MODELING AND ANALYSIS OF ON-DEMAND RIDE-SOURCING MARKETS

By

LITENG ZHA

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2017

To my family

ACKNOWLEDGMENTS

Ph.D. is a bitter sweet journey. After three years' academic training, what I have gained is not just the knowledge and skills, but the confidence and persistence when confronted with challenges. As the final deliverable, this dissertation could not be done by me alone.

First and foremost, I appreciate the support from my beloved parents Yuqiang Zha and Yuchen Chen as well as my dear girlfriend Yinan Zheng. Thank you for accompanying me through the ups and downs. You are the ones always there that I can turn to.

Next, my thanks go to my advisor Dr. Yafeng Yin, the most diligent (yes, not one of) professor I am fortunate to work with. I value the discussions and tips from Dr. Yin both in research and daily life. I appreciate Dr. Yin's guidance and generosity in my career design.

There are fantastic faculty members at UF from who I build up my knowledge set. I thank Dr. Steven Slutsky, Dr. Siva Srinivasan and Dr. Scott Washburn for being on my committee and offering constructive comments on this dissertation; Dr. Lily for the supervision on the FDOT project; Dr. Yongpei Guan, Dr. William Hager and Dr. J.P. Richard for the awesome courses on Operations Research. In particular, Dr. Slutsky's Microeconomics stimulates my frenzy on classical economic theory.

Last but not the least, I would like to pass my thanks to my friends at UF, without whom my Ph.D. life would become less colorful. I cherish the happy time with Zhibin Chen, Zhengtian Xu, Xiaotong Sun, Zibin Bian , Shanyue Guan, Don Watson and Stephen Spana. I also thank Ms. Ines, a lovely lady, for her consistent efforts in coordinating the transportation engineering group.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

MODELING AND ANALYSIS OF ON-DEMAND RIDE-SOURCING MARKETS

By

Liteng Zha

May 2017

Chair: Yafeng Yin
Major: Civil Engineering

Due to the innovative business mode and advanced dispatch technology, ride-sourcing companies such as Uber and Lyft have attracted many riders, eroding the traditional taxi market. Meanwhile, substantial controversies arise along their way of expansion. Among those often highlighted are dynamic pricing, service competition and regulation. In this dissertation, we propose both analytical and simulation frameworks to address these key issues. The proposed methodologies and tools could serve as important references for regulatory agencies in evaluating and managing the ride-sourcing markets.

Our first investigation is on the regulation and competition of the ride-sourcing market. An aggregate economic model is developed where the matching between customers and drivers are captured by an exogenous matching function. It is found that a monopoly ride-sourcing platform will maximize the joint profit with its drivers without any regulatory intervention. We establish the conditions for regulating only the amount of commission charged by the platform to achieve the second best. Then the model is extended into a duopoly setting and unveils that competition does not necessarily lower the price or improve the social welfare. In the latter case, regulators may rather encourage the merger of the platforms and regulate them directly as a monopolist.

The exploration of dynamic pricing is divided into two main thrusts. As drivers may adjust work schedules to cover more profitable periods, we first capture drivers' work hour choices under dynamic pricing utilizing a bi-level programming framework. The patterns of the market dynamics and the trade-offs among market players are then numerically discussed. Next, we extend the model into a spatially differentiated market. The spatial variations of market frictions and the welfare of involved agents under dynamic pricing are explicitly explored. Based on an empirical dataset, we demonstrate the equilibrium outcomes under dynamic pricing and evaluate the effectiveness of a commission cap regulation.

In parallel with the analytical models, an agent-based simulation is developed to better understand the ride-sourcing markets. We find that platform's matching technology exhibits increasing returns to scale property; and higher matching ranges lead to more transactions but lower magnitude of the returns to scale.

CHAPTER 1
INTRODUCTION

1.1 Background

The number of smart mobile devices in the U.S. has been rising steadily and a study suggests that nearly two-thirds of Americans now own at least one such device (Pew Research Center (2015)). These devices retrieve users' locations, enable ubiquitous communications and allow instant peer-to-peer interaction, giving rise to a new class of firms, on-demand companies, which aim to effectively bring together consumers and suppliers of resources (e.g., houses and parking spaces) and services (e.g., home cleaning and computer programming) with very low transaction costs. These companies are shaking up their industries and reshaping our daily lives.

On demand ride-sourcing companies, such as Uber, Lyft and Didi Chuxing—a typical example of on-demand economy—are transforming the way we travel in cities. The companies provide ride-hailing apps, which are real-time and internet-based platforms that intelligently source participating drivers to riders. A rider can monitor in real time the location of the coming vehicle and receive notification when it arrives. These apps are free to use but usually charge a commission for each transaction (20-25% of the fare paid to drivers). Thanks to their convenience and lower prices, on demand ride-sourcing services have successfully attracted many riders, eroding the traditional taxi market. For example, in 2013, the revenue of Uber, Lyft and Sidecar (now bankrupted) is $140M in San Francisco, half what the established cab companies made. Uber now operates in more than 200 cities in approximately 50 countries (The Economist, 2015).

Several terms exist for describing services provided by Uber-like companies, such as ridesharing, for-profit ridesharing, on-demand ridesharing and dynamic ridesharing, which may be partly attributed to the companies' marketing of their services being a type of ridesharing

(Anderson, 2014; Rayle et al., 2014). However, the services provided by Uber-like companies have distinct features and the term "ridesharing" may be confusing. Unlike taxis or jitneys, ridesharing is conventionally not for profit. For Uber-like services, however, it is largely the financial motivation that brings in private car owners to participate and some even become a full-time driver. Therefore, the California Public Utilities Committee proposed to use "Transportation Network Companies" to categorize Uber-like companies. Considering these companies are essentially sourcing a ride from a driver pool, Rayle et al. (2014) suggested the term "ride-sourcing" for their services. It is worth mentioning that major ride-sourcing companies also allow ride-splitting on their platforms to encourage multiple people to share a ride, e.g., UberPOOL and Lyft Line, which may further blur the line between ridesharing and ride-sourcing. However, we choose to use ride-sourcing (which is a reduced form from on-demand ride-sourcing for the convenience of presentation) to highlight its distinctive features, i.e., private car owners drive their own vehicles to provide on-demand taxi-like services for profit.

Since their advent in 2009, ride-sourcing companies have enjoyed huge success, but also created many controversies. The major opposition comes from the traditional or regular taxi industry, as ride-sourcing platforms become a growing threat for their profit. The regular taxi market is usually regulated in terms of price, entry and service quality while comparatively fewer regulatory requirements have been imposed on or proposed for ride-sourcing companies. Unfair competition is argued particularly by professional cab drivers and their employers, who have organized strikes and filed lawsuits around the world. Ride-sourcing companies have also brought headache to government officials and legislators, because they do not know how to deal with them. While many are still wondering what to do, some (e.g., cities in Netherlands and Japan) have decided to ban them or treat their services to be illegal; others (e.g., Stockholm,

Sweden; Salt Lake City, Utah) embrace them as a new type of transit provider or have passed ride-sourcing legislations and regulations. Although these legislations and regulations have some differences, they all essentially codify the insurance coverage, background check of drivers and inspection protocols that ride-sourcing companies already have in place. There is no intervention on the fare of their services or the affiliated fleet size (or the number of vehicles in service).

The success of ride-sourcing platforms casts doubts on the regulation of the taxi industry that is often criticized for limited supply (Badger, 2014). It also challenges some of the fundamentals for the traditional taxi regulation by significantly reducing information imperfection. The automatic matching algorithms offered by the platforms lead to significant savings in customers' waiting time and drivers' searching time. Together with the mutual rating mechanism (the reputation system), the interaction between customers and drivers is enhanced. Moreover, all trip-related fares are processed by built-in pricing and payment functions, which eliminate the chaos caused by soliciting and bargaining that often emerge after a taxi market is deregulated (ECMT, 2007). Last but not least, some argue that the competition among multiple ride-souring platforms may lower the prices and reduce the market power of a particular predominate platform. Should we simply deregulate the taxi industry and leave ride-sourcing companies alone, and then let the market decide the winners? What are the implications and welfare impacts of such a deregulation? Should we conduct a systematic reform of regulation of the ride-for-hire market? Presumably the reform needs to be tailored to specific cities considering their demographics, mobility options, patterns, and culture. Is there a unified theoretical framework to guide such a reform? Many such questions remain unanswered.

Besides, the improvement in the technology and the lack of regulatory intervention have stimulated new business modes and distinctive features in ride-sourcing markets. One of the

issues centered in the discussion is surge pricing, which is essentially a dynamic pricing scheme that adjusts trip fare in real time. Based on the market condition of a geographic area, the base trip fare is adjusted by a multiplier automatically generated from the platform's algorithm. The information of the surge multiplier (SM) is provided to both customers and drivers before a transaction happens. With the dual effect of suffocating demand and increasing supply, dynamic pricing is advocated to guarantee a reasonable amount of waiting time for customers. In most cases, the surge multiplier falls below 1.5 times. However, it can soar 7 times or higher without a cap limit (Curley, 2014; The Economists, 2014). Platforms generally benefit from the surges since commission is charged by a fixed percentage of the final trip fare. Therefore, platforms may surge unnecessarily high or more frequently to exploit customers, given that no regulation is imposed and the pricing algorithms are proprietary.

Despite the heated discussion over the social media, the temporal and spatial impacts of dynamic pricing have not been quantitatively modeled. The problem is manifested when we consider the flexibility ride-sourcing drivers enjoy in choosing their work hours and locations. On one hand, they may adjust work schedules to cover more profitable periods; and the temporal variation in the wage rate induced by dynamic pricing may further affect drivers' labor supply decisions, e.g., how long they would like to work (Chen and Sheldon, 2016). On the other hand, drivers are attracted to highly surged zones for better trip opportunities. Little is known for the induced variations of market frictions or the changes in agents' welfare in different geographic areas.

Unfortunately, related research on the ride-sourcing market is limited. To the best of our knowledge, 1) there is no analytical framework in studying the regulation and service competition in the ride-sourcing market; 2) although several studies empirically investigate the

labor supply under dynamic pricing (Hall and Krueger, 2015; Chen and Sheldon, 2016), there is an increasing need to explore the temporal and spatial impacts of dynamic pricing and provide guidance for management strategies.

1.2 Dissertation Objectives

This dissertation does not attempt to solve all the puzzles and controversies associated with the ride-sourcing companies. Specifically, we propose to establish methodologies and tools to analyze the structure, service competition and dynamic pricing of the ride-sourcing markets and then derive insights on the management policies. These topics focus on different analytical domains and the proposed frameworks are thus characterized by different modeling resolution and complexity. In a parallel effort, we develop an agent-based simulation to validate part of the assumptions and results.

1.2.1 Platform Regulation and Competition

We start our first model by treating the ride-sourcing platform as a two-sided market which matches users from both demand and supply sides (Rochet and Tirole, 2003; Armstrong, 2006; Rochet and Tirole, 2006). The matching process is featured by the negative same-side externality and the positive cross-side externality. That is, increasing one customer (driver) raises the average waiting (searching) time for the other customers (drivers) on the same side but reduces the average searching (waiting) time for the drivers (customers) cross the side. To capture such a market structure, we revisited an aggregate economic model with an exogenous matching function (Yang and Yang, 2011). The temporal and spatial variations of market dynamics are assumed away and we essentially focus on the static market equilibrium in the long run. Based on this, we first examine the economic outcomes under different market scenarios and investigate the optimal regulation strategy that requires the minimum regulatory burden. We next extend our analysis via incorporating platform competition, and explore the changes in trip fares

and social welfare to shed light on whether competition is beneficial to customers and the society at large.

1.2.2 Dynamic Pricing

Dynamic pricing may adjust the distributions of demand and supply both temporally (e.g., encourage drivers to work during more profitable periods) and spatially (e.g., induce drivers to transition to highly surged zones). Our investigation of dynamic pricing is divided into two main thrusts. To analyze the temporal effect of dynamic pricing, we develop a time-expanded network that outlines customers' demand distributions and drivers' work hour scheduling. Then we propose a bi-level programming framework with the upper level being specified with a revenue-maximizing surge scheme while the lower level capturing drivers' work hour choices. In the numerical experiments, we compare the patterns of the market dynamics under dynamic pricing and discuss the trade-offs among market players.

Subsequently, in a separate chapter, we extend the framework to the spatially differentiated market and explore the spatial heterogeneity of market frictions and welfare of market players under dynamic pricing. The matching process between the customers and drivers is further derived from a spatial point process. If market power is a concern, we present a regulation scheme to enhance market efficiency. Utilizing a public dataset from Didi Chuxing, we numerically explore the equilibrium outcomes under dynamic pricing and evaluate the performance of the proposed regulation.

1.2.3 Simulation Modeling

We develop an agent-based simulation in Netlogo to provide a better understanding of the working mechanism of ride-sourcing platforms and validate some of the previous results and key assumptions. The simulator attempts to represent various behaviors of the simulated agents, such as ride-sourcing drivers' searching and cruising behaviors, customers' willingness to pay

and sensitivity to waiting time as well as the pricing strategies and matching techniques of ride-sourcing platforms. Based on a simulation test bed, we explore the market dynamics under a variety of market scenarios that characterize different demand levels and matching ranges between customers and drivers. We also calibrate and discuss the employed matching functions using the simulated data.

1.3 Dissertation Outline

The remainder of this dissertation is organized as follows. Chapter 2 provides an overview of the literature on the analytical modeling of traditional taxi markets, some recent development on the regulation and dynamic pricing of the ride-sourcing markets. Chapter 3 presents an aggregate economic model and discusses platform regulation and competition. Chapter 4 models the work scheduling of the ride-sourcing drivers and investigates the temporal effects of dynamic pricing. Chapter 5 investigates the spatial heterogeneity under dynamic pricing and discusses the potential regulation policies. Chapter 6 introduces an agent based simulation and evaluates the platform's matching technology. Finally, the overall conclusions and recommendations for future work are provided in Chapter 7.

CHAPTER 2
LITERATURE REVIEW

The literature on ride-sourcing markets is rather thin as these services were launched into the market quite recently. Therefore, our review of literature interrelates some from the traditional taxi market which shares the similarity in the considerations of regulation and some modeling set-ups. More specifically, we summarize our review of literature on three aspects: 1) regulation of taxi and ride-sourcing markets; 2) modeling and analysis of taxi markets; and 3) dynamic pricing in the ride-sourcing markets.

2.1 Regulation of Taxi and Ride-Sourcing Markets

A regular taxi industry usually consists of dispatch, taxi stand and cruising segments. The regulation of the industry is to ensure the temporal and spatial stability and availability of taxi services and guarantee the public safety. It often takes the form of price, entry and service quality (Frankena and Pautler, 1986).

The rationale for regulating a cruising market is mainly to correct the information imperfection. It is unlikely for a rider to have a prior experience with the service quality of the cab during a random hailing (ECMT, 2007). Also, neither the cab driver nor the customer knows each other's ideal bid and ask for the trip. Consequently, a driver can potentially charge slightly higher than a customer's ask, considering the cost for her to wait for the next available cab if she turns down. For dispatch market, the reliable service quality (in terms of acceptable average waiting time and the 24 hour accessibility) needs to be guaranteed particularly in less popular areas (Frankena and Pautler, 1986; ECMT, 2007). On the other hand, regulation has been criticized for the lack of information and resources for regulators to take appropriate actions that maximize social benefit (Schaller, 2007). The most controversial aspect comes from the medallion system, a common tool for entry control. The value of a medallion has once topped 1

million in New York (Badger, 2014). The medallion system has long been blamed to create a cartel that operates for its own benefit rather than a simple tool to control the number of taxis (Schaller, 2007; Badger, 2014). Besides, possible occurrence of regulation capture and rent seeking is often cited by the opponents to taxi regulation (Rauch and Schleicher, 2015).

The majority of American cities still set limits on both the entry and price of the taxi markets. The advent of ride-sourcing companies has casted doubts on these regulations and caused sliding values of taxi medallions (Badger, 2014). Features such as real-time driver-rider matching and mutual rating possessed by ride-sourcing platforms can enhance the interaction among platform users and may reduce information imperfection. A reduction in searching and waiting time of drivers and riders has been reported in some preliminary studies on ride-sourcing companies (Rayle et al., 2014). The built-in pricing and payment functions have also eliminated soliciting and bargaining behaviors. Some have suggested that the ride-sourcing market is self-regulatory because the mutual rating system can lead to better service quality, and the competition among multiple platforms will lower the prices and reduce the market power of a predominant platform (Koopman et al., 2015). These statements need to be examined with caution. Competition can be socially inefficient given the dominance of the cross-side effect in a two-sided market (Wright, 2004; Hagiu, 2006).

The altitude towards regulation of ride-sourcing companies and the level of regulation vary among governments around the world. A detailed review on the process of regulation in major states in the U.S. can be found in Rayle et al. (2014). The core of the proposed or passed regulations is remarkably similar, codifying the insurance coverage, background check policy, and vehicle inspection protocol, in addition to other requirements such as voice-enabled apps and more detailed reporting of activities.

2.2 Modeling and Analysis of Taxi Markets

A number of analytical models have been developed (Vany, 1975; Frankena and Pautler, 1986; Cairns and Liston-Heyes, 1996; Salanova et al., 2011), after the seminal work of Douglas (1972). These models are all aggregate in nature, analyzing the properties of taxi market equilibrium. One of the most distinguishable features these models capture is the searching and waiting frictions, since the taxi market is not merely cleared by price. On the demand side, a customer will consider both the trip fare and waiting time when deciding whether to take a taxi. On the supply side, the profitability of a taxi will depend on its utilization rate (the fraction of time it is occupied) and operating cost. In a series of work, Yang and his collaborators developed the network equilibrium models to capture the spatial structure of a taxi market (Yang and Wong, 1998; Yang et al., 2002). They further expanded the models to incorporate congestion externality (Yang et al., 2005b), temporal variation of the demand and supply (Yang et al., 2005a), user heterogeneity and modal competition (Wong et al., 2008) and nonlinear pricing (Yang et al., 2010a).

The form of the average waiting time function reflects the mechanism of how a customer and a driver are "matched," which is eventually determined by the type of the taxi market. By applying the stochastic geometry techniques and the double-queuing model, Douglas (1972), Arnott (1996) and Matsushima and Kobayashi (2006) derived the average waiting time function for the stand, cruising and dispatching taxi market, respectively. Despite the difference in the analytical forms, they all confirm that the matching in the taxi market exhibits increasing returns to scale, a phenomenon quite common in the queuing of general transportation systems (Mohring, 1972). That is, customers' average waiting time will be reduced if the number of waiting customers and the number of vacant vehicles are doubled. The first-best solution is at deficit when the matching is increasing returns to scale, since the trip fare only covers the cost of

occupied taxi hours (Vany, 1975; Arnott, 1996; Cairns and Liston-Heyes, 1996). Accordingly, two remedies have been suggested. One is to subsidize the taxi service to obtain the first best; the other is to regulate both the fare level and fleet size to achieve the second best, which shares the same objective of maximizing the social welfare but subject to a break-even constraint of the industry profit (Cairns and Liston-Heyes, 1996; Yang et al., 2002).

Despite the insightful implication of an increasing-returns-to-scale matching function, previous derivations are somewhat flawed. The application of the stochastic geometry ignores the competition among customers over the same vehicle so that the average waiting time is only a function of the density of vacant vehicles (Yang et al., 2010b); direct employment of the queuing methods forces the first-in-first-out (FIFO) rule (for either the queue of waiting customers or the vacant vehicles) which may be too strict for the spatial queuing process in the general taxi market (with the exception of the taxi stand market). Therefore, Yang et al. (2010b) revisited the Cobb-Douglas function first introduced by Schroeter (1983) to delineate the bilateral search and meeting process in the taxi market. In fact, such an exogenous matching approach has long been considered for labor markets (e.g., see a recent short review by Moscarini and Wright (2010)). A static market equilibrium can be found where drivers' profit and social surplus are shown to depend critically on the returns to scale of the search and meeting technology (Yang and Yang, 2011).

Intuitively, the search friction still exists in the ride-sourcing or e-hailing platforms (e.g. Curb and Hailo, which serves exclusively professional cab drivers, Shaheen (2014)). However, its properties may be different since the matching technology is significantly improved. A few recent studies still rely on the Cobb-Douglas matching function (but with varying parameter values to differentiate the matching technology) to investigate the competition between

traditional street hailing and e-hailing services (He and Shen, 2015) as well as the optimal price perturbation for the platform (Wang et al., 2016). We also assume a Cobb-Douglas matching function in the discussion of platform competition and regulation in Chapter 3 due to its simple form; but will formally derive the form using a spatial point process in Chapter 5. The calibration of the Cobb-Douglas matching function and the discussion over the relationship between matching ranges and the returns to scale property are provided in Chapter 6.

2.3 Dynamic Pricing

Dynamic (surge) pricing is a commonly seen in the field of revenue management. Companies dynamically adjust prices corresponding to the variations in demand for profit maximization (Talluri and Van Ryzin, 2006). Two recent publications focus on the performances of dynamic pricing implemented by the ride-sourcing platforms. Cachon et al. (2015) found dynamic pricing generally worked well with the platform's current strategy of charging a fixed commission and yielded very close results to the optimal contract where the platform dynamically determined both the price (for the customers) and the wage (for the drivers). Compared to the static case, dynamic pricing increases the platform's profit; the results on the consumers' surplus is ambiguous and depends on the entry cost of the drivers (Cachon et al., 2015). Dynamic pricing increases the consumers' surplus when drivers' entry cost is high (supply is limited) but hurts it when the entry cost is low (supply is amber) (Cachon et al., 2015). Using an M/M(k)/1 queuing model, Banerjee et al. (2015) explored the effect of dynamic pricing and assumed the platform was either revenue or transaction maximizing. They modeled dynamic pricing via a threshold-based dynamic pricing approach where price was adjusted discretely and was exclusively based on the number of available vehicles in the queue. Interestingly, they found dynamic pricing under-performed the optimal static pricing under both surge schemes. The most

significant benefit from dynamic pricing is however its robustness, i.e., dynamic pricing performs better when the platform has limited information (Banerjee et al., 2015).

The studies mentioned above are both based on a simplified treatment in customers' demand, which is only assumed to be a function of price. In Cachon et al. (2015), a (random) rationing on either the supply or demand side would occur to guarantee the market equilibrium. Comparatively in Banerjee et al. (2015), customers were assumed to abandon if not served immediately. In our frameworks, customers are sensitive to both the price and the average waiting time. The form of the average waiting time is induced from the spatial matching process that is either captured in an aggregate matching function as in Chapter 3 or derived based on the spatial point process as in Chapter 5. In terms of dynamic pricing, we focus on its impact in adjusting drivers' work schedules in Chapter 4; in chapter 5, the discussion is extended to the spatial heterogeneity of market frictions and agents' welfare. We also explore the possible regulation policies of dynamic pricing if market power is a concern.

CHAPTER 3
ECONOMIC ANALYSIS OF ON-DEMAND RIDE-SOURCING MARKETS

This chapter makes a preliminary attempt to provide a quantitative analysis on the market structure of ride-scouring services and explores effective regulation policies. We consider hypothetical situations when ride-sourcing companies become self-sustainable and dominate the ride-for-hire market. We are subsequently interested in knowing whether and how to regulate the ride-sourcing market. Following Yang and Yang (2011), we develop an aggregate model with a Cobb-Douglas matching function to examine different market scenarios, properties and economic outcomes of a hypothetical ride-sourcing market with a single platform. In view of the potential market distortion, we investigate effective regulation policies that require minimal regulatory variables. The analysis is further extended to consider a duopoly market to examine the effects of platform competition. Analyzing the tradeoff in the pricing formula under the Nash equilibrium, we observe that competition does not necessarily lower the price level. Via a numerical analysis, we explore the conditions where competition is socially inefficient and a regulated monopoly market can be more efficient than a regulated duopoly one.

The chapter is organized as follows. The basic aggregate model for a hypothetical ride-sourcing market along with some comparative statics is presented in Section 3.1. Section 3.2 explores the pricing structure, solution properties of the single platform across different market scenarios as well as the regulation policies. In Section 3.3, the discussion is extended to a duopoly market with a particular focus on the changes in price and social welfare. The summary of research findings and policy implications are provided in Section 3.4.

3.1 Basic Model

This section introduces an aggregate model that captures a hypothetical ride-sourcing market with a focus on the matchings between customers and drivers. The model is established

by extending the work of Yang and Yang (2011). We firstly assume a hypothetical ride-sourcing market with a single platform, a group of customers and a group of affiliated drivers. The ride-sourcing platform serves as an intermediary that matches customers with potential drivers. The platform decides the fare paid by a customer, i.e., F, and charges a commission, i.e., P, from the payment by the customer at each transaction; the driver receives the remaining payment, i.e., $F - P$. It is assumed that the hypothetical ride-sourcing market is mature such that the platform will gain profit from providing the services. A few other things are worth noting here: 1) unlike ride-sourcing platforms, e-hailing platforms for traditional regulated taxi market do not have price-setting power (He and Shen, 2015); 2) although in practice ride-sourcing platforms charge commission as a percentage of the fare, such a distinction is immaterial in the context of this study; 3) congestion externality (caused by both ride-sourcing and regular vehicles) is not considered (Yang et al., 2005b).

3.1.1 Matching Function

The matchings between customers and drivers are completed via a matching algorithm implemented at the ride-sourcing platform. Taking Uber as an example, it dispatches one of the vehicles within a coverage radius of a requesting customer. The dispatching is made to minimize the estimated waiting time for the customer (Ranney, 2015). At the aggregate level, we assume that a matching function (a production function) can be used to characterize such a process. Note that aggregate matching functions have been calibrated for traditional street-hailing (e.g., Yang et al. (2014) or radio-dispatch taxi market (e.g., Schroeter (1983). Although being much more efficient with a larger matching area and more complete information of drivers and customers, the matching technology offered by a ride-sourcing platform is actually similar to the one adopted by radio-dispatch taxi companies. We thus assume that that aggregate matching functions may still be valid for representing the matching technology.

More specifically, we consider a stationary state where variables such as the numbers of waiting customers (N^c) and vacant ride-sourcing vehicles (N^v) are time invariant. The matching function then relates the rate of matchings (more precisely, meetings) per hour (m^{c-t}) to N^v and N^c at any instant (Yang and Yang, 2011). Note $N^v = w^v T^v$ where w^v is the average "searching" time for a driver before meeting a customer and T^v is the arrival rate of vacant vehicles per hour. $N^c = w^c Q$, where w^c is the average customer waiting time and Q is the customer demand per hour. The matching function can be formally written as:

$$m^{c-t} = M\left(N^v, N^c\right) = M\left(w^v T^v, w^c Q\right) \tag{3-1}$$

where $\partial M / \partial N^v > 0, \partial M / \partial N^c > 0$ are assumed so that the increase of either vacant vehicles or waiting customers will increase the meeting rate.

We define the elasticities of the matching function with respect to N^v and N^c as α_1 and α_2, respectively. The elasticities reflect the matching technology of the ride-sourcing platform. We hereinafter assume them to be constant and within the range of [0, 1]. This assumption leads to a Cobb-Douglas matching function:

$$m^{c-t} = A\left(N^v\right)^{\alpha_1}\left(N^c\right)^{\alpha_2} \tag{3-2}$$

where A is a scaling parameter, which depends on the unit of the meeting rate and encapsulates other factors in the matching technology that are not fully captured by α_1 and α_2. Moreover, the parameter can be interpreted to be related to the area of the ride-sourcing market and the cruising speed of vacant vehicles (we assume waiting customers remain stationary until being picked up). To see this, consider the following customers' average waiting time function, which is obtained by considering $m^{c-t} = Q = T^v$ at the stationary state together with Eqs. (3-1)- (3-2):

$$w^c = (Q)^{\frac{1-\alpha_1-\alpha_2}{\alpha_2}} (A)^{-\frac{1}{\alpha_2}} (w^v)^{-\frac{\alpha_1}{\alpha_2}} \tag{3-3}$$

Setting $\alpha_2 = 1$ yields:

$$w^c = \frac{1}{A(Qw^v)^{\alpha_1}} = \frac{1}{A(N^v)^{\alpha_1}} \tag{3-4}$$

The above implies that customers' waiting time only depends on the number of vacant vehicles. This is true particularly when the supply of vacant vehicles is more than sufficient to serve the customers and there is no competition among customers over a particular vehicle. Further assuming $\alpha_1 = 1$ and $\alpha_1 = 0.5$, Eq. (3-4) will be reduced to the waiting time functions derived by Douglas (1972) and Arnott (1996) for cruising and radio dispatching taxi market respectively where A represents the area of the market divided by the running speed of vacant vehicles.

The matching function is increasing, constant or decreasing returns to scale when the sum of α_1 and α_2 is larger than, equal to or smaller than one (Yang and Yang, 2011). By analyzing the radio-dispatch taxi market in Minneapolis, U.S. and the overall taxi market in Hong Kong, China, respectively, both Schroeter (1983) and Yang et al. (2014) reported increasing-returns-to-scale matching functions. The degree varies from 1.13 to 1.16. The increasing-returns-to-scale property is commonly seen in a queueing process that exists in many transportation systems (Mohring, 1972; Schroeter, 1983; Arnott, 1996; Yang and Yang, 2011). In a ride-sourcing market, drivers often cruise to "hotspots" in order for being matched early. Higher densities of both customers and drivers may increase the matching probability (Although this benefit may be marginal given the advanced matching technology ride-sourcing companies currently possess). After being matched, the matched driver will have to travel to pick up the customer, and the average travelling distance decreases with the increases in the numbers of customers and drivers.

Due to the economy of density, the average waiting times of customers and drivers is expected to decrease more than linearly if both the numbers of customers and drivers increase. Consequently, the meeting rate increases more than linearly with the numbers of customers and drivers. This observation has also been confirmed in an agent-based simulation study conducted in Chapter 6. Therefore, the analyses hereinafter primarily focus on cases with an increasing-returns-to-scale matching function.

3.1.2 Customer Demand

Consider a stationary state where the hourly demand of customers (passengers) is $\bar{Q}$. Each customer consumes exactly one trip and faces two transportation modes: the ride-sourcing service and alternative modes such as transit. The customer derives a deterministic utility V_0 from completing the trip while incurring a generalized cost μ for using the ride-sourcing service and C for the other options. Note μ is determined endogenously while C is exogenously given as a constant. With error terms ε_R, ε_O capturing unmeasurable attributes, the utility from utilizing each mode can thus be specified as follows:

$$U_R = V_0 - \mu + \varepsilon_R \tag{3-5}$$

$$U_O = V_0 - C + \varepsilon_O \tag{3-6}$$

where $\mu = F + \beta w^c + \tau l$, consisting of the trip fare (F), waiting time cost (βw^c) and in-vehicle travel time cost (τl). This specification implicitly assumes that customers are homogeneous in their values of waiting time (β) and in-vehicle travel time (τ). Moreover, w^c is determined endogenously in Eq. (3-3) while l represents the average trip time and is assumed constant. Each customer is assumed to choose the option that maximizes her utility. Consequently, the demand

for the ride-sourcing service (Q) will depend on the distribution of the error terms but can be specified as a decreasing function of the generalized cost:

$$Q = f\left(\mu\right) = f\left(F + \beta w^c + \tau l\right) \tag{3-7}$$

where $f' < 0$ over the domain $\mu > 0$.

3.1.3 Comparative Statics

At any instant of the stationary state, total vehicle equals the sum of the numbers of vacant vehicles N^v and occupied vehicles N^o, i.e., $N = N^v + N^o = Qw^v + Ql$. The matching function essentially dictates the form of the waiting time function $\overline{W^c}(Q, w^v)$, as defined in Eq. (3-3). So far, we have identified the following three equations:

$$w^c = \overline{W^c}(Q, w^v) \tag{3-8}$$

$$Q = f(F + \beta w^c + \tau l) \tag{3-9}$$

$$N = Q(w^v + l) \tag{3-10}$$

The unknowns are Q, F, w^v, w^c and N. Treating F and N as decision variables, we present some derivatives with respect to F and N as follows:

$$\frac{\partial Q}{\partial F} = \frac{f'}{1 - f'\beta\overline{W_1^c} + f'\beta\overline{W_2^c}\dfrac{N}{Q^2}} \tag{3-11}$$

$$\frac{\partial Q}{\partial N} = \frac{f'\beta\dfrac{1}{Q}\overline{W_2^c}}{1 - f'\beta\overline{W_1^c} + f'\beta\overline{W_2^c}\dfrac{N}{Q^2}} \tag{3-12}$$

where $\overline{W_1^c} = \dfrac{1-\alpha_1-\alpha_2}{\alpha_2}\dfrac{w^c}{Q}$, $\overline{W_2^c} = -\dfrac{\alpha_1}{\alpha_2}\dfrac{w^c}{w^v}$. It can be shown that

$$\partial Q/\partial F < 0, \quad \partial Q/\partial N > 0 \tag{3-13}$$

Similarly, we have:

$$\frac{\partial w^v}{\partial F} = -N\frac{1}{Q^2}\frac{\partial Q}{\partial F} \tag{3-14}$$

$$\frac{\partial w^c}{\partial F} = \left(\overline{W_1^c} - \overline{W_2^c}\frac{N}{Q^2}\right)\frac{\partial Q}{\partial F} \tag{3-15}$$

$$\frac{\partial w^v}{\partial N} = \frac{1}{Q^2} - \frac{N}{Q^2}\frac{\partial Q}{\partial N} \tag{3-16}$$

$$\frac{\partial w^c}{\partial N} = \overline{W_1^c}\frac{\partial Q}{\partial N} + \overline{W_2^c}\left(\frac{1}{Q^2} - \frac{N}{Q^2}\frac{\partial Q}{\partial N}\right) \tag{3-17}$$

where $\overline{W_1^c}, \overline{W_2^c}, \frac{\partial Q}{\partial F}, \frac{\partial Q}{\partial N}$ are defined in Eqs. (3-11) and (3-12).

3.2 Market Scenarios of Single Ride-Sourcing Platform

In this section, three market scenarios with a single ride-sourcing platform are examined. The monopoly scenario captures the platform's profit-maximizing behavior without any regulatory intervention. The first-best solution maximizes social welfare, but the platform and its drivers may be in deficit. We thus examine the second-best scenario by adding additional constraints to guarantee the reservation profits of the platform and drivers.

3.2.1 Monopoly Scenario

In this scenario, the ride-sourcing platform determines the trip fare and commission to attract both drivers and customers to the platform to facilitate their matchings in order to maximize the profit of the platform. Due to the static nature of our model, advanced pricing features such as surge pricing are not considered. The platform essentially provides a two-sided market and its decision making can be described as a leader-followers game where the monopoly platform serves as the leader who determines the trip fare F and the commission P to maximize its profit while customers and drivers are the followers. The former decides whether to use the ride-sourcing service while the latter decide whether to provide the service. With a free entry and sufficient labor supply, drivers will do so until their profit reaches zero. The platform's problem can thus be written as follows:

$$\text{(P1)} \quad \max_{F\geq 0, P\geq 0, N\geq 0} \pi_P = PQ - C_P(Q) \tag{3-18}$$

s.t.

$$\left(F - P\right)Q - cN = 0 \tag{3-19}$$

where $C_P(Q)$ is the cost function of the ride-sourcing platform and c captures the average hourly operation cost of a vehicle and the opportunity cost of the driver. For simplicity, we hereinafter do not consider non-negative constraints of decision variables and only focus on interior solutions. Define the Lagrangian function of P1 as follows:

$$L = \pi_P - \lambda\left[\left(F - P\right)Q - cN\right] \tag{3-20}$$

The first-order necessary conditions (FONCs) of P1 yield:

$$\lambda = -1 \tag{3-21}$$

$$F = C_P{}'\left(Q\right) + c(w^v + l) + \frac{1-\alpha_1-\alpha_2}{\alpha_1} cw^v - \frac{Q}{f'} \tag{3-22}$$

$$c = \beta \frac{\alpha_1}{\alpha_2} \frac{w^c}{w^v} \tag{3-23}$$

Eq.(3-21) indicates the profit of the drivers and that of the platform are substitutes while Eq. (3-23) shows w^c is proportional to w^v at optimality. Eq. (3-22) is the monopoly pricing formula. Define the price elasticity of demand as $\eta = -f'\frac{F}{Q} > 0$. It can be revised as:

$$F - C_P{}'\left(Q\right) - c(w^v + l) - \frac{1-\alpha_1-\alpha_2}{\alpha_1} cw^v = \frac{F}{\eta} \tag{3-24}$$

which follows the form of the Lerner formula (Lerner, 1934). The right-hand-side of the pricing formula in Eq. (3-22) consists of four terms: the marginal cost of the platform ($C_P{}'\left(Q\right)$), the average cost for a driver to serve a new customer ($c(w^v + l)$), a "matching externality" ($\frac{1-\alpha_1-\alpha_2}{\alpha_1} cw^v$) and the monopoly mark-up ($-\frac{Q}{f'}$). The "matching externality" is closely related

to the returns to scale of the matching function. As will be shown shortly, the realized waiting (or searching) frictions are lower if the matching function is increasing returns to scale. The incoming customer will thus be charged lower compared to the cases when the matching function is constant or decreasing returns to scale. Such an externality also exists in the optimal pricing formulae in other investigated scenarios.

When substituting P in the objective function using Eq. (3-19), we have the following equivalent formulation to P1:

$$\text{(P2)} \quad \max_{N\geq 0, F\geq 0} \pi = FQ - cN - C_P(Q) \tag{3-25}$$

s.t.

$$F - \frac{cN}{Q} \geq 0 \tag{3-26}$$

which indicates the platform maximizes its joint profit with its drivers. In fact, Eq. (3-26) can be safely dropped if we assume the resulting profit π is nonnegative at optimality. (See Appendix A for more information). This implies that, albeit not owning any vehicle, the platform essentially behaves like a traditional cab company that has a monopoly on the ride-for-hire market and thus determines its price and fleet size to maximize its profit.

3.2.2 First-Best Scenario

The first-best scenario represents an ideal case where a social planner or the platform maximizes total social welfare instead of its own profit by deciding the trip fare and fleet size. The commission does not impact social welfare but the revenue sharing between drivers and the platform. Therefore, the welfare formulation is similar to that in the traditional taxi literature except the additional cost incurred by the platform (Yang and Yang, 2011). The corresponding maximization problem can be written as:

$$\text{(P3)} \quad \max_{F\geq 0, N\geq 0} S = \int_0^Q f^{-1}(z)dz - Q\left(\beta w^c + \tau l\right) - cN - C_P(Q) \tag{3-27}$$

The FONC of the above problem leads to:

$$F - C_P'\left(Q\right) - c(w^v + l) - \frac{1-\alpha_1-\alpha_2}{\alpha_1} cw^v = 0 \tag{3-28}$$

$$c = \beta \frac{\alpha_1}{\alpha_2} \frac{w^c}{w^v} \tag{3-29}$$

In a regular taxi market where the taxi trip production is increasing returns to scale, the first-best pricing only covers the cost when a taxi is occupied and thus drivers' profit is negative (Douglas, 1972; Arnott, 1996; Cairns and Liston-Heyes, 1996). Utilizing Eq. (3-28), the joint hourly profit obtained by the ride-sourcing platform and its drivers is given as:

$$\pi^{f-b} = FQ - cN - C_P(Q) = \left(\frac{1-\alpha_1-\alpha_2}{\alpha_1} cN^v\right) + \left(C_P'(Q)Q - C_P(Q)\right) \tag{3-30}$$

Define the elasticity of the cost function as $\eta_c = C_P'(Q)\dfrac{Q}{C_P\left(Q\right)}$. If $\alpha_1 + \alpha_2 > 1$ and $\eta_c \le 1$, then $\pi^{f-b} < 0$. That is, if the matching function exhibits increasing returns to scale and the cost function of the platform shows economies of scale, the profits for the platform and drivers will be negative, making the first-best solution unsustainable.

3.2.3 Second-Best Scenario

Since the profits for the platform and its drivers may be in deficit in the first-best scenario, we consider the following second-best pricing.

$$\text{(P4)} \quad \max_{F\ge 0, N\ge 0, P\ge 0} S = \int_0^Q f^{-1}(z)dz - Q\left(\beta w^c + \tau l\right) - cN - C_P(Q) \tag{3-31}$$

s.t.

$$PQ - C_P(Q) \ge \pi_P^o \tag{3-32}$$

$$(F - P)Q - cN \ge 0 \tag{3-33}$$

where π_P^o represents the reservation profit of the platform and is nonnegative. Define $\pi_P = PQ - C_P(Q)$, $\pi_D = \left(F - P\right)Q - cN$, the Lagrangian function associated with P4 can be written as:

$$L = S + \lambda_P\left(\pi_P - \pi_P^o\right) + \lambda_D \pi_D \tag{3-34}$$

The FONC will then give:

$$\lambda_P = \lambda_D (= \lambda) \tag{3-35}$$

$$c = \beta \frac{\alpha_1}{\alpha_2} \frac{w^c}{w^v} \tag{3-36}$$

$$F - C_P{}'(\mathrm{Q}) - c(w^v + l) - \frac{1-\alpha_1-\alpha_2}{\alpha_1} cw^v = -\frac{\lambda}{1+\lambda}\frac{Q}{f'} \quad \left(\lambda \geq 0\right) \tag{3-37}$$

As shown in Eq. (3-35), the Lagrangian multipliers associated with the reservation profit constraints are the same. We can thus sum up Eqs. (3-32) and (3-33) to bound the joint profit of the platform and its drivers. Eq. (3-37) follows the Ramsey pricing (Oum and Tretheway, 1988; Yang et al., 2005b). It can be seen as a convex combination of the pricing formulas in the first-best and monopoly solutions. Further, the reservation profit constraints are binding given that the matching function exhibits increasing returns to scale. We summarize the results in the following proposition.

Proposition 1. If $\alpha_1 + \alpha_2 > 1$ and $\eta_c \leq 1$, then $\pi_P = \pi_P^o$, $\pi_D = 0$.

The proof of Proposition 1 is given in Appendix B. In our modeling framework, the commission P serves as an instrument to split the revenue obtained by the platform and its drivers. Given a solution of *F* and *N* (and thus *Q*), *P* is appropriately defined by the reservation profit constraints.

When the matching function shows increasing returns to scale, P4 is equivalent to the program that maximizes customers' demand subject to the same constraints (Douglas, 1972; Vany, 1975; Frankena, 1983; Frankena and Pautler, 1986; Yang et al., 2002).

3.2.4 Discussions on the Single Platform

3.2.4.1 Effects of homogeneous value of time

We have investigated the optimality conditions of three different scenarios. Despite the difference in the pricing formula, the optimality conditions all unveil the fact:

$$c = \beta \frac{\alpha_1}{\alpha_2} \frac{w^c}{w^v} \tag{3-38}$$

which means the average customer waiting time is proportional to the average driver searching time at optimality. Figure 3-1 illustrates an economic intuition, following Yang and Yang (2011). The meetings between customers and drivers can be viewed as the production of the ride-sourcing platform with inputs N^c and N^v. The monopolist and the social planner essentially differ in the chosen meeting rates (realized demand levels). At each demand level, the optimal matching is characterized by cost-minimizing production of the ride-sourcing platform. Therefore, the proportional relationship in Eq. (3-38) results from the tangency of the matching function and the total external cost curve. The assumption on the homogeneous value of time of the customers guarantees the external cost curve is a line, presenting a constant slope for all the scenarios examined[1]. When the heterogeneity of value of time is modelled, however, such a tangency condition generally does not hold (See Appendix C for a detailed discussion).

[1] The effect of homogeneous value of time is most evident when the matching function exhibits constant returns to scale. With the tangency condition, it is straightforward to show the average waiting and searching times are constant (Yang and Yang, 2011). The average waiting and searching times often serve as the "quality" measurement. Consequently, the monopolist and social planner will provide the same "service quality" if the customers are homogeneous in their value of time (Spence, 1975; Yang and Yang, 2011).

3.2.4.2 Contract solution set and properties

The solution set for all the investigated cases are conceptually displayed in Figure 3-2. The monopoly and first-best solutions are marked as the end points *M* and *S*, respectively along a contract curve. The remaining points on the curve correspond to the second-best solution of varying joint profit levels between the platform and its drivers. Each point on the contract curve is characterized by the triple tangency of the joint profit contour, consumers' surplus contour and social welfare contour (Spence, 1975). That is, for a given reservation profit, the associated point maximizes the total welfare and vice versa.

Define the average joint profit as $\bar{\pi} = \frac{FQ - cN - C_P(Q)}{Q}$. It is intriguing to investigate the state where the increase of Q will be mutually beneficial to all the participants, i.e., $\frac{d\bar{\pi}}{dQ} > 0$ and $\frac{dw^c}{dQ} < 0$ (Yang and Yang, 2011). Consider the process of gradually increasing Q from the monopoly to the first-best solution, it is generally expected the total joint profit keeps decreasing (and so does the average joint profit, i.e. $\frac{d\bar{\pi}}{dQ} < 0$) while the social welfare is strictly increasing[2]. In fact, the contract curve is characterized by Eqs. (3-8) - (3-10) together with Eq. (3-38). Total differentiating Eqs. (3-8) and (3-38) yields:

$$dw^c = \overline{W_1^c} dQ + \overline{W_2^c} dw^v = \frac{1-\alpha_1-\alpha_2}{\alpha_2} \frac{w^c}{Q} dQ - \frac{\alpha_1}{\alpha_2} \frac{w^c}{w^v} dw^v \tag{3-39}$$

$$\beta\alpha_1 dw^c = c\alpha_2 dw^v \tag{3-40}$$

[2] Parallel to Proposition 1 in Yang and Yang (2011), we can treat F, Q as decision variables and evaluate $\partial\bar{\pi}/\partial Q$ and $\partial w^c/\partial Q$. It can be shown the mutually beneficial situation occurs only when the matching function shows increasing returns to scale and the waiting time elasticity of demand is sufficiently large. However, such a solution is not on the contract curve and thus not of particular interest here.

Eliminating dw^t and re-arranging the terms, we have:

$$\frac{dw^c}{dQ} = \frac{1-\alpha_1-\alpha_2}{\left(\alpha_1+\alpha_2\right)}\frac{w^c}{Q} \tag{3-41}$$

which is negative given the matching function shows increasing returns to scale.

3.2.5 Regulation Policies

As previously shown, the monopoly ride-sourcing market is suboptimal in terms of social welfare and thus regulation may be necessary. This section seeks for regulation strategies by assuming the regulator has complete information. We examine possible combinations of regulatory variables to identify the most efficient regulation strategy that induces the second-best and requires the minimum number of regulatory variables. The potential regulation variables include P, F, N .

Let F^{SB}, N^{SB}, P^{SB} solve P4. Regulating all three variables definitely works. Since the break-even constraint for drivers is binding at the second-best solution as given in Proposition 1, together with Eqs. (3-8)-(3-10), the nonlinear equation system has only two degrees of freedom (four equations and six unknowns). Therefore, regulating any two of these three variables at the second-best level should work.

Below we show that properly regulating the commission charged by the platform can also yield the second best. To see this, recall that the second-best solution can also be achieved by maximizing the customer demand subject to the reservation profit constraint. If the commission is regulated, the proprietary ride-sourcing platform will maximize the customer demand, given that the cost function of the platform is assumed to exhibit economies of scale. Therefore, a proper choice of commission P by the regulator would trigger the platform to maximize its profit to the reservation level, leading the final solution to coincide with the second-best one. Mathematically, when only P is regulated at the second-best level, the FONC of P1 is given by

$c = \beta \frac{\alpha_1}{\alpha_2} \frac{w^c}{w^v}$ and $(F - P^{SB})Q - cN = 0$. Correspondingly we have $c = \beta \frac{\alpha_1}{\alpha_2} \frac{w^c}{w^v}$ and $(F - P^{SB})Q - cN = 0$ from the FONC of P4. Both systems of equations have two unknowns N, F and share the same equations. Therefore, they will have the same solution.

However, regulating *F* or *N* only will not work.

3.2.5.1 Regulating *F*

When only restricting $F = F^{SB}$, the FONC of P1 is given by $F^{SB} = C_P'(Q) + c / \frac{\partial Q}{\partial N}$ and $(F^{SB} - P)Q - cN = 0$. Correspondingly from the FONC of P4, we have $F^{SB} = C_P'(Q) + c / \frac{\partial Q}{\partial N} + \frac{1}{1+\lambda} \frac{Q}{f'}$ and $(F^{SB} - P)Q - cN = 0$. Both systems of equations have two unknowns: N, P, but they differ in one equation and thus generally admit different solutions.

3.2.5.2 Regulating *N*

When only restricting $N = N^{SB}$, the FONC of P1 is given by $F = C_P'(Q) - \frac{Q}{f'} + \beta \overline{W_1^c} Q - \beta \overline{W_2^c} \frac{N^{SB}}{Q}$ and $(F - P)Q - cN^{SB} = 0$. Correspondingly from the FONC of P4, we have $F = C_P'(Q) - \frac{\lambda}{1+\lambda} \frac{Q}{f'} + \beta \overline{W_1^c} Q - \beta \overline{W_2^c} \frac{N^{SB}}{Q}$ and $(F - P)Q - cN^{SB} = 0$. Following the same argument as above, we conclude that regulating *N* only may not achieve the second best.

In summary, we have the following proposition.

Proposition 2. Assuming customers are homogeneous in their value of time and the regulator has complete information, regulating the commission alone will achieve the second

best if the matching function exhibits increasing returns to scale and the cost function of the ride-sourcing platform shows economies of scale.

Admittedly, the discussion so far has been based on the assumption of homogeneous value of time for the customers. If a continuous distribution is adopted to capture the heterogeneity in the value of time, merely regulating the commission may not be enough. The main reason is that the tangency condition given in Eq. (3-38) may no longer hold. The regulator may then have to regulate two variables to achieve the second best.

3.3 Competing Platforms

The above analyses on a hypothetical single platform shed some light on the properties of the ride-sourcing market. However, the real-world situation is more complicated since several ride-sourcing platforms are often competing for the market (e.g., Uber and Lyft in the U.S.). Some proponents of ride-sourcing companies have stated that competition may lower down the price level and improve social welfare. In this section, we investigate such a statement in a duopoly ride-sourcing market. We consider that a driver will only work for a particular platform and a customer only uses one platform for a particular trip. It is not uncommon that a customer or driver installs more than one ride-sourcing apps. Such a multi-homing issue is not considered here (Rochet and Tirole, 2003; Armstrong, 2006).

With one more ride-sourcing platform, the utility function for a customer to choose each option is considered as:

$$U_{R_i} = V_0 - \mu_i + \varepsilon_{R_i},\ i \in 1,2 \tag{3-42}$$

$$U_O = V_0 - C + \varepsilon_O \tag{3-43}$$

All the other specifications are the same as the single platform case. The demand functions of the two platforms depend on the distribution of the error terms but can generally be

represented as $Q_1 = f^1(\mu_1, \mu_2)$ and $Q_2 = f^2(\mu_1, \mu_2)$. To facilitate the analysis, we make the following assumptions:

Assumption 1. The own-price effect is negative: $f_1^1 = \frac{\partial Q_1}{\partial \mu_1} < 0$, $f_2^2 = \frac{\partial Q_2}{\partial \mu_2} < 0$. The cross-price effect is positive and symmetric: $f_1^2 = \frac{\partial Q_2}{\partial \mu_1} = \frac{\partial Q_1}{\partial \mu_2} = f_2^1 > 0$. Further, $-f_1^1 > f_2^1$ and $-f_2^2 > f_1^2$.

The above assumption generally holds for the demand functions of two competing firms. Similarly, we assume a Cobb-Douglas matching function for each ride-sourcing platform. To simplify our analysis, we further assume the same parameter set $\{A, \alpha_1, \alpha_2\}$ as that in the single platform for each platform. Given the interpretation of the parameters in Section 3.1, we essentially assume that both platforms have the same geographic coverage of users, running speed of vehicles, and matching technology as the single platform. To summarize, we have the following assumption:

Assumption 2. The matching function for each ride-sourcing platform follows the Cobb-Douglas type with the same parameters as those being used for the single platform. Namely, $Q_1 = A\left(N^{v_1}\right)^{\alpha_1}\left(N^{c_1}\right)^{\alpha_2}$ and $Q_2 = A\left(N^{v_2}\right)^{\alpha_1}\left(N^{c_2}\right)^{\alpha_2}$.

3.3.1 Competition between Ride-Sourcing Platforms

As seen previously, the commission $P_i, i = 1, 2$ can be substituted into the objective function using the reservation profit constraint. It is positive with the assumption on the non-negativity of the joint profit between the platform and its drivers in a mature ride-sourcing

market. We further assume both platforms choose $\left(F_i, N_i\right)$, $i=1,2$ simultaneously for profit maximization and focus on the Nash equilibrium (NE). Let's consider Platform 1:

$$\text{(P5)} \max_{F_1 \geq 0, N_1 \geq 0} \pi_1 = F_1 Q_1 - cN_1 - C_{P_1}(Q_1) \tag{3-44}$$

By setting $\frac{\partial \pi_1}{\partial F_1} = 0$ and $\frac{\partial \pi_1}{\partial N_1} = 0$, we have:

$$\left[F_1 - C_{P_1}{}'\left(Q_1\right)\right]\frac{\partial Q_1}{\partial F_1} = -Q_1 \tag{3-45}$$

$$\left[F_1 - C_{P_1}{}'\left(Q_1\right)\right]\frac{\partial Q_1}{\partial N_1} = -c \tag{3-46}$$

Note that $\frac{\partial Q_1}{\partial F_1} = \frac{\delta_1}{1-\delta_1\Delta_1}$ and $\frac{\partial Q_1}{\partial N_1} = \frac{\delta_1 \beta \overline{W_2^{c_1}} / Q_1}{1-\delta_1\Delta_1}$, where $\delta_1 = \frac{f_1^1 - \Delta_2 f_1^1 f_2^2 + \Delta_2 f_1^2 f_2^1}{1-\Delta_2 f_2^2}$,

$\Delta_i = \beta\left[\overline{W_1^{c_i}} - \overline{W_2^{c_i}}\frac{N_i}{\left(Q_i\right)^2}\right], i=1,2$. Next, dividing Eq. (3-45) by Eq. (3-46) and substituting $\frac{\partial Q_1}{\partial F_1}$ and $\frac{\partial Q_1}{\partial N_1}$, we obtain:

$$c = \beta \frac{\alpha_1}{\alpha_2}\frac{w^{c_1}}{w^{v_1}} \tag{3-47}$$

Similar to the single-platform scenario, the tangency condition still holds, i.e., the cost-minimizing behavior of each competing platform yields a customer's waiting time being proportional to a driver's searching time.

Given Eq. (3-47), one can verify $\Delta_i = \frac{1}{Q_i}\left[\frac{1-\alpha_1-\alpha_2}{\alpha_1} c w^{v_i} + c\left(w^{v_i} + l\right)\right] > 0, i=1,2$. Define the price elasticity of $Q_1\left(\mu_1, \mu_2\right)$ as $\eta_1 = -f_1^1 \frac{F_1}{Q_1} > 0$. Eq. (3-45) can be spelled out as:

$$F_1 - C_{P_1}{}'\left(Q_1\right) - Q_1\Delta_1 = \frac{F_1}{\eta_1}\left(\frac{f_1^1}{\delta_1}\right) \tag{3-48}$$

It is expected that $\eta > 1$ and $\eta_1 > 1$ for the platform(s) to charge positive trip fares. The pricing formula in Eq. (3-48) is of similar form to that in Eq. (3-24), apart from being multiplied by a scaling factor f_1^1 / δ_1. It is strictly larger than one given $f_1^1 < 0$ and $\Delta_2 f_1^1 f_2^2 > \Delta_2 f_1^2 f_2^1 > 0$. The ratio of price levels under duopoly and monopoly is given as follows:

$$\frac{F_1}{F} = \frac{\overline{C}_P + \frac{1-\alpha_2}{\alpha_1} c w^{v_1} + cl}{\overline{C}_P + \frac{1-\alpha_2}{\alpha_1} c w^{v} + cl} \frac{\eta_1 \eta - \eta_1}{\eta_1 \eta - \eta \left(\frac{f_1^1}{\delta_1} \right)} \tag{3-49}$$

where the platform production cost is assumed linear, i.e. $C_{P_1}'(Q_1) = C_P'(Q) = \overline{C}_P$. For a symmetric NE, it is expected $Q_1 \le Q$. Similar to Eqs. (3-39)-(3-41), the average waiting and searching times increase as the decrease of the platform specific demand given the matching function has increasing returns to scale. It follows that $w^{v_1} > w^{v}$ ($w^{c_1} > w^{c}$), indicating more matching frictions. Accordingly, the first component in the RHS of Eqn. (3-49) is no less than one while the second component is related to the price elasticities of demand. Its value is unclear without fully specifying the demand function. If $\frac{\eta_1}{\eta} < \left(\frac{f_1^1}{\delta_1} \right)$, the price level under NE is strictly larger than that under the monopoly. Otherwise, the ratio of F_1 over F remains indeterminate. Generally, one needs to explore the change of the price elasticity of demand and that of the matching friction. The conventional wisdom that competition lowers the price level does not stand if the increase of the former is dominated by the latter.

3.3.2 Second Best for Dual Ride-Sourcing Platforms

This section directly investigates the second-best outcome with the dual platforms. The first-best solution can be obtained by setting the Lagrangian multiplier associated with the reservation profit constraint to zero in the optimality conditions of the second-best problem.

Define $V(\mu_1, \mu_2)$ as the customers' surplus from completing a trip. With certain regularity conditions (Sheffi, 1985), the following properties hold by construction:

$$\frac{\partial V}{\partial \mu_1} = -Q_1,\ \frac{\partial V}{\partial \mu_2} = -Q_2 \tag{3-50}$$

To simplify the derivation, we treat $(\mu_i, N_i),\ i = 1,2$ as the decision variables. Mathematically, $\mu_i \equiv F_i + \beta w^{c_i} + \tau l,\ i = 1,2$. The maximization program is formally written as:

$$\text{(P6)}\ \max_{\boldsymbol{\mu}>0, \mathbf{N}\geq 0} S = V(\mu_1, \mu_2) + \sum_{i=1}^{2}\left[F_i Q_i - cN_i - C_{P_i}(Q_i)\right] \tag{3-51}$$

s.t.

$$F_i Q_i - cN_i - C_{P_i}(Q_i) \geq \pi_{P_i}^o,\ i = 1,2 \tag{3-52}$$

Define its Lagrangian function as follows:

$$L(\boldsymbol{\mu}, \mathbf{N}) = V(\mu_1, \mu_2) + \sum_{i=1}^{2}(\lambda_i + 1)\left[F_i Q_i - cN_i - C_{P_i}(Q_i)\right] - \sum_{i=1}^{2}\lambda_i \pi_{P_i}^o \tag{3-53}$$

By setting $\frac{\partial L}{\partial N_i} = 0,\ i = 1,2$, it is immediate to show:

$$c = \beta \frac{\alpha_1}{\alpha_2}\frac{w^{c_i}}{w^{v_i}},\ i = 1,2 \tag{3-54}$$

which again follows the tangency condition. Next, setting $\frac{\partial L}{\partial \mu_i} = 0,\ i = 1,2$ yields:

$$(\lambda + 1)\left(\gamma_1 f_1^1 + \gamma_2 f_1^2\right) = -\lambda Q_1 \tag{3-55}$$

$$(\lambda + 1)\left(\gamma_1 f_2^1 + \gamma_2 f_2^2\right) = -\lambda Q_2 \tag{3-56}$$

where $\gamma_i = F_i - C_{P_i}'(Q_i) - Q_i \Delta_i,\ i = 1,2$. Substituting γ_2 in Eq. (3-55), we then focus on the pricing formula for platform 1:

$$\gamma_1 = F_1 - C_{P_1}'(Q_1) - Q_1 \Delta_1 = -\frac{\lambda_1}{\lambda_1 + 1}\frac{Q_1}{f_1^1}\sigma_1 \tag{3-57}$$

where $\sigma_1 = \frac{\left(Q_2 f_1^2\right)/\left(Q_1 f_2^2\right)-1}{\left(f_1^2 f_2^1\right)/\left(f_1^1 f_2^2\right)-1}$. For more insight, further assume a symmetric solution where $Q_1 = Q_2$ and $f_1^1 = f_2^2$. Then σ_1 is reduced to $\frac{f_1^1}{f_1^1 + f_1^2}$ which is strictly larger than one given Assumption 1. Define the price elasticity of $Q_1\left(\mu_1, \mu_2\right)$ as $\eta_1 = -f_1^1 \frac{F_1}{Q_1} > 0$, Eq. (3-57) can be rewritten as:

$$F_1 - C_{P_1}{}'\left(Q_1\right) - Q_1 \Delta_1 = \frac{\lambda_1}{\lambda_1 + 1} \frac{F_1}{\eta_1} \sigma_1 \tag{3-58}$$

The pricing formula for platform 2 can be obtained by symmetry. The formula structure is the same as that in the single platform case, except being adjusted by the factor σ_1, which has its roots in platform collusion (a case where the two platforms cooperatively maximize the total profit). The fact that it is greater than one is by the nature of the pricing for substitute goods[3].

The first-best pricing formula for the dual platforms can be obtained by setting $\lambda_i = 0,\ i = 1, 2$. Similarly, the first-best solution is not sustainable if the matching function is assumed to exhibit increasing returns to scale and the cost function of the platform has economies of scale. The second-best pricing formula can be thought of the convex combination of the first-best and the collusion formulas.

The solution sets obtained in the above scenarios can also be displayed over a contract curve that connects the social optimal (Point S) and NE solution (Point N) in Figure 3-3 (Point C

[3] Assume a market without friction or externality. Define $Q_1 = f^1\left(F_1, F_2\right),\ Q_2 = f^2\left(F_1, F_2\right)$ to be demand functions for two competing firms. Further assume $f_1^1 < 0,\ f_2^2 < 0,\ f_2^1 = f_1^2 > 0$; the total cost for production is linear: $cQ_1,\ cQ_2$. The collusion pricing formula under the symmetric solution is: $F_1 - c = \frac{F_1}{\eta_1} \frac{f_1^1}{f_1^1 + f_1^2}$.

represents the solution that two platforms collude to maximize the joint profit. It corresponds to the monopoly solution). The remaining points on the contract curve represent a continuum of the second-best solutions characterized by different regulated profit levels. Define the average joint profit for platform i $\bar{\pi}_i, i = 1, 2$, one can verify $\frac{d\bar{\pi}_i}{dQ_i} < 0, i = 1, 2$ for the solutions on the contract curve. $\frac{dw^{c_i}}{dQ_i}, i = 1, 2$ is strictly negative when the matching function shows increasing returns to scale.

3.3.3 Welfare Changes between Single and Dual Platforms

In this section, we attempt to investigate whether or not competition yields a more efficient market outcome and whether a regulated duopoly market is more efficient than a regulated monopoly market. To address the former, we compare the social welfare under the monopoly solution and duopoly solution. For the latter, we focus on the welfare change under the regulated second-best solutions.

To proceed, we specify the demand function by assuming that the error terms of the utility functions are identically and independently Gumbel distributed for both the single (Eqs. (3-5)-(3-6)) and dual platforms[4] (Eqs. (3-42)-(3-43)). Therefore, the demand functions for the single platform and the other transportation modes are:

$$Q = \frac{e^{-\theta\mu}}{e^{-\theta\mu} + e^{-\theta C}}\bar{Q},\ Q = \frac{e^{-\theta C}}{e^{-\theta\mu} + e^{-\theta C}}\bar{Q} \tag{3-59}$$

where $\theta > 0$. For consistency, we assume the total base demand remains the same for both scenarios and the demand for the dual platforms are given as:

[4] The error terms are likely dependent in the duopoly market. This assumption is made for the ease of welfare comparison.

$$Q_1 = \frac{e^{-\theta\mu_1}}{\Lambda}\bar{Q},\ Q_2 = \frac{e^{-\theta\mu_2}}{\Lambda}\bar{Q},\ Q_d = \frac{e^{-\theta C}}{\Lambda}\bar{Q} \tag{3-60}$$

where $\Lambda = e^{-\theta\mu_1} + e^{-\theta\mu_2} + e^{-\theta C}$.

In addition to specifying an increasing-return-to-scale matching function as highlighted in Assumption 2, we limit ourselves to the following aspects to further facilitate the comparison:

- The solutions for the dual ride-sourcing platforms are symmetric. The related variables are marked by the subscript *d*. e.g. $Q_d = Q_1 = Q_2$, $w^{v_d} = w^{v_1} = w^{v_2}$.
- Total cost incurred by the platform is linear: $\bar{C}_P Q = C_P(Q)$.

3.3.3.1 Analytical comparison

Utilizing the assumption on symmetric solution and linear platform cost, the welfare change between the single and dual platforms for a given market scenario is as follows:

$$\begin{aligned}\Delta S = S_d - S_s = &\left[\frac{1}{\theta}\ln\left(2e^{-\theta\mu_d} + e^{-\theta C}\right) - \frac{1}{\theta}\ln\left(e^{-\theta\mu} + e^{-\theta C}\right)\right]\bar{Q} + \\ &2Q_d\left[F_d - c\left(w^{v_d} + l\right) - \bar{C}_P\right] - Q\left[F - c\left(w^{v} + l\right) - \bar{C}_P\right]\end{aligned} \tag{3-61}$$

The first two terms measure the change in the consumers' surplus (Train, 2009) while the last two terms represent the change in producers' surplus, i.e., the joint profit between the platform and its drivers. For the duopoly and monopoly solutions, it is anticipated $2Q_d \geq Q \geq Q_d$. For welfare evaluation, one needs to trade off the potential increase of consumers' surplus and the decrease from the joint profit.

Consider the extreme case that Q_d approaches to $\frac{Q}{2}$ from the right. Then the change of consumers' surplus is approximately zero. Eq. (3-61) is reduced to:

$$\begin{aligned}\Delta S &= 2Q_d\left[F_d - F - c\left(w^{v_d} - w^{v}\right)\right] \\ &= 2Q_d\left[\ln 2 - \left(\frac{\alpha_1 + \alpha_2}{\alpha_1}\right)c\left(w^{v_d} - w^{v}\right)\right]\end{aligned} \tag{3-62}$$

One can verify $\Delta S \geq 0$ when $w^{v_d} - w^v \leq \frac{\alpha_1}{c(\alpha_1 + \alpha_2)} \ln 2$. That is, competition is welfare improving if the increased searching friction is bounded above by a constant. The same condition holds when evaluating the welfare change at the second-best solutions (e.g. at zero profit level), where the change of the social welfare is equal to that of the consumers' surplus. Unfortunately, the explicit relationship between w^{v_d} and w^v (or equally w^{c_d} and w^c) appears hard to obtain and neither has a closed form. We therefore resort to a numerical experiment to quantify the welfare changes.

3.3.3.2 Numerical analysis

Define $\Omega = \left\{\alpha_1, \alpha_2, A, C, \overline{Q}\right\}$ to be the experimental parameter set. α_1, α_2, A are directly included in the matching function while $\overline{Q}$ and C influence the base demand and realized demand splits. Since our focus is on the matching technology, the other parameters (e.g., c, β) are held constant. In fact, it can be shown $\frac{\partial w^v}{\partial \alpha_1} < 0, \frac{\partial w^v}{\partial \alpha_2} < 0, \frac{\partial w^v}{\partial A} < 0, \frac{\partial w^v}{\partial \overline{Q}} < 0, \frac{\partial w^v}{\partial C} < 0$ for all the solutions on the contract curve (The derivation for the sign of $\frac{\partial w^v}{\partial A}$ is given in Appendix D for illustration). Marginal increase of each of these parameters reduces the matching friction and thus indicates more efficient matching technology.

It is assumed $V_0 = 50$ (\$/trip), $\beta = 20$ (\$/hr), $c = 20$ (\$/hr), $\theta = 1$ (1/\$), $\tau = 6$ (\$/hr), $l = 0.3$ (hr), $\overline{C}_P = 2$ (\$/trip). For demonstration, we investigate two base demand levels: $\overline{Q} = 100,000$ and $50,000$ (trip/hr) and the outside transportation costs: $C = 19$ and 15 (\$/trip). For each combination of $\overline{Q}$ and C, a continuum of solutions are examined over a range of (α_1, α_2, A).

We explored three levels of returns to scale with $\alpha_1 + \alpha_2 = 1.6, 1.4, 1.2$, respectively. For symmetry, α_1, α_2 are set equal.

The welfare changes between duopoly and monopoly solutions are given in Figure 3-4. For a meaningful comparison, we assume the joint platform demand will account for 10%-90% of the total passenger demand in the equilibrium states. This in turn indicates that the difference between the generalized cost of using the ride-sourcing service and the outside options cannot be very large. Accordingly, two dot curves are introduced to bound the solution region. Each solid red curve represents the computed welfare change over a feasible range of *A* for a given elasticity pair (α_1, α_2). The little bar at the left end of the curve means the location where no solution is available for either problem P1 or P5 if we further reduce *A*. The welfare difference curve slopes upwards, indicating ΔS increases as the increase of A. For a given A, the increase of elasticities (α_1, α_2) generally yields larger ΔS.

In Figure 3-4 (a), the changes of the total social welfare are positive but for the left end in the case of $\alpha_1 = \alpha_2 = 0.7$. When the base demand increases from 50,000 to 100,000 (trip/hr), the total pattern moves upwards as seen in Figure 3-4 (b). All the stable solutions in the region are characterized by $\Delta S > 0$, indicating competition between the platforms is welfare improving. However, when C increases from 14 to 19 ($/trip) as shown in Figure 3-4 (c), some left portion of all curves fall in the solution region where $\Delta S < 0$. It should be pointed out that the welfare difference curve is plotted over the interval (0.12, 0.3) of A, which is different from Figure 3-4 (a) and (b). In the cases shown in Figure 3-4 (c), there are more matching frictions. Figure 3-5 shows the average searching times under duopoly and monopoly solutions when $\alpha_1 = \alpha_2 = 0.7$ and $\overline{Q} = 100,000$ (trip/hr). It can be observed that the average searching times for drivers are

larger in duopoly than those in monopoly due to the decrease of platform-specific demands. For the cases presented in Figure 3-4 (c) where competition is not efficient, the average search time is much higher, approximately 6-12 minutes.

The welfare comparison between the regulated dual and single platforms is given in Figure 3-6. The general tendency is similar to that in Figure 3-4. In many cases, competition increases social welfare. However in Figure 3-6 (c), all the meaningful solutions are characterized by $\Delta S < 0$. In those situations, a regulated monopoly platform will be superior to the regulated duopoly platforms in term of efficiency.

Although the numerical example is made from the arbitrary specification of function forms and parameter values, its interpretation is straightforward. Competition may not improve the social welfare when the matching technology is less efficient. Given the assumption of the increasing returns to scale of the matching function, its efficiency is not all about the advanced algorithms used by the ride-sourcing platform but also the size of the market, i.e., the number of users. To derive effective policies, the regulatory agency may need to obtain a good estimate on customers' demand function as well as the matching function. If the matching friction is large, the regulatory agency may rather encourage the merger of the ride-sourcing companies and then regulate them as a monopolist.

3.4 Summary

We analyze the ride-sourcing service using an aggregate model with the matchings between customers and drivers captured by an exogenous Cobb-Douglas matching function. We examine different market scenarios, solution properties and general economic outcomes of a hypothetical monopoly ride-sourcing market. It is found that without regulatory intervention the monopoly ride-sourcing platform would maximize the joint profits with its drivers. The first-best solution is not sustainable when the matching function is increasing returns to scale and the cost

incurred by the platform exhibits economies of scale. Therefore, we further analyze the second-best scenario with varying reservation profit levels for the ride-sourcing platform. In terms of market frictions, all the examined scenarios are characterized by a proportional relationship between the average waiting and searching times that implies the cost minimization of the matching production. In view of the market distortion, we demonstrate the feasibility of regulating two variables to achieve the second best. With the assumption of homogeneous value of time of customers, we further show that regulating only the commission should guarantee the second best.

To address the effects from platform competition, we extend our analysis by considering a duopoly setting. It is found that competition may not necessarily lower down the price levels when the increase of the matching friction overrides that of price elasticity of demand after the introduction of the competing platform. We further investigate the effects of competition on the welfare change when the matching function is increasing returns to scale. Based on the sensitivity analysis, we observe that competition can reduce social welfare when the matching technology is less efficient, and the increased matching friction for each platform dominates the surplus generated by having one additional option. In this case, the regulator may rather encourage the merger of the platforms and regulate them directly as a monopolist.

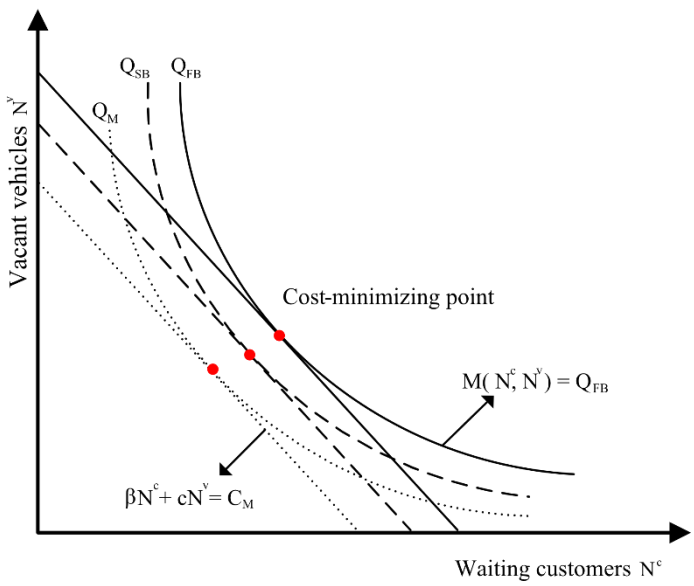


Figure 3-1. Cost Minimization of the Matching Technology.

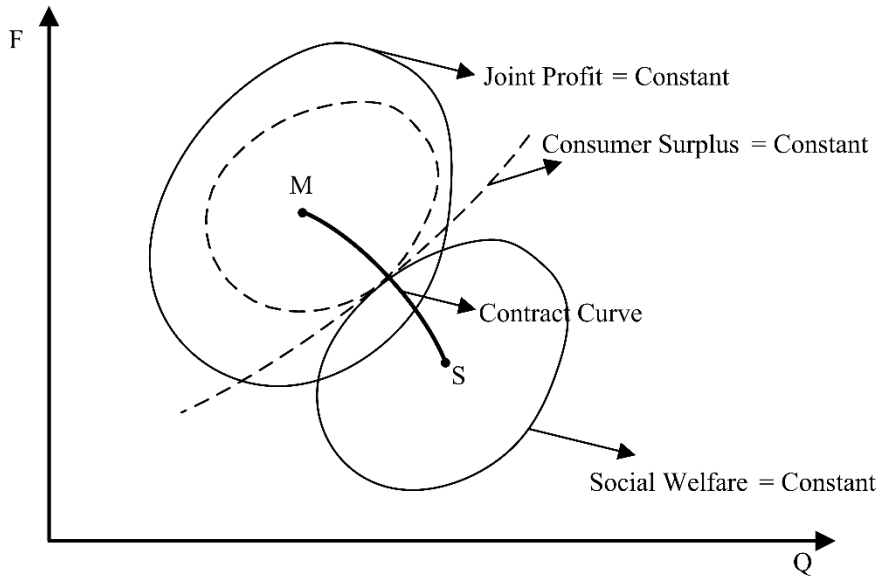


Figure 3-2. Contract Solution Set for Single Ride-Sourcing Platform.

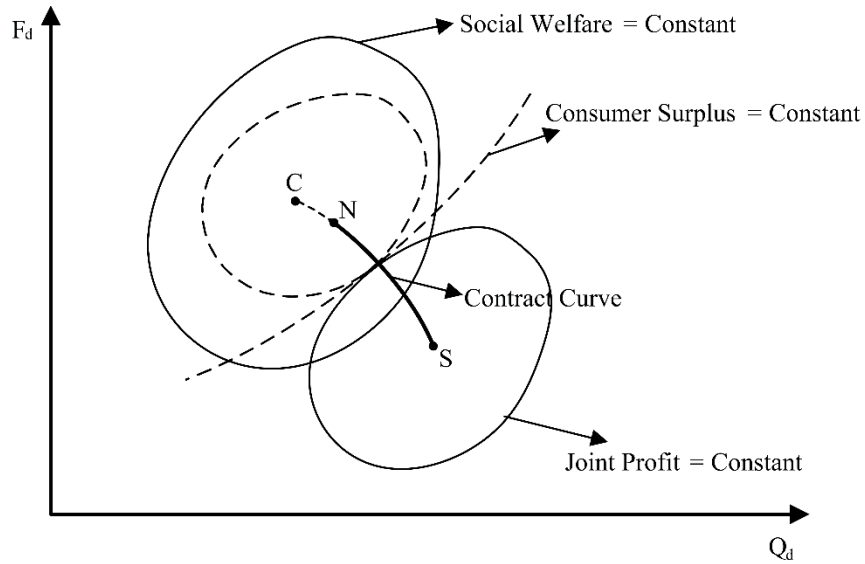


Figure 3-3. Contract Solution Set for Dual Ride-Sourcing Platforms.

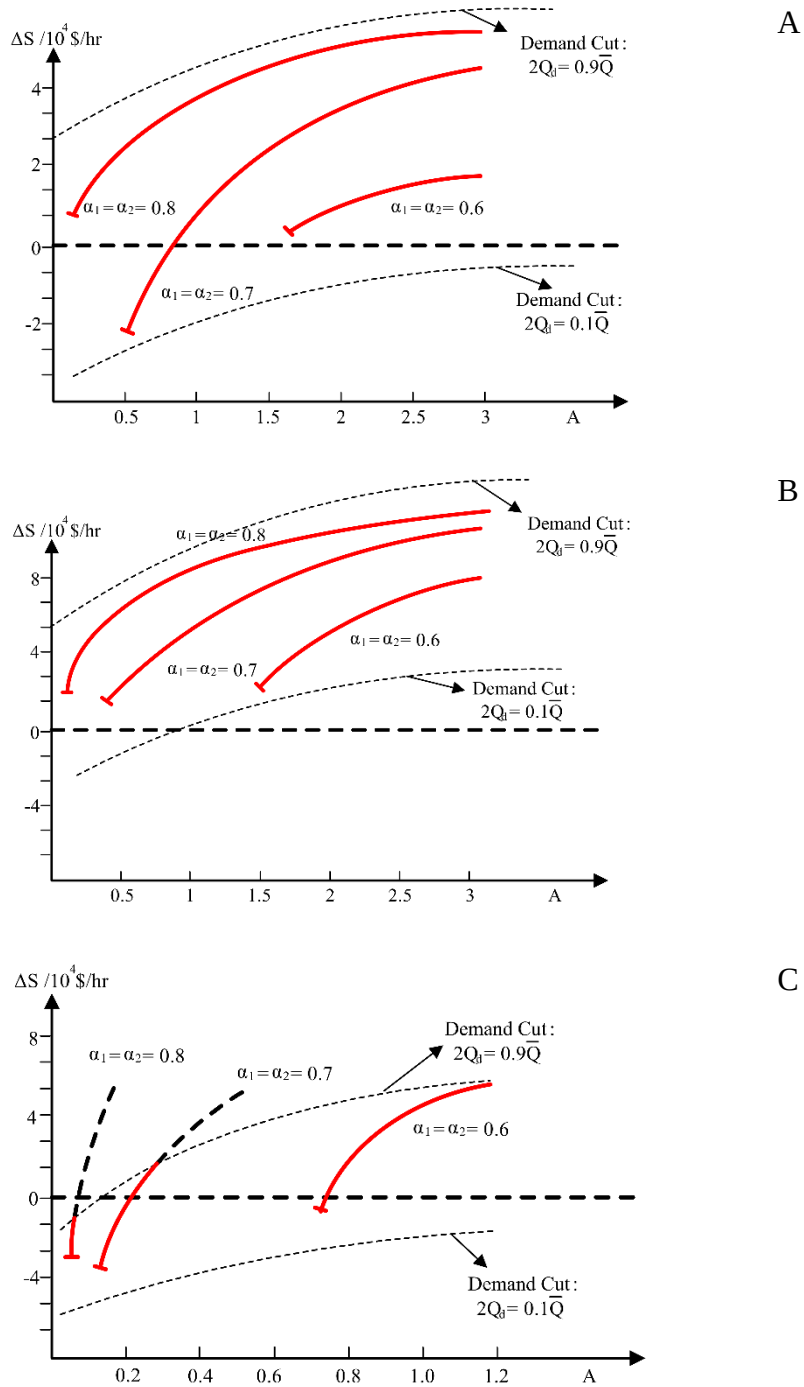


Figure 3-4. Change in Social Welfare between Duopoly and Monopoly Solutions. A) $C = 14$, $\overline{Q} = 50000$ (unit: \$/ trip, trip/hr), B) $C = 14$, $\overline{Q} = 100000$ (unit: \$/ trip, trip/hr), C) $C = 19$, $\overline{Q} = 100000$ (unit: \$/ trip, trip/hr)

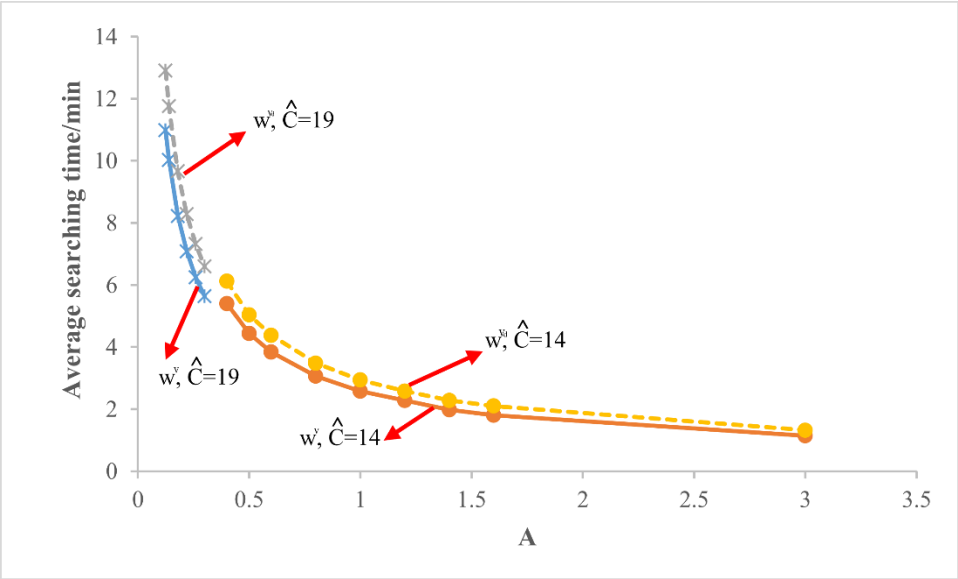


Figure 3-5. Average Searching Time ($\alpha_1 = \alpha_2 = 0.7$ and $\overline{Q} = 100000$).

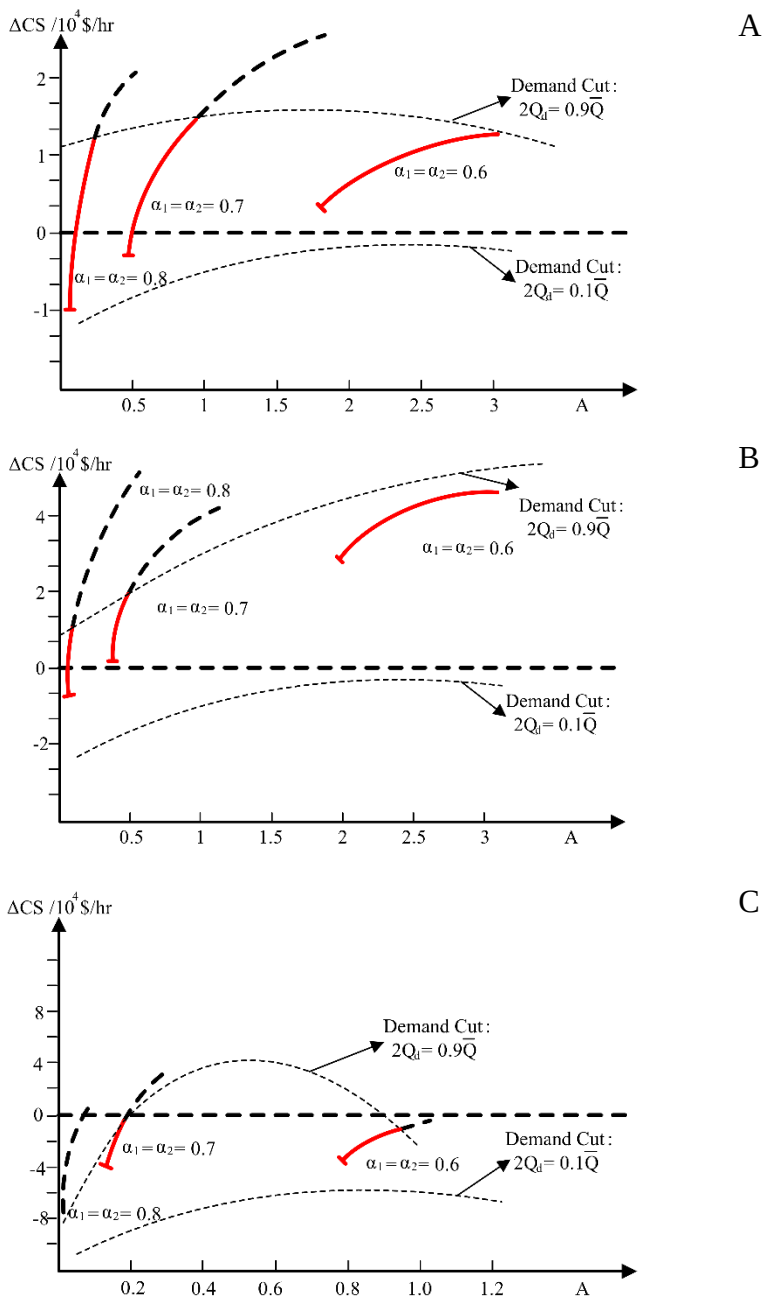


Figure 3-6. Change of Consumers' Surplus between Dual and Single Platforms at Second-Best Solutions (at Zero Profit Level). A) $C = 14$, $\overline{Q} = 50000$ (unit: \$/trip, trip/hr), B) $C = 14$, $\overline{Q} = 100000$ (unit: \$/trip, trip/hr), C) $C = 19$, $\overline{Q} = 100000$ (unit: \$/ trip, trip/hr)

CHAPTER 4
DYNAMIC PRICING AND LABOR SUPPLY IN ON-DEMAND RIDE-SOURCING MARKETS

In this chapter, we relax from the static setting in Chapter 3 and aim to analyze the temporal effects of dynamic pricing. Note that ride-sourcing companies provide flexibility for drivers to choose their work schedules. Some may work full time like professional taxi drivers while others only provide service for limited hours (e.g., on their way home from work). Such flexibility enables the use of drivers' fragmented time particularly with the existence of price surge. On one hand, drivers may adjust their work schedules to cover more profitable periods; on the other hand, temporal variation in the wage rate induced by dynamic pricing may further affect drivers' decision of how long they would like to work (Chen and Sheldon, 2016). Therefore, it is intriguing to investigate the effect of dynamic pricing by considering drivers' work hour choices in a dynamic context.

In the literature of labor economics, competing theories or hypotheses exist in understanding how a driver determines her shift length. The neoclassical theory expects drivers to work longer when their wage rate is higher, while the income-targeting theory speculates that drivers have target levels after which they are more likely to stop (Camerer et al., 1997; Farber, 2015). No consensus has yet been achieved on which theory better explains taxi drivers' labor supply decisions, even though empirical analyses have been carried out. Furthermore, the analyses mainly focus on inputs such as wage rate and work hours while structural information such as shift starting and ending times is largely overlooked.

In Section 4.1, we conduct a brief review of the literature in the labor supply. Subsequently in Section 4.2-4.3, we construct a time-expanded network to outline drivers' work scheduling under both labor supply assumptions. Given the variations in customers' demand and surge multipliers, drivers' equilibrium choices of start and end times, break and work durations

are all endogenously determined. Section 4.4 illustrates the equilibrium outcomes through numerical examples. In Section 4.5, we investigate the impact of dynamic pricing using a bi-level programing framework. The upper level specifies the platform's objective of revenue maximization while the lower level captures the equilibrium work hour choices of the drivers. A simple regulation scheme is presented when market power is a concern. Further insights on dynamic pricing and regulation outcomes are discussed via numerical experiments. Section 4.6 summarizes the chapter.

4.1 Literature on Labor Supply

Competing theories have been developed to explain employees' choice of work hours in response to the variations of hourly wage rate. The income-targeting theory was initially proposed by Camerer et al. (1997) to explain the phenomenon of negative labor supply elasticity found in a New York taxi dataset (i.e., an increased wage rate leads to shorter work hours). It can be seen as a special case of reference-dependent preferences (Kőszegi and Rabin, 2006) where a "kink" exists for the marginal utility of income at the reference income level. Drivers are loss averse in the sense that they suffer more when the target level is not achieved while the motivation of continuing working beyond the target vanishes. A dual reference approach was proposed by Crawford and Meng (2011) where drivers were assumed to have targets for both daily income and working hours. Their model confirms the role of reference dependence in labor supply and is flexible in accounting for the case when income effect is not significant.

Comparatively, the neoclassical theory suggests drivers will work longer given a higher wage rate, indicating a positive hourly wage rate elasticity of working hours (Farber, 2005; Farber, 2015). As often argued by the proponents of the neoclassical theory, the existence of negative elasticity is largely due to the "division bias" in regressing cumulative working hours against hourly wage rate or the improper use of instrument for wage rate (Farber, 2005; Farber,

2015). Most recently, Farber (2015) and Chen and Sheldon (2016) both reported positive labor supply elasticity using a large-scale New York taxi dataset and a high-resolution Uber dataset, respectively. Considering drivers may adjust their target income gradually based on their historical experience, Farber (2015) further argued that the income-targeting hypothesis may only make sense for temporary unanticipated wage variations.

Despite several studies from Uber's economists (Hall and Krueger, 2015; Chen and Sheldon, 2016), the research of labor supply in the ride-sourcing market is limited. As there is no clear evidence yet which theory would outperform, we will propose formulations based on both theories.

4.2 Basic Modeling Considerations

We report the basic modeling considerations in this section. A time-expanded network is first constructed in a similar spirit as that of Yang et al. (2005a) to delineate the work schedules of ride-sourcing drivers. Then we specify the demand form for customers, as well as the revenue and cost structures for drivers.

4.2.1 Time Expanded Network

A time-expanded network $G(V, A)$ is presented in Figure 4-1 where V and A represent the set of nodes and links, respectively. The nodes represent time points for an aggregate ride-sourcing market and no spatial structure of the market is explicitly specified. All drivers will start from the same hypothetic origin $O \in V$ and end at the hypothetic destination $D \in V$. For each driver, a path between the O-D pair defines a work schedule, which provides the complete information on a driver's daily labor supply. In contrast to the concept of shift adopted in traditional taxi studies (e.g., Yang et al., 2005), the work schedule here also contains the break periods between subsequent working sessions.

More specifically, we divide a whole day into *n* time periods of an equal length of *T*, and then denote V_1 the subset of work nodes *T*, 2*T*, …, *nT* and $V_1^{'}$ the subset of the associated auxiliary nodes. Without loss of generality, we set *T* to 1 hour and *n* to 24 as shown in Figure 4-1. So, the set of nodes in V_1 corresponds to clock hours. A driver at a work node can either traverse the subsequent work link in search of customers or temporarily take a rest by heading to the corresponding auxiliary node. When at an auxiliary node, the driver can travel to the adjacent work, auxiliary or destination node, which respectively corresponds to going back to work, continuing in the resting mode and ending the work completely.

The set of links A can be divided into five mutually exclusive and collectively inclusive subsets. A_0 - A_4 denotes the set of departure, work, transition, rest and end links. Note that passenger demand only arises on work links that belong to A_1, at which drivers can be matched with customers. Time is consumed only at "vertical links" that belong to $A_1 \cup A_3$, i.e., when a driver is at work or takes a rest. We thus use "time period" and "link" interchangeably for each link $b \in A_1 \cup A_3$. As an example, the path O→1→ 2→ 2'→ 3'→3→4→5→6→7→ 7'→D defines a work schedule in which a driver starts working at 0:00 a.m., takes a break for one hour between 2:00 and 3:00 a.m., goes back to work until 7:00 a.m. and then stops.

4.2.2 Demand and Revenue

The ride-sourcing market is not merely cleared by price, and each customer has to wait for some time until being served. Analytically, we assume that the passenger demand is a decreasing function of both average trip fare, $\bar{F}_b$ and average waiting time, w_b^c:

$$Q_b = q_b\left(\bar{F}_b, w_b^c\right),\ \forall b \in A_1 \tag{4-1}$$

where $\bar{F}_b = F_0 + \gamma_b F_b$. F_0 is the flag-drop fee; $F_b = \omega_b l_b$ is the time-based charge equal to a unit-time charge ω_b multiplied by the average trip time l_b; and γ_b is the surge multiplier that only applies to the time-based portion. We have $q_{b,1} = \partial Q_b / \partial \bar{F}_b < 0$ and $q_{b,2} = \partial Q_b / \partial w_b^c < 0$ by assumption.

Customers' average waiting time is closely related to the matching technology of the ride-sourcing platform, i.e., the algorithm that matches customers with drivers. Several recent studies adopted a Cobb-Douglas function to delineate the matching process from which the waiting time can be derived (He and Shen, 2015; Wang et al., 2016). For simplicity, we assume that the ride-sourcing platform matches a waiting customer with the closest vacant vehicle. Under this assumption, the average waiting time can be approximated as follows (Arnott, 1996):

$$w_b^c = \frac{k}{2v_b\sqrt{N_b^v / S}}, \ \forall b \in A_1 \tag{4-2}$$

where $k \approx \sqrt{\pi/2}$ is a scaling parameter that adjusts the Euclidean distance to the Manhattan distance; v_b is the average speed of the vehicles; N_b^v is the vacant vehicle hours and S is the area of the ride-sourcing market. The derivation of Eq. (4-2) does not account for the case where a vacant vehicle turns out to be the closest one for multiple customers and thus the formula provides a lower bound for the customers' average waiting time.

Considering that a ride-sourcing vehicle is either occupied or vacant when traversing a work link, the following relationship holds for an analysis period of one hour:

$$N_b^v + N_b^o = u_b, \ \forall b \in A_1 \tag{4-3}$$

where N_b^o is the occupied vehicle hours and u_b is the total vehicle hours. Let's denote $w_b^v = N_b^v / Q_b$ as the average searching (waiting) time until finding a customer. When the average service duration ($w_b^v + l_b$) is far less than *T*, we have:

$$N_b^o = Q_b l_b,\ \forall b \in A_1 \tag{4-4}$$

The nonlinear equation system of Eqs. (4-1)-(4-4) behaves properly. More specifically, for the given surge multipliers, the base trip fare as well as the total fleet size, the existence and uniqueness of the market equilibrium can be proven by means of the fixed point theorem as shown in Yang et al. (2005a). The ride-sourcing industry features a revenue-sharing structure. For each completed trip, the platform charges a fixed proportion η (e.g., 20%-25%) from the total fare as the commission while the driver keeps the remaining $1-\eta$. The average revenue for a driver in a working hour is then written as:

$$R_b = \frac{(1-\eta)\bar{F}_b Q_b}{u_b},\ \forall b \in A_1 \tag{4-5}$$

It is interesting to note:

$$\frac{\partial R_b}{\partial u_b} = \frac{1}{u_b^2}(1-\eta)\bar{F}_b Q_b(\varsigma_b - 1),\ \forall b \in A_1 \tag{4-6}$$

$$\frac{\partial w_b^c}{\partial u_b} = w_b^{c\prime}\frac{N_b^o}{u_b}\left(\frac{u_b}{N_b^o} - \varsigma_b\right),\ \forall b \in A_1 \tag{4-7}$$

where $\varsigma_b = \frac{\partial Q_b}{\partial u_b}\frac{u_b}{Q_b}$ is the elasticity of demand with respect to the vehicle hours and w_b^c is a function of N_b^v. In addition, $w_b^{c\prime} < 0$ and $\frac{\partial Q_b}{\partial u_b} = \frac{q_{b,2} w_b^{c\prime}}{1 + q_{b,2} w_b^{c\prime} l_b} > 0$.

Considering Eqs. (4-6)-(4-7), we have $\frac{\partial R_b}{\partial u_b} < 0,\ \frac{\partial w_b^c}{\partial u_b} < 0$ when $\varsigma_b < 1$. It is possible $\frac{\partial R_b}{\partial u_b} > 0,\ \frac{\partial w_b^c}{\partial u_b} < 0$ when $\varsigma_b > 1$. In the latter case, more vehicle hours would increase the average hourly revenue for the drivers and simultaneously reduce average waiting time of customers. However, such a Pareto-improving scenario may only occur with an unrealistically small vehicle size where the marginal reduction on the average waiting time stimulates significantly more demand (Yang et al., 2005a; Yang and Yang, 2011). For the analysis hereinafter, we limit our discussion to the following:

Assumption 1. Given the surge multipliers, the base trip fare and the total fleet size, the average hourly revenue of ride-sourcing drivers is strictly decreasing in vehicle hours, i.e., $\frac{\partial R_b}{\partial u_b} < 0,\ \forall b \in A_1$.

Without further referencing Eqs. (4-1)-(4-5), we denote the average revenue as a function of the total vehicle hours, i.e.,

$$R_b = R_b\left(u_b\right),\ \forall b \in A_1 \tag{4-8}$$

4.2.3 Cost for Drivers

A substantial level of heterogeneity exists in the work scheduling of ride-sourcing drivers. We define M the set of drivers with elements denoted by *m*. We assume drivers differ in their preferred start period and work duration. As previously mentioned, a path in the proposed network defines a work schedule. Following a path, a driver incurs two types of variable costs. The first cost is path dependent, denoted as c^{pm}, measuring the disutility from cumulative work hours, which is assumed to increase more than linearly with the cumulative

hours. The other is the link-specific cost c_a^m that captures the cost a driver experiences when traversing a specific link. It represents:

- a driver's preference in choosing the start hour as well as the average operating cost of driving (e.g., from home) to a normal business area (A_0);
- the operating cost during one work period (A_1) or driving back home (A_4);
- a sufficiently small modeling artifact to avoid unrealistic loops between work and rest links (A_2);
- or a driver's preference in choosing rest times (A_3).

We define P as the set of all possible simple paths between the OD pair. The total cost incurred by a driver of class m on path *p* is written as follows:

$$\overline{C}^{pm} = \sum_{a \in A} c_a^m \delta_a^p + c^{pm} = \sum_{a \in A} c_a^m \delta_a^p + \varpi_1^m \left(h^p\right)^{\varpi_2^m}, \ \forall p \in P, \ m \in M \tag{4-9}$$

where δ_a^p is the link-path incidence; ϖ_1^m is the cost associated with cumulative working hours $h^p = \sum_{b \in A_1} \delta_b^p$; and ϖ_2^m captures the level of aversion to long work durations and is assumed to be larger than 1.

For a simple quantification of available vehicles when the average service duration is rather small (much less than a modelling period *T*), we assume:

Assumption 2. All ride-sourcing vehicles that provide service during a time period will be available in the following period.

The relaxation of Assumption 2 is discussed in Section 4.3.3. We denote f^{pm} the number of drivers of type *m* choosing path *p*. Then the relationship between the link and path vehicle flow follows:

$$u_a = \sum_m \sum_p f^{pm} \delta_a^p, \ \forall a \in A \tag{4-10}$$

4.3 Equilibrium Models with Endogenous Labor Supply

We provide formulations and algorithms for the two labor supply assumptions. An enhanced treatment for long trip duration is also discussed.

4.3.1 Neoclassical Equilibrium Model

Given the surge multipliers and average trip duration, we consider a short-run market equilibrium with a fixed total fleet size. It is reflected by a multiclass network flow equilibrium on the time-expanded network, formally defined as follows:

Definition 1. At equilibrium, for each driver class $m \in M$, all paths that carry positive vehicle flows yield equal "profit" (total revenue minus all cost components), which is no less than that of any unused path.

The equilibrium defined above can be viewed as a consequence of the neoclassical labor supply theory (Farber, 2015). However, we incorporate the interactions among different drivers in competing for trip opportunities. Note this definition, together with our specification of hourly demand, implies that drivers and customers make decisions at different time scales, a modeling feature also highlighted by Banerjee et al. (2015). Drivers typically consider profits at a daily level while customers are sensitive to the price and waiting time only when they request rides.

We are now ready to present the following mathematical program whose solution describes the equilibrium of the ride-sourcing market:

(ME-N)

$$\min_{\mathbf{f}} \; Z = \sum_{b \in A_1} \int_0^{u_b} -R_b(w)\,dw + \sum_{m \in M} \sum_{a \in A} c_a^m u_a^m + \sum_{p \in P} \sum_{m \in M} c^{pm} f^{pm} \tag{4-11}$$

s.t.

$$\sum_{p \in P} f^{pm} \le N^m, \; \forall m \in M \tag{4-12}$$

$$f^{pm} \ge 0, \; \forall p \in P, \; m \in M \tag{4-13}$$

where u_a, $\forall a \in A$ is defined in Eq. (4-10). Given Assumption 1, it is straightforward to verify that ME-N is convex in $\mathbf{f}$ with linear constraints. However, multiple solutions may exist as strict convexity does not hold. The optimality conditions of ME-N yield the following system of complementarity equations:

$$f^{pm} : \lambda^m - \pi^{pm} \geq 0,\ f^{pm}\left(\lambda^m - \pi^{pm}\right) = 0,\ f^{pm} \geq 0,\ \forall p \in P,\ m \in M \tag{4-14}$$

$$\lambda^m : N^m - \sum_{p \in P} f^{pm} \geq 0, \left(N^m - \sum_{p \in P} f^{pm} \right) \lambda^m = 0,\ \lambda^m \geq 0,\ \forall m \in M \tag{4-15}$$

where $\pi^{pm} = \sum_{b \in A_1} R_b \delta_b^p - \sum_{a \in A} c_a^m \delta_a^p - c^{pm}$ represents the average profit of the chosen path. Let $\boldsymbol{\lambda}$ be the vector of λ^m, $m \in M$, the Lagrangian multipliers associated with constraints (4-12). It is nonnegative and can be interpreted as the equilibrium profit. The complementary slackness condition in constraint (4-14) corresponds to our definition of market equilibrium. We make the following assumption to limit our discussion to a reasonable profit domain.

Assumption 3. The equilibrium profit $\boldsymbol{\lambda}$ is strictly positive.

Assumption 3 avoids the unrealistic situation where the maximum possible profit that a driver could earn is zero. It implies that constraints (4-12) are always binding. Besides, the gradients of constraints (4-12) and those binding in constraints (4-13) are linearly independent. This regularity condition guarantees the uniqueness of $\boldsymbol{\lambda}$ (Bertsekas, 1999).

The formulation of ME-N bears some similarity with the one in Yang et al. (2005). The key difference is that our formulation considers that individual drivers drive their own cars to provide for-hire rides and explicitly models each driver's choice of work hours. The formulation of ME-N has two distinctive features. Namely, $R_b\left(u_b\right)$, $b \in A_1$ is not explicit in u_b and path costs are not link-additive. The former makes it cumbersome to evaluate the objective function while the latter precludes the possibility of using efficient link-based solution algorithms for network

equilibrium flow problems (Patriksson, 2015). Instead, various path-based solution algorithms can be applied to solve ME-N, e.g., the one based on a gap function proposed by Lo and Chen (2000), which directly aims to solve the nonlinear complementarity problem (NCP) characterized by Eqs. (4-14)-(4-15). The gap function is defined as:

$$G(\boldsymbol{\lambda},\mathbf{f}) = \sum_p \sum_m \phi\left(f^{pm}, \lambda^m - \pi^{pm}\right) + \sum_m \phi\left(\lambda^m, N^m - \sum_p f^{pm}\right) \tag{4-16}$$

where $\phi(x,y) = 0.5\left[\sqrt{x^2+y^2} - (x+y)\right]^2$. Note that $G(\boldsymbol{\lambda},\mathbf{f}) \geq 0$ and $G(\boldsymbol{\lambda},\mathbf{f})$=0 if and only if Eqs. (4-14)-(4-15) hold (Lo and Chen, 2000). The solution to ME-N can be obtained by solving an unconstrained minimization program GAP-N whose optimal value is zero, i.e., $\min\ G(\boldsymbol{\lambda},\mathbf{f}) = 0$ and $\left(\boldsymbol{\lambda}^*,\mathbf{f}^*\right) = \arg\min\ G(\boldsymbol{\lambda},\mathbf{f})$ (Lo and Chen, 2000).

Recognizing eventually only a small fraction of paths carry positive vehicle flow, we apply a column generation scheme to gradually expand a pre-defined small path set. To ensure the validity of Assumption 1, the initial path set should be chosen properly to cover every work link. The column generation approach requires iteratively solving a shortest path finding problem with a non-additive path cost conditional on the current optimal link/path flows for each driver class:

(SP)

$$\min_{\mathbf{x}^m}\ \overline{C}^m = \varpi_1^m \left(\sum_{b\in A_1} x_b^m\right)^{\varpi_2^m} + \sum_{a\in A} c_a^m x_a^m - \sum_{b\in A_1} R_b^* x_b^m \tag{4-17}$$

s.t.

$$EX^m = d \tag{4-18}$$

$$x_a^m \in \{0,1\},\ a \in A \tag{4-19}$$

where $E \in R^{|V|\times|A|}$ is the node-link incidence and $d \in R^{|A|}$ is a column vector with element 1 and -1 corresponding to the origin and destination nodes and 0 for the other intermediate nodes. At each iteration, at most one path (with profit level larger than the current equilibrium profit) is added to the path set. We denote this growing path set as P^e. The detailed steps of the solution procedure are summarized as follows:

- Step 0: Select several paths that cover the work links $b \in A_1$ and construct the path set P^e.
- Step 1: Given P^e, solve GAP-N and obtain the optimal solution $\mathbf{f}^*$ and $\boldsymbol{\lambda}^*$.
- Step 2: For each $m \in M$, solve SP and obtain the optimal path p^{m*} with the minimum cost $\overline{C}^{m*}$. If $-\overline{C}^{m*} > \lambda^m$, add p^{m*} into P^e. If $-\overline{C}^{m*} = \lambda^m$ for all $m \in M$, stop; otherwise, go to Step 1.

4.3.2 Income-Targeting Equilibrium Model

The income-targeting theory assumes that drivers work until reaching a daily target, beyond which the gain from working is more likely to be offset by the disutility of working long hours. Specifically, drivers of class $m \in M$ are assumed to choose their work schedules to maximize their utility, which is defined as follows (Farber, 2015):

$$U^{pm} = \begin{cases} \left(1+\rho^m\right)\left(R^p - I^m\right) - \overline{C}^{pm} + U_0, \ R^p < I^m \\ \left(1-\rho^m\right)\left(R^p - I^m\right) - \overline{C}^{pm} + U_0, \ R^p \geq I^m \end{cases}, \forall p \in P, \ m \in M \tag{4-20}$$

where I^m is the target income level; R^p represents the total revenue of a chosen work schedule; ρ^m controls the degree of loss aversion and is assumed to vary between $(0,1)$. The above specification can be viewed as the consumption part of the utility $R^p - I^m - \overline{C}^{pm} + U_0$ augmented with the gain-loss component $\pm\rho^m\left(R^p - I^m\right)$ (Farber, 2015). For the work hour choice problem on the proposed network, a constant U_0 is added to guarantee $U^{pm} > 0$. Drivers gain more from

making money when their target level is not achieved and gain relatively less if they continue working after passing the target. A "kink" therefore occurs in the marginal utility of daily revenue (Farber, 2015). Note that that Eq. (4-21) can be further reduced to:

$$U^{pm} = R^p - \rho^m \left| R^p - I^m \right| - \overline{C}^{pm} + U_0^m,\ \forall p \in P,\ m \in M \tag{4-21}$$

where $U_0^m = U_0 - I^m$.

To shed light on the income-target assumption, we deviate to parameterize $R^p = w^p h^p$ which makes h^p explicit; and assume that $\overline{C}^{pm}\left(h^p\right)$ is a differentiable function of h^p. Note w^p is the wage rate that equals the average drivers' revenue per time period (Farber, 2015). The plot in Figure 4-2 assumes U^{pm} is concave in h^p and treats it as a continuous variable for demonstration purposes. For sufficiently low or high average hourly wages (e.g., w_4^p and w_1^p, respectively), the motivation for utility maximization yields the neoclassical fashion of labor supply. In the former, drivers stop working earlier because of bad trip opportunities, while in the latter they work longer because the high wage rate overrides the disutility from cumulative work hours. However, for the intermediate wage levels (e.g., w_2^p and w_3^p), the optimal decision is to stop right after hitting the target income level. This is because the wage rate is high enough to motivate drivers to work longer when $R^p < I^m$ but too low to offset the disutility from the increased work hours when $R^p > I^m$. For example, an increase of wage rate from w_3^p to w_2^p reduces the work hours from I^m / w_3^p to I^m / w_2^p, indicating a negative elasticity of labor supply. Although in our model h^p is implicit with the resolution equal to one modeling period, we expect the general intuition still holds.

The definition of equilibrium is similar to Definition 1, with "profit" replaced by "utility." We assume I^m is exogenously given and remains fixed. However, the target income

level can be endogenized as the mean of daily revenue when the stochasticity in drivers' work hour choices is considered (Crawford and Meng, 2011; Farber, 2015). See Xu et al. (2011) for a similar treatment to develop a prospect-based user equilibrium model with endogenous reference points.

We define the set of feasible path flow as $\Omega = \left\{ \sum_{p \in P} f^{pm} = N^m,\ f^{pm} \geq 0, \forall m \in M, \forall p \in P \right\}$. The equilibrium condition can be captured by the following NCP:

(ME-I)

$$\left(U^m - U^{pm}\right) f^{pm} = 0,\ \forall m \in M,\ p \in P \tag{4-22}$$

$$U^m - U^{pm} \geq 0,\ \forall m \in M,\ p \in P \tag{4-23}$$

where $U^m = \max_{p \in P}\left(U^{pm}\right) > 0$ and $f^{pm} \in \Omega$. When $\rho^m = 0$, the equilibrium condition degrades to that of the neoclassical case. It is well-known that the solution to the above NCP, i.e., $\left(\mathbf{U}\left(\mathbf{f}^*\right), \mathbf{f}^*\right)$ in the vector form, can be characterized by the following variational inequality (Nagurney, 2013):

$$\sum_p \sum_m \left(-U^{pm*}\right)\left(f^{pm} - f^{pm*}\right) \geq 0,\ \forall m \in M,\ p \in P \tag{4-24}$$

where $f^{pm} \in \Omega$. Note that U^{pm} is continuous in R^p and is further continuous in f^{pm}. The set Ω is convex and compact. Therefore, a solution always exists. However, path flows may not be unique, as U^{pm} is not strictly monotone in f^{pm}.

The procedure for solving the income-targeting model is similar to the one for solving the neoclassical model, with some modifications to the sub problems of flow equilibration and shortest path finding. The absolute operator in Eq. (4-21) poses a challenge for classical derivative-based algorithms in minimizing the gap function. We thus provide a reformulation

based on the fact that $\left|R^p - I^m\right|$ is equivalent to the sum of two non-negative auxiliary variables, i.e., $y_1^{pm} + y_2^{pm}$, where $y_1^{pm} - y_2^{pm} = R^p - I^m$, $y_1^{pm} y_2^{pm} = 0$, $y_1^{pm} \geq 0$, $y_2^{pm} \geq 0$. The complementary feature of this reformulation fits well in the adopted gap function. If we denote the vector of U^m, f^{pm} and y^{pm} as $\mathbf{U}, \mathbf{f}$ and $\mathbf{y}$, then the program for the income-targeting model can be reformulated:

(GAP-I)

$$\min_{\mathbf{U},\mathbf{f},\mathbf{y}_1} G\left(\mathbf{U},\mathbf{f},\mathbf{y}_1\right) = \sum_p \sum_m \phi\left(f^{pm}, U^m - U^{pm}\right) + \sum_m \phi\left(U^m, N^m - \sum_p f^{pm}\right) + \sum_p \sum_m \phi\left(y_1^{pm}, y_2^{pm}\right) \tag{4-25}$$

where $y_2^{pm} = y_1^{pm} - \left(R^p - I^m\right)$ is by construction. An optimal equilibrium solution $\left(\mathbf{U}^*, \mathbf{f}^*, \mathbf{y}^*\right)$ is characterized by $G\left(\mathbf{U}^*, \mathbf{f}^*, \mathbf{y}^*\right) = 0$. For the shortest path finding, we have:

(SP-I)

$$\min_{\mathbf{x}^m,\mathbf{y}^m} \overline{C}^m = \varpi_1^m \left(\sum_{b \in A_1} x_b^m\right)^{\varpi_2^m} + \sum_{a \in A} c_a^m x_a^m - \sum_{b \in A_1} R_b^* x_b^m + \rho^m \left(y_1^m + y_2^m\right) + \phi\left(y_1^m, \mathrm{y}_2^m\right) - U_0^m \tag{4-26}$$

s.t.

Eqs. (4-18)-(4-19)

$$y_1^m - y_2^m = \sum_{b \in A_1} R_b^* x_b^m - I^m, \ \forall m \in M \tag{4-27}$$

4.3.3 Enhanced Equilibrium Models with Long Trip Duration

The previous two formulations rely on Assumption 2 and are valid if $w_b^v + l_b << T$, $\forall b \in A_1$. The assumption can be violated in hours of heavy congestion. When l_b is close to or more than *T*, a vehicle will not be available in the beginning of the next time period after picking up a customer. The key challenge of modelling potential long trip duration is to properly trace the availability of a vehicle, which is jointly determined by how long the vehicle is

occupied in the previous period(s) and its status in the current period (e.g., to continue working or not). The reason for the former is obvious since a vehicle previously occupied with long trip duration may still be occupied in the current period. The latter is imposed to "force" the movement of the vehicle consistent with its path: a driver can never take a trip longer than one period if she is about to exit the market. The detailed formulations and discussions are given in Appendix E for interested readers.

4.4 Numerical Examples

We first specify the functional forms and parameter values necessary for the discussion. Then, we numerically investigate the equilibrium solutions of the proposed formulations and demonstrate their implication on the labor supply elasticity.

4.4.1 Set-up

We set T equal to one hour and assume an exponential (hourly) demand function for each work link:

$$Q_b = \bar{Q}_b \exp\left(-\theta_b \left(\bar{F}_b + \beta_b w_b^c + \tau_b l_b\right)\right), \forall b \in A_1 \tag{4-28}$$

where $\bar{Q}_b$ is the base demand; θ_b is the demand sensitivity parameter; β_b and τ_b are the value of waiting time and in-vehicle traveling time, respectively. The specification above assumes customers are homogeneous in the value of time for a given time period, and the temporal variations are captured by $\left(\bar{Q}_b, \theta_b, \beta_b, \tau_b\right)$.

Daily distribution of the base demand is depicted in Figure 4-3, with two peak periods at 7:00-9:00 and 17:00-19:00. Accordingly, we set β_b and τ_b at \$33/hr and \$16/hr for peak hours while \$20/hr and \$10/hr for the remaining hours. We let θ_b be 0.03 and 0.05 for peak and off-peak hours, respectively. These specifications reflect the assumption that customers value time more during peak hours and become less sensitive to price (Small, 2012).

We consider four driver classes, each of which consists of 2,000 drivers. Specifically, we denote drivers who prefer to start late and work long/short hours as class $m = 1/2$, and those preferring to start early and work long/short hours as class $m = 3/4$. Let $H(a)$ denote the index of the starting node of link a. For a driver of $m = 1,\ 2$ that starts on link $a \in A_0$, we set $c_a^m = 1$ when $15 \le H(a) \le 19$ and 10 otherwise. For a driver of $m = 3,\ 4$ on link $a \in A_0$, we set $c_a^m = 1$ when $7 \le H(a) \le 10$ and 10 otherwise. For the work, transition, rest and end links of all driver types, we have $c_a^m = 5,\ \forall a \in A_1$; $c_a^m = 0.1,\ \forall a \in A_2$ and $c_a^m = 1,\ \forall a \in A_3 \cup A_4$, respectively. To represent drivers' heterogeneity of cumulative working hours, we set $\varpi_1^m = 2$ for $m = 1,\ 3$ and $\varpi_1^m = 3$ for $m = 2,\ 4$. All costs are in the unit of \$/hour. The income target levels are set at \$300, \$200, \$300 and \$200 for *m* = 1 to 4. Values for all parameters are summarized in Table F-1 of Appendix F for the convenience of the readers.

4.4.2 Equilibrium Solutions

We present the equilibrium solutions to ME-N and ME-I in Table 4-1 and Table 4-2, respectively. For ME-N, the equilibrium profits for all chosen paths (work schedules) are \$107.4, \$65.3, \$107.4, \$74.3 for class 1 to 4, respectively, with average work hours equal to 8.52, 4.83, 7.58 and 4.95 respectively. It can be observed that although drivers of *m* = 1 tend to start later than those of *m* = 3, the distinction among the starting hour is not significant for those preferring shorter work hours (*m* = 2, 4). Another set of sufficiently differentiated link costs may be used to better demonstrate such heterogeneity. Large variations exist for the rest (break) hours. For example, some drivers of class 2 do not even take a rest while others may have a break time as long as 7 hours. Our modeling framework clearly shows the capability of capturing the flexibility of ride-sourcing drivers' work scheduling.

For ME-I, all paths that finally carry positive flow are characterized by the same utility levels of \$281.3, \$350.9, \$281.3 and \$359.9 for classes 1-4. Note that the presented values are all added by a constant $U_0 = \$500$. The interpretation of work/rest hours as well as starting time choice is similar to that for ME-N. Due to the existence of the income target, one can easily verify that for each driver class the profit of used paths may not necessarily be equal. Moreover, several chosen paths for m = 1 and 3 have revenue around \$293, which is very close to the specified target. Given the complex structure of the proposed network, however, cases where drivers stop exactly at their target levels are rarely seen.

4.4.3 Labor Supply Elasticity

To investigate the sign of the labor supply elasticity under both assumptions, we present the relationship between average work hours and hourly revenue in Figure 4-4. For demonstration purposes, we simply increase the surge multipliers for all work hours from 1 to 2 every 0.25 unit.

Although average hourly revenue increases with SMs, different patterns in average work hours are uncovered. Higher hourly revenue corresponds to longer work hours in the neoclassical assumption while in the income-targeting theory, this relationship only holds when SMs are larger than 1.25. In this case, the hourly revenue is high enough to offset the disutility from continuously working even after the income target has been achieved.

The income-target fashion of labor supply occurs when SMs increase from 1 to 1.25: an increase in hourly revenue comes with a reduction in average work hours. Further examination of the output when SMs are equal to 1.25 reveals that the average revenue is \$314.6, \$191.9, \$305.1 and \$193.1 for *m* = 1 to 4, respectively, which is close to the corresponding target, i.e., \$300, \$200, \$300 and \$200.

4.5 Dynamic Pricing and Its Regulation

Based on the proposed network that characterizes drivers' work hour scheduling, we introduce a bi-level programming framework to study dynamic pricing. A simple regulation scheme to enhance market efficiency is also presented. To facilitate the presentation, we adopt the neoclassical assumption to characterize the labor supply in this section.

4.5.1 Modeling Framework

We now model the effect of dynamic pricing using bi-level programming where the upper-level problem represents the behavior of the platform while the lower-level problem captures the response of the ride-sourcing system to the decision made by the platform. In the short run, dynamic pricing has a limited impact on attracting additional drivers to register to the platform. We thus assume that the total fleet size remains fixed. We further assume that the base fare structure is given and surge multipliers (SMs) become the only control variable for the platform, with the objective to maximize its daily revenue, i.e., $J = \sum_{b \in A_1} \eta \bar{F}_b Q_b$. The operation cost of the platform is not considered for simplicity. In general, the platform solves the following program:

(BI)

$$\max_{\boldsymbol{\lambda}, \mathbf{f}, \boldsymbol{\gamma} > \mathbf{0}} J \tag{4-29}$$

s.t.

$$G(\boldsymbol{\lambda}, \mathbf{f} \mid \boldsymbol{\gamma}) = 0 \tag{4-30}$$

where J is the control objective and $G(\boldsymbol{\lambda}, \mathbf{f} \mid \boldsymbol{\gamma})$ is the gap function previously discussed, with $\boldsymbol{\lambda}$ being the vector of the equilibrium profit.

For a given path set, BI is difficult to solve with Eq. (4-30) being a constraint. A relaxation of it to an allowable error, i.e., $G(\boldsymbol{\lambda}, \mathbf{f} \mid \boldsymbol{\gamma}) \leq \varepsilon_1$ will make the problem more solvable by commercial nonlinear solvers. A good solution can be obtained by iteratively reducing ε_1. If paths cannot be enumerated, it remains feasible to apply a column generation scheme to determine the effective path set where the shortest path finding problems are solved based on the currently optimized SMs. It is worth mentioning that this type of bi-level programming problem has been extensively studied. Various locally convergent algorithms proposed in the literature may also apply here (Marcotte and Zhu, 1996; Meng et al., 2001).

4.5.2 Commission Cap Regulation

If empirical evidence confirms significant mark-up in platform's price setting, further intervention from the regulatory agency is needed. Below we present a simple but insightful regulation scheme that suggested in Chapter 3: capping the amount of commission charged by the platform.

More specifically, under such a scheme, the monopoly ride-sourcing platform charges at most the capped amount of commission from each transaction. For revenue maximization, it will maximize the total transaction (i.e., realized demand) which is positively related to consumers' surplus. Whenever surge pricing is in place, the surged part completely goes to the drivers. Different choices of the commission cap essentially provide varying profit margins for the ride-sourcing company. In a static setting with strict homogeneity assumptions, we showed that capping commission alone could achieve the second best. Given the heterogeneity among trip distance and duration in practice, the proposed regulation may be implemented in the form of distanced-based charge, or time-based charge, or the combination of both.

The market outcomes under the proposed regulation can be captured by the following program:

(BI-R)

$$\max_{\boldsymbol{\pi},\boldsymbol{\lambda},\mathbf{f},\boldsymbol{\gamma}>\mathbf{0},\hat{\boldsymbol{\eta}}\geq\mathbf{0},\mathbf{N}\geq\mathbf{0}} \hat{J} = \sum_{b\in A_1} \hat{\eta}_b Q_b \tag{4-31}$$

s.t.

$$G\left(\boldsymbol{\pi},\boldsymbol{\lambda},\mathbf{f} \mid \boldsymbol{\gamma},\hat{\boldsymbol{\eta}},\mathbf{N}\right) = 0 \tag{4-32}$$

$$\boldsymbol{\lambda} \geq \boldsymbol{\pi}_{\mathbf{R}} \tag{4-33}$$

$$\hat{\eta}_b \leq \bar{\eta},\ \forall b \in A_1 \tag{4-34}$$

where $\boldsymbol{\pi}_{\mathbf{R}}$, $\hat{\boldsymbol{\eta}}$ and $\mathbf{N}$ are respectively the vectors of drivers' reservation profit, the commission charged from each trip and the fleet size N^m. $\bar{\eta}$ is the commission cap set by the regulatory agency. The notations for the other variables are the same as previously given. BI-R is similar to BI except that 1) the direct controls under the platform are surge multipliers and commission; 2) fleet size is endogenously determined by the entry constraints (34). The consideration of drivers' entry is necessary since otherwise the platform may set an extremely low trip fare (to attract more customers) while leaving drivers in deficit. The analysis is in the longer term as drivers' entry decision is captured. Note that we do not consider the platform's subsequent strategies, if any, to exploit either side when faced with the proposed regulation.

4.5.3 Numerical Experiments

The tolerance rate ε_1 is initially set at 1 for constraint (4-30) and can be reduced to 0.01. Values for the other parameters remain the same as in Section 4.4. We first explore the effect of dynamic pricing. The corresponding plots on SMs, average waiting times and searching times are given in Figure 4-5. In Figure 4-5 (a), a revenue maximizing platform surges in peak periods when customers value time more and become less sensitive to price. Figure 4-5 (b) illustrates

that changes in average waiting and searching times are generally opposite each other. Dynamic pricing suffocates customers' demand and attracts more drivers. Due to less competition for vacant vehicles, customers on average enjoy lower waiting times. In return, it becomes more difficult for a driver to be matched with a customer, and so their average searching time increases.

To answer the question whether or not dynamic pricing is beneficial, we construct a static pricing counterpart where a uniform SM is applied across all periods for revenue maximizing. The market outcomes under such (optimal) static scenario (with the SM equal to 2.02) are used for comparison. The difference is computed as the metric of interest under dynamic pricing minus that under the static one. Formally, we define consumers' surplus to measure the welfare of the passengers, which is calculated under a hypothetical market demand curve where passengers' average waiting time is fixed at the equilibrium level (Cairns and Liston-Heyes, 1996; Yang et al., 2002). Producers' surplus is used to measure the welfare for both the platform and the drivers. It is the sum of the platform's revenue and the drivers' profit[5]. The total social surplus is the sum of the consumers and producers' surplus.

In Figure 4-6, the changes in consumers' surplus and the changes in the joint revenue between the platform and the drivers are plotted[6]. With dynamic pricing, the platform tends to charge less during off-peak hours but more during peak hours. Figure 4-5 shows that the joint revenue increases for all hours. However, changes in customers' surplus are mixed. Customers are better off during the off-peak hours, while during the peak hours, they have to pay significantly more and are thus worse off.

[5] When the income-targeting model is adopted at the lower level, it remains valid to use the total revenue for the platform's surplus while the accurate measure of drivers' surplus should follow Eq. (20).

[6] The calculation of hourly profit is infeasible as it requires decomposing drivers' cumulative cost onto each hour.

It is interesting to note though dynamic pricing has the potential to simultaneously increase consumers' surplus and the joint revenue (e.g., when prices are adjusted below the optimal static pricing level as in off-peaks). However, the net change of the total social surplus aggregated in a day can be positive or negative, depending on the demand pattern and other factors.

Finally, the results are similar when we employ the income-targeting assumption in the lower level problem. Note the income-targeting assumption relies on exogenous parameters such as I^m, the value of which affects the equilibrium flow distribution and eventually the market dynamics. Therefore, a head-to-head comparison of these two assumptions on market dynamics is not meaningful.

We now investigate the performance of the proposed regulation scheme when the market power of the ride-sourcing platform is a concern. We assume the reservation profit levels are \$120, \$80, \$120 and \$80 for driver types 1-4. The operation cost of the platform, and the sunk costs of the platform and the drivers are not considered in quantifying the corresponding surplus.

The left vertical axis of Figure 4-7 (a) represents the ratio of the platform's revenue to the joint revenue (with the drivers). The platform grabs more from the drivers' side as the commission cap increases. The right vertical axis of Figure 4-7 (a) shows the ratio of the social surplus under different regulated commissions to that of a (quasi) second-best case where the platform maximizes the total social surplus while considering drivers' entry. Total social surplus keeps decreasing as the commission cap increases. The commission cap corresponding to the maximum social surplus is very close to zero, which is mainly due to the fact that the platform's costs are not specified in our analysis. Market efficiency is enhanced when compared to the

unregulated monopoly (as the commission cap goes to infinity), which only accounts for approximately 70% of the second-best social surplus.

Figure 4-7 (b) shows the percentage of the platform's revenue from the social surplus, a rough estimate of market power. With the increase of the commission cap, the platform obtains a larger share from the social surplus. In summary, the proper choice of commission cap essentially provides a healthy profit margin for the development of the ride-sourcing company while limiting its market power.

4.6 Summary

By structurally incorporating drivers' work schedule choice, this study has proposed formulations under different behavioral assumptions of labor supply to investigate the effects of dynamic pricing in the ride-sourcing industry. A time-expanded network is first constructed to represent the work scheduling of ride-sourcing drivers. Based on such a network representation, formulations and algorithms are presented to describe the equilibrium of the ride-sourcing market for both the neoclassical and income-targeting assumptions of labor supply. Through numerical experiments, we demonstrate that our models generate outcomes consistent with the definitions of market equilibrium, and are able to uncover the work schedule for multiple driver classes with different break durations, start and end times. We further show that a higher average revenue rate corresponds to longer work hours in the neoclassical assumption but may lead to a reduction in work hours in the income-targeting one.

We next investigate the impact of dynamic pricing based on a bi-level programing framework. With the lower level capturing the equilibrium work schedule choices, the upper level is tailored to represent the scenario where a platform aims to maximize daily revenue. Patterns of average waiting and searching times are generally opposite each other. Our numerical results indicate that the platform and drivers enjoy higher revenue while customers may be at a

loss during highly surged periods. We investigate a commission cap regulation where the platform can charge at most a fixed amount from each transaction. The proposed regulation demonstrates the potential to increase market efficiency and limits the market power of the monopoly platform. In Chapter 5, we will evaluate the effect of dynamic pricing and the proposed commission cap regulation using an empirical dataset.

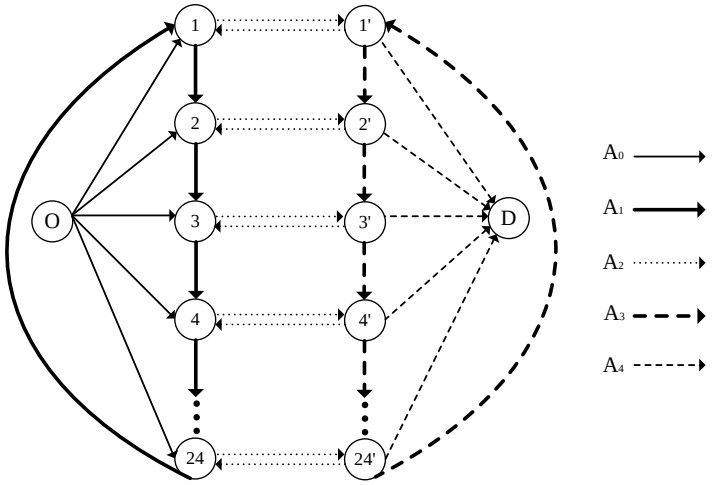


Figure 4-1. Time-Expanded Network for Work Hour Choices.

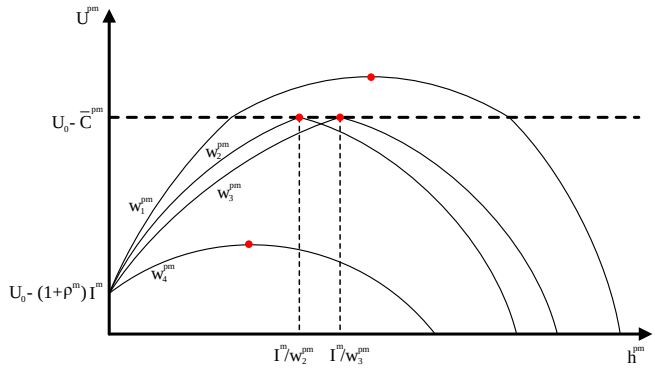


Figure 4-2. The Relationship between Utility and Working Hours with Varying Wage Rates.

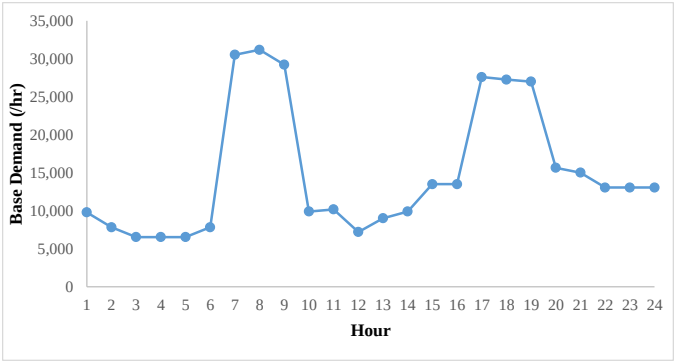
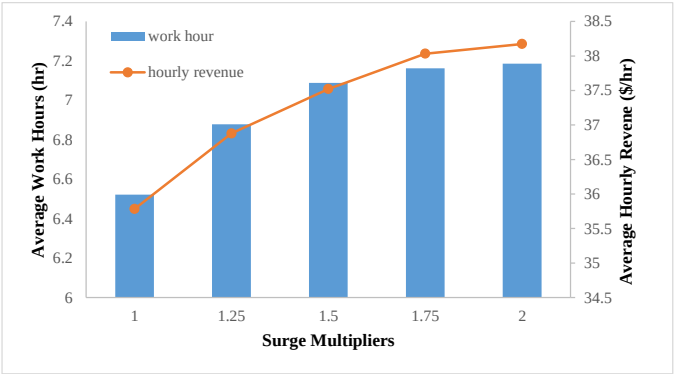
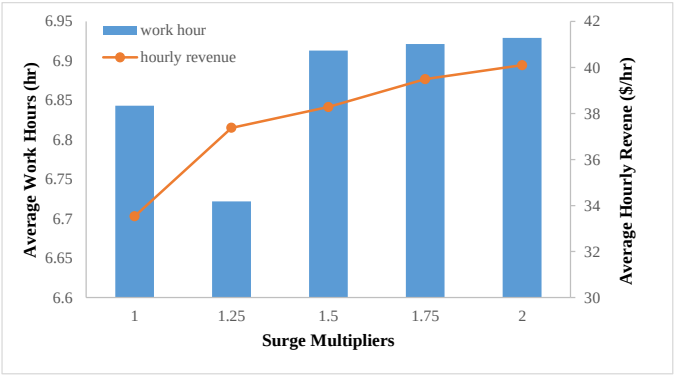


Figure 4-3. Distribution of Daily Base Demand.

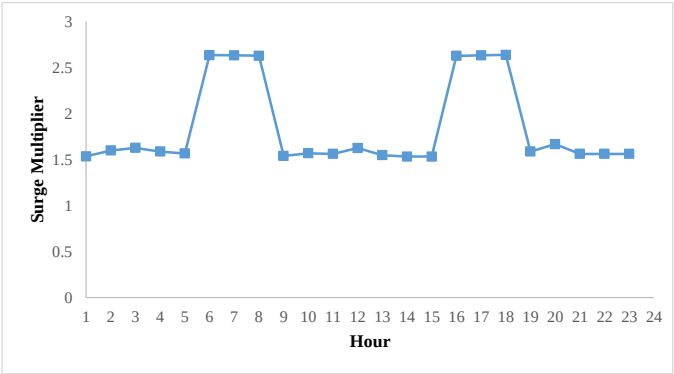


A

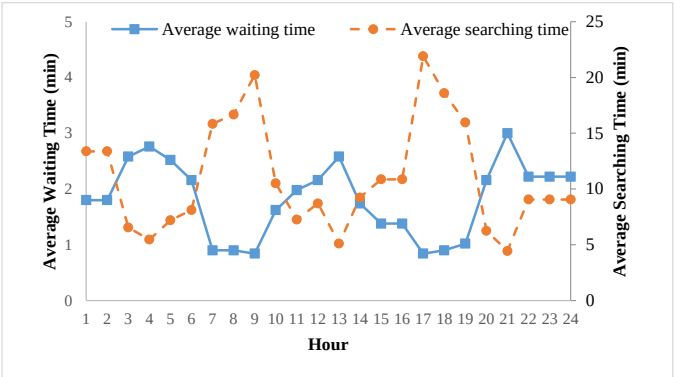


B

Figure 4-4. Equilibrium Work Hours and Hourly Revenue with Varying Surge Multipliers. A) Neoclassical, B) Income-Targeting



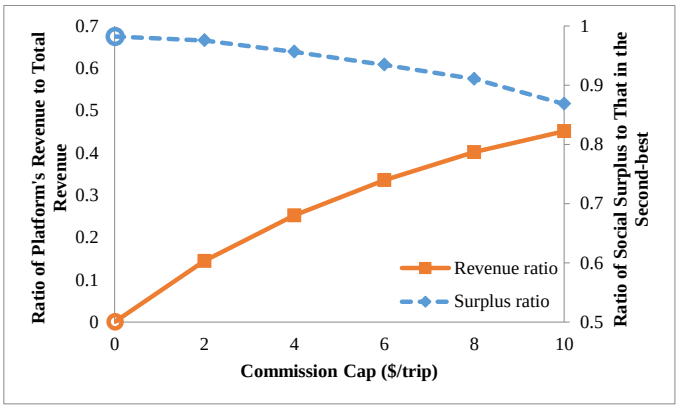
A



B

Figure 4-5. Market Outcomes under Dynamic Pricing. A) Surge Multipliers, B) Market Frictions

Figure 4-6. Changes in Consumers' Surplus and Joint Revenue under Dynamic Pricing.



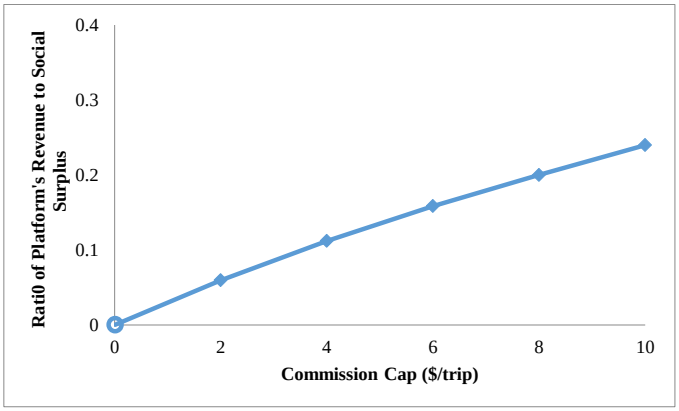


Figure 4-7. Market Outcomes under Different Commission Caps. A) Share of Platforms' Revenue (Left) and Level of Market Efficiency (Right), B) Ratio of Platform's Revenue to Social Surplus

Table 4-1. Equilibrium Flow Distributions under ME-N.

Start time	End time	Work hours (hr)	Rest hours (hr)	Revenue ($)	Cost ($)	Profit ($)	Path flow(veh)
Class 1							
15	1	9	1	317.7	210.3	107.4	1182
1	9	8	0	286.5	179.1	107.4	818
Class 2							
1	7	6	0	214.4	149.1	65.3	49
7	12	5	0	176.4	111.1	65.3	546
10	22	5	7	183.8	118.5	65.3	380
10	20	5	5	181.8	116.5	65.3	337
6	12	5	1	177.6	112.3	65.3	125
10	22	5	7	183.8	118.5	65.3	66
7	15	5	3	179.6	114.3	65.3	497
Class 3							
7	17	9	1	317.7	210.3	107.4	268
9	19	9	1	317.7	210.3	107.4	248
1	9	8	0	286.5	179.1	107.4	266
8	18	9	1	317.7	210.3	107.4	289
7	20	6	7	218.7	111.3	107.4	930
Class 4							
7	12	5	0	176.4	102.1	74.3	345
7	14	5	2	178.6	104.3	74.3	755
6	11	5	0	177.0	102.7	74.3	110
9	22	5	8	184.8	110.5	74.3	302
10	22	5	7	183.6	109.3	74.3	488

Table 4-2. Equilibrium Flow Distributions under ME-I.

Start time	End time	Work hours (hr)	Rest hours (hr)	Revenue ($)	Cost ($)	Utility ($)	Path flow(veh)
Class 1							
15	1	9	1	293.0	210.3	281.3	1294
1	9	8	0	267.0	179.1	281.3	706
Class 2							
8	20	5	7	174.3	118.3	350.9	1482
7	12	5	0	168.3	111.1	350.9	518
Class 3							
7	17	9	1	293.0	210.3	281.3	840
1	9	8	0	267.0	179.1	281.3	547
8	17	9	1	293.0	210.3	281.3	613
Class 4							
8	20	5	7	174.3	109.3	359.9	210
7	22	5	10	177.0	112.5	359.9	381
6	22	5	11	178.1	113.9	359.9	41
7	15	5	3	171.1	105.5	359.9	488
7	22	5	10	177.0	112.5	359.9	631
7	22	5	10	177.0	112.5	359.9	249

CHAPTER 5
MODELLING SPATIAL EFFECTS OF DYNAMIC PRICING IN RIDE-SOURCING MARKETS

In Chapter 4, we capture drivers' work scheduling decisions under dynamic pricing. In addition, drivers are free to choose where to search for customers upon completing a trip. Given the ever-changing market conditions, it is often the case that ride-sourcing drivers may cruise to/near geographic areas which they deem the most profitable. In this sense, dynamic pricing is also likely to stimulate the redistribution of vehicles among locations. In this chapter, we focus on the induced spatial variations in market frictions, and further the changes in agents' welfare in different geographic areas under dynamic pricing.

This chapter is organized as follows. In Section 5.1, we first propose a novel spatial model that features the equilibration of demand and supply, while explicitly capturing the matching technology from a geometric probability perspective. The description of an empirical data set and the algorithms for parameter estimation are given in Section 5.2. Section 5.3 demonstrates the equilibrium outcomes of the proposed model. Subsequently, in sections 5.4 and 5.5, we introduce and numerically explore the framework for studying dynamic pricing and its regulation. Summary of this chapter is given in Section 5.6.

5.1 Spatial Equilibrium Model

This section presents our model of the ride-sourcing market. For simplicity, we only consider a single ride-sourcing platform, and each vehicle can only pick up one customer during a trip. We first model the platform's matching of the customer-vehicle pairs from a geometrical probability perspective, based on which the waiting and searching frictions are endogenously determined. Then, we specify customers' demand and the distribution of vacant vehicles across a

spatial market. The mathematical formulation of market equilibrium and the solution algorithms are discussed subsequently.

5.1.1 Geometrical Matching

Below we describe a matching scheme in the spatial market. The framework is general enough to depict the matching technologies of both ride-sourcing and street-hailing services and to differentiate the features in their returns to scale property. We consider a urban region divided into different geographic zones indexed by $i \in I$ with the area of A_i. The market conditions in each zone are assumed homogeneous. We discretize the total study horizon into periods of equal length, each indexed by $t \in T$. Without loss of generality, we assume the length of one study period is one hour.

5.1.1.1 Technology

The matchings between customers and drivers are completed automatically by the ride-sourcing platform. The algorithm can be simplified as matching a requesting customer to her closest vehicle within a coverage radius r (Ranney, 2015). To delineate such a process, we discretize one analysis period into smaller time steps of equal length δ. For each time step, we consider the following procedure:

- At the beginning of each time step, unmatched and newly arrived customers send their requests to the platform and remain stationary at their current locations.
- The platform randomly loops through all requesting customers. For each customer, it checks whether a vacant vehicle is available within distance *r* from the customer. If so, the platform assigns the closest vehicle to her; otherwise, it does nothing.
- The matched pairs disappear from the market while those who remain unmatched wait until the next step.

We assume the area of each zone is sufficiently large compared to the matching radius (one can always merge areas to satisfy this requirement if necessary), and thus both the customer and driver in a successful matching come from the same zone. We assume the arrivals of

customers and vehicles follow the spatial Poisson Point Process (SPPP) (Chiu et al., 2013) and the market is stable. Therefore, the dynamics in each matching step have the same mean statistics through one study period. In contrast to Xu et al. (2017), our approach is customer oriented and relies on an exogenous matching radius that limits the matching within a study zone.

We define customers' average waiting time w_{it}^{c} as the expected time from the moment a customer requests a ride until she is picked up. It can be decomposed into two parts:

$$w_{it}^{c} = w_{it}^{cm} + m_{it}, \ \forall i \in I, t \in T \tag{5-1}$$

where w_{it}^{cm} is the matching time, more specifically, the expected time for the requesting customer to be matched; m_{it} represents the meeting time, i.e., the expected travel time for the driver to pick up the requesting customer after a successful matching. By symmetry, drivers' average searching time until picking up a customer is:

$$w_{it}^{v} = w_{it}^{vm} + m_{it}, \ \forall i \in I, t \in T \tag{5-2}$$

where w_{it}^{vm} is the expected time for a vacant vehicle to be matched. A vehicle is considered to be either occupied or vacant. A vacant vehicle is either waiting to be matched or en route to pick up a customer. Therefore, w_{it}^{vm} consists of the contribution of the vehicles that have searched in zone *i* during period *t*, regardless whether they are successfully matched.

The specific forms of the waiting and searching frictions are induced by the proposed matching procedure. Let N_{it}^{vm} and N_{it}^{cm} denote the intensities of unmatched vehicles and customers. Given the assumptions above, the probability of the existence of *n* unmatched vehicles in a circle of radius *r* centered at a requesting customer is:

$$P\{n\} = \frac{e^{-\left(\pi r^2 N_{it}^{vm}\right)} \left(\pi r^2 N_{it}^{vm}\right)^n}{n!}, \ n = 0,1,2,..., \ \forall i \in I, t \in T \tag{5-3}$$

We denote by x the distance of the closest unmatched vehicle with a cumulative distribution function $G(\cdot)$ and density function $g(\cdot)$:

$$G(x) = 1 - P\{0\} = 1 - e^{-\pi x^2 N_{it}^{vm}},\ x \geq 0,\ \forall i \in I,\ t \in T \tag{5-4}$$

$$g(x) = 2\pi x N_{it}^{vm} e^{-\pi x^2 N_{it}^{vm}},\ x \geq 0,\ \forall i \in I,\ t \in T \tag{5-5}$$

We consider a thinning process to "label" the potential drivers that could be matched to a waiting customer. The average number of potential drivers is approximately $N_{it}^{vm} / N_{it}^{cm}$ given the sufficiently large matching radius. The probability of a customer not being matched in a time step equals that of zero potential drivers being "labeled" to her, i.e., $1 - \exp\left(-N_{it}^{vm} / N_{it}^{cm}\right)$. Her expected matching time can be roughly estimated as δ times the number of steps she is going to wait:

$$w_{it}^{cm} = \frac{\delta}{1 - \exp\left(-N_{it}^{vm} / N_{it}^{cm}\right)},\ \forall i \in I,\ t \in T \tag{5-6}$$

Note that $w_{it}^{cm} > \delta$ by construction and δ is a modeling parameter. We interpret it as the system response time to be estimated in Section 5.2.2. Conditional on a successful matching, m_{it} is a function of N_{it}^{vm}, denoted as $M\left(N_{it}^{vm}\right)$:

$$\begin{aligned} m_{it} &= \frac{k}{v} \int_0^r x g\left(x \mid x \leq r\right) dx \\ &= \left[-\frac{kr}{v} e^{-\pi r^2 N_{it}^{vm}} + \frac{k}{2v\sqrt{N_{it}^{vm}}} erf\left(\sqrt{\pi N_{it}^{vm}}\, r\right) \right] \Big/ \left(1 - e^{-\pi r^2 N_{it}^{vm}}\right) \\ &= M\left(N_{it}^{vm}\right),\ \forall i \in I, t \in T \end{aligned} \tag{5-7}$$

where $k \approx \sqrt{\pi/2}$ is a scaling parameter that adjusts the Euclidean distance to the Manhattan distance (Arnott, 1996); v is the average speed of the vehicle and $erf(x) = 2/\sqrt{\pi} \int_0^x e^{-t^2} dt$ is the Gauss error function. When $\pi r^2 N_{it}^{vm}$ is sufficiently large, m_{it} degrades to $k / 2v\sqrt{N_{it}^{vm}}$, the

waiting time formula used by Daganzo (1978) and Arnott (1996) for the demand responsive transit and radio-dispatched taxi market, respectively.

The average matching time for vacant vehicles is still missing. By Little's law, we have the following relationship at the steady state $N_{it}^{vm} = Q_{it} w_{it}^{vm} / A_i$ and $N_{it}^{cm} = Q_{it} w_{it}^{cm} / A_i$. Following Eq. (5-6) we have:

$$w_{it}^{vm} = -w_{it}^{cm} \ln\left(1 - \delta / w_{it}^{cm}\right), \ \forall i \in I, \ t \in T \tag{5-8}$$

The above scheme also applies to traditional street hailing, although the matching radius r_s and the duration of the matching step δ_s may be technology-specific. It is reasonable to argue that r_s is much smaller than that of the ride-sourcing technology. Consequently, the average number of potential vehicles identified in the labeling process is approximately $\pi r_s^2 N_{it}^{vm}$. Given that the meeting time is often negligible, the average waiting time contains only the matching time:

$$w_{its}^{c} \approx w_{its}^{cm} = \frac{\delta_s}{1 - \exp(-\pi r_s^2 N_{it}^{vm})} \approx \frac{1}{\left(\pi r_s^2 / \delta_s\right) N_{it}^{vm}} \tag{5-9}$$

which is of the same form as provided by Douglas (1972).

5.1.1.2 Returns to scale

The waiting time formulae above demonstrate the trade-offs between components for both technologies. In what follows, we discuss the returns to scale properties assuming a constant operating speed of vehicles. The argument still holds as long as traffic congestion is not severe.

In the matching stage described by Eqs. (5-6) and (5-9), the ride-sourcing technology exhibits constant returns to scale while street hailing technology exhibits increasing returns to scale. Specifically, doubling N_{it}^{vm} and N_{it}^{cm} simultaneously leaves the matching time unchanged for the former but reducing the latter by half. Such a tremendous reduction in matching time for

street-hailing implies the information imperfection confined by the limited matching radius. As indicated by Eq. (5-7), however, the ride-sourcing technology enjoys a significant reduction in meeting time due to the economies of density. Despite the differences in specific components, both technologies show increasing returns to scale at the aggregate level, which is commonly seen in a spatial queuing system (Yang and Yang, 2011). One can further observe that street hailing technology generally has higher returns to scale than ride sourcing, an observation also verified by our extensive simulation in Chapter 6.

5.1.2 Customer Demand

We aggregate passengers' arrival from each origin and focus on the hourly specific demand. It is well-recognized that the ride-sourcing market is not only cleared by price, since each customer has to wait for some time before getting served (Douglas, 1972; Arnott, 1996; Taylor, 2016). We assume that travel demand is a function of the average trip fare $\bar{F}_{it}$, average waiting time w_{it}^{c}, and other control variables $\boldsymbol{X}$ in the following exponential form:

$$Q_{it} = \theta_{0t} e^{\theta_{1t}\bar{F}_{it} + \theta_{2t} w_{it}^{c} + \boldsymbol{\theta}_{\mathbf{3t}}\mathbf{X}}, \ \forall i \in I, t \in T \tag{5-10}$$

where $\bar{F}_{it} = F_0 + \gamma_{it} F_{it}$. F_0 is the flag-drop fee, and γ_{it} is the surge multiplier adjusted by the platform. The surge multiplier is the same for all customers in zone *i* at period *t*.

$\boldsymbol{\theta} = \{\theta_{0t}, \theta_{1t}, \theta_{2t}, \boldsymbol{\theta}_{\mathbf{3t}}\}$ is the vector of parameters to be estimated. It is anticipated that $\theta_{0t} > 0$, $\theta_{1t} < 0$ and $\theta_{2t} < 0$.The temporal or spatial correlations among the parameters are not considered.

5.1.3 Vehicle Supply and Spatial Distribution

When completing a trip, a driver obtains $(1-\eta)\times 100\%$ of the trip fare while the remaining serves as the commission to the ride-sourcing platform. We consider a fixed fleet size

N in the short run and each driver works for an average hour *h*. We assume *h* is given exogenously, although serval research in labor economics discusses drivers' incentives in supplying work hours with respect to wage variations. See Farber (2015) and Chen and Sheldon (2016) for more details.

To account for the temporal variations in vehicle supply, we impose the constraint that the percentage of on-line vehicles at period *t* is fixed:

$$n_t = N_t \left(N \times h \right)^{-1}, \ \forall t \in T \tag{5-11}$$

where N_t is the number of vehicles during period *t*. Given that the length of one period is one hour, N_t is also the total vehicle hours during the period. The justification for the assumption above is that drivers' structural decisions of work scheduling (e.g., starting and ending times) are likely subject to latent budget constraints that tend to be stable, at least in the short run. This results in similar day-to-day patterns (e.g., in weekdays) often observed in empirical data. Notice that $\sum_t n_t = 1$ and n_t can be easily estimated as the average of the observations. For the conservation of vehicle hours:

$$N \times h = \sum_t N_t = \sum_t \sum_i N_{it} \tag{5-12}$$

where N_{it} is the number of vehicles in zone *i* at period *t*. Given that a vehicle is either occupied or vacant, the following conservation equation holds at the steady state:

$$N_{it} = Q_{it} \left(w_{it}^{v} + l_{it} \right), \ \forall i \in I, \ t \in T \tag{5-13}$$

Note that when an occupied vehicle traverses several zones, it only contributes to its originated zone at each matching process along the journey.

We assume that ride-sourcing drivers search for customers in zones where they perceive to offer the highest payoff. The drivers are further assumed to have perfect information on the

profitability of each zone. Consequently, the distribution of the vacant vehicles among all zones at equilibrium is to make each zone equally profitable (Lagos, 2000):

$$\left(1-\eta\right)\bar{F}_{it}\left(w_{it}^{v}+l_{it}\right)^{-1}=\pi_{t},\ \forall i\in I,t\in T \tag{5-14}$$

where $\left(1-\eta\right)\bar{F}_{it}$ and $w_{it}^{v}+l_{it}$ can be respectively treated as the average reward for completing a trip and the service duration for completing a trip; π_{t} can be interpreted as the equilibrium hourly revenue.

The quantification above assumes drivers are well positioned so that each zone is equally attractive at equilibrium. Our approach does not consider the transition cost in drivers' zonal choice and is likely to overestimate the total vehicle supply. In contrast, Yang et al. (2010b) and Buchholz (2015) explicitly specify the transition cost in a logit-type zonal choice model; and they assume drivers are not able to provide rides when traversing to their desired destination zones. Such an assumption may be reasonable for traditional street hailing technology. Ride-sourcing drivers, however, can be matched to customers as long as their apps are on. A direct application of their approach will underestimate the total vehicle supply. Given that the current dataset does not contain the information on zonal adjacency or drivers' log data, our approach seems to be the simpler and feasible one.

5.1.4 Market Equilibrium

For a ride-souring market where trip fares and fleet size are given, market equilibrium is defined by the following set of nonlinear equations:

$$w_{it}^{v}=w_{it}^{vm}+m_{it},\ \forall i\in I,t\in T \tag{5-15}$$

$$w_{it}^{c}=w_{it}^{cm}+m_{it},\ \forall i\in I,t\in T \tag{5-16}$$

$$w_{it}^{vm}=-w_{it}^{cm}\ln\left(1-\delta/w_{it}^{cm}\right),\ \forall i\in I,\ t\in T \tag{5-17}$$

$$m_{it} = M\left(Q_{it} w_{it}^{vm}\right),\ \forall i \in I,\ t \in T \tag{5-18}$$

$$Q_{it} = \theta_{0t} e^{\theta_{1t}\bar{F}_{it} + \theta_{2t} w_{it}^{c} + \boldsymbol{\theta}_{3t}\mathbf{X}},\ \forall i \in I, t \in T \tag{5-19}$$

$$\sum_i N_{it} = n_t\left(N \times h\right),\ \forall t \in T \tag{5-20}$$

$$N_{it} = Q_{it}\left(w_{it}^{v} + l_{it}\right),\ \forall i \in I,\ t \in T \tag{5-21}$$

$$\left(1-\eta\right)\bar{F}_{it}\left(w_{it}^{v} + l_{it}\right)^{-1} = \pi_t,\ \forall i \in I, t \in T \tag{5-22}$$

The nonlinear system has a total $7|I||T|+|T|$ unknowns and equations. Eq. (5-15) and Eq. (5-22) may pose challenges to the existence of a solution. The former requires w_{it}^{v} to be bounded below by w_{it}^{vm}, while the latter may force w_{it}^{v} to be extremely small (even less than w_{it}^{vm}) in zones with short trip durations. The following proposition guarantees the existence of equilibrium under sufficient labor supply. The detailed proof is provided in the Appendix G.

Proposition 1. There exists a constant $k > 0$, such that when $N \cdot h > k$, the above nonlinear system admits at least one solution.

We adopt a partially augmented Lagrangian method to solve for the equilibrium solution (Bertsekas, 1999). We denote $\mathbf{z} \in R^{7|\mathrm{I}||\mathrm{T}|+|\mathrm{T}|}$ the vector of variables, $\boldsymbol{\lambda} \in R^{|\mathrm{I}||\mathrm{T}|}$ the vector of multipliers, the functional mapping $\mathbf{f}\left(\cdot\right): R^{7|\mathrm{I}||\mathrm{T}|+|\mathrm{T}|} \to R^{|\mathrm{I}||\mathrm{T}|}$ which corresponds to Eq. (5-15) and $\mathbf{G}\left(\cdot\right): R^{7|\mathrm{I}||\mathrm{T}|+|\mathrm{T}|} \to R^{6|\mathrm{I}||\mathrm{T}|+|\mathrm{T}|}$ which represents the remaining constraints. The method solves a sequence of problems of the following form:

(RP)

$$\min_{\mathbf{z}^k, \boldsymbol{\lambda}^k} = L^k\left(\mathbf{z}^k, \boldsymbol{\lambda}^k\right) = \mathbf{f}^T\left(\mathbf{z}^k\right)\boldsymbol{\lambda}^k + \frac{c^k}{2}\left\|\mathbf{f}\left(\mathbf{z}^k\right)\right\|^2 \tag{5-23}$$

s.t.

$$\mathbf{G}\left(\mathbf{z}^k\right) = \mathbf{0} \tag{5-24}$$

where $c^k > 0$ is the penalty parameter updated via $c^{k+1} = \rho c^k$; $\boldsymbol{\lambda}^k$ is updated via $\boldsymbol{\lambda}^{k+1} = \boldsymbol{\lambda}^k + c^k \mathbf{f}(\mathbf{z})$. The choice of ρ requires some trial and error (Bertsekas, 1999). Under some regularity assumptions (Bertsekas, 1999), $\mathbf{z}^k, \boldsymbol{\lambda}^k$ will converge respectively to the solution of the original nonlinear system and the Lagrangian multipliers associated with Constraint (5-15). The augmented Lagrangian approach is quite stable and usually converges to at least a local minimum (Bertsekas, 1999). We terminate the iterative process when $\left\|\mathbf{f}\left(\mathbf{z}^k\right)\right\|$ is less than a tolerance value ε.

5.2 Model Estimation

We first introduce a dataset from Didi Chuxing, based on which we describe the procedure that jointly estimates the system response time and customers' average waiting time. Finally, we estimate the proposed log-linear demand function as well as the mean statistics in labor supply.

5.2.1 Dataset

The original dataset contains 8,540,614 order requests from Jan. 1st -21st , 2016 in a Chinese city and is publicly available by Didi Chuxing (2016a). Didi anonymously divides the study region into 66 zones and we follow the default indexing of zones. Each request records the customer's ID, driver's ID, trip fare, the timestamp when the request occurs as well as the IDs of the origin and destination. In addition, there are supplementary datasets that contain other information such as weather, PM2.5 indices and the number of community facilities in a zone. After removing the observations with missing values and duplicate IDs, the master dataset retains 7,816,328 request records. Overall, approximately 83% of the time a customer succeeds in being matched to a driver in one request.

Despite the high-resolution order data, some key information is missing which includes the adjacency of the study zones, trip duration (distance) and customers' average waiting time. Intentionally, the proposed framework does not rely on zonal adjacency. The raw trip duration is calculated using the fare structure of 3.86 Yuan/mile+0.35 Yuan/min (Didi Chuxing, 2016b). Considering potential price surging, we approximate the trip duration between each OD pair based on the median of the corresponding raw trip durations. Zone 39 is identified as the airport area with an average trip duration of 1.03 hours. We filter all trips originating from Zone 39 since the matching mechanism may be different in such a "spot" market.

The master dataset is used to estimate system response time and average waiting time in Section 5.2.2. Two subsets are constructed from the master dataset. One summarizes observations of hourly requests in the daytime (6:00 a.m.-17:00 p.m.) for all workdays. It is used for demand estimation in Section 5.2.3. The other summarizes the work hour information of drivers under the same restrictions, from which we obtain the mean statistics presented in Section 5.2.4.

5.2.2 Estimation of System Response Time and Average Waiting Time

Note that the estimation of δ is non-trivial. It is often observed that customers whose ride requests failed may wait for several minutes before requesting again. We account for these lagged times in determining δ. We use the hat notation to denote the estimate of a parameter. In the estimation process, $\hat{\delta}$ is first numerated from an ordered candidate set Δ. Conditional on the estimated $\hat{\delta}$, key variables that include $\hat{w}_{it}^{cm}, \hat{m}_{it}, \hat{w}_{it}^{c}$ are then estimated mainly based on the probability of successful matching. We settle with the $\hat{\delta}$ that yields the minimum error. To elaborate the process, let's define a counter *n* so that $\hat{\delta}_n$ refers to the nth element from Δ and *n* is initialized to be 1. The process is described as follows:

Step 1. Choose $\hat{\delta}_n \in \Delta$. Compute the benchmark average matching time w_{it}^{cm} as follows: when a customer is successfully matched to a driver in her first trial, the customer incurs a matching time of $\hat{\delta}_n$; otherwise the matching time is obtained as the elapsed time from her first trial until the latest trial of being matched.

Step 2. Compute the proportion of requests that is successfully matched, $\hat{P}r_{it} = Q_{it}/q_{it}$ where q_{it} is the rate of requests from zone *i* during period *t*. Then, we have

$$\hat{N}_{it}^{cm} = \begin{cases} Q_{it}\delta \Big/ \left(\hat{\mathrm{Pr}}_{it}\ A_i\right), \hat{\mathrm{Pr}}_{it} > 0 \\ q_{it}/A_i, \hat{\mathrm{Pr}}_{it} = 0 \end{cases}, \ \forall i \in I,\ t \in T \tag{5-25}$$

Step 3. Estimate the intensity of the unmatched vehicles given the fact that $\hat{\mathrm{Pr}}_{it} = 1 - \exp\left(-\hat{N}_{it}^{vm} \big/ \hat{N}_{it}^{cm}\right)$:

$$\hat{N}_{it}^{vm} = \begin{cases} -\hat{N}_{it}^{cm} \ln\left(1 - \hat{\mathrm{Pr}}_{it}\right), \hat{\mathrm{Pr}}_{it} < 1 \\ \psi \hat{N}_{it}^{cm}, \hat{\mathrm{Pr}}_{it} = 1 \end{cases}, \ \forall i \in I,\ t \in T \tag{5-26}$$

where ψ should be significantly larger than 1 to represent sufficient vehicle supply.

Step 4. Calculate $\hat{w}_{it}^{c}, \hat{w}_{it}^{cm}, \hat{m}_{it}$ according to Eqs. (5-1), (5-6), (5-7), when $\hat{\mathrm{Pr}}_{it} > 0$ (with some abuse of notations); otherwise set them to NAs.

Step 5. Compute the augmented percentage error:

$$\varepsilon = \frac{\left|\sum_i \sum_t Q_{it}\left(\hat{w}_{it}^{cm} - \hat{\delta}_n\right) - \sum_i \sum_t Q_{it}\left(w_{it}^{cm} - \hat{\delta}_n\right)\right|}{\sum_i \sum_t Q_{it}\left(w_{it}^{cm} - \hat{\delta}_n\right)} \tag{5-27}$$

and set $n = n+1$; go to Step 1.

In Step 5, the calculation of the estimation error excludes δ from each successful matching and thus focuses exclusively on the scenario where customers' first request is not filled, hence the term "augmented."

We let $\Delta = \{10, 11..., 20\}$ seconds since the system response time is typically short based on the authors' user experience. We choose $\psi = 10$ and a matching radius $r = 1$ mile a priori, which may be subject to certain biases. A rigorous estimation of these two parameters is left for future study, particularly when the geographic information of study zones and the traces of drivers are available.

The estimated outcomes are given in Figure 5-1. Figure 5-1 (a) shows the curve for ε when δ is increased from 10 seconds to 20 seconds. The optimal δ of 13 seconds corresponds to a minimum error rate of 0.28. The rapid change of ε is due to its "augmenting" nature. The distributions of average waiting time and matching time are given in Figure 5-1 (b) and (c). The average waiting time ranges from 0.33 min to 7.52 min with a mean of 3.61 min while the average matching time varies between 0.21 min to 2.93 min with a mean of 0.38 min.

5.2.3 Demand Estimation

Consider the log-transformed regression formula from Eq. (5-10) with the Gaussian error ε_{it}:

$$\log Q_{it} = \log \theta_{0t} + \theta_{1t} \bar{F}_{it} + \theta_{2t} w_{it}^{c} + \theta_{3t} X + \varepsilon_{it}, \ \forall i \in I, t \in T \tag{5-28}$$

where we choose X to be the number of community facilities in the zone to capture a certain level of spatial heterogeneity (Didi Chuxing, 2016a). The estimation of demand may be subject to endogeneity problems particularly with the existence of dynamic pricing. Specifically, both trip fare and average waiting time are likely correlated to the error term due to the factors not explicitly specified. In fact, obtaining the "orthogonal effect" of regressors on demand is one of the most challenging tasks in demand estimation.

To partially alleviate the endogeneity problem, we implement a two-stage least squares (2SLS) approach (Hansen, 2017). We instrument $\bar{F}_{it}$ by the fare in the closest zone to i, $\bar{F}_{ivt}$ in

the first stage. Note that $\bar{F}_{ivt}$ is positively related to $\bar{F}_{it}$ but is less likely to be related to ε_{it}. In the second stage, we replace $\bar{F}_{it}$ with the exogenous components identified in the first stage for demand estimation. Due to the limited data, we do not model the probable variation of θ_{1t} with trip durations which may require an origin-destination demand specification or a segmentation of trip durations (Buchholz, 2015).

The 2SLS estimation is listed in Model III of Table 5-1. The estimated results for the intercept terms and variable *X* are omitted given the limited space. We also report the ordinary least square (OLS) results via gradually modifying the regressors: without instrumenting trip fares (Model II); further removing the average waiting time (Model I). The estimated parameters are all negative and significantly different from zero at the 0.05 significance level. The differences in the magnitude of parameters associated with trip fares (θ_{1t}) between Model I and II imply a certain level of correlation between $\bar{F}_{it}$ and w_{it}^{c}. The difference in θ_{1t} between Model II and III is possibly due to the endogeneity bias. It is also interesting to note that the price elasticity of demand varies by time of day and is relatively smaller in early morning (e.g., period from 6:00-7:00). We utilize the results from the 2SLS estimation for the remaining analysis.

5.2.4 Labor Supply

We estimate drivers' average working hour *h*, total fleet size *N*, as well as the percentage of on-line drivers in each time period n_t using their mean values.

As drivers' break time may bias the estimation of working hours, the first step in estimating *h* is to define a working session (Chen and Sheldon, 2016). We formally define a working session for a driver as a collection of trips with the inter-trip time no longer than $\left(1+l_{jt}\right)$ hours, assuming the driver's most recent trip occurs in zone *j* during period *t*. Working

hour is then calculated as the length of the working session[7]. Other subroutines are also developed to handle extreme cases (e.g., when a driver only takes one ride in one day). Total fleet size N represents the number of drivers who provide services during the study period and is simply computed as the average of the chosen weekdays. The display of n_t across the days is given in Figure 5-2. The percentage of on-line drivers at each hour is quite stable, which supports the assumptions in Eq. (5-11) of Section 5.1.3. The corresponding mean values are summarized in Table 5-2.

5.3 Equilibrium Results

We demonstrate the equilibrium outcomes when the surge multipliers are set to one (so that total trip fares are given). To facilitate the presentation, we choose Zone 2 and Zone 23 as two representatives with an average trip duration of 4.1 min and 12.2 min, respectively. We also evaluate the changes of equilibrium outcomes when reducing the fleet size from 31,097 to 25,097. Other parameters remain the same as estimated from Section 5.2. The fare structure of 3.86 Yuan /mile +0.35 Yuan/min suggests an hourly rate of 81 Yuan assuming a constant speed of 15.6 mile/hr (Didi Chuxing, 2016c). At each iteration, RP is solved in GAMS with the nonlinear solver CONOPT (Drud, 1994). For the augmented Lagrangian scheme, we update the penalty with $\rho = 5$ and set the tolerance rate to 0.01. The method often converges within 10 iterations.

We outline the components in searching and waiting frictions in Figure 5-3 with different combinations of zone ID and fleet size. Combinations I and II correspond respectively to the case with the fleet size of 31,097 and 25,097 in Zone 2; combinations III and IV correspond respectively to the case with the fleet size of 31,097 and 25,097 in Zone 23. The average waiting

[7] The calculation is different from Chen and Sheldon (2016), as we do not have access to drivers' log data.

time and its matching time portion are presented in Figure 5-3 (a)-(d). Customers in Zone 2 incur a higher average waiting time than those in Zone 23, as more vehicles are attracted to the latter (for longer trip duration once occupied). For the same zone, the reduction of fleet size raises the average waiting time. The average searching time and its matching time portion are given in Figure 5-3 (e)-(h), which shifts in an opposite direction to their waiting time counterparts. As an example, it is easier for vehicles in Zone 2 to find a customer; and the reduction in fleet size further decreases the average searching time.

Under sufficient supply (e.g., N=31,097, h=2.7), the scenario that customers are not able to be matched in their first trial does not occur (Figure 5-3 (a) and (c)). But as we decrease the fleet size, it becomes harder for customers in zones with shorter trip durations to hail ride-sourcing vehicles (Figure 5-3(b)). On the other hand, the matching time portion accounts for a relatively larger share in vehicles' average searching time, particularly in zones with longer average trip duration (e.g., Zone 23).

The patterns of drivers' equilibrium hourly revenue rate are similar and are plotted in Figure 5-4 (a)-(b). Apart from the early morning period 6:00-7:00 a.m., the equilibrium hourly revenue seems stable, ranging from 21.9–25.4 Yuan/hr and 27.1–31.2 Yuan/hr respectively for the fleet size of 31,097 and 25,097. A reduction in fleet size corresponds to a higher wage rate as drivers on average enjoy more trip opportunities.

5.4 Dynamic Pricing

We present a framework for studying dynamic pricing in this section and then evaluate its impacts on the market dynamics and the welfare of involved market players.

5.4.1 Formulation

We consider the market outcomes under dynamic pricing in the short run where the total fleet size is fixed. It is assumed that the platform adjusts period-specific surge multipliers $\boldsymbol{\gamma}$ for revenue maximizing, subject to the equilibration of demand and supply:

(DP)

$$\max_{\gamma,\mathbf{z}} = \sum_i \sum_t \eta\left(F_0 + \gamma_t F_{it}\right) Q_{it} \tag{5-29}$$

s.t.

$$\mathbf{T}\left(\mathbf{z} \mid \boldsymbol{\gamma}\right) = \mathbf{0} \tag{5-30}$$

where $\mathbf{z} \in R^{7|\mathrm{I}||\mathrm{T}|+|\mathrm{T}|}$ is the vector of variables and $\mathbf{T}\left(\mathbf{z} \mid \boldsymbol{\gamma}\right) = \mathbf{0}$ represents the nonlinear equation system characterized by Eqs. (5-15)-(5-22). For simplicity, we do not consider the platform's operation cost. Note that DP may be best seen as an approximation of the pricing practice for recurrent market patterns as it assumes that the platform has perfect anticipation. Other possible pricing heuristics that are passively trigged by random market events are beyond the scope of DP.

5.4.2 Numerical Experiments

To evaluate the performance of dynamic pricing, we numerically compare the equilibrium outcomes under dynamic pricing with those of optimal static pricing. The formulation of the latter is a trivial adaptation from DP. The proposed formulations can be reformulated using the augmented Lagrangian scheme and solved in a similar fashion as in Section 5.1.4. A fixed fleet size of 31,097 is assumed.

The surge multipliers are uniformly set at 1.21 in the optimal static pricing. Figure 5-5 (a)-(b) display the period-specific surge multipliers in two selected zones under dynamic pricing. The patterns of these two share great similarity as the same sensitivity parameters (i.e., θ_{1t}, θ_{2t}) are used in the demand function. However, the platform has to increase the surge multipliers

more aggressively in Zone 2 to attract enough drivers. As long as the waiting cost is not severe, the objective function of DP suggests that the platform tries to set the total trip fare negative to the inverse of the price sensitivity parameter (i.e., $\bar{F}_{it} \sim -1/\theta_{it}$). Zones with shorter average trip durations achieve this optimality condition at higher surge multipliers. Although θ_{it} is fixed among trips from zone *i* at period *t*, the implication still holds if customers' sensitivity to price won't vary too much with trip durations (Buchholz, 2015). As will be seen, this observation has fundamental effects on the variations of market frictions and changes in the welfare of the market players.

We measure the spatial variations in waiting and searching frictions using the coefficient of variation (CV), which is computed respectively as σ_t^v / w_t^v and σ_t^c / w_t^c . w_t^c and w_t^v are the average waiting and searching times across all study zones at period *t*; σ_t^v and σ_t^c are the corresponding standard deviations.

The temporal displays of the CVs are given in Figure 5-6 (a)-(b). Compared to static pricing, dynamic pricing induces much smaller CV in average searching time but slightly larger CV in average waiting time. Eq.(5-14) implies that drivers' average searching time in a zone is positively related to the average trip duration when there is no price surge. Under dynamic pricing, surge multipliers change to the opposite direction of the average trip duration as is seen in Figure 5-5. This in return reduces the spatial variation of the average searching time.

The spatial variations in average waiting time are affected by both customers' demand and trip durations. A few notations are introduced for better explanation. We denote the possible combinations of pricing types and trip durations as $\{S, D\} \times \{s, l\}$. *S* and *D* are short for static and dynamic (pricing) while *s* and *l* are short for short and long (trip durations). We retain the

basic notation w^c for average waiting time with elements from $\{S, D\} \times \{s, l\}$ as subscripts for distinction. For example, w^c_{Sl} represents the average waiting time for a zone characterized by short trip duration under static pricing. A qualitative presentation of customers' average waiting time is demonstrated in Figure 5-6 (c). Dynamic pricing still gives the average waiting time with a smaller spread at a given demand level. However, our data simply has fewer zones characterized by the combinations of high demand with long trip duration or low demand with short trip duration (highlighted in the shaded areas). Therefore, we tend to see w^c_{Dl} and w^c_{Ds} at the extremes, which explains the slightly larger variation under dynamic pricing.

For the analysis of the trade-offs for each market player, we formally define consumers' surplus to represent the welfare of passengers from taking ride-sourcing services (see Cairns and Liston-Heyes (1996) on the computation of consumers' surplus when the average waiting time enters demand function); and the (quasi-) producers' surplus as the joint revenue between the platform and the drivers. Total social surplus is the sum of consumers' surplus and producers' surplus. We define the welfare change as the metric of interest under dynamic pricing minus that under static pricing.

Figure 5-7 and Figure 5-8 present the welfare changes both temporally (at two selected zones) and spatially. With some abuse of unit, the average adjusted surge multipliers are deployed in Figure 5-8. Three observations are worth highlighting. First, there exists temporal and spatial heterogeneity in the change of consumers' and producers' surplus. Next, producers are in general better off while customers are worse off in highly surged periods. For example, the changes in consumers' surplus are negative during 11:00-16:00 in Zone 23 and throughout the whole study period in Zone 2. Customers in Zone 2 suffer more from dynamic pricing than those from Zone 23, a source of inefficiency due to spatial heterogeneity. Finally, dynamic pricing

may have the potential to simultaneously increase consumers' and producers' surplus. A certain fraction of zones demonstrate this win-win situation when price is dynamically adjusted below its static counterpart. Note that these zones are all characterized by relatively longer trip durations, which corresponds again to the observation in Figure 5-5. In reality, trip durations vary even from the same zone. The current practice of setting a zonal-specific surge multiplier may hurt those customers with longer trip durations and subsequently the revenue as well.

As we pointed out previously in Chapter 4, the net outcome of the welfare changes depends on demand patterns and other factors. In the current setting, static pricing yields almost the same social surplus as the dynamic one: a rough estimate of 1.098×10^{7} Yuan/weekday and 1.089×10^{7} Yuan/weekday, respectively. In general, the comparisons on dynamic pricing and static pricing are mixed, particularly for customers. In addition, equilibrium outcomes under both strategies are likely to deviate from the socially optimal levels.

5.5 Commission Cap Regulation

We evaluate the performance of the proposed commission cap regulation based on the empirical data. Under the proposed regulation, a revenue-maximizing platform has the incentive to maximize transactions (realized demand) which are positively related to consumers' surplus. The additional revenue from price surging completely goes to the drivers. In Chapter 3, we showed that the commission cap regulation could achieve the second best under ideal homogeneity assumptions. In the current spatial market, a strict second best is barely achievable. However, the equilibrium market outcomes may still be superior to those from an unregulated monopoly, a statement yet to be verified.

5.5.1 Formulation

Given that no congestion is considered and trip duration varies, we implement the proposed regulation as the time-based charge. As the platform may lower trip fare to raise demand, we shall model drivers' entry decision when evaluating the performance of the regulation. Mathematically, market outcomes are given by solving the following problem:

(DP-R)

$$\max_{\boldsymbol{\gamma},\mathbf{z},\hat{\boldsymbol{\eta}},N} = \sum_i \sum_t \hat{\eta}_{it} l_{it} Q_{it} \tag{5-31}$$

s.t.

$$\mathbf{T}\left(\mathbf{z} \mid \boldsymbol{\gamma}, \hat{\boldsymbol{\eta}}, N\right) = \mathbf{0} \tag{5-32}$$

$$\hat{\boldsymbol{\eta}} \le \bar{\eta} \tag{5-33}$$

$$R \ge \pi_R \tag{5-34}$$

where N is the fleet size that is endogenously determined; $R = \left(\sum_t \sum_i \pi_t Q_{it} \Big/ \sum_t \sum_i Q_{it} \right) h$ is the average revenue obtained in a working section; $\hat{\boldsymbol{\eta}}$ is the vector of commission rate per unit time; $\bar{\eta}$ and π_R are the cap of commission rate per unit time and driver's reservation income level per working section, respectively; l_{it} is the average trip duration from zone i during period t. Note that Constraint (5-34) implicitly assumes that vehicle supply is perfectly elastic in the long run.

5.5.2 Numerical Experiments

The equilibrium outcomes under different commission caps are presented in Figure 5-9. We assume the annual reservation income is 52,935 Yuan (Didi Chuxing, 2016c), which makes the hourly reservation income around 25 Yuan. Note that when the commission cap goes to infinity, the platform becomes an unregulated monopoly with a commission rate of 97 Yuan/hour; the commission rate will be bounded at this value for any larger commission cap. For a

commission rate of less than 97 Yuan/hour, the increase in commission cap corresponds to an increase in a platform's revenue split, as illustrated in Figure 5-9 (a). Figure 5-9 (b) gives the ratio of the total social surplus to that under a (quasi-) second-best case (i.e., the platform maximizes the social welfare subject to the same constraints as in DP-R). The socially optimal commission cap is very close to zero given that the operation cost for the platform is not specified. Notice that the social surplus by the unregulated monopoly is only approximately 62% of that under the second best; and the proposed regulation improves market efficiency for a wide range of commission caps. The platform is able to exploit the drivers with high commission caps, which, as shows in Figure 5-9 (c), results in a reduction of the fleet size in the long term.

In general, the equilibrium outcomes under the proposed regulation are consistent with what we've obtained in Chapter 4. However, the proper choice of the commission cap for practical implementation should fully consider the cost components of the platform which is beyond the scope of this dissertation.

5.6 Summary

This chapter presents a spatial equilibrium model for studying dynamic pricing in the ride-souring market. We first present a deductive approach for depicting the technology adopted by a ride-sourcing platform for matching customers and drivers, and discuss its differences from traditional street hailing technology. We then specify both the demand and supply sides; and give the conditions for the existence of the equilibrium when price is fixed. Utilizing a public dataset from Didi Chuxing, we estimate the key modeling parameters and discuss the patterns for market frictions and equilibrium revenue levels.

We evaluate the effect of dynamic pricing assuming a revenue-maximizing platform. It is found that the platform will set surge multipliers opposite to the average trip durations of a study zone. For this reason, drivers under dynamic pricing will see less variation in their average

searching time when compared to the optimal static pricing. In the welfare analysis, both the platform and the drivers are better-off under dynamic pricing. However, the changes of consumers' surplus are negative during highly surged periods. Interestingly, certain zones demonstrate an increase in both consumers' and producers' surplus when price is dynamically adjusted below its static counterpart. In general, the comparisons on dynamic and static pricing are mixed.

We then evaluate the performance of the suggested commission cap regulation. It reaps the responsive feature of dynamic pricing while incentivizing the platform to maximize the total transactions for revenue maximization. Our experiments demonstrate that such a regulation policy generates higher social surplus for a wide range of commission caps compared to an unregulated monopoly. A good choice of commission cap balances the welfare between the platform and its users.

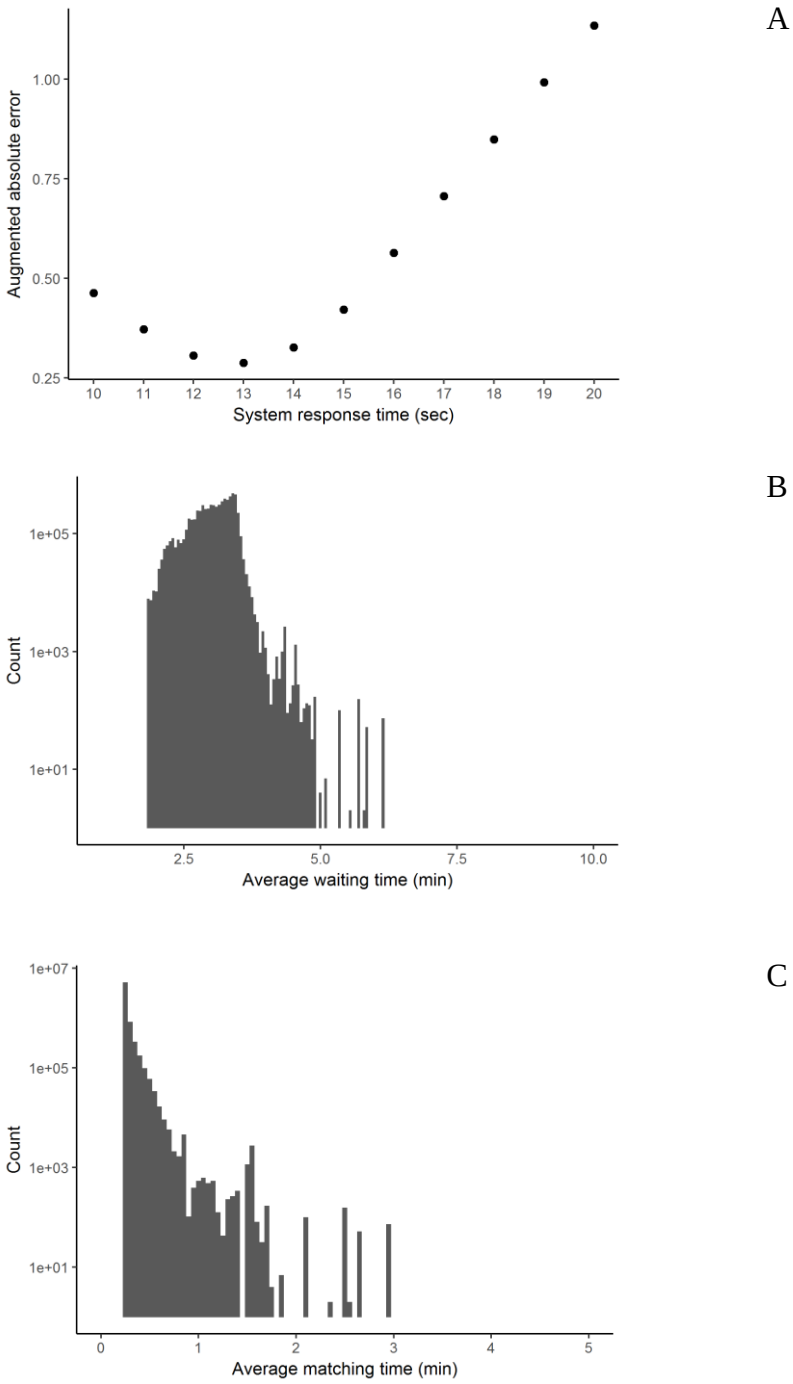


Figure 5-1. Estimation of The System Response Time, Customers' Average Waiting and Matching Time. A) Augmented Absolute Error Rate versus System Response Time, B) Distribution of Customers' Average Waiting Time, C) Distribution of Customers' Average Matching Time

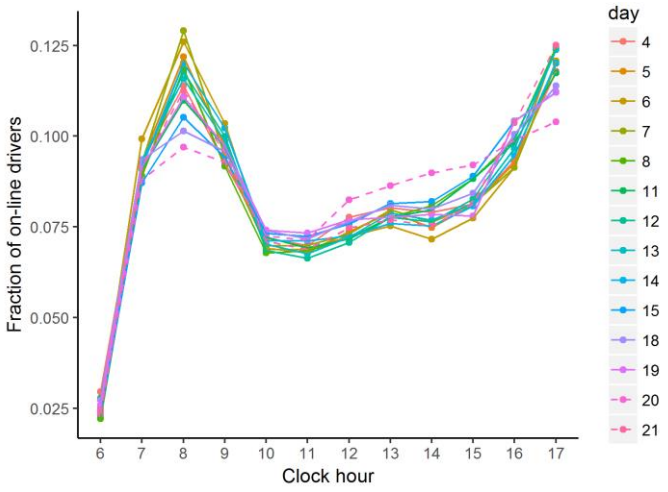
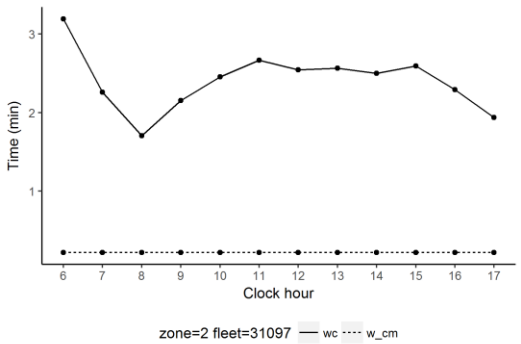
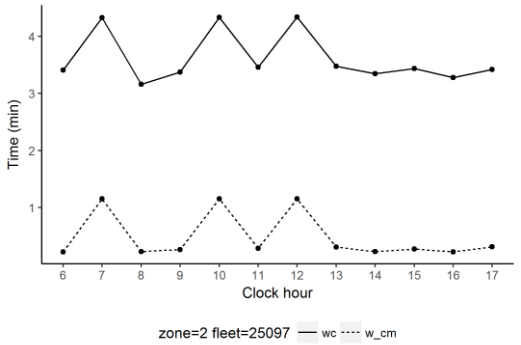


Figure 5-2. Display of the Fraction of The On-Line Drivers by Hour of Day.

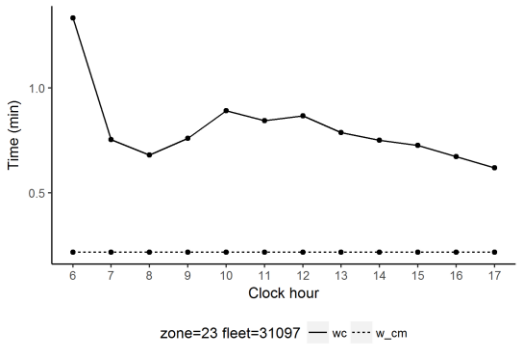
A


Time (min)
Clock hour
zone=2 fleet=31097
wc
w_cm

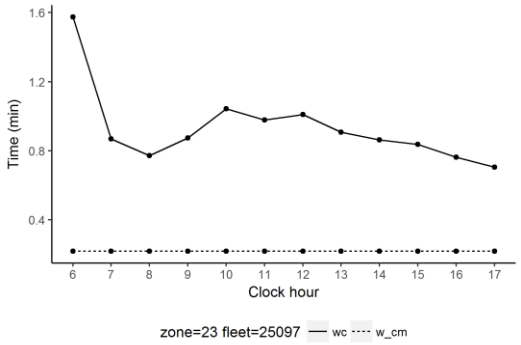
B

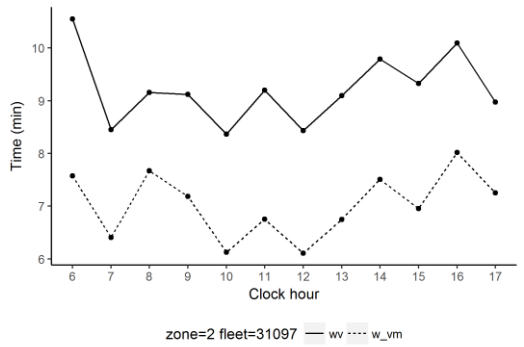

Time (min)
Clock hour
zone=2 fleet=25097
wc
w_cm

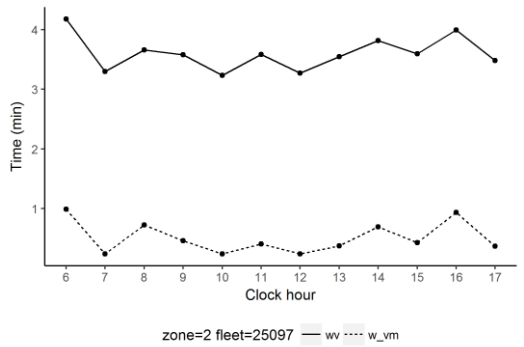
C

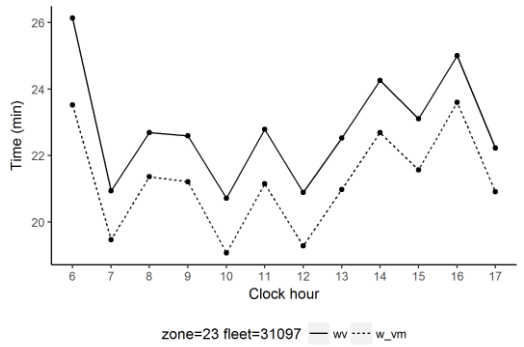

Time (min)
Clock hour
zone=23 fleet=31097
wc
w_cm

D


Time (min)
Clock hour
zone=23 fleet=25097
wc
w_cm







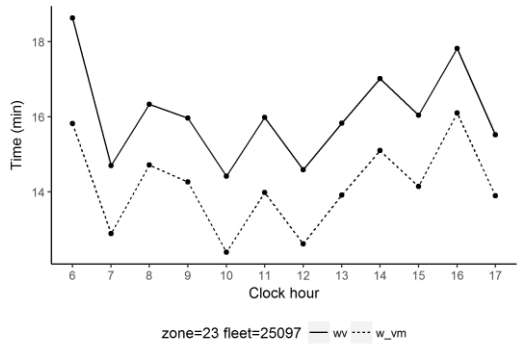
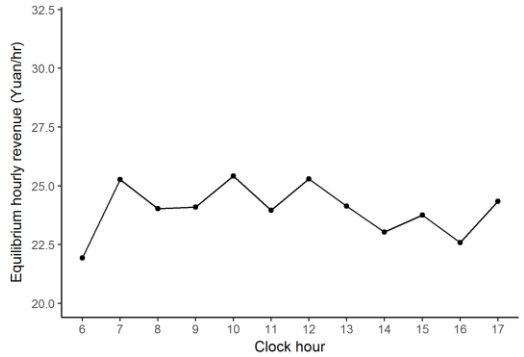


Figure 5-3. Market Frictions with Different Zone and Fleet Size Combinations. A) Waiting Time and The Matching Time Portion (I), B) Waiting Time and The Matching Time Portion (II), C) Waiting Time and The Matching Time Portion (III), D) Waiting Time and The Matching Time Portion (IV), E) Searching Time and The Matching Time Portion (I), F) Searching Time and The Matching Time Portion (II), G) Searching

Time and The Matching Time Portion (III), H) Searching Time and The Matching Time Portion (IV)



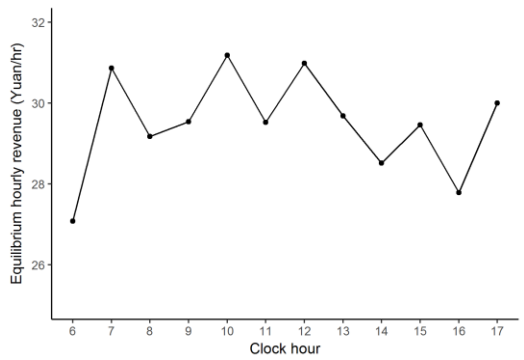
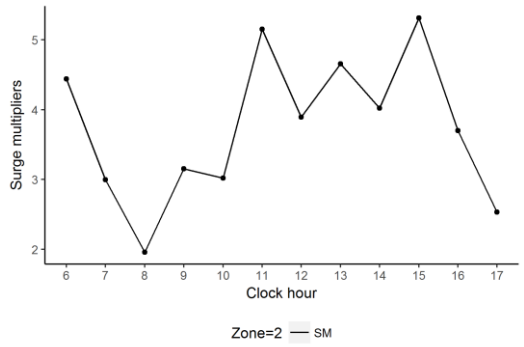


Figure 5-4. Display of Drivers' Equilibrium Revenue Rate under Varying Fleet Sizes. A) Equilibrium Revenue with Fleet = 31097, B) Equilibrium Revenue with Fleet = 25097



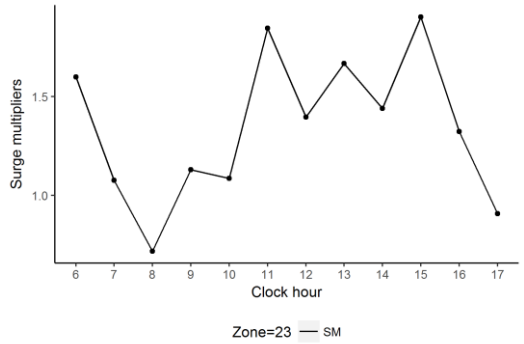


Figure 5-5. Display of Surge Multipliers under Dynamic Pricing. A) Surge Multipliers in Zone 2, B) Surge Multipliers in Zone 23

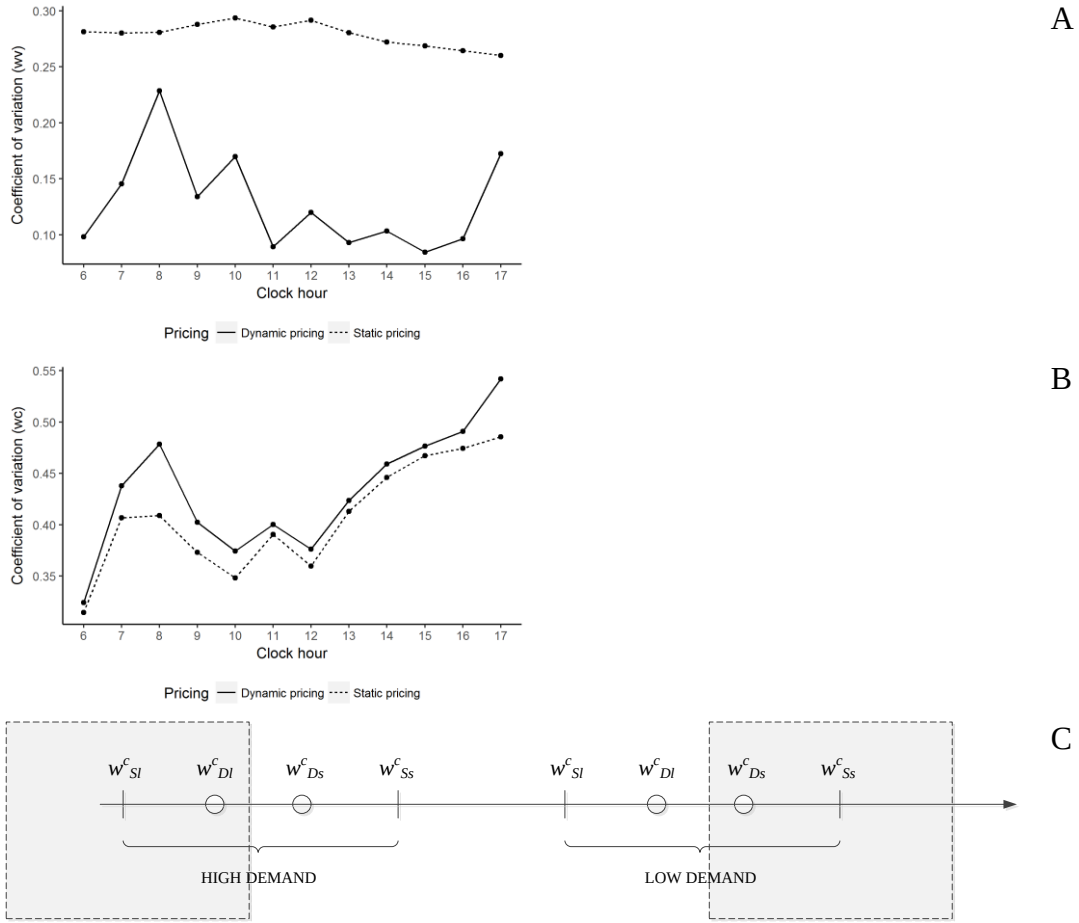


Figure 5-6. Spatial Variations of Average Searching and Waiting Times. A) Temporal Display of CV for Average Searching Time, B) Temporal Display of CV for Average Waiting Time, C) Spreads of Average Waiting Times under Dynamic and Static Pricing

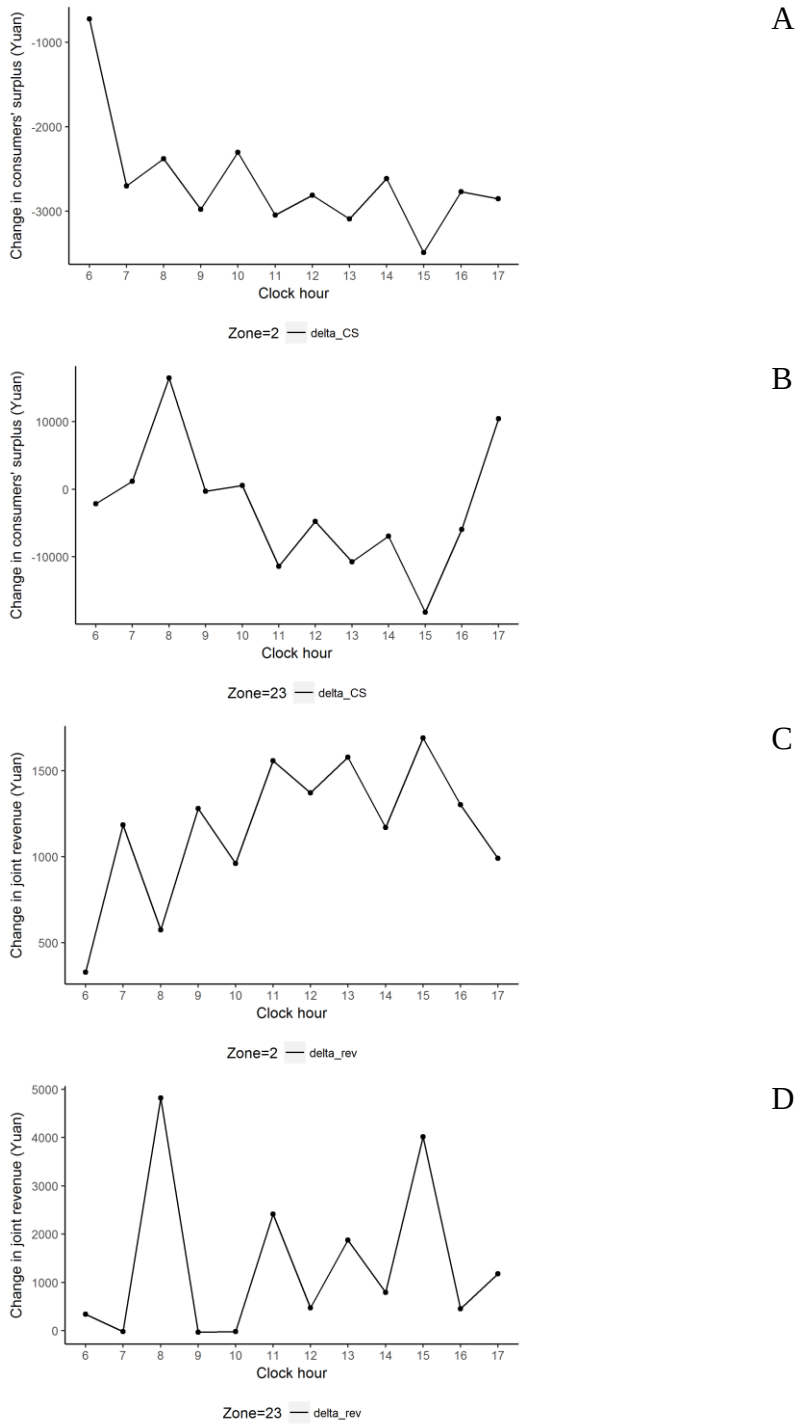


Figure 5-7. Temporal Display of Welfare Changes under Dynamic Pricing. A) Change of CS in Zone 2, B) Change of CS in Zone 23, C) Change of Joint Revenue in Zone 2, D) Change of Joint Revenue in Zone 23

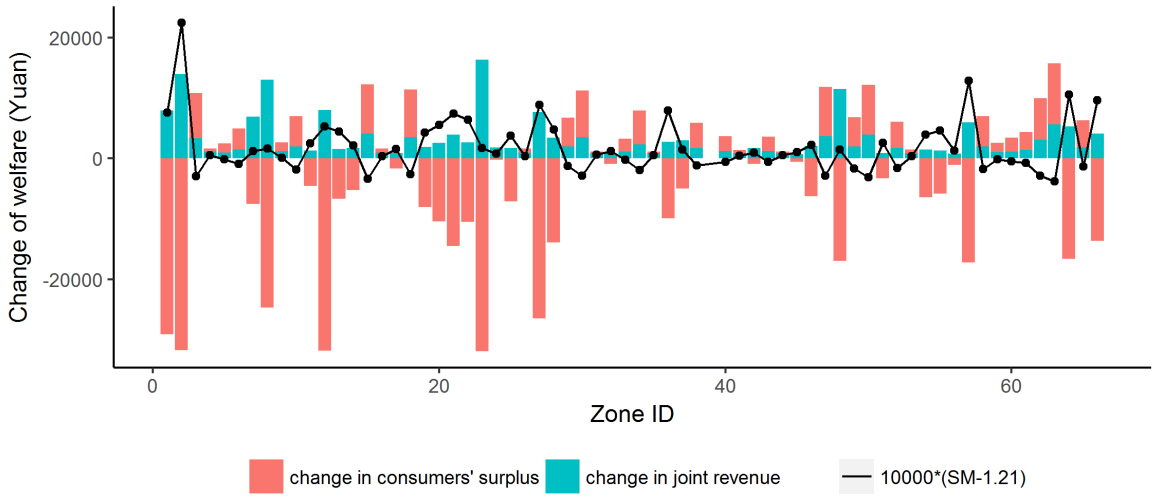


Figure 5-8. Spatial Display of Welfare Changes under Dynamic Pricing.

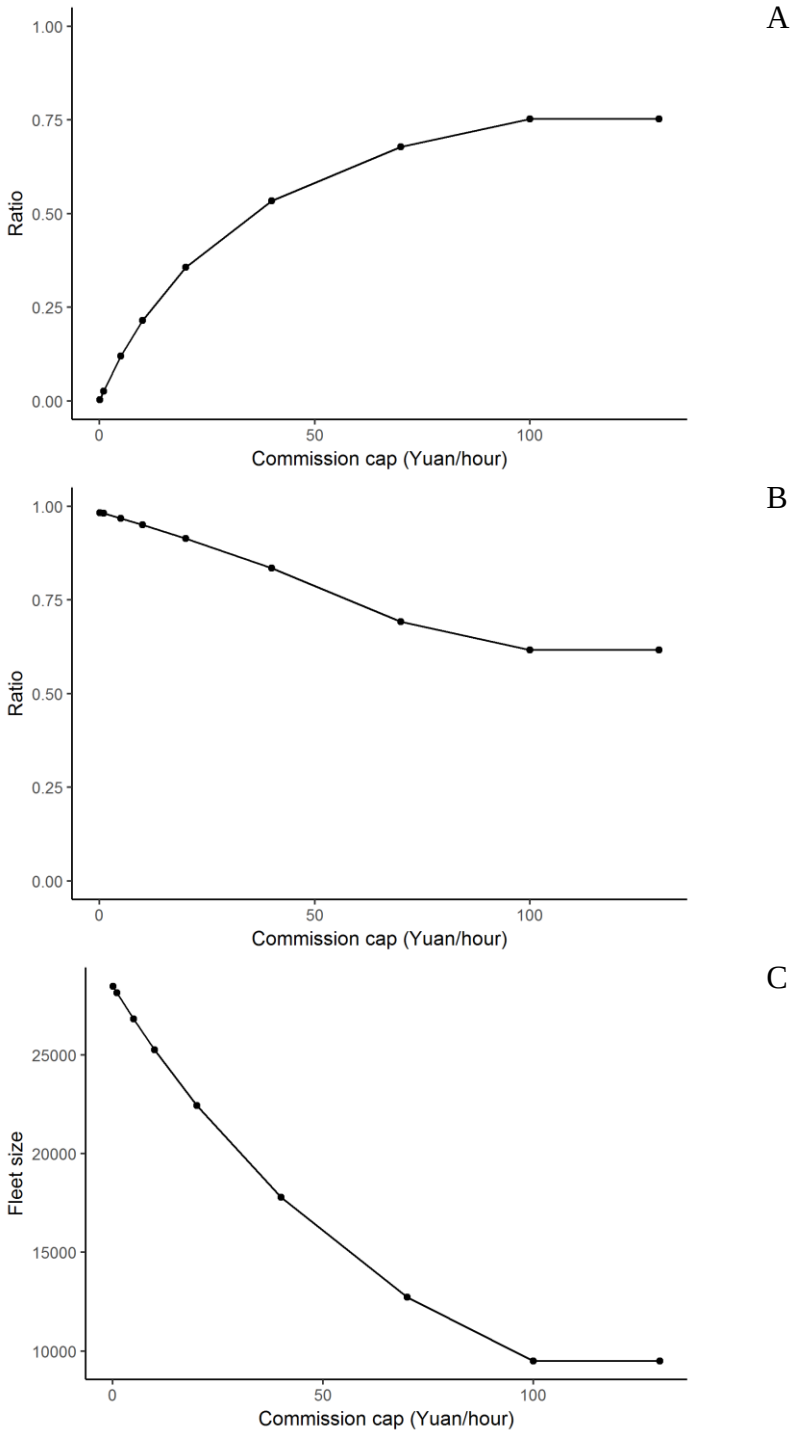


Figure 5-9. Equilibrium Outcomes under the Commission Cap Regulation. A) Share of Platform's Revenue, B) The Ratio of The Social Welfare under The Proposed Regulation to That of The (Quasi-) Second Best, C) Fleet Size versus Commission Cap

Table 5-1. Demand Estimation for Different Parameterizations.

Variables	Model I Parameter	S.E.	Model II Parameter	S.E.	Model III Parameter	S.E.*
Fare6	-0.033	0.004	-0.008	0.003	-0.033	0.008
Fare7	-0.087	0.006	-0.026	0.004	-0.052	0.009
Fare8	-0.088	0.006	-0.034	0.004	-0.081	0.012
Fare9	-0.049	0.005	-0.021	0.003	-0.071	0.015
Fare10	-0.042	0.004	-0.009	0.002	-0.068	0.015
Fare11	-0.048	0.005	-0.011	0.003	-0.045	0.01
Fare12	-0.049	0.005	-0.014	0.003	-0.06	0.015
Fare13	-0.049	0.005	-0.011	0.003	-0.039	0.009
Fare14	-0.048	0.005	-0.009	0.003	-0.043	0.02
Fare15	-0.042	0.005	-0.011	0.003	-0.029	0.016
Fare16	-0.050	0.005	-0.013	0.003	-0.042	0.01
Fare17	-0.048	0.005	-0.015	0.003	-0.065	0.018
Waiting time6	NA	NA	-1.618	0.074	-1.894	0.093
Waiting time7	NA	NA	-3.556	0.128	-3.649	0.348
Waiting time8	NA	NA	-2.899	0.127	-2.743	0.355
Waiting time9	NA	NA	-2.476	0.089	-2.677	0.183
Waiting time10	NA	NA	-1.97	0.064	-2.199	0.13
Waiting time11	NA	NA	-2.023	0.066	-2.351	0.133
Waiting time12	NA	NA	-2.341	0.074	-2.616	0.161
Waiting time13	NA	NA	-2.189	0.072	-2.494	0.14
Waiting time14	NA	NA	-2.067	0.075	-2.322	0.196
Waiting time15	NA	NA	-2.291	0.08	-2.649	0.167
Waiting time16	NA	NA	-2.95	0.105	-3.171	0.277
Waiting time17	NA	NA	-2.677	0.103	-2.738	0.315
Intercept	Yes		Yes		Yes	
X	Yes		Yes		Yes	

*robust standard errors are reported.

Table 5-2. Summary of the Parameters for the Supply Side.

Parameter	Value	Parameter	Value
H (hr)	2.7	n_{11}	0.070
N (veh)	31,097	n_{12}	0.074
n_6	0.025	n_{13}	0.079
n_7	0.092	n_{14}	0.078
n_8	0.114	n_{15}	0.083
n_9	0.097	n_{16}	0.098
n_{10}	0.071	n_{17}	0.119

CHAPTER 6
AN AGENT-BASED SIMULATION FOR ON-DEMAND RIDE-SOURCING MARKETS

Different analytical models have been proposed in previous three chapters, which simplify the behavioral interactions among market players and rely on the hypothetical equilibrium states to generate insights. Realistically, each market participant should have specific attribute and preference. The interactions among the involved agents and the evolution of market dynamics can be well captured by the agent-based simulation (Wilensky, 2015). The goal of this chapter is two-fold. We first develop a simulation tool and demonstrate its capability of analyzing the dynamics of the ride-sourcing market. Using the simulation output, we next calibrate and discuss the matching function employed in Chapter 3.

In Section 6.1, we build a simulation test bed in NetLogo while specifying customers' and drivers' behavioral rules in its Application Programming Interface (API). Section 6.2 presents a simulation evaluation framework where different scenarios are created via modifying the zonal demand levels and the matching ranges. In Section 6.3, we demonstrate the market equilibrium, calibrate the Cobb-Douglas matching function for a range of matching radii and compare their efficiencies and returns to scale properties. Conclusions are summarized in Section 6.4. A detailed introduction on the simulation interface and its functionality is attached in Appendix H for interested readers. Note that some of the notations in this chapter may be different from previous chapters.

6.1 Simulation Framework

This section mainly introduces the configuration of the simulation test bed, the specifications of the behavioral heuristics of drivers and customers as well as the concept of market equilibrium.

6.1.1 Simulation Test Bed

Figure 6-1 shows our simulation test bed that is based on the Manhattan network. The total network is divided into 13 zones. We index 12 of them as feasible zones where customers and ride-sourcing vehicles are interacted (The center zone is the Center Park area and is treated as the infeasible zone in our simulation). As no local streets or roads is modeled, our simulation is to the resolution of the zonal level.

6.1.2 Basic Agents Module

Three types of agents are considered, namely, a ride-sourcing platform, potential customers and the affiliated drivers. The role of the ride-sourcing platform is to determine the prices (e.g., trip fare and the commission) and the matching algorithms. We assume all these components are exogenously given. Therefore, our focus is on the interactions between the costumers and the drivers, represented by the blue and orange "dots" in Figure 6-1. Both agents are assumed to be homogeneous in their attributes to be mentioned.

We assume a predetermined origin-destination (OD) trip demand table for customers that use the ride-sourcing services. Customers are constantly generated according to the OD table and randomly distributed in the feasible zones. Locations of the customers are fixed once generated; they keep sending requests until being matched with the ride-sourcing vehicles nearby. All customers are assumed to wait up to a maximum time w , above which they give up the service (e.g., switch to taxi or public transportation) and "disappear" from the network. Customers' demand is thus elastic.

We assume a total number of N ride-sourcing vehicles throughout the simulation. The value of N shall be determined endogenously at the market equilibrium, given that the ride-sourcing market is entry free. A random realization of vehicles' location (in all feasible zones) is used when the simulation is initialized. When being vacant, ride-sourcing vehicles search in each

zone with a constant speed v. To specify the local searching behavior, we define an allowable "cone" Δ that represents the potential rotation from their previous direction. When they get close to the boundary of an unsearchable area (i.e., Center Park zone and areas outside of the simulation network), the ride-sourcing vehicles are forced to face the center of the zones they are currently in and rotates an angle δ for their new direction. Drivers will consider to transition to other zones periodically for better trip opportunities. Their decision is triggered either right after successfully finishing a transaction or remaining vacant after certain time ΔT spent in the local search. The specification of the decision rules will be explained in the Zonal Choice Module.

6.1.3 Transaction Module

At each time step, the platform loops through the queue of the waiting customers; and for each customer, the platform matches a vacant vehicle closest to her within a matching radius *r*. If no vehicle is available, the customer will wait for the next time step until being matched or the maximum waiting time W is reached, whichever comes first. When a customer is successfully matched, the matched pair "disappears" from the network. After some time period (i.e., the trip duration), the customer is assumed to be served and is removed from the simulation. The vehicle "appears" in the network again and determines which zone to go for customers. We assume the travel distance table for inter-zone trips is given which induces the inter-zone travel times. Trip length for intra-zone trips is randomly drawn with a minimum value $d_{\min}$.

We assume that vehicles are always running either occupied or not. Drivers incur an average cost *c* per vehicle per unit time, which represents the operating cost and the opportunity cost. Let's define I, J as the set of origination and destination zones. The price charged to a customer traveling from zone $i \in I$ to zone $j \in J$ is given by:

$$F_{ij} = F_0 + \omega \times l_{ij}, \forall i \in I, j \in J \tag{6-1}$$

where F_0 is the flag drop fee, ω represents the incremental fee per unit time, l_{ij} is the corresponding trip time. We do not consider the distance-based charge as it duplicates the time-based charge given that traffic congestion is not considered here. ω is fixed and there is no surge in price. As the rule of operation, a certain percentage η is changed by the ride-sourcing company from the final trip fare. Therefore, the driver's net profit per transaction is $(1-\eta)F_{ij} - cl_{ij}$.

Given the purposes of this study, there is no need to capture the dynamics in the temporal dimension. The presentation focuses on the spatial domain.

6.1.4 Zonal Choice Module

Drivers will consider to transition to other zones for better trip opportunities periodically. They have the perceived information about the market dynamics across all zones. A driver in zone $j \in J$ will choose zone $i \in I$ to maximize her utility. The utility is assumed positively related to the average profit $\bar{\pi}_i$; and negatively related to the sum of transition time l_{ji} and the average searching time $\bar{w}_i^v$ (Yang et al., 2010b). The perception error is captured by the random variable ε_i. Together, we have:

$$U_{ji} = \bar{r}_i - c\left(l_{ji} + w_i^v\right) + \varepsilon_i = V_{ji} + \varepsilon_i,\ \forall j \in J, i \in I \tag{6-2}$$

where $\bar{r}_i = (1-\eta)\bar{F}_i - c\bar{l}_i$, $\bar{F}_i = \sum_{k\in J} Q_{ik}F_{ik} \Big/ \sum_{k\in J} Q_{ik}$ and $\bar{l}_i = \sum_{k\in J} Q_{ik}l_{ik} \Big/ \sum_{k\in J} Q_{ik}$ is the average trip fare and trip time from zone i, adjusted by the demand rate from zone i, $\sum_{k\in J} Q_{ik}$; w_i^v is the average drivers' searching time in zone *i*. If we assume the error term is independently and identically Gumbel distributed, the probability of a driver at zone *j* choosing to travel to zone *i* can be represented in the following multinomial logit form:

$$P_{ji} = \frac{\exp\left(\theta V_{ji}\right)}{\sum_{m \in I} \exp\left(\theta V_{jm}\right)}, \quad \forall i \in I, j \in J \tag{6-3}$$

where θ represents the uncertainty of the driver's perception and should be calibrated using the real-world data (Yang et al., 2010b). Note that if $i = j$, the driver will stay in the current zone and continue the local search for another period of ΔT.

Drivers always follow this probabilistic approach for zonal choice. It is still possible, however, for them to be matched with any customers on the way when transitioning. This occurs as long as a driver happens to be the closest to a waiting customer within the matching radius. The increasing probability of being matched essentially reflects the improvement of the matching technology over the traditional street hailing. Given no street system has been specified, we predefine a set of seudo-paths and project the zonal crossing movement of the ride-sourcing vehicles onto these seudo-paths. In this manner, the location of vehicles during transitioning can be explicated traced at each time step. Note that such a restriction is only initiated in the Zonal Choice Module while drivers still follow the random search behavior locally as is specified in the Basic Agent Module

6.1.5 Market Entry and Equilibrium

Our simulation is capable of delineating the evolution of market dynamics. For the purpose of this study, we only capture a stable state that approximates the static equilibrium. To explain the concept of the equilibrium, we temporally omit the spatial dimension and introduce a fluid queuing process in Figure 6-2. As is shown in the left reservoir of Figure 6-2, the incoming rate of customers Q equals to the rate of the served customers T^o (otherwise the queue of the waiting customers grows). A circular flow is used in the right reservoir to represent the conservation between the rate of the occupied and vacant ride-sourcing vehicles, represented as T^o and T^v respectively (otherwise the queue of the vacant vehicles grows). The red dashed line

indicates $Q = T^v$ which means the equilibrium between demand and supply. Therefore, $Q = T^v = T^o$ must hold at the equilibrium (Yang and Yang, 2011).

Levels of the fluid are proportional to the number of waiting customers (N^c) and searching vehicles (N^v) in the system. Both N^c and N^v remain constant at the equilibrium; the corresponding average waiting and searching times (w^v and w^c, respectively) can be obtained through Little's formula:

$$N^v = T^v w^v = Q w^v \tag{6-4}$$

$$N^c = T^o w^c = Q w^c \tag{6-5}$$

Multiple equilibria can be achieved with different fleet size *N*. We assume an exogenous reservation profit level π^o for the drivers, which reflects their preferences and the competition level in the labor market. Drivers are implicitly assumed to be perfectly elastic, which is likely true in the long run: if the average profit is larger than π^o, more drivers will be attracted to enter; otherwise a random portion of the drivers will exit. The fleet size is adjusted dynamically in this matter until the equilibrium profit matches the reservation one. Therefore, we are essentially considering a long-run equilibrium with free entry.

6.2 Simulation Evaluation Framework

This section explains the simulation evaluation framework. We first clarify the simulation tasks based on which we develop the simulation scenarios. The parameters used in the simulation are also specified.

To examine the effectiveness of the agent-based simulation in characterizing the market equilibrium, we are interested in verifying the queuing status highlighted in Eqs.(6-4)-(6-5). To better understand the impact of the improved technology on the market dynamics, we analyze the patterns of matching frictions and the success rate of the zonal choice for different combinations

of demand levels and matching radii. Based on the data collected at the equilibrium state, we estimate the matching function for each matching radius and discuss its returns to scale property.

We consider 4 scenarios by increasing *r* from 1 to 15 unit distance, with the lowest and highest values representing street hailing and ride-sourcing technologies respectively. To represent a wide range of market conditions, three demand levels are designed for each *r*. The fleet size also changes accordingly to achieve the same reservation profit across all simulation runs. In total, we have a 12 simulation runs.

Table 6-1 provides the base demand used in the analysis. Three demand levels are obtained via multiplying by an adjusting parameter $\lambda \in \{1,\ 2,\ 2.67\}$. We thus have the low, median and high demand scenarios for each *r*. The trip length matrix corresponding to the OD pairs is summarized in Table 6-2 which is assumed symmetric. All the other parameters used in the simulation are listed in Table 6-3. To be consistent with the real world operations, most of the parameters are estimated based on the Uber/taxi data from the Taxi & Limousine Commission (TCL) at New York (TLC, 2016). Note that in our simulation, one unit distance and time approximately equals 0.06 mile and 1/6 min.

6.3 Results and Discussions

This section contains the discussion of the simulation outputs which mainly demonstrate the market dynamics; and the regression analysis of the matching function.

6.3.1 Analysis of the Output Data

6.3.1.1 Equilibrium property

When a static equilibrium is achieved, all variables of interest are asymptotically constant. Dividing Eq. (6-4) by Eq. (6-5), we obtain:

$$\frac{w^v}{w^c} = \frac{N^v}{N^c} \tag{6-6}$$

Therefore, the ratio of w^v/w^c to N^v/N^c should equal one ideally. Note that all variables in the equation above are computed as the average across the network, as no spatial dimension is captured in characterizing this relationship. This ratio is plotted for each simulation run. As is given in Figure 6-3, all the points fluctuate closely around one, confirming the equilibrium state. In fact, this ratio is monitored during the simulation and serves as a supplementary criterion to determine the equilibrium state, a point when we terminate the simulation. The whole procedure that converges to the equilibrium state is demonstrated in the Appendix H under the item "Equilibrium Check".

6.3.1.2 Matching frictions

In the spatial ride-sourcing market, matching frictions always exist: both drivers and customers have to wait for certain amount of time to be served. Following Chapter 5, we decompose the matching process into two stages: 1) matching stage: the platform randomly matches a customer to her closet available vehicle within *r* when possible; 2) meeting stage: the vehicle comes and picks up the customer conditional on a successful matching. Accordingly, customers' average waiting time w_i^c can be decomposed into two parts:

$$w_i^c = w_i^{cm} + w_i^m,\ i \in I \tag{6-7}$$

where w_i^{cm} is the average time for the requesting customer being matched; w_i^m represents the average travel time for the driver to pick up the requesting customer after a successful matching. By symmetry, drivers' average searching time w_i^v is:

$$w_i^v = w_i^{vm} + w_i^m,\ i \in I \tag{6-8}$$

where w_i^{vm} is the average time for a vacant vehicle to be matched.

We now analyze the patterns of the matching frictions. Note that the average trip distances (and trip times) are approximately equal for the investigated scenarios. As vehicles are either occupied or vacant, the average searching time for the drivers must be about the same so that the same reservation profit level can be achieved. For all the investigated simulation scenarios, w^v is around 4.2 min. Our focus is thus on customers' average waiting time. Figure 6-4 displays the customers' average waiting time under varying demand levels and matching radii. It is found that w^c decreases as the increase of the demand for each matching radius. To explain, the increasing demand attracts the same proportional drivers to enter which generally raises the number of vacant vehicles (N^v) and customers (N^c) at the equilibrium. The simultaneous increase of N^v and N^c will reduce the distance between the customers and drivers, which leads to lower market frictions.

For each demand level, w^c decreases as the increase of *r* but the mechanism differs. For the lower matching radius (e.g., street hailing), the majority of the waiting time is spent on the matching time due to the high transaction cost (or more precisely the bilateral searching cost). Comparatively, the meeting time is negligible since the matched pair is quite close. At *r* = 1 (0.06 mile) and $\lambda = 1$, for example, the average value for the matching time and meeting time is around 1.8 min and 0.2 min, respectively. In contrast, for the case of high matching radius (e.g. ride-sourcing), the matching time is reduced significantly while the customer has to wait an increasing amount of time for a vehicle that can be farther away (but is still the closet available vehicle within the matching radius). The net outcome is the dominance of the reduction of the matching time. When *r* = 15 (0.9 mile) and $\lambda = 1$, the average value of the two portions are around 0.05 min and 0.6 min, respectively.

6.3.1.3 Success rate of zonal transition

As has been discussed in the Zonal Choice Module, not all the vacant vehicles are able to transition to their ideal zones successfully. We provide the rate of successful zonal transition for the low demand case under r =1 and r = 15 in Table 6-4. The rate of success decreases as vehicles travel to zones further away. Besides, higher matching radius leads to lower success rate, i.e., ride-sourcing vehicles are more likely to be matched than the traditional taxis when transitioning.

As customers' demand increases, both matching radii see a reduction on the success rate and the difference between the two is nil. The specific matrixes are not provided due to the limited space. Overall, it is clear that the success rate is not high from our simulation. The assumption of perfect success rate in previous analytical studies can be problematic even for the traditional street hailing market (Yang and Wong, 1998; Yang et al., 2010b; Nicholas, 2015).

6.3.2 Regression Analysis

We calibrate the Cobb-Douglas matching function assumed in Chapter 3. To be consistent, we regress the realized demand Q against the number of vacant vehicles (N^v) and waiting customers (N^c):

$$Q = A \times (N^v)^{\alpha_1} \times (N^c)^{\alpha_2} \tag{6-9}$$

The parameters (α_1, α_2, A) measures the efficiency of the matching technology. Take logarithm of both sides:

$$\ln(Q) = \ln(A) + \alpha_1 \ln(N^v) + \alpha_2 \ln(N^c) \tag{6-10}$$

With observations of (Q, N^v, N^c) at the equilibrium, the associated parameters can be estimated using standard multivariate regression technique.

To generate enough samples for the regression, we use the data at the zonal level. Therefore, a total 36 observations are utilized for each simulation. The estimated results are presented in Table 6-5 and the standard error for each estimate is given in the associated parenthesis. All the parameters are found significant at the 95% confidence interval. As the goodness-of-fit measure, the R^2 values are close to 1, indicating the explanatory variables capture the majority of the variance in the data.

Two observations are worth mentioning. First, the estimated results confirm the increasing-returns-to-scale property (i.e., $\alpha_1 + \alpha_2 > 1$) for all matching radii. Second and perhaps more interesting, the increase of the matching radius is associated with a decreasing magnitude of the returns to scale but increasing intercept term $\ln(A)$. The efficiency of the matching technology in large matching radius is mainly reflected by the parameter *A*.

To explain, the obstacle for the street hailing market is the high search cost due to the limited matching radius (e.g., 0.06 mile is assumed). Therefore, an increase of both N^v and N^c significantly increases the matching probability via reducing the distance between the customers and drivers. This indicates a matching function of high returns to scale. On the other hand, the matching radius in the ride-sourcing market has already been sufficiently large (e.g., 0.9 mile is assumed). The benefit in increasing the matching probability is thus marginal. However, both sides still enjoy the reduced distance because the matched pair can be realized more quickly (i.e., economies of density). This is why the returns to scale is always above one.

To shed more light on the impact of the matching technology, we explicitly account for the matching radius as:

$$M = Q = A \times r^{\alpha_3} \times (N^v)^{\alpha_1} \times (N^c)^{\alpha_2} \tag{6-11}$$

The previous 144 observations are jointly used for the parameterization above. We are only interested in the value of α_3 which turns out to be 0.65 (with the standard error 0.03) and is significantly larger than 0. In the investigated matching ranges, an increase in the matching radius leads to larger customers' demand (or total transaction).

6.4 Summary

We develop an agent based simulation for the ride-sourcing market with special focus on demonstrating the market dynamics at equilibrium and calibrating the matching function. A simulation test-bed is built in NetLogo while customers' and drivers' behavioral heuristics are specified in its Application Programming Interface (API). Three major modules are developed. The Basic Agent Module specifies the basic attributes and moving kinematics of the simulated agents. The Transaction Module reproduces the algorithms in the current matching technology of the ride-sourcing applications. The Zonal Choice Module consists of the utility maximization rule of how a driver determines the location for seeking the customers. Different scenarios are created via modifying the zonal demand levels and the matching radii. We conclude that:

- Our simulation demonstrates the capability of delineating the evolution of the market dynamics as well as the equilibrium state often assumed in previous analytical studies (He and Shen, 2015; Wang et al., 2016).
- The rate of successful zonal transitioning decreases as vehicles travel further away; a higher matching radius leads to a lower success rate.
- An increase in the matching radius in general leads to lower average waiting time for the customers. However, a trade-off exists in the matching time and the meeting time: the reduced search cost significantly lowers the matching time while the increasing possibility of matching a vehicle further away increases the meeting time.
- The proposed Cobb-Douglas matching function accounts for the majority of the variance in the simulated data and the matching technologies exhibits increasing returns to scale. Lower matching radius (e.g. street hailing) implies higher returns to scale but smaller intercept term. That is, the technology with lower matching radius yields lower matching rate; however, a simultaneous increase of the number of vacant vehicles and waiting customers gives more significant reduction of customers' average waiting time.

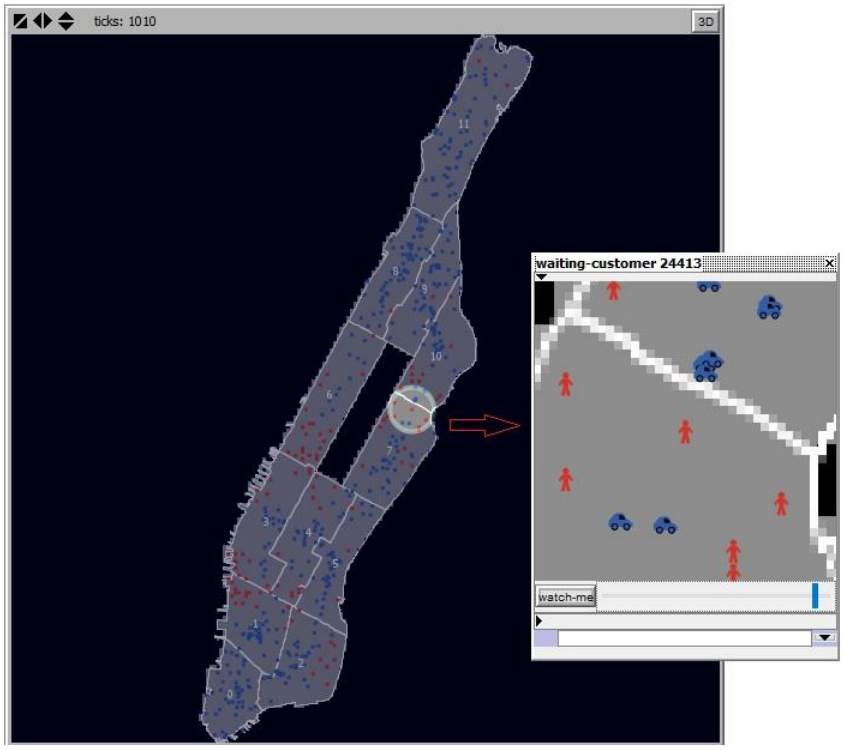


Figure 6-1. Simulation Configuration.

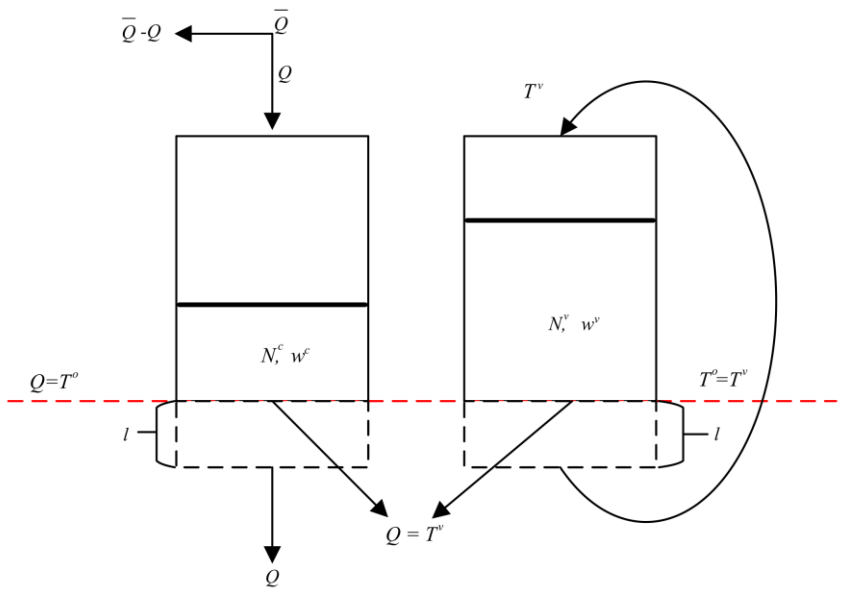


Figure 6-2. Demand and Supply Equilibrium in the Ride-Sourcing Market.

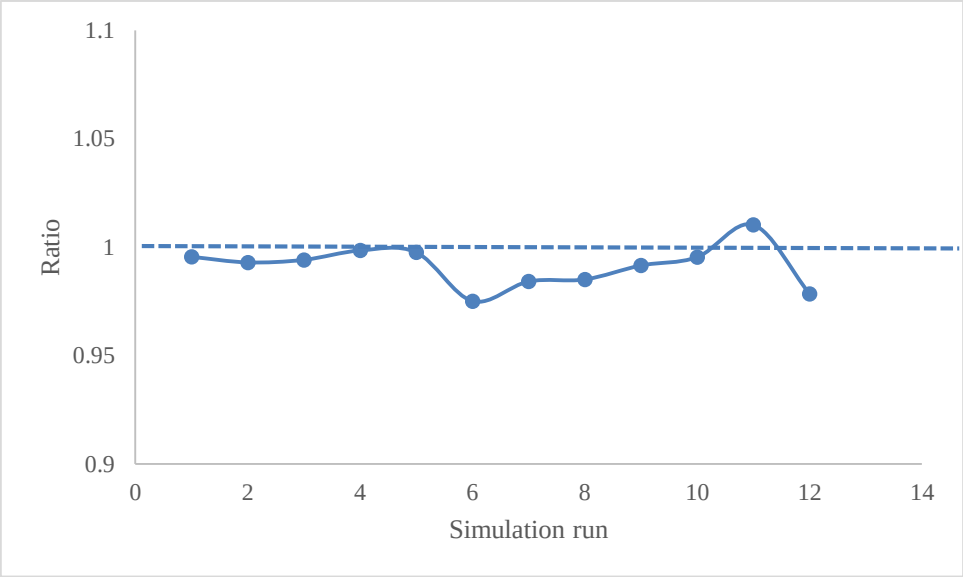


Figure 6-3. Display of the Ratio of (w^v/w^c) to (N^v/N^c) for Each Simulation Run.

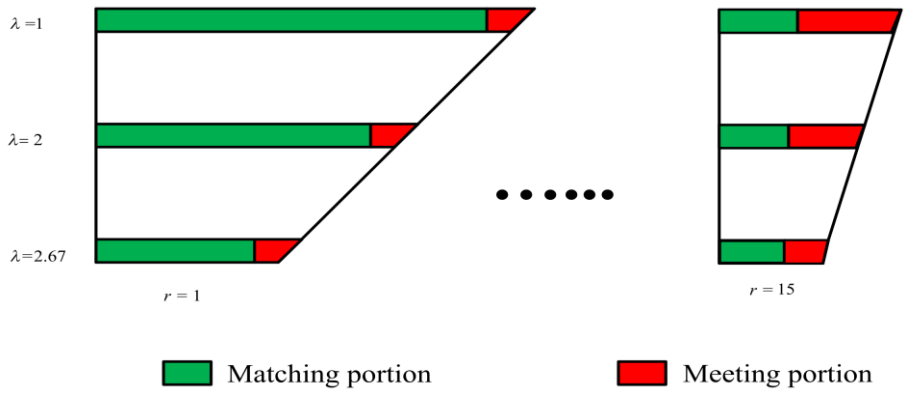


Figure 6-4. Customers' Average Waiting Time under Varying Demand Levels and Matching Radii.

Table 6-1. Distributions of Customers' Demand (/hour).

O/D	0	1	2	3	4	5	6	7	8	9	10	11
0	30	30	30	30	30	30	30	30	30	30	30	30
1	30	30	30	30	30	30	30	30	30	30	30	30
2	30	30	30	30	30	30	30	30	30	30	30	30
3	30	30	30	30	30	30	30	30	30	30	30	30
4	30	30	30	30	30	30	30	30	30	30	30	30
5	30	30	30	30	30	30	30	30	30	30	30	30
6	30	30	30	30	30	30	30	30	30	30	30	30
7	30	30	30	30	30	30	30	30	30	30	30	30
8	30	30	30	30	30	30	30	30	30	30	30	30
9	30	30	30	30	30	30	30	30	30	30	30	30
10	30	30	30	30	30	30	30	30	30	30	30	30
11	30	30	30	30	30	30	30	30	30	30	30	30

Table 6-2. Distributions of Trip Distances (mile).

O/D	0	1	2	3	4	5	6	7	8	9	10	11
0	N/A	1.3	1.3	3.0	3.1	3.1	5.5	5.2	7.8	8.0	7.1	10.6
1	1.3	N/A	1.0	1.8	1.8	2.5	4.2	3.8	6.6	6.8	5.7	9.4
2	1.3	1.0	N/A	2.8	2.5	1.8	5.2	3.9	7.6	6.9	5.7	9.8
3	3.0	1.8	2.8	N/A	0.7	1.5	2.4	2.7	4.8	5.4	4.5	7.6
4	3.1	1.8	2.5	0.7	N/A	0.7	3.2	2.0	5.6	5.0	3.8	7.9
5	3.1	2.5	1.8	1.5	0.7	N/A	3.9	2.1	6.3	5.1	4.0	8.0
6	5.5	4.2	5.2	2.4	3.2	3.9	N/A	6.0	2.4	3.0	4.1	5.2
7	5.2	3.8	3.9	2.7	2.0	2.1	6.0	N/A	3.6	3.0	1.8	5.9
8	7.8	6.6	7.6	4.8	5.6	6.3	2.4	3.6	N/A	0.6	1.7	2.8
9	8.0	6.8	6.9	5.4	5.0	5.1	3.0	3.0	0.6	N/A	1.2	2.9
10	7.1	5.7	5.7	4.5	3.8	4.0	4.1	1.8	1.7	1.2	N/A	4.1
11	10.6	9.4	9.8	7.6	7.9	8.0	5.2	5.9	2.8	2.9	4.1	N/A

Table 6-3. Suggested Value of the Parameters.

Parameter	Description	Value
r	Matching radius	Varies
π^o	Reservation profit for the drivers	$ 19.6/hr
v	Average speed	13 mph
Δ	Allowable angle of left and right rotation with each movement	$\pm 15°$
δ	Allowable angle of left and right rotation when a vehicle encounters the boundary	180°
$d_{\min}$	Minimum distance required for intra-zone trip	0.6 mile
c	Average cost	$ 11/hr
F_0	Flag drop fee	$ 6
ω	Incremental fee	$ 0.5/min
η	Percentage charged by the platform from the final trip fare	20%
θ	Uncertainty associated with drivers' zonal choice	1
ΔT	Amount of time a vehicle searches before deciding if it should leave	8.3-16.6 min
W	Maximum waiting time	10 min

Table 6-4. Success Rate for Zonal Choice.

A) r = 1 (0.06 mile)

O/D	0	1	2	3	4	5	6	7	8	9	10	11
0	N/A	0.34	0.35	0.10	0.14	0.18	0.03	0.00	0.00	0.00	0.00	0.00
1	0.32	N/A	0.40	0.20	0.20	0.12	0.03	0.01	0.00	0.00	0.00	0.00
2	0.35	0.39	N/A	0.13	0.14	0.21	0.03	0.04	0.00	0.00	0.00	0.00
3	0.12	0.15	0.10	N/A	0.39	0.31	0.05	0.03	0.02	0.00	0.06	0.00
4	0.12	0.19	0.17	0.44	N/A	0.44	0.05	0.10	0.00	0.03	0.02	0.00
5	0.12	0.14	0.23	0.32	0.42	N/A	0.05	0.05	0.00	0.00	0.03	0.00
6	0.01	0.06	0.04	0.08	0.07	0.00	N/A	0.00	0.07	0.06	0.03	0.03
7	0.02	0.04	0.02	0.04	0.10	0.05	0.00	N/A	0.07	0.07	0.12	0.01
8	0.00	0.00	0.00	0.00	0.08	0.00	0.05	0.04	N/A	0.48	0.22	0.14
9	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.07	0.47	N/A	0.29	0.10
10	0.00	0.00	0.00	0.03	0.07	0.05	0.03	0.10	0.28	0.30	N/A	0.07
11	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.00	0.28	0.19	0.13	N/A

B) r = 15 (0.9 mile)

O/D	0	1	2	3	4	5	6	7	8	9	10	11
0	N/A	0.25	0.25	0.08	0.09	0.06	0.00	0.00	0.00	0.00	0.00	0.00
1	0.27	N/A	0.34	0.14	0.12	0.09	0.02	0.00	0.00	0.00	0.00	0.00
2	0.24	0.31	N/A	0.10	0.07	0.12	0.00	0.02	0.00	0.00	0.00	0.00
3	0.06	0.13	0.05	N/A	0.30	0.24	0.02	0.03	0.00	0.00	0.00	0.00
4	0.06	0.12	0.09	0.36	N/A	0.39	0.01	0.03	0.00	0.00	0.00	0.00
5	0.04	0.08	0.11	0.25	0.37	N/A	0.01	0.02	0.00	0.00	0.00	0.00
6	0.01	0.02	0.01	0.01	0.02	0.01	N/A	0.00	0.04	0.04	0.04	0.01
7	0.00	0.02	0.00	0.03	0.03	0.02	0.03	N/A	0.02	0.03	0.04	0.01
8	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	N/A	0.51	0.27	0.10
9	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.02	0.48	N/A	0.26	0.09
10	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.05	0.18	0.23	N/A	0.06
11	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.03	0.16	0.18	0.14	N/A

Table 6-5. Summary of Regression Results.

	$\ln A$	α_1	α_2	$\alpha_1 + \alpha_2$	R^2
r = 1	0.94 (0.46)	0.86 (0.02)	0.80 (0.09)	1.66	0.86
r = 5	2.76 (0.19)	0.58 (0.03)	0.81 (0.06)	1.39	0.94
r = 10	2.83 (0.18)	0.55 (0.03)	0.83 (0.06)	1.38	0.94
r = 15	2.86 (0.19)	0.69 (0.05)	0.60 (0.10)	1.29	0.93

CHAPTER 7
CONCLUSIONS AND FUTURE RESEARCH

7.1 Conclusions

In this dissertation, we develop both analytical and simulation tools for analyzing emerging (on-demand) ride-sourcing markets. Our investigation focuses on the key issues that include dynamic pricing, platform competition and regulation. An aggregate and static model is first introduced to understand the effects of platform competition in trip fares and the social welfare. An exogenous Cobb-Douglas matching function is used to delineate the matching technology the platform offers to match the customers with the drivers. Possible regulation policies are then analytically explored and compared.

To explore the effects of dynamic pricing, we extend our modeling framework in two ways. First, we adopt a deductive approach to describe the matching process which features the platform's matching technology. Next, we incorporate temporal and spatial dimensions in the proposed framework. Accordingly we are capable of capturing drivers' work hour scheduling decisions, the spatial variations of market frictions and the change of agents' welfare under dynamic pricing. In addition, we present a simple regulation scheme for dynamic pricing if market power is a concern.

Lastly, an agent-based simulation is developed to validate the properties of the matching function. It also demonstrates the potential for the applications in other domains of the markets. Overall, the proposed methodologies are characterized by different modeling focuses and resolutions, which can be used by regulatory agencies to better understand and manage the ride-sourcing markets.

The main results and policy implications are summarized as follows:

- The matching of the ride-sourcing services exhibits increasing returns to scale property: a simultaneous increase in the intensities of unmatched vehicles and requesting customers

leads to a reduction of the customers' average waiting time. The reason behind is that ride-sourcing technology benefits from the reduced distance between the matched pairs (so the meeting time for a vehicle to pick up a requesting customer decreases); given the sufficient matching radius, however, the time for matching a vehicle to a requesting customer (matching time) is nearly unchanged. Although the matching in traditional street-hailing is also increasing returns to scale, the mechanism differs. Due to the limited matching range, the reduction of the average waiting time is mostly from the matching time while the meeting time is negligible. In fact, more advanced matching technology (with larger matching radius) comes with lower returns to scale, an observation verified by our simulation.

- The increasing-returns-to-scale matching process implies that for the ride-sourcing market, more competition doesn't necessary generate more efficient outcomes. Increasing the number of ride-sourcing companies probably better caters the preferences of the customers. Market frictions (average waiting and searching times) increase as well. If the increased friction is a dominant factor, competition can neither lower down the trip price nor improve the social welfare. The optimal decision for the regulatory agency is then to encourage the merger of the platforms and regulate them as a monopolist.

- The impact of dynamic pricing on the welfare of the involved market agents is mixed when compared with the optimal static pricing. In general, dynamic pricing benefits the platform and the drivers due to the current revenue sharing structure (e.g., the platform takes 20% of the final fare as the commission while the remaining goes to the drivers). However, customers may be worse off during highly surged periods. Dynamic pricing has the potential to create a win-win situation when trip fare is adjusted dynamically below its static counterpart. This phenomenon occurs mostly in geographic zones that are characterized by longer average trip duration. As trip price is proportional to trip duration (when traffic congestion is not considered), a revenue-maximizing platform tries to "realize" the demand from such zones even at the price of lowering down the surge multipliers. Subsequently, the increase of realized demand boosts consumers' surplus.

- Lastly, in case market power is a concern, we propose a simple regulation scheme that can further improve the efficiency of the market: capping the amount of commission from each transaction. Under the proposed regulation, a revenue-maximizing platform has the incentive to maximize total transactions. The additional revenue from price surging completely goes to the drivers. We prove that such a regulation scheme can achieve the second best based on our static model; its performance is also numerically confirmed under dynamic pricing. Given the heterogeneity of the trip durations, the proposed scheme may be implemented as the distance-based charge, time-based charge or the combination of both. However, a proper choice of commission cap should fully explore the cost structure of sourcing companies, as we must balance a healthy profit margin of a ride-sourcing company and the welfare of its users.

7.2 Future Research

Despite the efforts we have made, there are tremendous opportunities to be explored.

Future research can be done in at least the following directions.

- Calibration of the current framework. For example, the application of the temporal model in Chapter 4 requires the estimation of drivers' preferences of the link on the time-expanded network, which is possible given drivers' log data. Besides, the geographic matching framework applies for both street hailing and ride-sourcing technologies. It is therefore necessary to estimate the corresponding matching radius as well as the duration of a matching step. Based on this, a head-to-head comparison on market frictions will provide convincing evidences on the trade-offs of both technologies.

- Integrating with an event-based dynamic pricing. Our current formulations approximate the market behavior under recurrent demand patterns (so that the platform knows customers' sensitivity to price and average waiting time; drivers know exactly the market conditions). In reality, there are scenarios which involve unexpected random shocks of demand. Besides, estimating a temporally and spatially differentiated demand function may be cumbersome. Therefore, incorporating the event triggered (e.g., based on the matching time) dynamic pricing may give a comprehensive evaluation.

- Exploring (near) Pareto-optimal policies (where no agent is worse-off compared with the optimal static pricing). Note that the revenue-maximizing dynamic pricing may make customers worse-off in highly surged periods. One possible strategy is a user-specific dynamic pricing with the commission cap regulation initiated only for trips with short durations. Such a hybrid policy synchronizes the observed win-win features under dynamic pricing and the customer-favored pricing under the proposed regulation.

- Capturing the effect of traffic congestion. The level of traffic congestion (either caused by the ride-sourcing vehicles or conventional vehicles) may change some of the findings. For example, the increasing-returns-to-scale-matching process is only valid without severe traffic congestion; otherwise, the scaling of either requesting customers or unmatched drivers will make the transportation network more congested which may reduce the returns to scale.

APPENDIX A
RELAXATION OF P2

Let F, N, P solves P1 and the optimal platform profit is $\pi_P(F, N, P)$. Considering the equivalence of P1 and P2, we know F, N solves P2 with the optimal joint profit $\pi(F, N)$. Note $\pi_P(F, N, P) = \pi(F, N)$ by construction. Let F_R, N_R solves the relaxation of P2 in which Eq. (3-26) is dropped and the optimal joint profit is $\pi_R(F_R, N_R)$. It follows that $\pi(F, N) \le \pi_R(F_R, N_R)$. Define $P_R = F_R - c\frac{N_R}{Q_R}$. Using the fact $F_R = C_P'(Q_R) + \frac{1-\alpha_1-\alpha_2}{\alpha_1} cw_R^v + c(w_R^v + l) - \frac{Q_R}{f_R'}$, we can spell out P_R as:

$$P_R = \frac{1-\alpha_1-\alpha_2}{\alpha_1} cw_R^v - \frac{Q_R}{f_R'} + C_P'(Q_R) \tag{A-1}$$

As $\pi(F, N)$ is assumed to be nonnegative, we have:

$$\pi_R(F_R, N_R) = \frac{1-\alpha_1-\alpha_2}{\alpha_1} cw_R^v Q_R - \frac{Q_R^2}{f_R'} + Q_R C_P'(Q_R) - C_P(Q_R) \ge 0 \tag{A-2}$$

which leads to:

$$\frac{1-\alpha_1-\alpha_2}{\alpha_1} cw_R^v - \frac{Q_R}{f_R'} \ge -C_P'(Q_R) + \frac{C_P(Q_R)}{Q_R} \tag{A-3}$$

Substituting the above inequality into Eq. (A1), we obtain $P_R \ge C_P(Q_R)/Q_R \ge 0$. Clearly, F_R, N_R, P_R satisfies the constraints of P1 and thus is feasible. It follows that $\pi_P(F_R, N_R, P_R) = \pi(F_R, N_R) = \pi_R(F_R, N_R)$. Therefore, the optimal solution of the relaxation of P2 solves P1.

APPENDIX B
PROOF OF PROPOSITION 1

If $\lambda > 0$, then $\frac{\partial L}{\partial \lambda_P} = \pi_P - \pi_P^o = 0$ and $\frac{\partial L}{\partial \lambda_D} = \pi_D = 0$ from the complementary slackness conditions $\lambda_P \frac{\partial L}{\partial \lambda_P} = 0, \lambda_D \frac{\partial L}{\partial \lambda_D} = 0$. If $\lambda = 0$, then $\pi = \frac{1-\alpha_1-\alpha_2}{\alpha_1} cw^t Q + QC_P{}'(Q) - C_P(Q) < 0$ which conflicts with $\pi \geq \pi_P^o$. Therefore, $\pi_P = \pi_P^o$, $\pi_D = 0$.

APPENDIX C
EFFECTS FROM CUSTOMERS' HETEROGENEOUS VALUE OF TIME

We can assume a continuous distribution of β in the utility functions defined in Eqs. (3-5)-(3-6). Generally F and w^c in the demand function does not present a linear relationship and thus can be written as:

$$Q = f\left(F, w^c\right) \tag{C-1}$$

where $f_1 < 0,\ f_2 < 0$. Substituting this demand function to all the investigated formulations (P1~P4), one can verify that the tangency condition for the monopoly, first-best and second-best scenarios are replaced respectively by:

$$\frac{f_1}{f_2} = \frac{1}{c}\frac{\alpha_1}{\alpha_2}\frac{w^c}{w^v} \tag{C-2}$$

$$-\frac{Q}{I} = \frac{1}{c}\frac{\alpha_1}{\alpha_2}\frac{w^c}{w^v} \tag{C-3}$$

$$\frac{\psi_1}{\psi_2} = -\frac{f_2}{f_1}\frac{Q}{I} + \bar{\Delta} f_2\left(\frac{f_2}{f_1}\frac{Q}{I} + 1\right) \tag{C-4}$$

where $I = \int_F^{+\infty} f_2\left(x, w^c\right)dx$, $\bar{\Delta} = \frac{\partial w^c}{\partial Q}$, $\psi_1 = F - C_P'\left(Q\right) + c\Big/\left(\frac{\partial Q}{\partial N}\right)$ and

$\psi_2 = F - C_P'\left(Q\right) + Q\Big/\left(\frac{\partial Q}{\partial F}\right)$. Note when calculating the consumers' surplus, we fix the average waiting time at the equilibrium level and integrate under a hypothetical market demand curve (Cairns and Liston-Heyes, 1996; Yang et al., 2002).

If $f_2 = \beta f_1$ (the case of homogeneous value of waiting time), Eqs. (C-2)-(C-4) will reduce to the tangency condition.

APPENDIX D
SENSITIVITY ANALYSIS

Expressing the Cobb-Douglas type matching function utilizing the tangency condition, we have:

$$Q = AQ^{\alpha_1+\alpha_2}\left(w^v\right)^{\alpha_1}\left(\frac{c}{\beta}\frac{\alpha_2}{\alpha_1}w^v\right)^{\alpha_2} = A^{-(1-\alpha_1-\alpha_2)}\xi_2\left(w^v\right)^{\frac{\alpha_1+\alpha_2}{1-\alpha_1-\alpha_2}} \tag{D-1}$$

where $\xi_2 = \left(\frac{c}{\beta}\frac{\alpha_2}{\alpha_1}w^v\right)^{\frac{\alpha_2}{1-\alpha_1-\alpha_2}}$. Further, at equilibrium,

$$\theta\mu + \ln Q = \theta C + \ln\left(\overline{Q} - Q\right) \tag{D-2}$$

Without loss of generosity, we assume the dispersion parameter to be 1. Differentiating the above equations w.r.t A:

$$\frac{\partial Q}{\partial A} = \frac{1}{1-\alpha_1-\alpha_2}\frac{Q}{A} + \frac{\alpha_1+\alpha_2}{1-\alpha_1-\alpha_2}\frac{Q}{w^t}\frac{\partial w^v}{\partial A} \tag{D-3}$$

$$\frac{\partial \mu}{\partial A} = -\frac{\overline{Q}}{Q\left(\overline{Q}-Q\right)}\frac{\partial Q}{\partial A} \tag{D-4}$$

The specification of the generalized cost μ depends on the investigated scenario.

Case (1): Zero-profit second-best.

$\mu = c\left(w^v + l\right) + \overline{C}_P + c\frac{\alpha_2}{\alpha_1}w^v + \tau l$. Together with Eqs. (D-3) and (D-4):

$$\left(1+\frac{\alpha_2}{\alpha_1}\right)c\frac{\partial w^v}{\partial A} = -\frac{\overline{Q}}{Q\left(\overline{Q}-Q\right)}\left(\frac{1}{1-\alpha_1-\alpha_2}\frac{Q}{A} + \frac{\alpha_1+\alpha_2}{1-\alpha_1-\alpha_2}\frac{Q}{w^v}\frac{\partial w^v}{\partial A}\right) \tag{D-5}$$

Re-arranging the terms, we obtain:

$$\frac{\partial w^v}{\partial A} = \frac{-w^v\frac{1}{\alpha_1+\alpha_2}\frac{\overline{Q}}{A\left(\overline{Q}-Q\right)}}{\frac{1-\alpha_1-\alpha_2}{\alpha_1}cw^v + \frac{\overline{Q}}{\left(\overline{Q}-Q\right)}} \tag{D-6}$$

It is straightforward to see that the numerator is negative. Next, we will show the sign of the denominator is positive. Given the assumption that the joint profit for the monopoly solution is nonnegative:

$$F^*Q^* - cN^* - C_P(Q^*) = Q\left(\frac{1-\alpha_1-\alpha_2}{\alpha_1}cw^{v*} - \frac{Q^*}{f^{*\prime}}\right) \geq 0 \tag{D-7}$$

The asterisk denotes the monopoly solution. For the specified demand function $\frac{Q^*}{f^{*\prime}} = -\frac{\overline{Q}}{\overline{Q}-Q^*}$. Therefore we have:

$$\frac{1-\alpha_1-\alpha_2}{\alpha_1}cw^{v*} \geq -\frac{\overline{Q}}{\overline{Q}-Q^*} \tag{D-8}$$

When the demand increases from the monopoly to the second-best level, $w^v < w^{v*}, Q > Q^*$.

Therefore, at the second-best:

$$\frac{1-\alpha_1-\alpha_2}{\alpha_1}cw^{v} > \frac{1-\alpha_1-\alpha_2}{\alpha_1}cw^{v*} \geq -\frac{\overline{Q}}{\overline{Q}-Q^*} > -\frac{\overline{Q}}{\overline{Q}-Q} \tag{D-9}$$

That is, $\frac{1-\alpha_1-\alpha_2}{\alpha_1}cw^{v} + \frac{\overline{Q}}{\left(\overline{Q}-Q\right)} > 0$ which gives $\frac{\partial w^v}{\partial A} < 0$ at the second-best. This result still holds for other second-best solutions with varying reservation profit levels.

Case (2) Monopoly solution.

$\mu^* = c\left(w^{v*}+l\right) + \overline{C}_P + \frac{1-\alpha_1-\alpha_2}{\alpha_1}cw^{v*} + \frac{\overline{Q}}{\left(\overline{Q}-Q^*\right)} + c\frac{\alpha_2}{\alpha_1}w^{v*} + \tau l$. Then Eq. (D-4) can be rewritten as:

$$\frac{\partial w^{v*}}{\partial A} = -\frac{\alpha_1}{c}\frac{\overline{Q}^2}{Q\left(\overline{Q}-Q^*\right)^2}\frac{\partial Q^*}{\partial A} \tag{D-10}$$

Substituting $\frac{\partial Q^*}{\partial A}$ as specified in Eq. (D-3) into Eq. (D-10) leads to:

$$\frac{\partial w^{v*}}{\partial A} = \frac{-\frac{1}{\left(1-\alpha_1-\alpha_2\right)}\frac{1}{A}}{\frac{c}{\alpha_1}\frac{\left(\overline{Q}-Q^*\right)^2}{\overline{Q}^2}+\frac{\alpha_1+\alpha_2}{1-\alpha_1-\alpha_2}\frac{1}{w^{v*}}} \tag{D-11}$$

The numerator of $\frac{\partial w^{v*}}{\partial A}$ is positive. From Eq. (D-8):

$$\frac{1}{w^{v*}} \geq c\frac{\alpha_1+\alpha_2-1}{\alpha_1}\frac{\overline{Q}-Q^*}{\overline{Q}} \tag{D-12}$$

Then

$$\frac{c}{\alpha_1}\frac{\left(\overline{Q}-Q^*\right)^2}{\overline{Q}^2}+\frac{\alpha_1+\alpha_2}{1-\alpha_1-\alpha_2}\frac{1}{w^{v*}} \leq \frac{c}{\alpha_1}\frac{\overline{Q}-Q^*}{\overline{Q}}\left(\frac{\overline{Q}-Q^*}{\overline{Q}}-\left(\alpha_1+\alpha_2\right)\right) < 0 \tag{D-13}$$

which indicates $\frac{\partial w^{v*}}{\partial A} < 0$ at the monopoly solution. Following a similar procedure, one can verify the sign of $\frac{\partial w^{v}}{\partial \kappa}$, $\forall \kappa \in \Omega$.

APPENDIX E
EQUILIBRIUM MODEL WITH LONG TRIP DURATION

We consider long trip duration in formulating the equilibrium model in this appendix.

E.1 Model Development

Due to the potential long trip duration, the specification of N_b^o in Eq. (4-4) needs to be revised accordingly. With the additional assumption that demand is constantly realized over a time period, we have the following approximation:

$$N_b^o \approx \left[\frac{1}{2} Q_b \frac{\min\{l_b, T\}}{T} + Q_b \frac{\max\{0, T - l_b\}}{T} \right] \min\{l_b, T\},\ b \in A_1 \tag{E-1}$$

which can be spelled out as:

$$N_b^o \approx \begin{cases} Q_b l_b \left(1 - 0.5 l_b / T\right),\ l_b < T \\ 0.5 Q_b T,\ l_b \geq T \end{cases},\ b \in A_1 \tag{E-2}$$

When l_b is far less than T, $N_b^o \approx Q_b l_b$. When l_b is close to T, $N_b^o \approx 0.5 Q_b T$.

To facilitate the discussion, we denote the departure time as $k \in K$ associated with a path $p \in P$. The index of a link in A_1 is the same as that of its starting node. Take the network in Figure 4-1 as an example. The indexes are 1, 2,…, 23, 24 for links 1-2, 2-3, …, 23-24, 24-1. For a path $p \in P$, we define links in A_1 that proceed a transition link as divergent links. Note that the definition of the divergent link is path-dependent. For a path highlighted in Figure A-1, divergent links are 8-9 and 11-12. Drivers traversing a divergent link "must" leave the market.

For vehicles at link a that depart in period k along path p, we denote $\Delta_a^p(k)$ a 0-1 variable that represents their availability and an auxiliary variable $t_a^p(k)$ that represents the service/break time (ceiling-valued to be consistent with the modeling resolution).

For $\forall a \in A \setminus A_1$, $\Delta_a^p(k) = \delta_a^p$ since there is no customer demand. For $\forall a \in A_1$, $\Delta_a^p(k)$ is endogenously determined. Whenever a vehicle has been previously occupied but is unavailable

in the current period, $t_a^p(k)=0$ (e.g., link 7-8 in Figure A-1). If a vehicle is available at a non-divergent link, $t_a^p(k)=\lceil w_a^v+l_a \rceil$ (e.g., links 6-7, 10-11). If $\lceil w_a^v+l_a \rceil>1$ at a divergent link, the vehicle is assumed unable to provide service to ensure flow integrity but the corresponding $t_a^p(k)$ is set to T (e.g., link 8-9). We also assume $t_a^p(k)=T$ for $a\in A_3$ on a feasible path (e.g., link 9'-10'). The specifications of $t_a^p(k)$ help trace the availability of a vehicle. Without loss of generality, we assume $t_a^p(k)\in\{0,\ T,\ 2T\}$.

We formally define the availability of a vehicle at a non-divergent link $a\in A_1$ as:

$$\Delta_a^p(k)=\begin{cases}1,\ k+T_{-a}^p=aT\\ 0,\ k+T_{-a}^p\neq aT\end{cases},\ \forall a\in A_1, k\in K, p\in P \tag{E-3}$$

where $T_{-a}^p(k)=\sum_{\substack{i<a\\ i\ on\ p}} t_i^p(k)$ measures the cumulative time from the departure hour. $-a$ denotes the preceding link of a on path p of interval k. A vehicle is available at link a only when $k+T_{-a}^p=aT$. $t_a^p(k)$ at a non-divergent link is in return related to its availability:

$$t_a^p(k)=\begin{cases}\lceil w_a^v+l_a \rceil,\ \Delta_a^p(k)=1\\ 0,\ \Delta_a^p(k)=0\end{cases},\ \forall a\in A_1, k\in K, p\in P \tag{E-4}$$

For a divergent link $a\in A_1$, the issue of flow integrity is the dominant factor under consideration. The availability of a vehicle is explicitly determined by $\lceil w_a^v+l_a \rceil$ and implicitly affected by $T_{-a}^{pm}(k)$ through $t_a^{pm}(k)$. Together, we have:

$$\Delta_a^p(k)=\begin{cases}t_a^p(k),\ \lceil w_a^v+l_a \rceil=T\\ 0,\ \lceil w_a^v+l_a \rceil>T\end{cases},\ \forall a\in A_1, k\in K, p\in P \tag{E-5}$$

$$t_a^p(k)=\begin{cases}T,\ k+T_{-a}^p=aT\\ 0,\ k+T_{-a}^p\neq aT\end{cases},\ \forall a\in A_1, k\in K, p\in P \tag{E-6}$$

The following equation delineates the relationship between link and path flows:

$$u_a^{pm}(k) = f^{pm}(k)\Delta_a^p(k),\ a \in A, p \in P, m \in M, k \in K \tag{E-7}$$

Unlike the previous two formulations, $\Delta_a^p(k)$ is endogenously determined via Eqs. (E-3) -(E-6) and Eq. (E-7) is thus non-linear and non-convex.

E.2 Equilibrium definition and formulation

For simplicity, drivers are assumed to behave as the neoclassical theory suggests, and thus the equilibrium conditions are the same as in Definition 1. When the availability of vehicles is explicitly traced, the average path profit is revised as:

$$\pi^{pm}(k) = \sum_{b \in A_1} R_b \Delta_b^p(k) - \sum_{a \in A} c_a^m \delta_a^p - c^{pm},\ \forall k \in K, p \in P, m \in M \tag{E-8}$$

Similarly, the equilibrium condition can be characterized by the following VI:

(ME-L)

$$\sum_p \sum_m \left(-\pi^{pm*}(k)\right)\left(f^{pm}(k) - f^{pm*}(k)\right) \geq 0,\ \forall k \in K, p \in P, m \in M \tag{E-9}$$

where $f^{pm} \in \Omega$. However, the existence of a solution is not always guaranteed, as $\pi^{pm}(k)$ may not be continuous in $f^{pm}(k)$ due to the complexity of $\Delta_b^{pm}(k)$. If exist, there may be multiple solutions for path flows as strict monotonicity does not hold.

A flow-swapping algorithm can be adopted to solve the VI (Huang and Lam, 2002; Lu et al., 2009). For a given path set, the algorithm consists of two loops. The inner loop aims to load the vehicles to the network and determines $\Delta_a^p(k)$. The outer loop adjusts the path flow from less profitable paths to those with maximal profit.

E.3 Numerical Results

To be consistent with our parameterization, we assume an average trip distance of 4.5 miles, and the period-specific speed v_b is given in Figure E-2. The corresponding minimal and maximal trip duration is 0.16 and 1.2 hours, respectively.

When the availability of drivers is explicitly traced, $\Delta_a^p(k)$ may not always be equal to δ_a^p. The ceiling of service durations ($w_b^v + l_b,\ \forall b \in A_1$) are two hours for periods 7:00-9:00 a.m. and one hour for the remaining periods. For a path used by a driver of Class 2 (O→1→ 2→ 3→4→5→6→7→8→ 8'→D), she stops working after 8:00 a.m. and is not available on link 7→ 8. Generally, drivers are less likely to be available on a divergent link or on a subsequent link of ones with long service duration.

We next compare the outcomes of the enhanced formulation (ME-L) with the original one (ME-N). We plot distributions of average searching times and waiting times in Figure E-3 and Figure E-4, respectively. A longer average searching time is found during peak hours (with longer trip time), because more drivers are attracted to those links so that they can enjoy a higher fare once occupied. The competition among the drivers raises the average searching time. Customers see a similar tendency in terms of the average waiting time yet at a smaller magnitude, which mainly results from the longer time for the vehicle to come and pick up the customer as the average speed drops. Despite a similar daily pattern, ME-N tends to overestimate average searching time and underestimate average waiting time, particularly at peak hours. This is because the available vehicle hours in ME-N are generally larger than those in MN-L due to Assumption 2.

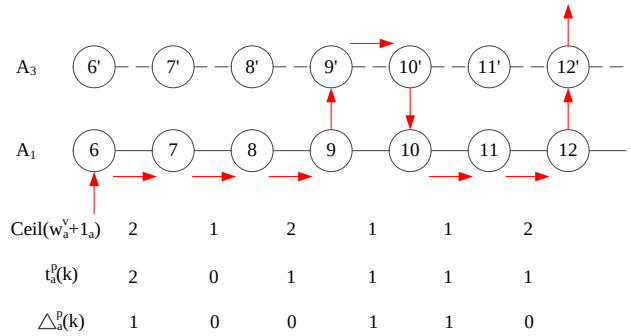


Figure E-1. Display of A Sample Work Schedule with Corresponding Modeling Components.

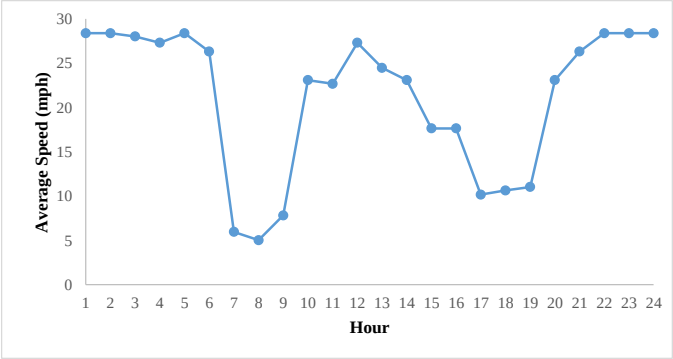


Figure E-2. Distribution of Average Speed.

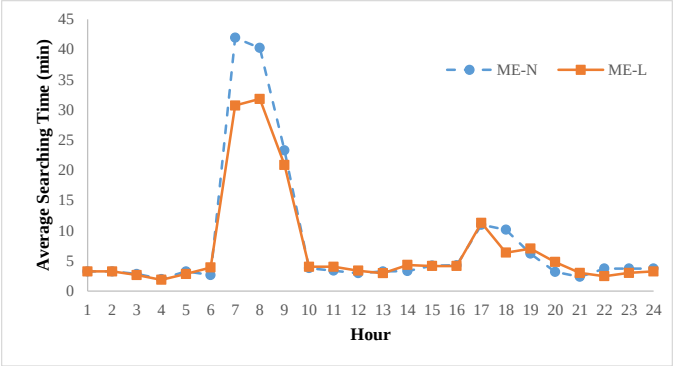


Figure E-3. Average Searching Time and Service Duration with Time-Varying Trip Durations.

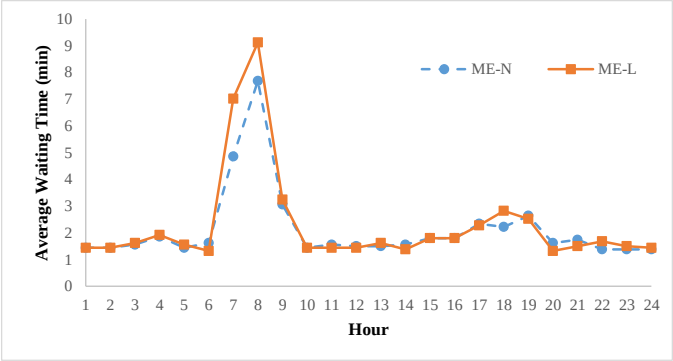


Figure E-4. Average Waiting Time with Time-Varying Trip Durations.

APPENDIX F
NOTATIONS

Table F-1. Summary of Parameters.

Parameter	Description	Value
S	Area of the ride-souring market (mile^2)	300
$\beta_b,\ b \in A_1$	Value of customers' waiting time ($/hr)	See Chapter 4
$\tau_b,\ \forall b \in A_1$	Value of customers' in-vehicle travel time ($/hr)	See Chapter 4
$\theta_b,\ \forall b \in A_1$	Demand sensitivity of generalized cost	See Chapter 4
$c_a^m,\ \forall a \in A, m \in M$	Link-specific cost of driver class *m* ($/hr)	See Chapter 4
$c^{pm},\ \forall p \in P, m \in M$	Path-dependent cost for a driver of of class *m* on path *p* ($)	See Chapter 4
$\bar{C}^{pm},\ \forall p \in P, m \in M$	Total cost incurred by a driver of class *m* choosing path *p* ($)	See Chapter 4
$l_b,\ b \in A_1$	Average trip time (hr)	0.3
$v_a,\ a \in A$	Average speed of vehicles (mile/hr)	15
$\varpi_1^m,\ \forall m \in M$	Unit cost of cumulative working hours of driver class *m*	See Chapter 4
$\varpi_2^m,\ \forall m \in M$	Level of aversion to cumulative work hours of driver class *m*	2
$N^m,\ \forall m \in M$	Fleet size of driver class *m* (veh)	2000
F_0	Flag drop fee per trip ($/trip)	2
$\omega_b,\ \forall b \in A_1$	Time based charge per hour ($/hr)	40
η	Proportion charged by the platform per completed trip	20%
$\bar{Q}_b,\ \forall b \in A_1$	Time-dependent hourly base demand (trip/hr)	See Chapter 4
$I^m,\ \forall m \in M$	Income target for driver class *m* ($)	See Chapter 4
U_0	Constant in the reference-dependent utility ($)	500
$\rho^m,\ \forall m \in M$	Degree of loss aversion for driver class *m*	0.2
$\delta_a^p,\ \forall a \in A, p \in P$	Link-path incidence element	See Chapter 4
$h^{pm}, \forall p \in P^e, m \in M$	Cumulative working hours of driver class *m* on path *p* (hr)	See Chapter 4
E	Node-link incidence matrix	See Chapter 4
d	Column vector in shortest path finding	See Chapter 4
π_R^m	Reservation profit levels for driver class *m* ($)	See Chapter 4
ε_1	Initial tolerance used for solving bi-level problems	1

APPENDIX G
EXISTENCE OF EQUILIBRIUM

The proof of the existence of equilibrium utilizes Brouwer's fixed point theorem (De la Fuente, 2000). We first construct a convex and closed set Ω for the vector of demand $\mathbf{Q}$. Then we will argue there exists a continuous function mapping from Ω to itself.

We define $\bar{Q}_{it} = \theta_{0t} e^{\theta_{1t}\bar{F}_{it} + \theta_{2t}\delta + \boldsymbol{\theta}_{3t}\mathbf{X}}$ and $Q_{it} = 0$ when $w_{it}^{c} = \infty$. A valid set for demand is defined as $\Omega = \left\{ \mathbf{Q} = \left(Q_{it}, i \in I, t \in T \right) \mid 0 \le Q_{it} \le \bar{Q}_{it} \right\}$ which is clearly closed and convex.

Given $\mathbf{Q}$, Eqs. (5-20) -(5-22) imply that $\pi_t(\mathbf{Q}) = \left(\sum_{j \in I} (1-\eta) \bar{F}_{jt} Q_{jt} \right) \Big/ \left(Nhn_t \right)$. Using Eq. (5-22) again, we solve for $w_{it}^{v}(\mathbf{Q}) = (1-\eta)\bar{F}_{it} / \pi_t(\mathbf{Q}) - l_{it}$. Clearly, $\pi_t(\mathbf{Q})$ and $w_{it}^{v}(\mathbf{Q})$ are continuous in $\mathbf{Q}$. $w_{it}^{v}(\mathbf{Q})$ is guaranteed to be greater than a positive constant $w_{\min}^{v}$ if $Nh > k$ where $k = \max_{i \in I, t \in T} \left\{ \left(l_{it} + w_{\min}^{v} \right) \sum_{j \in I} \bar{F}_{jt} \bar{Q}_{jt} \Big/ \left(\bar{F}_{it} n_t \right) \right\}$. Eq. (5-15) implies a possible way of choosing $w_{\min}^{v}$ is via $w_{\min}^{v} = \max_{i \in I, t \in T} \min_{Q_{it} \in \Omega} \left\{ w_{it}^{vm}(\mathbf{Q}) + m_{it}\left(\mathbf{Q}, w_{it}^{vm} \right) \right\}$.

There may be an identification problem when we inversely solve for $w_{it}^{vm}(\mathbf{Q})$ given $w_{it}^{v}(\mathbf{Q})$. A graphic illustration is presented in Figure G-1 where $w_{it}^{v}(\mathbf{Q})$ may intercept the curve twice (Point A, B in Figure G-1). In this case, we choose the right point (Point B) to make sure the continuity condition holds. Subsequently, we obtain $w_{it}^{cm}(\mathbf{Q})$, $m_{it}(\mathbf{Q})$ via Eq. (5-17) and Eq.(5-18), respectively. As shown in Figure G-2, $w_{it}^{vm}(\mathbf{Q})$ is continuously decreasing in $w_{it}^{cm}(\mathbf{Q})$. $w_{it}^{c}(\mathbf{Q})$ is then the sum of $w_{it}^{cm}(\mathbf{Q})$ and $m_{it}(\mathbf{Q})$, both continuous in $\mathbf{Q}$.

Finally, we define the mapping: $\mathbf{Q} = \Gamma(\mathbf{Q})$, where $\Gamma_{it}(\mathbf{Q}) = \theta_{0t} e^{\theta_{1t}\bar{F}_{it} + \theta_{2t} w_{it}^{c}(\mathbf{Q}) + \boldsymbol{\theta}_{3t}\mathbf{X}}$, $\forall i \in I, t \in T$. It is straightforward to see Γ is continuous. Given the way we chose $w_{\min}^{v}$ and k together with the condition that $Nh > k$, we claim the nonlinear system admits at least one solution.

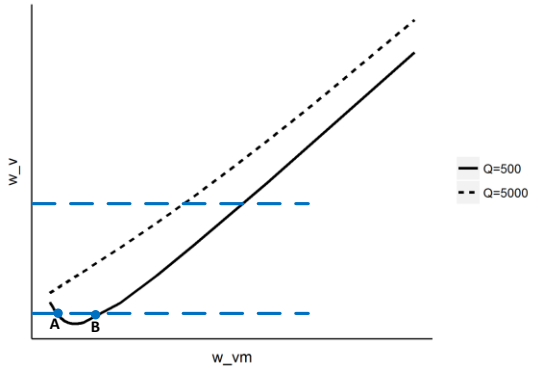


Figure G-1. Average Searching Time and Its Matching Time Portion.

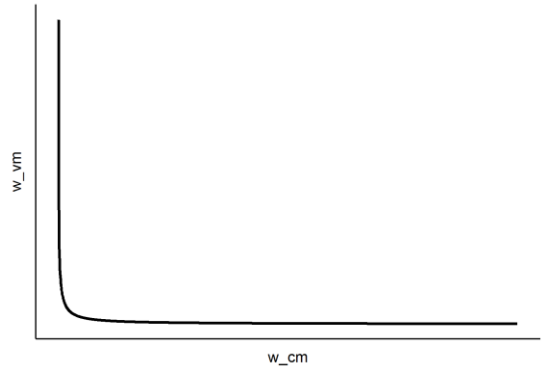


Figure G-2. Matching Frictions.

APPENDIX H
OVERVIEW OF THE SIMULATION PACKAGE

This section gives an overview of the simulation interface. The general outlook is presented Figure H-1. It contains 9 parts as been labeled. The functionality of each part is explained subsequently.

1. Setup/Go

- Setup: Restores the simulation to its initial state.
- Go-once: Moves the simulation forward 1 time step.
- Go: Runs the simulation continuously when first pressed; stops the simulation when pressed again.

2. Sliders

- number-vehicles: Adjusts the total number of vehicles to be used in the simulation. This number remains fixed throughout the simulation.
- communication-range: Adjusts the matching radius over which customers can summon vehicles.

3. Simulation Space

- As discussed in Section 6.1.

4. Simulation Summary

- Assemble the evolution of all the variables of interest at the network level. At equilibrium, each plot remains approximately constant.

5. Revenue and Operating Cost

- Plots the average revenue gained by each vehicle (per time step) and the operating cost that each vehicle incurs (per time step), which guarantees the same reservation profit at the equilibrium.

6. Waiting Customers and Vacant Vehicles by Zone

- Plots the number of waiting customers and vacant vehicles in each zone. At the equilibrium, each plot remains approximately constant.

7. Average Search Time by Zone

- Plots the average searching time for vehicles in each zone. At the equilibrium, each plot remains approximately constant.

8. Demand by Zone

- Plots the realized demand of customers in each zone. At the equilibrium, each plot remains approximately constant.

9. Average Waiting Time by Zone

- Plots the average waiting time for vehicles in each zone. At the equilibrium, each plot remains approximately constant.

10. Equilibrium check

- As discussed in Section 6.3.1.

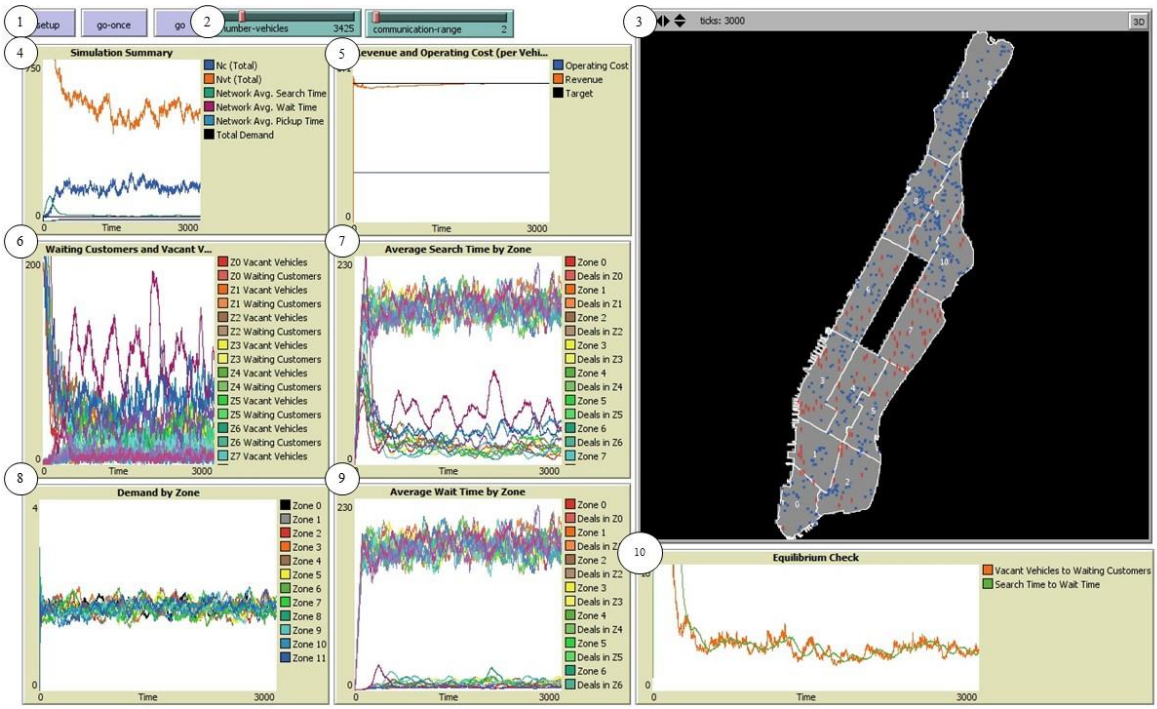


Figure H-1. Display of the Simulation Interface.

LIST OF REFERENCES

Anderson, D., 2014. "Not Just a Taxi"? For-Profit Ridesharing, Driver Strategies, and VMT. *Transportation* 41, 1099-1117.

Armstrong, M., 2006. Competition in Two-Sided Markets. *The RAND Journal of Economics* 37, 668-691.

Arnott, R., 1996. Taxi Travel Should Be Subsidized. *Journal of Urban Economics* 40, 316-333.

Badger, E. *Taxi Medallions Have Been the Best Investment in America for Years. Now Uber May Be Changing That.* The Washington Post. http://www.washingtonpost.com/blogs/wonkblog/wp/2014/06/20/taxi-medallions-have-been-the-best-investment-in-america-for-years-now-uber-may-be-changing-that/. Accessed Jan 4, 2015.

Banerjee, S., Johari, R., Riquelme, C., 2015. Pricing in Ride-Sharing Platforms: A Queueing-Theoretic Approach, *Proceedings of the Sixteenth ACM Conference on Economics and Computation*. ACM, pp. 639-639.

Bertsekas, D.P., 1999. *Nonlinear Programming*. Athena Scientific Belmont.

Buchholz, N., 2015. Spatial Equilibrium, Search Frictions and Efficient Regulation in the Taxi Industry, Working Paper at the University of Texas at Austin.

Cachon, G.P., Daniels, K.M., Lobel, R., 2015. The Role of Surge Pricing on a Service Platform with Self-Scheduling Capacity. *Available at SSRN*.

Cairns, R.D., Liston-Heyes, C., 1996. Competition and Regulation in The Taxi Industry. *Journal of Public Economics* 59, 1-15.

Camerer, C., Babcock, L., Loewenstein, G., Thaler, R., 1997. Labor Supply of New York City Cabdrivers: One Day at a Time. *The Quarterly Journal of Economics* 112, 407-441.

Chen, M.K., Sheldon, M., 2016. Dynamic Pricing in a Labor Market: Surge Pricing and Flexible Work on the Uber Platform, UCLA Anderson School of Management.

Chiu, S.N., Stoyan, D., Kendall, W.S., Mecke, J., 2013. *Stochastic Geometry and Its Applications*. Wiley.

Crawford, V.P., Meng, J., 2011. New York City Cab Drivers' Labor Supply Revisited: Reference-Dependent Preferences with Rational Expectations Targets for Hours and Income. *The American Economic Review* 101, 1912-1932.

Curley, B. *A Deep Look at Uber's Dynamic Pricing Model*. http://abovethecrowd.com/2014/03/11/a-deeper-look-at-ubers-dynamic-pricing-model/. Accessed Nov. 17, 2014.

Daganzo, C.F., 1978. An Approximate Analytic Model of Many-To-Many Demand Responsive Transportation Systems. *Transportation Research* 12, 325-333.

De la Fuente, A., 2000. *Mathematical Methods and Models for Economists*. Cambridge University Press.

Didi Chuxing. *Didi Algorithm Competition*. http://research.xiaojukeji.com/competition/main.action?competitionId=DiTech2016. Accessed Jan. 14, 2017.

Didi Chuxing. *Didi Chuxing Inc.* http://www.xiaojukeji.com/index/index. Accessed Jan. 16, 2017.

Didi Chuxing. *Report on Residents' Traveling Data: Case of Hang Zhou*. http://www.useit.com.cn/thread-12885-1-1.html. Accessed Feb., 18, 2017.

Douglas, G.W., 1972. Price Regulation and Optimal Service Standards: The Taxicab Industry. *Journal of Transport Economics and Policy* 6, 116-127.

Drud, A.S., 1994. CONOPT—A Large-Scale GRG Code. *ORSA Journal on Computing* 6, 207-216.

ECMT, 2007. *De/Regulation of the Taxi Industry*. OECD Publishing.

Farber, H.S., 2005. Is Tomorrow Another Day? The Labor Supply of New York City Cabdrivers. *Journal of political Economy* 113, 46-82.

Farber, H.S., 2015. Why You Can't Find a Taxi in the Rain and Other Labor Supply Lessons from Cab Drivers. *The Quarterly Journal of Economics* 130, 1975-2026.

Frankena, M.W., 1983. The Efficiency of Public Transport Objectives and Subsidy Formulas. *Journal of Transport Economics and Policy* 17, 67-76.

Frankena, M.W., Pautler, P.A., 1986. Taxicab Regulation: An Economic Analysis. *Research in Law and Economics* 9, 129-165.

Hagiu, A., 2006. Proprietary vs. Open Two-Sided Platforms and Social Efficiency. *AEI-Brookings Joint Center Working Paper*.

Hall, J.V., Krueger, A.B., 2015. An Analysis of the Labor Market for Uber's Driver-Partners in the United States, Princeton University Industrial Relations Section.

Hansen, B., 2017. *Econometrics*, University of Wisconsin.

He, F., Shen, Z.-J.M., 2015. Modeling Taxi Services with Smartphone-Based E-Hailing Applications. *Transportation Research Part C: Emerging Technologies* 58, Part A, 93-106.

Huang, H.-J., Lam, W.H.K., 2002. Modeling and Solving the Dynamic User Equilibrium Route and Departure Time Choice Problem in Network with Queues. *Transportation Research Part B: Methodological* 36, 253-273.

Koopman, C., Mitchell, M., Thierer, A., 2015. The Sharing Economy and Consumer Protection Regulation: The Case for Policy Change. Mercatus Center at George Mason University.

Kőszegi, B., Rabin, M., 2006. A Model of Reference-Dependent Preferences. *The Quarterly Journal of Economics*, 1133-1165.

Lagos, R., 2000. An Alternative Approach to Search Frictions. *Journal of Political Economy* 108, 851-873.

Lerner, A.P., 1934. The Concept of Monopoly and the Measurement of Monopoly Power. *The Review of Economic Studies* 1, 157-175.

Lo, H.K., Chen, A., 2000. Traffic Equilibrium Problem with Route-Specific Costs: Formulation and Algorithms. *Transportation Research Part B: Methodological* 34, 493-513.

Lu, C.-C., Mahmassani, H.S., Zhou, X., 2009. Equivalent Gap Function-Based Reformulation and Solution Algorithm for The Dynamic User Equilibrium Problem. *Transportation Research Part B: Methodological* 43, 345-364.

Marcotte, P., Zhu, D.L., 1996. Exact and Inexact Penalty Methods for the Generalized Bilevel Programming Problem. *Mathematical Programming* 74, 141-157.

Matsushima, K., Kobayashi, K., 2006. Endogenous Market Formation with Matching Externality : An Implication for Taxi Spot Markets. *Structural Change in Transportation and Communications in the Knowledge Society*, 313-336.

Meng, Q., Yang, H., Bell, M., 2001. An Equivalent Continuously Differentiable Model and a Locally Convergent Algorithm for The Continuous Network Design Problem. *Transportation Research Part B: Methodological* 35, 83-105.

Mohring, H., 1972. Optimization and Scale Economies in Urban Bus Transportation. *The American Economic Review* 62, 591-604.

Moscarini, G., Wright, R., 2010. Introduction to Search Theory and Applications. *Journal of Economic Theory* 145, 1319-1324.

Nagurney, A., 2013. *Network Economics: A Variational Inequality Approach*. Springer Science & Business Media.

Nicholas, B., 2015. Spatial Equilibrium, Search Frictions and Efficient Regulation in the Taxi Industry, *Department of Economics*. University of Texas at Austin.

Oum, T.H., Tretheway, M.W., 1988. Ramsey Pricing in The Presence of Externality Costs. *Journal of Transport Economics and Policy*, 307-317.

Patriksson, M., 2015. *The Traffic Assignment Problem: Models and Methods*. Courier Dover Publications.

Pew Research Center. *U.S. Smartphone Use in 2015*. http://www.pewinternet.org/2015/04/01/us-smartphone-use-in-2015/. Accessed Sep 2, 2015.

Ranney, M. *Scaling Uber's Real-time Market Platform.* QCON London. http://www.slideshare.net/InfoQ/scaling-ubers-realtime-market-platform-50758747. Accessed Nov., 23, 2015.

Rauch, D., Schleicher, D., 2015. Like Uber, But For Local Government Policy: The Future of Local Regulation of The Sharing Economy.

Rayle, L., Shaheen, S., Chan, N., Dai, D., Cervero, R., 2014. App-Based, On-Demand Ride Services: Comparing Taxi and Ridesourcing Trips and User Characteristics in San Francisco, p. 18.

Rochet, J.-C., Tirole, J., 2003. Platform Competition in Two-Sided Markets. *Journal of the European Economic Association* 1, 990-1029.

Rochet, J.C., Tirole, J., 2006. Two-Sided Markets: A Progress Report. *The RAND Journal of Economics* 37, 645-667.

Salanova, J.M., Estrada, M., Aifadopoulou, G., Mitsakis, E., 2011. A Review of The Modeling of Taxi Services. *Procedia - Social and Behavioral Sciences* 20, 150-161.

Schaller, B., 2007. Entry Controls in Taxi Regulation: Implications of US and Canadian Experience for Taxi Regulation and Deregulation. *Transport Policy* 14, 490-506.

Schroeter, J.R., 1983. A Model of Taxi Service under Fare Structure and Fleet Size Regulation. *The Bell Journal of Economics* 14, 81-96.

Shaheen, S. *Transportation Network Companies and Ridesourcing.* California Public Utilities Commission (CPUC) En Banc. http://www.cpuc.ca.gov/NR/rdonlyres/5C961222-B9C8-4E53-A54D-FC2A89C0A30C/0/RidesourcingCPUCShaheen_Final_v2.pdf. Accessed Jan 3, 2015.

Sheffi, Y., 1985. *Urban Transportation Network: Equilibrium Analysis with Mathematical Programming Methods.* Prentice Hall.

Small, K.A., 2012. Valuation of Travel Time. *Economics of Transportation* 1, 2-14.

Spence, A.M., 1975. Monopoly, Quality, and Regulation. *The Bell Journal of Economics* 6, 417-429.

Talluri, K.T., Van Ryzin, G.J., 2006. *The Theory and Practice of Revenue Management*. Springer Science & Business Media.

Taylor, T., 2016. On-Demand Service Platforms. *Available at SSRN 2722308*.

The Economist. *There's An App for That*. http://www.economist.com/news/briefing/21637355-freelance-workers-available-moments-notice-will-reshape-nature-companies-and. Accessed Sep 2, 2015.

The Economists. *Pricing the Surge*. http://www.economist.com/news/finance-and-economics/21599766-microeconomics-ubers-attempt-revolutionise-taxi-markets-pricing-surge. Accessed Nov.17, 2014.

TLC. *2016 Taxi & Limousine Commission Factbook.* NYC Taxi & Limousine Commission. http://www.nyc.gov/html/tlc/downloads/pdf/2016_tlc_factbook.pdf. Accessed Sep. 12, 2016.

Train, K.E., 2009. *Discrete Choice Methods with Simulation*. Cambridge University Press.

Vany, A.S.D., 1975. Capacity Utilization under Alternative Regulatory Restraints: An Analysis of Taxi Markets. *Journal of Political Economy* 83, 83-94.

Wang, X., He, F., Yang, H., Gao, H., 2016. Pricing Strategies for A Taxi-Hailing Platform. *Transportation Research Part E: Logistics and Transportation Review* 93, 212-231.

Wilensky, U. *NetLogo.* Center for Connected Learning and Computer-Based Modeling., Northwestern University, Evanston, IL. https://ccl.northwestern.edu/netlogo/index.shtml. Accessed July,23, 2015.

Wong, K., Wong, S., Yang, H., Wu, J., 2008. Modeling Urban Taxi Services with Multiple User Classes and Vehicle Modes. *Transportation Research Part B: Methodological* 42, 985-1007.

Wright, J., 2004. One-sided Logic in Two-sided Markets. *Review of Network Economics* 3.

Xu, H., Lou, Y., Yin, Y., Zhou, J., 2011. A Prospect-Based User Equilibrium Model with Endogenous Reference Points and Its Application in Congestion Pricing. *Transportation Research Part B: Methodological* 45, 311-328.

Xu, Z., Yin, Y., Zha, L., 2017. Optimal Parking Provision for Ride-sourcing Services. *Submitted to Trasnportation Research Part B.*

Yang, H., Fung, C.S., Wong, K.I., Wong, S.C., 2010a. Nonlinear Pricing of Taxi Services. *Transportation Research Part A: Policy and Practice* 44, 337-348.

Yang, H., Leung, C.W.Y., Wong, S.C., Bell, M.G.H., 2010b. Equilibria of Bilateral Taxi–Customer Searching and Meeting on Networks. *Transportation Research Part B: Methodological* 44, 1067-1083.

Yang, H., Wong, S.C., 1998. A Network Model of Urban Taxi Services. *Transportation Research Part B: Methodological* 32, 235-246.

Yang, H., Wong, S.C., Wong, K.I., 2002. Demand–Supply Equilibrium of Taxi Services in a Network under Competition and Regulation. *Transportation Research Part B: Methodological* 36, 799-819.

Yang, H., Yang, T., 2011. Equilibrium Properties of Taxi Markets with Search Frictions. *Transportation Research Part B: Methodological* 45, 696-713.

Yang, H., Ye, M., Tang, W.H.-C., Wong, S.C., 2005a. A Multiperiod Dynamic Model of Taxi Services with Endogenous Service Intensity. *Operations Research* 53, 501-515.

Yang, H., Ye, M., Tang, W.H., Wong, S.C., 2005b. Regulating Taxi Services in The Presence of Congestion Externality. *Transportation Research Part A: Policy and Practice* 39, 17-40.

Yang, T., Yang, H., Wong, S.C., Sze, N.N., 2014. Returns to Scale in The Production of Taxi Services: An Empirical Analysis. *Transportmetrica A: Transport Science* 10, 775-790.

BIOGRAPHICAL SKETCH

Liteng Zha obtained his Bachelor of Science (2012), Master of Science (2014) and Ph.D. (2017) degrees respectively from the Southeast University, Texas A&M University and the University of Florida, all majored in transportation engineering.

Liteng's primary research interest is in the applications of the emerging technologies in the field of transportation. His work covers the safety applications of connected vehicle technologies at the signalized intersections as well as the economic analysis of recent booming of the ride-sourcing markets.

Liteng won the System Optimum Award offered by the University of Florida Transportation Institute (UFTI) in 2015. In 2014, his paper: "Advancing Safety Performance Monitoring at Signalized Intersections Using Connected Vehicle Technology" was selected as the Young Researcher Award for the Safety Data, Analysis and Evaluation Committee at the 93rd Transportation Research Annual Meeting (TRB, 2014). He was the recipient of the prestigious Eisenhower Fellowship from the Federal Highway Administration (FHWA) in 2013.