network, each job posting k is connected to qualified job candidates (i) who are identified through LinkedIn connections of employees (j). The company may identify multiple candidates for one job opening. Each candidate i may also be connected to multiple employees. There are 9,089 candidates (i) – employee (j) unique pairs in my referral network. Figure 4.3 illustrates my referral network structure. Based on this referral network, I analyze how the tie characteristics of candidate (i) – employee (j) influence the likelihood of an employee referring a candidate to a job.

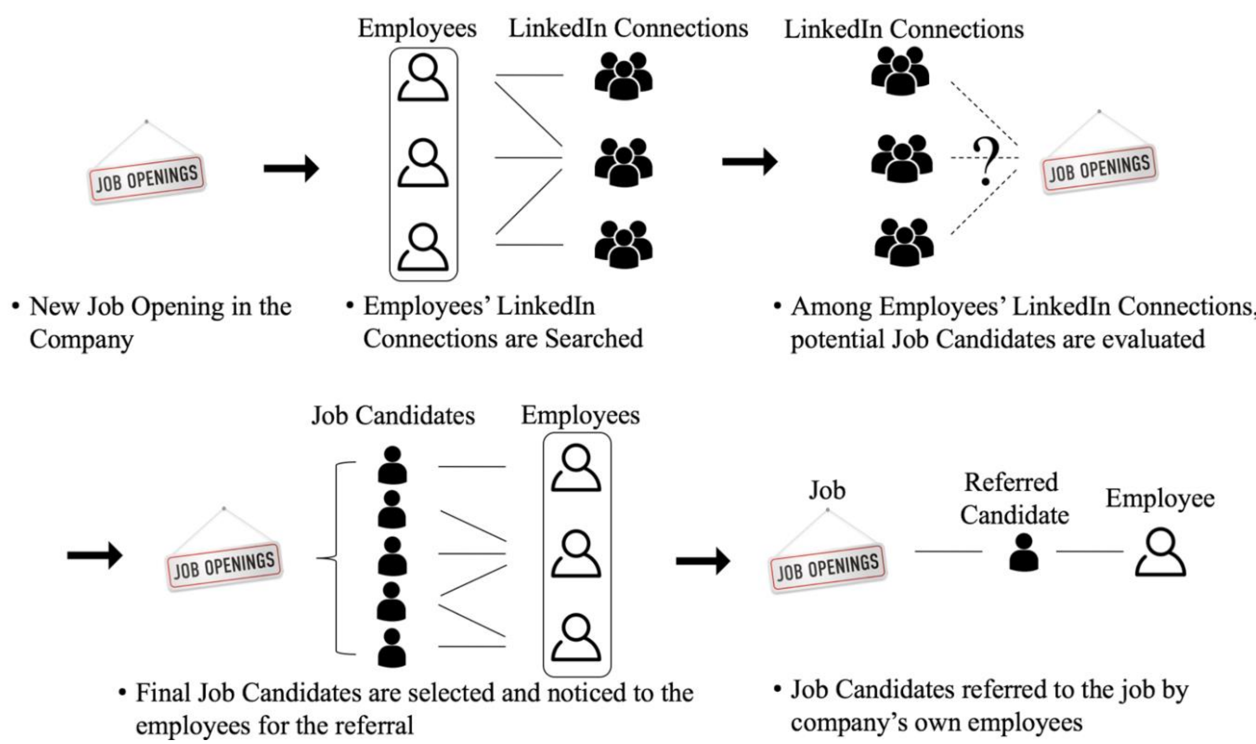

Figure 4.2. Social Hiring Process in My Empirical Context

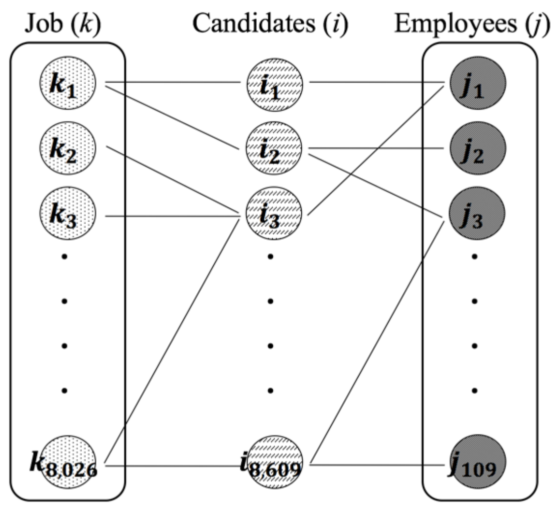

Figure 4.3. Referral Network Structure

To answer my main research question, the effect of job similarity between a candidate and a referrer on referral likelihood, I need to quantify how similar an employee's and a candidate's job functions are. For instance, in Figure 3, if employee j_1 is a software engineer, candidate i_1 is a data scientist, and candidate i_3 is a human resource manager, I need to quantify the level of job similarity between employee j_1 and candidate i_1 as well as that between employee j_1 and candidate i_3 to examine the effect of job similarity on the likelihood of employee j_1 to refer i_1 and/or i_3. As my data provide information about candidates, jobs, and employees in an unstructured text format, it is challenging to classify and quantify the job similarities between an employee and a job candidate using human coders. So, I use a recent natural language modeling technique, word embedding (Mikolov et al. 2013), to process the unstructured data.

Word embedding maps words or phrases from the vocabulary to vectors of real numbers, allowing us to handle the intensive amount of plain text or strings and extract information. Key to this technique is to use a densely distributed representation for each word: Each word is represented by a real-valued vector, in the high level of dimensions. Methods similar to word embeddings, such as Latent Semantic Analysis (LSA) or Indexing (LSI), have existed in various forms (Dumais 2004). However, recent breakthroughs in word embedding models enable researchers to incorporate relevant information about word contexts from highly local windows of surrounding terms rather than an entire document. Thus, high-quality word embeddings can capture the syntactic as well as semantic similarity between words (Mikolov et al. 2013). I trained a word embedding using the GloVe algorithm. The GloVe is an unsupervised learning algorithm for obtaining vector representations for words. In GloVe, the training process is performed on aggregated global word co-occurrence statistics from a corpus, and the result represents linear substructures of the word vector space (Pennington et al. 2014). The GloVe is structured so that vector differences between two word-vectors capture as many different relationships as possible that describe two words. This is very useful because it overcomes some of the limitations of traditional ways of interpreting word relationships—for example, similarity metrics used for nearest neighbor evaluations, in which only a single scalar is used to quantify the relationship between two words. Such simplicity can be problematic, because two given words may convey more complex relationships beyond one type of relationship that is shown by a single scalar. For instance, “boy” and “girl” may be the same type of words, as they both describe human beings. On the other hand, the two words are often considered opposites because they represent different genders of humans. Therefore, the model needs to relate more than a single number to the pair of words to capture the nuanced differences between boy and girl quantitatively.

By employing the word embedding technique, I quantified job similarity between a candidate and an employee. Specifically, I followed the four steps. First, I merged the entire corpus of job qualification descriptions and cleaned up the dataset so that it comprises only English. All texts were split into one sentence per line and converted to lowercase. I then created a set of words from this corpus, which I want to learn word vectors from. Second, I constructed the term co-occurrence matrix. The co-occurrence matrix presents how frequently words co-occur with one another in a given corpus. In the main analysis, I constrain a given corpus to a window of five terms.[14] I trained the GloVe algorithm on these non-zero entries of a global term co-occurrence matrix (TCM). Third, with a vector representation of words constructed from the training data set, I predicted the word vectors for terms in my final documents. In my setting, a document refers to job titles of candidates and employees. I created a document-term matrix (DTM) in which each row represents an individual document and each column represents terms that exist in the term co-occurrence matrix. Thus, each cell in the matrix informs us whether the document has a certain term or not: If a document incorporates a certain term, the cell value is 1 and zero otherwise. I created a product of a document-term matrix and word vectors, resulting in final word vectors for each job titles.

Candidate(i)'s Job Title	**Employee(j)'s Job Title**	$\boldsymbol{Job\ Similarity_{ij}}$	**Comparison**
Quality Analyst	Lead Quality Assurance Engineer	0.412	More Similar
Mobile Web Developer	Soft Consultant	0.287	
…	…	…	
Sourcing Recruiting Consultant	Software Engineer Technology	-0.208	
Software Engineer	Strategy Consultant	-0.326	Less Similar

Figure 4.4. Candidate-Employee Job Similarity

Finally, using the word vectors, I calculated the job similarity measure for every one of the unique 9,089 candidates (i) – employee (j) pairs in my referral network. I used a cosine similarity function to calculate similarities between word vectors for candidate i and word vectors for employee j. The range of cosine similarity is -1 to 1. The closer the cosine similarity value is to 1, the higher the job similarity between a candidate and an employee, and the closer the similarity value is to -1, the lower the job similarity between a candidate and an employee. Figure 4.4 presents

[14] For robustness check, I also used differing length of windows, and the results are qualitatively similar.

examples of candidate – employee job similarity measures. I followed the same procedure to calculate candidate i ’s fitness for a job k.

4.3.3 *Measures*

Table 1 lists and describes all model variables.

Table 4.14. Model Variable Descriptions

Covariates	Type*	Description		Measure
$Referred_{ijk}$	DV	Whether candidate *i* is referred to job *k* by employee *j*	1: 0:	If the candidate *i* is referred to job *k* by employee *j* Otherwise
Dyad-Level Covariates				
$JobSimilarity_{ij}$	IV	A similarity between candidate *i*’s and employee *j*’s job functions		Cosine similarity between word vectors of candidate *i*’s job title and employee *j*’s job title
$HierarchicalDifference_{ij}$	IV	Difference between job level of candidate *i* and employee *j*.		Candidate *i*’s job level subtracted from employee *j*’s job level
$GenderDifference_{ij}$	IV	Gender similarity between candidate *i* and employee *j*.	1: 0:	If candidate *i* and employee *j* are different gender Otherwise
$JobSim_{ij} \times HierarchicalDiff_{ij}$	IV	Interaction term between $HierarchicalDifference_{ij}$ and $JobSimilarity_{ij}$		
$JobSim_{ij} \times GenderDiff_{ij}$	IV	Interaction term between $GenderDifference_{ij}$ and $JobSimilarity_{ij}$		
$CandidateFitnesstoaJob_{ik}$	C	The extent to which candidate *i* is qualified for the job *k*		Cosine similarity between word vectors of job *k*’s job title and candidate *i*’s job title
$GeographicalDistance_{ik}$	C	The geographical distance (miles) between candidate *i*’s location and the location of job *k*		
$ConnectionTime_{ij}$	C	Duration of candidate *i* and employee *j*’s LinkedIn connection		The number of months candidate *i* and employee *j* have been connected on LinkedIn
Individual-Level Covariates				
$EmployeeReferralBehavior_j$	C	Employee *j*’s activeness in making referrals		Total number of candidates employee *j* referred during our empirical period
$JobLevel_k$	C	Job level of open position *k*		Job level ranges from 1 to 7 in the company. Level 1 is the lowest, level 7 the highest
$TimeJobPosted_k$	C	Time job *k* has been posted. Monthly time index		

*Type: DV (Dependent Variable), IV (Independent Variable), C (Control)

4.3.3.1 Dependent Variable

$Referred_{ijk}$. The dependent variable of this study is whether candidate i is referred to job k by employee j. I consider that the candidate is referred if the state of “candidate” i has changed to "referred candidate" in the provided data set. Approximately 23% of employees made at least one referral through the social hiring platform.

4.3.3.2 Independent Variables

$Job\,Similarity_{ij}$. My measure of job similarity between an employee and a candidate is based on their job titles in the company. By using the word embedding technique that I described in §4.2, I retrieved vectors for each of the words that compose job titles. I then aggregated word vectors to represent each job title. I computed cosine similarity using the final word vectors of employees' and candidates' job titles. The cosine similarity measure determines whether the vectors representing job titles of employees and candidates are pointing to the same direction or not by calculating the cosine angle of the two different combinations of word vectors. Previous studies have adopted the cosine similarity measure to quantify the expertise similarity between two participants in a knowledge forum (Hwang et al. 2015, Zhang et al. 2007). The formula for the job similarity of candidate i and employee j is defined as below:

$$Job\,Similarity_{ij} = \frac{\vec{a}\cdot\vec{b}}{\|\vec{a}\|\|\vec{b}\|} \tag{4.4}$$

Where $\vec{a} = (a_1, a_2, a_3, ..., a_n)$ and $\vec{b} = (b_1, b_2, b_3, ..., b_n)$ are two different word vectors representing job titles of candidate i and employee j, a_n and b_n are the components and n is the dimension of the word vector. The resulting similarity ranges from -1 to 1, in which 1 indicates that the two vectors have the same orientation, whereas -1 indicates that two vectors are antithetical. The greater overlap indicates that candidates are from similar job functions as the employees. The mean value of 9,089 candidates (i) – employee (j) pairs in my referral network is 0.06. My preliminary analysis shows that the mean value of the job similarity of referred candidates is lower (*M* = 0.019, *SD* = 0.133) than that of non-referred candidates (*M* = 0.063, *SD* = 0.142), as depicted in Figure

4.5, and the difference is statistically significant ($t(121,454) = 2.54$, $p < 0.05$). This descriptive statistic presents model-free evidence of the proposed negative effect of job similarity on the likelihood of referral.

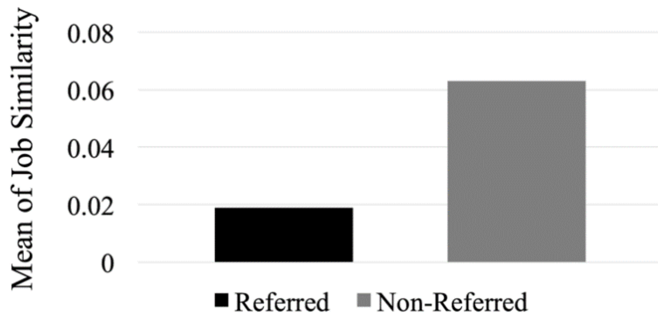

Figure 4.5. Mean Job Similarity of Non-referred and Referred Candidates

$Hierarchical\ Difference_{ij}$. Job levels of all candidates and employees were coded from 1 to 7 following the official hierarchical job level structure of the company where 1 is the lowest level in the job hierarchy (e.g., an analyst) and 7 is the highest (e.g., a senior manager). The difference between a candidate's and an employee's job level was calculated by subtracting a candidate's level from an employee's level. Positive $Hierarchical\ Difference_{ij}$ indicates that employee j's level is higher than candidate i's and vice versa. Table 2 presents the distribution of the values of $Hierarchical\ Difference_{ij}$.

Table 4.15. Statistics on Hierarchical Difference

$\text{Hierachical Difference}_{ij}$	Freq.	Percent (%)
-4	5,446	4.46
-3	10,752	8.81
-2	20,705	16.97
-1	13,856	11.36
0	18,399	15.08
1	12,033	9.86
2	12,494	10.24
3	28,313	23.21
Total	121,998	100

$Gender\ Difference_{ij}$. I coded $Gender\ Difference_{ij}$ as 1 if an employee-candidate pair has different gender, and 0 otherwise. For instance, a Male-Female pair was coded as 1, whereas a Male-Male pair was coded as 0. Table 3 presents statistics of candidate-employee pair gender combinations.

Table 4.16. Statistics on Gender Combinations

Gender Combination (Candidate (i) – Employee (j))		Freq.	Percent (%)
$Same\ Gender_{ij}$	$Male\text{-}Male_{ij}$	50,545	46.85
	$Female\text{-}Female_{ij}$	11,308	10.48
$Different\ Gender_{ij}$	$Male\text{-}Female_{ij}$	24,609	22.81
	$Female\text{-}Male_{ij}$	21,434	19.87
Total		107,896	100

4.3.3.3 Control Variables

The decision to refer a candidate for a job can also be driven by other factors. To tease out the proposed effect of job similarity on the likelihood of referral, I incorporated a number of control variables. First, an employee's decision to refer a candidate is likely to be influenced by the qualification of a candidate. To control for this effect, I included the variable $Candidate\ Fitness\ to\ a\ Job_k$. As I constructed the job similarity measure, I employed the word embedding technique to compute the cosine similarity between candidate i's LinkedIn profile and job *k*'s description. The value closer to 1 indicates that a candidate is more qualified for job k, and the value closer to -1 indicates that a candidate is less qualified for job *k*.

Second, it is also plausible that the likelihood of referral is affected by the geographical proximity of candidate i to job k. If offered a job, candidates who are geographically proximate would incur less relocation cost, which would increase their probability of accepting a job offer. This higher chance of offer acceptance may influence an employee's decision to refer a candidate to a certain job. To control for this potential influence of geographical proximity on referral decision, I incorporated the variable $\mathit{Geographical\ Distance}_{ik}$. It is the shortest distance between candidate i's location and the location of job *k*.

Third, I incorporated a proxy for tie strength between candidate i and employee j. Previous studies have found that the strength of a tie influences the instrumentality of the tie in a job referral process (Liu and Duff 1972; Granovetter 1973; Yakubovich 2005; Marin 2012; Gee et al. 2017; Garg and Telang 2017). The empirical results are mixed so far: Some studies found that weak ties are more useful in eliciting referrals than strong ties (e.g., Granovetter 1973, Yakubovich 2005), whereas others documented the opposite pattern in the online platform (e.g., Garg and Telang 2017; Burke and Kraut 2013; Gee et al 2017). Although the effect from tie strength is not the focus of this study, I controlled for its effect on referral likelihood by incorporating the variable $\mathit{Connection\ Time}_{ij}$ in my main model. $\mathit{Connection\ Time}_{ij}$ serves as a proxy for the tie strength between a candidate and an employee and is measured in terms of the amount of time candidate i and employee j have been connected to each other in LinkedIn. As a robustness check, I incorporated an alternative tie strength measure, an inverse of total LinkedIn connections of employees, into my model. The details of the variable $\mathit{Inverse\ of\ LinkedIn\ Connections}_{i}$ and results of the robustness check are reported in §6.

Fourth, I controlled for the hierarchical level of open position k, $\mathit{Job\ Level}_{k}$. It is plausible to say that the frequency of job openings differs across different job levels: Certain positions in the firms' hierarchy are more, or less, susceptible to referrals than other hiring strategies (Marsden 1994). For instance, senior-level jobs are less likely to be advertised and follow a formal job hiring process than junior-level jobs. Instead, they are often scouted unofficially from other firms via top managers' interim network ties. This practice happens because senior executives tend to have a huge impact on the future and esprit of the organization (Fernández-Aráoz 2001; Williamson and

Cable 2003). To control for such heterogeneities of the hiring process across job levels, I controlled for the job level of open positions.

Fifth, I further controlled for the heterogeneity of each employee's referral behavior by incorporating the variable $Employee\,Referral\,Behavior_j$. Some employees may be quite active in referring through this social hiring platform, whereas others may not. By controlling for the referral frequency of each employee within a given time span, this variable controls for such heterogeneity. Sixth, to control for the unobservable temporal shocks, my model incorporated $Time\,Job\,Posted_k$, the number of months that job k was posted. Seventh, I included random effects for each employee and a candidate in an employee-candidate pair to control for other unobserved heterogeneities that are not captured by model variables. Lastly, I also incorporated random effects for each employee-candidate pair to control for uncaptured unobserved dyad-specific heterogeneity such as their offline tie strength.

4.3.4 *Model Specification*

I estimated parameters by proposing the following equation. The unit of analysis of the empirical model is candidate (i) – employee (j) – job (k) level, where i indicates a job candidate, j refers to an employee and k refers to a job opening. For estimation, I mean-centered all continuous independent and control variables in order to minimize potential multicollinearity issue (Gelman and Hill 2007).

$$\begin{aligned} Referred_{ijk} = {} & \beta_1 Job\,Similarity_{ij} + \beta_2 Hierarchical\,Difference_{ij} + \beta_3 Gender\,Difference_{ij} \\ & + \beta_4 JobSim_{ij} \times HierarchicalDiff_{ij} + \beta_5 JobSim_{ij} \times GenderDiff_{ij} + \delta_{1ik} Candidates\,Firness\,to\,a\,Job_{ik} \\ & + \delta_{2ik} Geographical\,Distance_{ik} + \delta_{3ij} Connection\,Time_{ij} + \delta_{4j} Employee\,Referral\,Behavior_j \\ & + \delta_{5k} Job\,Level_k + \delta_{6k} Time\,Job\,Posted_k + \alpha + \gamma_i + \tau_j + d_{ij} + \varepsilon_{ijk} \end{aligned} \tag{4.5}$$

where α is intercept term $\beta_{ij} = (\beta_1, \beta_2, \beta_3, \beta_4, \beta_5)$, where β_1, β_4, and β_5 are the estimates for my hypotheses. $\delta = (\delta_{1ik}, \delta_{2ik}, \delta_{3ij}, \delta_{4j}, \delta_{5k}, \delta_{6k})$ consists of a candidate (i) – employee (j), candidate (i) – job (k), employee (j) and job (k) specific control variables that may influence referral outcomes. Control variables include candidate i 's fitness for a new job k ($Candidate\ Fitness\ to\ a\ Job_k$), geographical distance between candidate i and job opening k ($Geographical\ Distance_{ik}$),time the candidate and employee have been connected via LinkedIn ($Connection\ Time_{ij}$), employee j 's referral behavior ($Employee\ Referral\ Behavior_j$), and the job level of job k ($Job\ Level_k$). $Time\ Job\ Posted_k$ captures unobserved time shocks.

In my empirical setting, observations are not independent, because the same candidate can be referred by multiple employees. In addition, the same employee may refer multiple candidates. That is, each candidate or employee can be a member of multiple dyads, which makes the error terms to be correlated across observations. If I do not account for such dependence, my model will produce artificially reduced standard errors. Among many solutions to the problem (Simpson 2001), I included random effects of each individual in each dyad: γ_i for candidate i and τ_i for employee j, to allow my dependent variable $Referred_{ijk}$ to vary across multiple observations on a single individual that is nested within dyads (Greene 2011). I conducted a Hausman test to ensure that a random effects model is unbiased and adequate compared to the fixed effects model (Hausman 1987). The test statistics ($\chi^2(4) = 0.03,\ p > 0.05$) indicate that the random effects model is unbiased. In addition, the unobserved candidate (i) – employee (j) dyad-specific characteristics are captured by using a dyad-specific unobserved random effect, d_{ij}. Given that my dependent variable is dichotomous, logistic regression is used to estimate my parameters.

4.4 RESULTS

Table 4 presents the summary statistics of the model variables based on the raw values before mean-centering. The correlations of the model variables are reported in Table 5.

Table 4.17. Summary Statistics

Variable	Mean	Std.Dev	Min	Max
$JobSimilarity_{ij}$	0.06	0.14	-0.22	0.41
$Hierarchical\ Difference_{ij}$	0.10	2.21	-4	3
$Gender\ Difference_{ij}$	0.43	0.49	0	1
$JobSim_{ij} \times HierarchicalDiff_{ij}$	0.03	0.38	-1.64	1.23
$JobSim_{ij} \times GenderDiff_{ij}$	0.03	0.09	-0.22	0.41
$Candidate\ Fitness\ to\ a\ Job_{ik}$	0.78	0.08	0.59	0.99
$Geographical\ Distance_{ik}$	1121.53	699.66	12.83	2983.86
$Connection\ Time_{ij}$	10.45	6.84	1	23
$Employee\ Referral\ Behavior_j$	2.16	3.02	0	11
$Job\ Level_k$	4.42	0.85	2	6
$Time\ Job\ Posted_k$	9.45	2.48	8	23

Table 4.18. Correlation among Variables

	1	2	3	4	5	6	7	8	9	10	11
$Job\ Similarity_{ij}$	1.00										
$Hierarchical\ Difference_{ij}$	0.05	1.00									
$Gender\ Difference_{ij}$	0.02	-0.21	1.00								
$JobSim_{ij} \times HierarchicalDiff_{ij}$	0.02	0.59	-0.12	1.00							
$JobSim_{ij} \times GenderDiff_{ij}$	0.60	-0.08	0.34	-0.15	1.00						
$Candidate\ Fitness\ to\ a\ Job_{ik}$	-0.02	0.01	0.00	0.00	0.00	1.00					
$Geographical\ Distance_{ik}$	0.13	0.01	0.06	-0.04	0.12	0.02	1.00				
$Connection\ Time_{ij}$	0.18	0.24	-0.02	0.14	0.05	0.01	0.06	1.00			
$Employee\ Referral\ Behavior_j$	0.28	-0.18	0.06	-0.22	0.22	0.01	0.11	0.11	1.00		
$Job\ Level_k$	-0.15	-0.40	0.05	-0.16	-0.03	0.01	0.02	-0.24	-0.10	1.00	
$Time\ Job\ Posted_k$	0.09	0.13	-0.02	0.07	0.00	0.01	0.00	0.55	0.08	-0.25	1.00

Table 4.19. Candidate-Employee Job Similarity on Referral Dynamics

Model	Model 1a	Model 1b	Model 1c	Model 1d	Model 1e
		Logit with Three-Way Random Effects			
Dyad-Level Covariates					
Job Similarity$_{ij}$	-	-6.934 ***	-4.933 **	-7.088 ***	-5.143 **
		(1.894)	(2.235)	(2.152)	(2.437)
Hierarchical Difference$_{ij}$	-	-0.475 ***	-0.478 ***	-0.474 ***	-0.475 **
		(0.152)	(0.153)	(0.152)	(0.153)
Gender Difference$_{ij}$	-	-0.949 *	-0.955 *	-0.938 *	-0.937 *
		(0.489)	(0.497)	(0.494)	(0.502)
JobSim$_{ij}$×HierarchicalDiff$_{ij}$	-	-	1.623 *	-	1.630 *
			(0.942)		(0.940)
JobSim$_{ij}$×GenderDiff$_{ij}$	-	-	-	0.520	0.749
				(3.402)	(3.473)
Candidate Fitness to a Job$_{ik}$	3.695	4.698	4.505	4.705	4.509
	(4.259)	(4.394)	(4.403)	(4.389)	(4.394)
Geographical Distance$_{ik}$	-0.766 ***	-0.760 ***	-0.765 ***	-0.760 ***	-0.764 **
	(0.121)	(0.127)	(0.129)	(0.127)	(0.129)
Connection Time$_{ij}$	-0.004	0.038	0.038	0.038	0.037
	(0.040)	(0.039)	(0.039)	(0.039)	(0.039)
Individual-Level Covariates					
Employee Referral Behavior$_{j}$	2.284 ***	2.019 ***	2.084 ***	2.016 ***	2.077 **
	(0.400)	(0.394)	(0.417)	(0.394)	(0.416)
Job Level$_{k}$	-0.583 ***	-1.092 ***	-1.135 ***	-1.089 ***	-1.130 **
	(0.220)	(0.283)	(0.285)	(0.284)	(0.285)
Time Job Posted$_{k}$	-1.323 ***	-1.161 ***	-1.151 ***	-1.160 ***	-1.147 **
	(0.389)	(0.399)	(0.377)	(0.399)	(0.376)
Constant	-3.022	-3.507	-3.659	-3.510	-3.657
	(3.141)	(3.385)	(3.219)	(3.382)	(3.208)
Notes:					
No. of observations	109,712	82,548	96,913	82,548	96,913
No. of candidates	8,305	7,402	7,494	7,402	7,494
No. of employees	8,774	7,767	7,863	7,767	7,863
No. of candidate-employee dyads	8,774	7,767	7,863	7,767	7,863
Mean VIF	1.18	1.21	1.21	1.32	1.32

Standard errors are in parentheses

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table 6 displays the results of logistic regression analysis with three-way random effects that examine how job similarity between candidates and employees affects referral decision. Variables are introduced across the columns. Model 1a incorporates only control variables. The main interest variables $Job\,Similarity_{ij}$, $Hierarchical\,Difference_{ij}$, and $Gender\,Difference_{ij}$ are included in model 1b to examine the impact of candidate (i) – employee (j) dyadic characteristics on referrals. Models 1c and 1d incorporate two-way interaction terms of $Job\,Similarity_{ij}$ with $Hierarchical\,Difference_{ij}$ and $Gender\,Difference_{ij}$, respectively. As the direction and significance of all coefficients are consistent across the models, I use the complete model specification (model 1e) to discuss the results.

In Hypothesis 1, I proposed that candidates are less likely to be referred by employees who are doing similar jobs. The results support this hypothesis, as the likelihood of referral decreases as the job similarity between an employee and a candidate increases ($\beta_{Job\,Similarity} =$ -5.143, $p < 0.05$). As proposed in Hypothesis 2, I further find that this negative effect of job similarity weakens as the hierarchical difference between an employee and a candidate increases: The interaction effect between $Job\,Similarity_{ij}$ and $Hierarchical\,Difference_{ij}$ is positive and statistically significant ($\beta_{JobSim \times HierarchicalDiff} = 1.630$, $p < 0.1$). To aid the interpretation of the significant interaction effect between $Job\,Similarity_{ij}$ and $Hierarchical\ \ Difference_{ij}$, a plot of the relationship between $Job\,Similarity_{ij}$ and the likelihood of referral, given different levels of $Hierarchical\,Difference_{ij}$, is displayed in Figure 4.6. The upward slope of the plot suggests that as $Hierarchical\,Difference_{ij}$ increases, the negative effect of $Job\,Similarity_{ij}$ on referral becomes weaker.

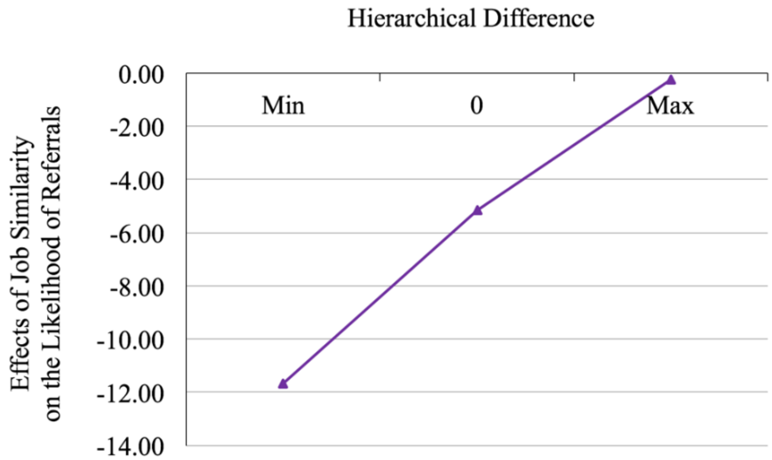

Figure 4.6. Interaction Effects of Job Similarity and Hierarchical Difference on the Likelihood of Referrals

In my third hypothesis, I proposed that the negative effect of job similarity on referral is likely to be reduced if a candidate is the same gender as an employee. To examine how the effect of job similarity changes according to their gender match, I included the interaction term between $Job\,Similarity_{ij}$ and $Gender\ Difference_{ij}$ in my model. Although I expected that gender homophily would positively moderate the relationship between job similarity and referral likelihood, results reveal that the interaction term does not statistically significantly moderate the effect of $Job\,Similarity_{ij}$ on referral likelihood ($\beta_{JobSim \times GenderDiff} = 0.749,\ p > 0.1$). This finding suggests that gender homophily, at least in this context, is not strong enough to overcome the negative effect generated by job similarity between a candidate and an employee.

The results of some control variables are interesting. The negative and statistically significant estimate of the variable $Hierarchical\,Difference_{ij}$ (-0.475, $p < 0.05$) indicates that as an employee's hierarchical level gets higher than that of a candidate, referrals are less likely to be made. The negative and statistically significant estimate of $Gender\,Difference_{ij}$ (-0.937, $p < 0.1$) indicates that employees are less likely to refer a candidate with a different gender, suggesting the presence of gender homophily. Interestingly, the parameter estimates of $Candidate\ Fitness\ to\ a\ Job_{ik}$ are statistically insignificant across all the models. Further investigation reveals that the job candidates who are chosen by the company for potential referrals are all highly qualified for their matched jobs, as they have been prescreened based on their objective qualifications. The negative and statistically significant coefficient $\textit{Geographical Distance}_{ik}$ indicates that candidates are less likely to be referred for a job if they are physically far away from the working environment. $\textit{Employee Referral Behavior}_{j}$ controls for employees' engagement in referral behavior. Positive and statistically significant results across the models indicate that candidates are likely to be referred by employees who are active in the platform. The coefficients of $\textit{Job Level}_{k}$ are negative and statistically significant in the models, indicating that the higher an open job position is, the less likely candidates are to be referred.

4.5 ROBUSTNESS CHECKS

Although the results are consistent with my predictions, there could be alternative explanations for my findings. I took several steps to investigate. First, as outlined above, I incorporated the following control variables to tease out the effect of job similarity on the likelihood of referrals from other plausible factors: $\textit{Candidate Fitness to a Job}_{ik}$, $\textit{Geographical Distance}_{ik}$, $\textit{Connection Time}_{ij}$ $\textit{Employee Referral Behavior}_{j}$, $\textit{Job Level}_{k}$, and $\textit{Time Job Posted}_{k}$. In addition, I included random effects for each candidate and employee in a candidate-employee pair to control for their unobservable individual heterogeneities that may affect referral outcome. Further, I included a random effect for each candidate-employee pair to control for additional unobservable dyad-specific heterogeneities (e.g., offline tie strength) that may impact referral outcome. My main results stay consistent after incorporating the control variables and random effects.

Table 4.20. Candidate-Employee Job Similarity on Referral Dynamics: Falsification Tests

Model	Model 1b	Model 2a	Model 2b
Logit with Three-Way Random Effects			
Dyad-Level Covariates			
Job Similarity$_{ij}$	-6.934 **	-6.794 ***	-7.718 ***
	(1.894)	(2.247)	(3.111)
Hierarchical Difference$_{ij}$	-0.475 **	-0.371 **	-0.241
	(0.152)	(0.155)	(0.253)
Gender Difference$_{ij}$	-0.949 *	-0.456	-0.702
	(0.489)	(0.499)	(0.691)
Employees' Knowledge on Jobs$_{jk}$	-	5.620	-
		(4.788)	
Candidate Fitness to a Job$_{ik}$	4.698	0.338	8.295
	(4.394)	(2.171)	(6.695)
Geographical Distance$_{ik}$	-0.760 **	-0.789 ***	-0.501 **
	(0.127)	(0.150)	(0.209)
Connection Time$_{ij}$	0.038	0.066 *	0.0578
	(0.039)	(0.040)	(0.064)
Individual-Level Covariates			
Inverse of LinkedIn Connections$_{i}$	-	-	-5.480
			(12.950)
Employee Referral Behavior$_{j}$	2.019 **	1.961 ***	2.522 ***
	(0.394)	(0.447)	(0.804)
Job Level$_{k}$	-1.092 **	-0.704 **	-0.899 **
	(0.283)	(0.304)	(0.425)
Time Job Posted$_{k}$	-1.161 **	-1.067 ***	-0.955 ***
	(0.399)	(0.369)	(0.365)
Constant	-3.507	-4.443	-4.442
	(3.385)	(3.231)	(3.258)
Notes:			
No. of observations	82,548	91,529	44,049
No. of candidates	7,402	7,416	1,481
No. of employees	7,767	7,781	1,655
No. of candidate-employee dyads	7,767	7,781	1,655
Mean VIF	1.21	1.27	1.16
Standard errors are in parentheses			
$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$			

Second, the negative effect of job similarity on referral probability may also be driven by the criticality mechanism. The criticality mechanism is that employees who are knowledgeable about the posted job can be more accurate (and hence more critical) in judging a candidate's qualification than other employees who are less knowledgeable about the posted job. In addition to the competition mechanism, this criticality of judgement driven by job similarity is also likely to lower candidates' chances of getting referrals. In order to tease out this effect, I further controlled for employees' knowledge about an open job position by incorporating the variable, $\textit{Employees' Knowledge on Jobs}_{jk}$ into my model. The variable $\textit{Employees' Knowledge on Jobs}_{jk}$ is measured by the similarity between the employees' and the open position's job titles using the word embedding technique described in §4.2. The value of $\textit{Employees' Knowledge on Jobs}_{jk}$ ranges from -1 to 1, in which 1 indicates that the employees' and the open positions' job titles are similar, whereas -1 indicates that two are antithetical. The more similar an employee's job is to an open position, the more likely that she has higher knowledge of the job. After incorporating this variable, I reran the complete specification model (model 1e in Table 6). The results indicate that the negative effect of job similarity still holds after controlling for an employee's knowledge about the job, suggesting that the competition mechanism still holds after controlling for the criticality mechanism. The results are presented in model 2a in Table 7.

Third, in the main model, I controlled for candidate-employee tie strength by incorporating the duration of the LinkedIn connection and dyad-specific random effect. As an additional robustness check, I incorporated an alternative tie strength measure: an inverse of total LinkedIn connections of employees, $\textit{Inverse of LinkedIn Connections}_i$. The rationale for this alternative proxy is that the relative strength of each LinkedIn tie for an individual with 500+ connections will be weaker than that of individuals with only 10 connections. My results (model 2b in Table 7) show the negative effects of $Job\,Similarity_{ij}$ still hold after incorporating the variable $Inverse\ of\ \ LinkedIn\,Connections_{i}$.

Last, I checked for a multicollinearity issue. Substantial multicollinearity problems lead to imprecision in the estimator, as they increase the variance of the coefficient estimates and make the estimates sensitive to small changes in the model (Greene 2011). I find no evidence of multicollinearity among the independent variables in this study: The mean-variance inflation factors (VIFs) of all models are below 2. A VIF of 10 and above is suggested to indicate a multicollinearity problem (O'brien 2007).

4.6 Discussion and Conclusion

In the 21st century, online professional social networks are one of the main routes where job candidates look for jobs. Likewise, a growing number of companies rely on professional networking platforms to search for high-quality candidates with low search cost. Despite the popularity of social hiring practices, only a limited number of scientific studies have explored how online professional network connections come into play in the job search process.

My major goal in this study is to advance my understanding of the characteristics of professional connections that help (or harm) job candidates obtain a job referral from incumbent employees. Connections with a target company's employees are believed to be instrumental in getting referrals. Consequently, many job candidates make efforts to build such connections. Yet, my empirical analysis reveals that job seekers are actually less likely to be referred by employees who are doing similar jobs. This indicates that despite the information benefits that employees in a similar field can offer, they are not the best people from whom to expect referrals. I attribute this adverse effect of job similarity on job referral to a potential competition threat. From the perspective of an employee, qualified job candidates in similar fields are considered as potential competitors who can threaten their status and promotion opportunities by raising the bar.

However, forming connections with employees in a target company is hard for candidates to give up, because the connections provide access to rich information about the job and new openings. Hence, I further explored the conditions that can relieve the observed adverse effect of job similarity. I find that a difference in hierarchical levels relieves the negative impact of job similarity. A larger difference in hierarchical levels indicates that candidates are less likely to threaten employees, as roles and tasks vary by levels within an organization (Pavett and Lau 1983). Further, if employees are in higher hierarchical positions than job candidates, they feel less threatened, as they have greater power over job candidates; thus, employees' higher job levels increase the likelihood of referrals by alleviating the negative effects of competition driven by job similarity. Although I expected that gender homophily, the tendency to prefer same-gender candidates, may weaken the competition effect, my results reveal that gender homophily is not strong enough to cancel out the competition effects generated by job similarity.

This study has several practical implications. First, online professional network ties with employees working in a similar field are not helpful for job referral outcomes, even though they may still be helpful in terms of information access. This finding would be disappointing to many job candidates as they strive to make LinkedIn connections with employees in their target companies and target fields, with the expectation of obtaining job referrals. Based upon my results, I suggest that candidates should be aware that employees may consider them as potential competitors for tasks and promotions. Therefore, I further recommend that job candidates be careful not to provoke unnecessary competition dynamics. Second, my study provides not only compelling findings from the perspective of job candidates, but also useful insights for firms who utilize technologies in the labor market. From a managerial point of view, an employee referral based on an online social network is a cost-efficient way to find qualified candidates. However, firms should be mindful that this method can unexpectedly eliminate the opportunity to find the best candidate for a new job position, due to potential competition dynamics between their employees and job candidates.

My study has a few limitations that suggest interesting avenues for future research. First, in the current study, I focused on the employees' intrinsic incentives to refer job candidates. It would be intriguing to examine how extrinsic incentives (e.g., financial incentives) influence employees' referral behavior. In my setting, the company did not provide any extrinsic rewards to employees for referrals, which was an ideal setting for us to observe untainted employee referral behavior and its consequences for job candidates. Yet, it would be interesting to investigate how extrinsic incentives can be used to induce desirable referral behavior from employees. Second, the current study is conducted based on the social hiring data from one global consulting company. I ask readers to be cautious in extrapolating my findings to organizations operating in different contexts. I hope that subsequent research investigates how organizational culture influences the observed referral dynamics. My empirical setting is a Fortune 500 consulting company, which has a competitive culture. Just like most for-profit companies, the company has a pyramid organizational structure where fewer and fewer employees can be promoted as they climb up the hierarchy ladder. It would be meaningful to conduct a similar study in another organization where collaboration is more valued than competition.

In sum, this paper investigates whether a common strategy of job candidates making online professional connections with employees in the target company and the target job function practically helps candidates obtain referrals. I found that the common strategy in fact tends to hurt the likelihood of referrals because of the competitive force at work. Job candidates may strategically aim to make connections with higher-level employees to circumvent these unwanted competition dynamics that forestall referrals.

Chapter 5. CONCLUDING REMARKS

Technology development is rapidly reshaping how economic value is created in the digital markets. In this dissertation, I explored different aspects of digital markets to gauge the social and economic values and whether unfairness problems in digital markets can be addressed. Chapter 3 dealt with a study on digital markets that has significant contributions to both, transportation market literature and practitioners. To the best of my knowledge, this is one of the early studies to isolate the economic value of information sharing for transportation network companies (TNC) and the urban transportation systems. I also contribute to the digital market design literature by considering information sharing design that maintains both the efficiency of TNC markets and incorporates the concern on the fairness of passengers. I complement the research stream that highlights the importance of the unique feature of TNC and transportation digital markets, hence provide a set of actionable insights on the economic value of information sharing in the TNC market. Further,

In chapter 4, I examine the characteristics of professional connections that help (or harm) job candidates obtain a job referral from incumbent employees in the online labor markets. This essay advances theoretical knowledge in labor digital markets in several ways. It advances my understanding of the antecedents of referral-based hiring. Also, by examining various influencing factors (i.e., job similarity, hierarchical difference, gender homophily, tie strength, individual characteristics), the study offers a more nuanced understanding of the factors that influence online referral decisions. In addition, this study joins a small number of recent studies investigating the impact of online network ties on job search outcomes. Practically, the study informs managers about the potential drawback of utilizing employee referrals: The common strategy of job seekers making connections with employees in the target company and job function may not be instrumental in obtaining referrals. In a job seekers perspective, the study informs how the online professional connection may not be equally beneficial to the job seekers; Instead, benefits are likely to differ according to the characteristics of their connections.

This dissertation attempts to deliver both theoretical and practical insights into the economic values created in the digital markets. Further, it addresses digital inequality problems that arise due to the unfairness in the access and the use of digital technology. The digital markets constantly change with the introduction of new technology and redirect the users' consumption patterns along with the managers' and policymakers' strategies. Therefore, it is important that the scholars address new questions accordingly and maintain up to date to the fast changes of the digital markets. For instance, customers have shifted to the transportation digital markets or online labor markets from traditional transportation or the labor markets. This dissertation underlines the importance of such studies on the digital markets is to first understand the nature of the digital economy and the unique feature of the marketplace that is created by the introduction of digital technology. This is because the demand structure and consumers' consumption patterns differ from traditional markets as the digital market offers new features introduced from technology. Another take away is that there is a rising concern of customers or the users who do not take fair advantage of such development of digital markets. Business managers, as well as focal governments, take action to resolve such a phenomenon by adjusting or regulating the new digital technology offers. This dissertation proposes future research to develop and investigate new economies created by digital technology. Being in the era of big data and AI, I look forward to the studies that gauge social and economic value created from upheavals in the digital markets. Simultaneously, exploring the dark side of digital markets such as uneven advantages among users would be very interesting and give insights to the practitioners to prepare for the upcoming digital era.

BIBLIOGRAPHY

[1] Arnott R (2001) The economic theory of urban traffic congestion: a microscopic research agenda. Workshop on Environmental Economics and the Economics of Congestion: Coping with Externalities, Venice International University, Venice Summer Institute, San Servolo, Italy.

[2] Badger E (2017). Is Uber helping or hurting mass transit? The New York Times (October 2017).

[3] Bailey JP, Bakos Y (1997) An exploratory study of the emerging role of electronic intermediaries. Int. J. Electron. Comm. 1(3):7-20.

[4] Bakos Y (1997) Reducing buyer search costs: Implications for electronic marketplaces. Manag. Sci. 43(12): 1676-1692.

[5] Bakos Y, Katsamakas E (2008). Design and ownership of two-sided networks: Implications for Internet platforms. J. Manag. Inform. Sys. 25(2): 171-202.

[6] Balakrishna R, Wen Y, Ben-Akiva M, Antoniou C (2008) Simulation-based framework for transportation network management in emergencies. Trans. Res. Rec. 2041(1): 80-88.

[7] Bapna R, Ramaprasad J, Shmueli G, Umyarov A (2016) One-way mirrors in online dating: A randomized field experiment. Manage. Sci. 62(11):3100–3122.

[8] Bertsimas D, Farias VF, Trichakis N (2011) The price of fairness. Oper. Res. 59(1):17–31.

[9] Bertsimas D, Farias VF, Trichakis N (2012) On the efficiency-fairness trade-off. Manage. Sci. 58(12):2234–2250.

[10] Böcker L, Dijst M, Faber J (2016) Weather, transport mode choices and emotional travel experiences. Tran. Res. Part A: Pol. Prac. (94):360–373.

[11] Brynjolfsson E, Hu Y, Smith MD (2003) Consumer surplus in the digital economy: Estimating the value of increased product variety at online booksellers. Manag. Sci. 49(11): 1580-1596.

[12] Brynjolfsson E, Smith MD (2000) Frictionless commerce? A comparison of Internet and conventional retailers. Manage. Sci. 46(4):563–585.

[13] Burtch G, Ghose A, Wattal S (2013) An empirical examination of the antecedents and consequences of contribution patterns in crowd-funded markets. Inform. Syst. Res. 24(3):499–519.

[14] California Public Utilities Commission (2013) Decision adopting rules and regulations to protect public safety while allowing new entrants to the transportation industry (November 2013), http://docs.cpuc.ca.gov/PublishedDocs/Published/G000/M077/K112/77112285.PDF

[15] Chan J, and Ghose A (2014) Internet's dirty secret: Assessing the impact of online intermediaries on the outbreak of sexually transmitted diseases. MIS Quart. 38(4):955–976.

[16] Chen MK, Chevalier JA, Rossi PE, Oehlsen E (2017) The value of flexible work: Evidence from Uber drivers. Working Paper, National Bureau of Economic Research, Cambridge, MA.

[17] Cheng A, Pang MS, Pavlou PA (2019) Mitigating traffic congestion: The Role of intelligent transportation systems. Inform. Syst. Res. (Forthcoming).

[18] Cohen P, Hahn R, Hall J, Levitt S, Metcalfe R (2016) Using big data to estimate consumer surplus: The case of Uber. Working Paper, National Bureau of Economic Research, Cambridge, MA.

[19] DePillis L (2016) One reason you might be better off driving for Uber than in a taxi. The Washington Post (March 2016), https://www.washingtonpost.com/news/wonk/wp/2016/03/15/one-reason-you-might-be-better-off-driving-for-uber-than-in-a-taxi/?utm_term=.9d3d4b653c01.

[20] Dickenson, G (2018). How the world is going to war with Uber. The Daily Telegraph (June 2018).

[21] DiMaggio P, Hargittai E, Celeste C, Shafer S (2004) Digital inequality: From unequal access to differentiated use. In Social inequality, Rus. Sag. Found. 355–400.

[22] Farahani RZ, Miandoabchi E, Szeto WY, Rashidi H (2013) A review of urban transportation network design problems. Eur. J. Oper. Res. 229(2): 281-302.

[23] Gan J, An B, Wang H, Sun X, Shi Z (2013) Optimal pricing for improving efficiency of taxi systems. Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence, 2811–2818, IJCAI'13 (AAAI Press), ISBN 978-1-57735-633-2.

[24] Gong J, Greenwood BN, Song Y (2017) Uber might buy me a Mercedes Benz: An empirical investigation of the sharing economy and durable goods purchase. Working Paper, Temple University, Philadelphia.

[25] Greenwood B, Agarwal R (2016) Matching platforms and HIV incidence: An empirical investigation of race, gender, and socio-economic status. Manage. Sci. 68(2):2281–2230.

[26] Greenwood BN, Wattal S (2017) Show me the way to go home: An empirical investigation of ride-sharing and alcohol-related motor vehicle fatalities. MIS Quart. 41(1):163–187.

[27] Hagiu A, Halaburda H (2014) Information and two-sided platform profits. Int J of Ind Organ 34 (1): 25-35.

[28] Hagiu A, Wright J (2014) Marketplace or reseller? Manag. Sci. 61(1):184–203.

[29] Hall JV, Horton JJ, Knoepfle DT (2017) Labor market equilibration: Evidence from Uber. Working Paper, Uber Technologies, San Francisco, CA.

[30] Hu Y (2017) Shanghai demands taxi apps to conceal passengers' destinations from drivers (December 2017), http://urbanfamily.thatsmags.com/shanghai/post/2142/shanghai-demands-taxi-apps-to-conceal-passengers-destinations-from-drivers.html

[31] Kvasny L (2002) A conceptual framework for examining digital inequality. AMCIS Proc. 246.

[32] LaFrance A, Eveleth R (2015). Are taxis safer than Uber? The Atlantic. Archived from the original on April 5, 2017. Retrieved June 8, 2017.

[33] Lam C, Liu M (2019) Toward Inclusive Mobility: Ridesharing Mitigates Geographical Disparity in Transportation. Working paper, Clemson University, Clemson, SC.

[34] Langfitt F (2014) What a ban on taxi apps in Shanghai says about China's economy. National Public Radio (April 2014), http://www.npr.org/sections/parallels/2014/04/10/301049584/what-a-ban-on-taxi-apps-in-shanghai-says-about-chinas-economy.

[35] Li Z, Hong Y, Zhang Z (2016) Do ride-sharing services affect traffic congestion? An empirical study of Uber entry. Working Paper, Arizona State University, Phoenix.

[36] McPhate M (2016) Uber and Lyft End Rides in Austin to Protest Fingerprint Background Checks. The New York Times (May 2016), https://www.nytimes.com/2016/05/10/technology/uber-and-lyft-stop-rides-in-austin-to-protest-fingerprint-background-checks.html.

[37] Ozkan E, Ward A (2017) Dynamic matching for real-time ridesharing. Working Paper, Available at SSRN 2844451.

[38] Park J, Kim J, Pang MS, Lee B (2017) Offender or guardian? An empirical analysis of ride-sharing and sexual assault. Working paper, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea.

[39] Rayle L, Dai D, Chan N, Cervero R, Shaheen S (2016) Just a better taxi? A survey-based comparison of taxis, transit, and ridesourcing services in San Francisco. Trans. Pol. (45):168–178.

[40] Rochet JC and Tirole J (2004) Two-sided markets: An overview. Working Paper, Institut d'Economie Industrielle (IDEI), Toulouse, France.

[41] Rogers B (2015) The social costs of Uber. University of Chicago Law Review 82.

[42] Saneinejad S, Roorda MJ, Kennedy C (2012) Modelling the impact of weather conditions on active transportation travel behavior. Trans. Res. Part D: Trans. Env. 17(2):129–137.

[43] Tuss A (2016) Uber drivers stiff passengers after finding out final destination. NBC Washington (December 2016), http://www.nbcwashington.com/news/local/Uber-Drivers-Stiff-Passengers-After-Finding-Out-Final-Destination-408432315.html

[44] Van Dijk J, Hacker K (2003) The digital divide as a complex and dynamic phenomenon. The Inf. Soc. 19(4): 315–326.

[45] Wallsten S (2015) The competitive effects of the sharing economy: How is Uber changing taxis? Tech. Pol. Ins.

[46] Wang W, Zheng J, Kannan K, Ren F (2019) Estimating the value of destination disclosure: A dynamic structural model in a transportation network. Working Paper, Purdue University, West Lafayette, IN.

[47] Wang X, Agatz N, Erera A (2018) Stable matching for dynamic ride-sharing systems. Transport. Sci. 52(4):850–867.

[48] White M (2015) Companies adapt to Uber and Lyft rides on the expense report. New York Times (March 2015), https://www.nytimes.com/2015/03/24/business/companies-adapt-to-uber-and-lyft-rides-on-the-expense-report.html

[49] Wolfe S (2018) "Uber and Lyft are creating more traffic and congestion instead of reducing it, according to a new report." Business Insider (July 2018).

[50] Yan H (2015) Shanghai extends traffic restriction hours. China Daily (April 2015), http://www.chinadaily.com.cn/business/motoring/2015-04/02/content_19982620.htm

[51] Yue M (2015) New taxi rules put limits on app use (April 2015), https://archive.shine.cn/metro/public-services/New-taxi-rules-put-limits-on-app-use/shdaily.shtml

[52] Zhan X, Qian X, Ukkusuri SV (2014). Measuring the efficiency of urban taxi service system. In The Third International Workshop on Urban Computing (UrbComp'14).

[53] Zheng J, Ren F, Tan Y, Chen X (2016) Optimizing two-sided promotion for transportation network companies: A Structural model with conditional Bayesian Learning. Working Paper, Purdue University, West Lafayette, IN.

[54] Acquisti A, Fong CM (2015) An experiment in hiring discrimination via online social networks. Working Paper, Carnegie Mellon University, Pittsburgh, PA.

[55] Beaman L, Magruder J (2012) Who gets the job referral? Evidence from a social networks experiment. Amer. Econ. Rev. 102(7), 3574-93.

[56] Bendersky C, Hays NA (2012) Status conflict in groups. Organ. Sci. 23(2): 323-340.

[57] Berger J (1995) Were You Referred by a Man or a Woman? Gender of Contracts and Labor Market Outcomes (Princeton University Press, Princeton, NJ).

[58] Bersin J (2012) Facebook vs. LinkedIn - what's the difference? Forbes (May 21), https://www.forbes.com/sites/joshbersin/2012/05/21/facebook-vs-linkedin-whats-the-difference/#2d76388f6af5

[59] Bielby WT, Baron JN (1986) Men and women at work: Sex segregation and statistical discrimination. Amer. J. Sociol. 91(4): 759-799.

[60] Brown M, Setren E, Topa G (2016) Do informal referrals lead to better matches? Evidence from a firm's employee referral system. J. Labor. Econ. 34(1): 161-209.

[61] Burke M, Kraut R (2013) Using Facebook after losing a job: Differential benefits of strong and weak ties. Proc. Conf. Comp. Supp. Coop. (ACM Press), 1419-1430.

[62] Burks SV, Cowgill B, Hoffman M, Housman M (2015) The value of hiring through employee referrals. Quart. J. Econ. 130(2): 805-839.

[63] Castilla EJ (2005) Social networks and employee performance in a call center. Amer. J. Sociol. 110(5): 1243-1283.

[64] Chan J, Wang J (2018) Hiring preferences in online labor markets: Evidence of a female hiring bias. Manag. Sci. 64(7): 2973-2994.

[65] Darrow, B. (2017) LinkedIn claims half a billion users. Fortune (April 20), http://fortune.com/2017/04/24/linkedin-users/

[66] Dumais ST (2004) Latent semantic analysis. Annu. Rev. of Inform. Sci. Tech. 38(1): 188-230.

[67] Dustmann C, Glitz A, Schönberg U, Brücker H (2016) Referral-based job search networks. Rev. Econ. Stud. 83(2): 514-546.

[68] Fernandez RM, Castilla EJ, Moore P (2000) Social capital at work: Networks and employment at a phone center. Amer. J. Sociol. 105(5): 1288-1356.

[69] Fernandez RM, Sosa ML (2005) Gendering the job: networks and recruitment at a call center. Amer. J. Sociol. 111(3): 859-904.

[70] Fernández-Aráoz, C (2001) The challenge of hiring senior executives. Emot. Intel. Workplace, 182-206.

[71] Garcia SM, Tor A, Schiff TM (2013) The psychology of competition: A social comparison perspective. Pers. Psych. Sci. 8(6): 634-650.

[72] Garg R, Telang R (2017) To be or not to be linked: Online social networks and job search by unemployed workforce. Manag. Sci. 64(8): 3926-3941.

[73] Gee LK, Jones J, Burke M (2017) Social Networks and labor markets: How strong ties relate to job finding on Facebook's social network. J. Labor. Econ. 35(2): 485-518.

[74] Gelman A, Hill J (2007) Data Analysis Using Regression and Multilevel/Hierarchical Models (Cambridge University Press, Cambridge, UK).

[75] Gilchrist K (2018) 3 Ways LinkedIn can help MBA grads bag their dream job. CNBC (August 9), https://www.cnbc.com/2018/08/09/mba-careers-3-ways-to-use-linkedin-to-find-a-graduate-job.html

[76] Goethals GR, Darley JM (1977) Social comparison theory: An attributional approach. Suls JM, Miller RL, eds. Social Comparison Processes: Theoretical and Empirical Perspectives, (Washington, DC: Hemisphere), 259–278.

[77] Granovetter MS (1973). The strength of weak ties. Amer. J. Sociol. 78(6): 1360-1380.

[78] Greene W (2011) Econometric Analysis. 7th ed. (Prentice Hall, Upper Saddle River, NJ).

[79] Hambrick DC (1994) Top management groups: a conceptual integration and reconsideration of the “team” label. Res. Organ. Behavior (16): 171-213.

[80] Hausman JA (1987) Specification tests in econometrics. Econometrica. 46(6):1251-1271.

[81] Hwang EH, Singh PV, Argote L (2015) Knowledge sharing in online communities: Learning to cross geographic and hierarchical boundaries. Organ. Sci. 26(6): 1593-1611.

[82] Jacobs P (2009) What is social recruiting? Hum. Res. Magazine. 14(5): 2-3.

[83] Jovanovic B (1979) Job matching and the theory of turnover. J. Polit. Econ. (87): 972–990.

[84] Jovanovic B (1984) Matching, turnover, and unemployment. J. Polit. Econ. (92): 108–122.

[85] Kalleberg AL, Knoke D, Marsden PV, Spaeth JL (1996) Organizations in America: Analysing their Structures and Human Resource Practices (Sage Publications, Thousand Oaks, CA).

[86] Kilduff GJ, Elfenbein HA, Staw BM (2010) The psychology of rivalry: A relationally dependent analysis of competition. Acad. Manag. J. 53(5): 943-969.

[87] Leicht KT, Marx J (1997) The consequences of informal job finding for men and women. Acad. Manag. J. 40(4): 967-987.

[88] Liu WT, Duff RW (1972) The strength in weak ties. Public. Opin. Quart. 36(3): 361-366.

[89] Makra K (2015) Case study: How to make connections in your field and tap into the hidden job market. Workopolis (September 09), https://careers.workopolis.com/advice/case-study-how-to-make-connections-in-your-field-and-tap-into-the-hidden-job-market/

[90] Manant M, Pajak S, Soulié N (2014) Do recruiters ‘like' it? online social networks and privacy in hiring: A pseudo-randomized experiment. Working Paper, University of Paris Sud, France.

[91] Marin A (2012) Don’t mention it: Why people don’t share job information, when they do, and why it matters. Soc. Networks. 34(2): 181-192.

[92] Marsden PV (1994) The Hiring Process: Recruitment Methods. Am. Behav. Sci. 37(7): 979-991.

[93] McPherson M, Smith-Lovin L, Cook JM (2001) Birds of a feather: Homophily in social networks. Annu. Rev. Sociol. 27(1): 415-444.

[94] Mikolov T, Chen K, Corrado G, Dean J (2013) Efficient estimation of word representations in vector space. Proc. ICLR.

[95] Montgomery JD (1992) Job search and network composition: Implications of the strength-of-weak-ties hypothesis. Am. Sociol. Rev. 57(5): 586-596.

[96] Mouw T (2003) Social capital and finding a job: do contacts matter? Amer. Sociol. Rev. 68(6): 868-898.

[97] Obukhova E (2012) Motivation vs. relevance: Using strong ties to find a job in urban China. Soc. Sci. Res. 41(3): 570-580.

[98] O’brien RM (2007) A caution regarding rules of thumb for variance inflation factors. Qual. Quant. 41(5): 673-690.

[99] Ollington N, Gibb J, Harcourt M (2013) Online social networks: An emergent recruiter tool for attracting and screening. Personn. Rev. 42(3): 248-265.

[100] Overbeck JR, Correll J, Park B (2005) Internal status sorting in groups: The problem of too many stars. Thomas-Hunt MC, eds. Res. Managing Groups and Teams, Status and Groups, (Elsevier, Oxford, UK), (7): 171- 202.

[101] Pallais A, Sands EG (2016) Why the referential treatment? Evidence from field experiments on Referrals. J. Polit. Econ. 124(6): 1793-1828.

[102] Pavett CM, Lau AW (1983) Managerial work: The influence of hierarchical level and functional specialty. Acad. Manag. J. 26(1): 170-177.

[103] Pennington J, Socher R, Manning CD (2014) Glove: Global vectors for word representation. Conf. Emp. Meth. Natural. Lang. Proc. (EMNLP), (14): 1532-1543.

[104] Perez S (2018) LinkedIn now lets job seekers ask for referrals with a click of a button. TechCrunch (March 19),

[105] Petrone P (2016) 7 stats that prove your employees are your secret recruiting weapon. LinkedIn (February 23), https://business.linkedin.com/talent-solutions/blog/employer-brand/2016/7-stats-that-prove-your-employees-are-your-secret-recruiting-weapon

[106] Pieper JR, Greenwald JM, Schlachter SD (2018) Motivating employee referrals: The interactive effects of the referral bonus, perceived risk in referring, and affective commitment. Hum. Resour. Manag. (57):1159–1174.

[107] Rees A (1966) Information networks in labor markets. Amer. Econ. Rev. (55): 559–566.

[108] Salancik GR, Pfeffer J (1977) Who gets power—and how they hold on to it: A strategic-contingency model of power. Organ. Dyn. 5(3): 3-21.

[109] Schlachter SD (2018) An empirical and theoretical investigation of the complexities of employee referrals in organizations. Working Paper, University of Nebraska-Lincoln, Lincoln, NE.

[110] Schmutte IM (2014) Job referral networks and the determination of earnings in local labor markets. J. Labor. Econ. 33(1): 1-32.

[111] Simon CJ, Warner JT (1992) Matchmaker, matchmaker: The effect of old boy networks on job match quality, earnings, and tenure. J. Labor Econ. 10(3): 306-330.

[112] Simpson WB (2001) QAP: The quadratic assignment procedure. Presented in North American STATA Users' Group Meeting. March 20-21, 12-13

[113] Smith SS (2005) "Don't put my name on it": Social Capital Activation and Job-Finding Assistance among the Black Urban Poor. Amer. J. Sociol. 111(1):1–57.

[114] Stadd A (2013) 55% of recruiters turn to Twitter, compared to Facebook's 65% and LinkedIn's 94%. Adweek (September 10), http://www.adweek.com/socialtimes/recruiters-twitter/490440

[115] Verbrugge L (1977) The structure of adult friendship choices. Soc. Forces. (56): 576–597.

[116] Vecchio RP (1995) The impact of referral sources on employee attitudes: Evidence from a national sample. J. Manag. (21): 953–965.

[117] Williamson IO, Cable DM (2003) Organizational hiring patterns, interfirm network ties, and interorganizational imitation. Acad. Manag. J. 46(3): 349-358.

[118] Yakubovich V (2005) Weak ties, information, and influence: How workers find jobs in a local Russian labor market. Amer. Sociol. Rev. 70(3): 408-421.

[119] Zhang J, Ackerman MS, Adamic L (2007) Expertise networks in online communities: Structure and algorithms. Proc. 16t.h Int. Conf. WWW. 221–230.

APPENDIX A

A1. Heterogeneous Vertical/Mixed Effect of Technology Restriction Policy

To check whether the heterogeneity among temporal spillover effects of the technology restriction policy will affect my results, I add four weekday-specific "vertical effects" (i.e. Monday-dummy interacted with the "vertical effects") to the main model to check whether my results are robust to additional heterogeneity control. Shown in Table A1, I find that the results of weekday-specific "vertical effects" are mixed: some are significant whereas the rest are less or not. However, the signs of all of the heterogeneous effects are consistent with the main effects but the scales of the heterogeneous ones are smaller than those of the main effects. All the main effects and other coefficients of interest are also consistent with my main results. Therefore, qualitatively speaking, these results are consistent with my main results.

Table A.1. Heterogeneous Vertical/Mixed Effect of Tech Restrict Policy Results

	Model 9a	Model 9b	Model 9c	Model 9d	Model 9e
Method					
Dependent Variable	**Taxi**	**Metro**	**Bus**	**Ferry**	**Parking**
Main/Horizontal Effect	-0.013 ***	0.392 ***	0.270 ***	0.203 ***	0.396 ***
	(0.002)	(0.016)	(0.014)	(0.010)	(0.042)
Vertical/Mixed Effect	0.180 ***	0.439 ***	0.415 ***	0.181 ***	0.706 ***
	(0.021)	(0.026)	(0.023)	(0.037)	(0.110)
Vertical/Mixed Effect *Mon	**0.135** ***	**0.433** ***	**0.286** ***	**0.145** ***	**0.705** ***
	(0.027)	**(0.004)**	**(0.040)**	**(0.027)**	**(0.119)**
Vertical/Mixed Effect *Tue	**0.097** ***	**0.360** ***	**0.227** ***	**0.206** ***	**0.499** ***
	(0.027)	**(0.010)**	**(0.050)**	**(0.022)**	**(0.127)**
Vertical/Mixed Effect *Wed	**0.075** *	**0.255** ***	**0.204** ***	**0.122** ***	**0.294** *
	(0.029)	**(0.021)**	**(0.024)**	**(0.027)**	**(0.114)**
Vertical/Mixed Effect *Thu	**0.036**	**0.183** ***	**0.127** ***	**0.081** *	**0.439** ***
	(0.022)	**(0.023)**	**(0.025)**	**(0.034)**	**(0.120)**
Rush Hour Policy	-0.063 ***	0.155 ***	0.087 ***	0.097 ***	0.042
	(0.007)	(0.008)	(0.014)	(0.010)	(0.040)
Vertical/Mixed Effect of Rush Hour	-0.014 ***	0.179 ***	0.131 ***	0.123 ***	0.165 ***
	(0.002)	(0.013)	(0.001)	(0.011)	(0.018)
Holiday	-0.108 ***	-0.021	-0.071 ***	-0.483 ***	-0.108
	(0.001)	(0.029)	(0.015)	(0.036)	(0.167)
Precipitation	-0.030 **	-0.173 ***	-0.124 ***	-0.129 ***	-0.147 ***
	(0.011)	(0.003)	(0.001)	(0.011)	(0.008)
Temperature	-0.004 ***	-0.001	-0.002 ***	0.003 *	-0.002
	(0.001)	(0.007)	(0.001)	(0.001)	(0.012)
Constant	8.516 ***	0.408	6.508 ***	2.776 ***	-0.714 **
	(0.017)	(1.185)	(0.025)	(0.021)	(0.218)
Half-Hour Fixed Effects	Yes	Yes	Yes	Yes	Yes
Day Fixed Effects	Yes	Yes	Yes	Yes	Yes
No. of Observations	1410	1410	1410	1410	1410

Note. Robust standard errors are in parentheses.
* $p < 0.05$, ** $p < 0.01$,*** $p < 0.001$

VITA

Kyung Sun (Melissa) Rhee is an Assistant Professor of Information Systems at Warrington College of Business, University of Florida. Her current research interests center around economics in digital markets. Topics vary across transportation networks, social media, online labor market and economics of information technology. Melisa mainly applies econometric models and machine learning techniques to answer research questions. Prior to joining Warrington College of Business, Melissa earned Ph.D. in Information Systems area from the Foster School of Business, University of Washington. Melissa received her M.A. in the Operations Management and B.A. in the Business Administration, both from Seoul National University in Seoul, South Korea.