Search Frictions, Network Effects and Spatial Competition: Taxis versus Uber

Bo Bian*
Shanghai University of Finance and Economics

September 22, 2022

Abstract

This paper models the search and matching process among passengers, taxis and Uber drivers in New York City to analyze matching efficiency taking into account network effects and supply competition. Drivers make dynamic spatial search decisions across locations for passengers and the latter make static choice decisions between taxi and Uber. Network effects exist if increased participation of one side impacts searches of the other side. Spatial allocation of matched trips are affected by network effects. This paper finds existence of network effects. Then, the estimates are used to analyze frictions as spatial mismatches between drivers and passengers in three counterfactual scenarios: restricting supply of Uber, improving traffic condition, and eliminating surge multiplier. The results show that regulating Uber increases mismatches of taxis. Eliminating the surge multiplier increases mismatches of Uber. Traffic improvement increases matching efficiency. Most importantly, ignoring network effects will lead to incorrect welfare conclusions.

Key Words: spatial competition, network effects, search and matching, regulation. (*JEL: C73, D83, L90, R12*)

*Assistant Professor, College of Business, Shanghai University of Finance and Economics. Mailing address: Office 322, College of Business, 100 Wudong St., Shanghai, China 200433. Email: **bian.bo@mail.shufe.edu.cn**. I would like to thank Mark Roberts, Peter Newberry, Charles Murry, Paul Grieco, Joris Pinkse, Daniel Grodzicki and participants in PSU brownbag seminar for their comments and discussions. I also thank Alon Eizenberg for helpful discussion in International Industrial Organization Conference 2018. I am solely responsible for any errors.

1 Introduction

Frictions play an important role in explaining the failure of market clearing. When there exists information imperfections, coordination failure, and congestion, some potential traders on one side of the market may not successfully contact potential traders on the other side, leaving some buyers and sellers unable to trade. The early search and matching literature use a reduced form matching function to capture effects of frictions on equilibrium outcomes of bilateral trades[1]. In more recent literature, the microfoundation of the matching function is studied. Lagos (2000) builds a model of taxis' spatial search for passengers and finds that even without information imperfections and a random search assumption, aggregate mismatches over locations arise endogenously as outcomes of drivers' optimal search decisions. Specifically, when one location has relatively higher expected profit conditional on picking up a passenger, taxis may overcrowd that location leaving another location undersupplied. Similar to the idea in Lagos (2000), Buchholz (2022) empirically studies search frictions in the taxi industry as a consequence of price regulation which creates heterogeneity among locations and fails to coordinate demand and supply across locations. Both Lagos (2000) and Buchholz (2022) emphasize the prices as a source of spatial mismatches and mainly focus on the supply side of the matching process[2].

This paper studies matching efficiency in the taxi industry by accounting for both price and non-price factors. The non-price factors introduced in the model are network effects and supply competition. The presence of network effects is due to interdependence between demand and supply decisions as function of the participation of the other side. To be specific, drivers prefer locations, ceteris paribus, with higher demand and passengers prefer ride option with higher supply due to higher matching probabilities and lower waiting time. Therefore, each geographic location forms a two-sided market with indirect network effects. Moreover, agents of the same side could also affect each other through direct network effects such as congestion effect or informative effect[3].

[1] Blanchard and Diamond (1989), Pissarides (1990), and Mortensen and Pissarides (1994) are examples.

[2] Taxi industry is well known for existence of matching frictions such that some areas have excess demand whereas some have excess supply. This industry is ideal for analyzing search and matching frictions for several reasons. First, search decisions are made by decentralized individuals without coordination. Second, the taxi market in many cities is highly regulated in fares and medallions. Third, there is no preference heterogeneity among drivers and passengers.

[3] For the rest of this paper, I refer to indirect network effects as network effects without explicitly stating and use location, firm and platform interchangeably.

The relative magnitude of bilateral responses between demand and supply causes mismatches within a location leaving unmatched drivers or passengers. Furthermore, the misallocation of drivers leaving some locations oversupplied and other locations undersupplied leads to mismatches across locations. Thus, quantifying network effects and understand how they affect matching efficiencies are very important.

Instead of modelling only taxi drivers' search decisions as in Buchholz (2022) or only Uber drivers as in Castillo (2022), I introduce competition between taxi and Uber. They compete for passengers in each location by providing differentiated products with different prices and qualities (i.e. waiting time, matching technology). The model distinguishes competitions from drivers within the same firm and those from the opponent. These two competitions have different mechanisms. Within-firm competition may have positive effect due to indirect network effects (or spillover effect). However, the cross-firm competition not only negatively affect demand, which is further exaggerated via network effects. By providing a differentiated product, it makes passengers travelling different trips choose different firms and the distribution of passengers' destinations becomes firm specific. Thus, the conditional expected profit determined by the trip route, is impacted by inside and outside firm competition. Since drivers endogenously crowd locations with high profitability, competition would affect supply allocation and matching efficiency. Understanding the effects of competition on matching friction is important to evaluate regulatory policies on market structure.

The empirical model estimates demand and supply in this market using data on trip records of taxis and Uber from the New York City Taxi and Limousine Commission (TLC). This dataset provides detailed information on taxis' trips including pickup/dropoff locations and timestamps. Uber's trips include only pickup zones and timestamps[4]. The trip data is limited since it does not provide the number of potential passengers and drivers in the matching process. I use this data and the strategy of Buchholz (2022) to estimate equilibrium demand and supply in a dynamic spatial search model. I extend Buchholz's strategy by nesting a discrete choice model in the dynamic search to allow network effects and incorporating competition between taxi and Uber. I apply a new identification strategy to estimate the model. Due to the large number of drivers in this dynamic game, the concept of nonstationary oblivious

[4]The data used in this paper is from 2016. After 2017, TLC upgrade ridesharing data by including dropoff information. Due to the restriction on matching trip data with my collected Uber surge multiplier, this paper does not use updated TLC data.

equilibrium is used to solve the equilibrium[5]. The estimation results show significant network effects and the substitution between taxi and Uber is relatively strong given the estimate of nested logit parameter being 0.38 (out of range (0,1)).

I simulate several counterfactuals to both study factors influencing matching efficiency and policy issues in the real world. Uber has been a disruptive force in the taxi industry and impacted congestion in cities. This had led to restriction on Uber. In the first counterfactual, I simulate a regulatory policy that restricts Uber's supply and analyze how this change of competition affects matching efficiency and taxi profits. In the second counterfactual, I improve the traffic condition to study the magnitude of traffic congestion's effect on matching efficiency. Since traffic condition is not endogenously modelled, this improvement could come from an exogenous infrastructure improvement. In the last one, Uber's surge multiplier is eliminated to study the efficiency of flexible pricing. In each simulation, predicted market outcomes are compared with and without network effects.

The counterfactuals show that after decreasing Uber's supply by 30%, taxis' demand and pickups increase due to less competition. However, it only increases taxis' pickups (profits) by 1,660 ($ 34,300) compared to a 7,004 ($ 119,320) loss for Uber in a representative day shift. The cross-location mismatches of taxi and Uber increase by 5.81% and 36.97% respectively. After improving traffic conditions, matching efficiencies increase for both firms. Supplies of both firms increase due to less travel time and taxi demand increases due to network effects by 5.86%. Finally, after eliminating Uber's surge multiplier, Uber's cross-location mismatches increase by 3,152 (170.38%) which indicates the success of surge pricing. On the contrary, taxis' mismatches decrease by 2,573 (17.46%) resulting from increased price competition of Uber. Passengers are worse off after the regulatory policy but better off in the last two scenarios. All simulation outcomes indicate the importance of network effects on the policy conclusions.

The rest of this paper is organized as follows. Section 2 discusses literature in detail. Section 3 presents a simplified model with calibration to show why the network effects and competition matter for matching efficiency. Section 4 and 5 introduce industry background and the data separately. Section 6 presents the empirical model. Estimation strategy is discussed in section 7. The results and counterfactuals are in

[5]The concept is proposed by proposed by Weintraub, Benkard and Jeziorski (2008). The idea of OE is that, instead of competing with each other, taxi and Uber drivers are atomistic and compete against deterministic paths of distribution of other drivers in equilibrium.

section 8 and 9. Section 10 concludes this paper.

2 Literature

This paper is built upon two streams of literature, network effects and search and matching. This paper contributes to the network effects literature by modelling both direct and indirect network effects. A direct network effect measures the externalities of other agents from the same side of market. In this case, it is the impact of other riders on the demand for rides. An indirect network effect measures the cross-side externalities. In this case, it is the impact of the number of drivers on demand or impact of the number of riders on supply. This interdependence between demand and supply is reflected in the matching probability or waiting time. The theoretical literature on network effects begins with Katz and Shapiro(1985) and is followed by Farrell and Saloner (1986), Chou and Shy (1990), Church and Gandal (1992), Rochet and Tirole(2003, 2006) and Amstrong (2006).

The empirical work on network effects begins with Gandal (1994), Saloner and Shepard (1995) and has grown rapidly in recent years. Gandal, Kende, and Rob (2000) develop a dynamic model of consumer's adoption of CD player and software entry to estimate the feedback in CD industry. Ohashi (2003) and Park (2004) study network effects in the U.S. home VCR market. Nair, Chintagunta, and Dube(2004) quantify the network effects in the PDA market. Clements and Ohashi (2005), Corts and Lederman (2009) focus on network effects in the video game industry. Rysman (2004) estimates the network effects in Yellow Pages market and how it is related to market concentration and antitrust policy. Ackerberg and Gowrisankaran (2006) estimate the importance of network effect in the ACH banking industry. Dubé, Hitsch, and Chintagunta (2010) study network effects in video game market and its tipping effects. Lee (2013) also studies game industry but focus on software exclusivity. Liu and Luo (2022) studies network effects in smartphone industry and how it affects carriers' dynamic penetration pricing strategy.

Most of these empirical works focus on the indirect network effect of a two-sided platform and ignore the direct network effects, or they use network size of one side to estimate joint effects of indirect network effect and direct network effect. Goolsbee and Klenow (2002) is one paper focusing on only direct network effect in the diffusion of home computers. Chu and Manchanda (2016) is one recent paper that estimates and distinguishes both direct and indirect network effects in e-commerce platform (Alibaba). This paper contributes to the literature by quantifying both direct and indirect network effects. I allow the sizes of agents from both sides of the market to affect decisions of individuals on either side.

This paper also contributes to the search and matching literature by adding network effects to the model. Early search and matching models use a reduced form matching function to introduce frictions that prevent the market from clearing (Blanchard and Diamond(1989), Pissarides(1984), Mortensen and Pissarides (1999)). Microfoundations of the matching function are introduced, for example, as coordination failures (Butters (1977) and Burdett, Shi and Wright (2001)). Lagos (2000) develops a spatial search model of taxis without imperfect information and random search assumptions showing that frictions arise in the aggregate matching function endogenously as the outcomes of drivers' search decisions. Specifically, when prices are fixed and one location is more lucrative than other locations, drivers will overcrowd that location leaving other locations undersupplied. Coexistence of excess demand and excess supply reflect frictions in the aggregate matching function. Buchholz (2022) extends Lagos (2000) and builds an empirical model with non-stationary drivers' dynamics and price-sensitive demand. He shows that price regulation of NYC taxis leads to inefficient matching because drivers making dynamic search decisions prefer searching locations with high profitability. The fixed pricing structure for taxis prevents the market from clearing on prices.

This paper follows the approach of Buchholz (2022) and extend his model in two directions. First, the demand model is not only sensitive to prices, but also sensitive to supply/demand to incorporate network effects. Frechette, Lizzeri and Salz (2016) also includes supply in passenger's demand function in the form of a simulated waiting time, but they do not model drivers' location choices. Second, I incorporate cross-firm competition between taxi and Uber drivers for passengers. With these extensions, I can study impacts of non-price factors on matching efficiency and simulate regulation of Uber on social welfare. One similar work to my paper by Shapiro (2018) focuses on Uber's welfare contribution to the New York City, however, with less emphasis on network effects. Leccese (2022) considers the two-sided market feature of ridesharing industry, but focuses on price elasticity and tax pass-through. Castillo (2022) studies matching frictions of Uber with emphasis on surge pricing and its welfare outcome. He does not model taxi and Uber competition. The network effect feedback loop is not explicitly emphasized. Castillo et al. (2022) studies the matching friction due to dispatch protocol of Uber, which is different to spatial matching friction that my paper focuses on.

This paper also contributes to the empirical literature with dynamic oligopoly models. When there are a large number of firms within the market, Weintraub et al.(2007,2008) propose the concept of oblivious equilibrium (OE) to approximate Markov-perfect equilibrium in order to avoid the curse of dimensionality. In oblivious equilibrium, the firm is assumed to make the decision based only on its own state and deterministic average industry state rather than states of other competitors. In this paper, I assume drivers compete with the distribution of other drivers throughout the day. Under the OE assumption, only the distribution path at equilibrium is calculated. There are empirical papers using stationary OE (Xu(2008), Saeedi(2014)) and nonstationary OE (Qi(2013), Buchholz (2022)) to solve equilibrium of a model with a large number of agents[6].

3 A Simple Model with Calibration

This section develops a model of drivers searching for passengers among two islands in one period as an exposition of the main ideas. The model is a simplified version of Lagos (2000) but with drivers from both taxi and Uber, and provides insights on the relevance of network effects and competition. I solve the equilibrium and use comparative statics to show the influences of network effects and competition on matching efficiency especially for taxis. This simple model can be deemed as one slice of the full empirical model which has a multi-period dynamic game between taxi and Uber drivers. Two comparative statics are analyzed after calibrating the parameters in this model. The first is how network effects influence matching efficiency by changing the parameter of indirect network effect. The second is how competition affects matching by changing the total number of Uber cars. These exercises help to understand the mechanisms underlying the dynamic structural model of this paper.

[6]Finally, there are many works related to the taxi and ride-sharing industry. Early work studying NYC taxi industry include Farber (2005, 2008) and Crawford and Meng (2011) which study taxy drivers' labor supply decisions. Frechette, Lizzeri and Salz (2016) study taxi drivers' labor supply decisions with matching frictions. In recent years, research on ride-sharing industry has grown. For example, Chen et al.(2019) study flexible labor supply of Uber drivers and Chen and Sheldon (2015) study surge pricing of Uber. Other work related to traffic conditions and government regulation is by Kreindler (2018) which studies road congestion pricing policy in India.

3.1 Model Environment

There are a fixed number of taxi and Uber drivers, denoted as N_y and N_x where y stands for yellow taxis and x stands for Uber for the rest of this paper. Drivers are searching for passengers among two isolated islands $i = 1, 2$. Passengers need rides if and only if they travel across islands. Thus, the fare in each island of each firm is fixed and denoted as p_{fi} for $f = y, x, i = 1, 2$. The matching only happens in one period. Both drivers and passengers are assumed to have perfect information when making their decisions. Supply and demand of each firm-island combination is denoted as v_{fi} and u_{fi}. The supply v_{fi} aggregates all drivers of firm f choosing to search island i. For demand, I use a linear aggregate demand function as a reduced form of the aggregation of choices over passengers. The demand functions are:

$$\begin{aligned} u_{yi} &= \alpha u_{yi} + \beta v_{yi} + \theta v_{xi} + d_{yi}, \forall i = 1, 2 \\ u_{xi} &= \alpha u_{xi} + \beta v_{xi} + \theta v_{yi} + d_{xi}, \forall i = 1, 2 \end{aligned} \tag{3.1}$$

where the parameter α and β account for direct and indirect network effects respectively. In specific, α measures the impact of additional riders on the demand for rides and β measures the impact of supply in the same market on demand. The parameter θ measures substitution effect of competitor's supply on demand[7]. The term d_{fi} in the equations represents firm-island specific shocks such as population size and an idiosyncratic demand shock. Since price is fixed for each firm-island combination, it is fully captured in d_{fi}.

At the beginning of this period, each driver chooses which island to search for passengers in order to maximize his expected profit. The optimization problem is:

$$i^* = \arg\max_i \frac{m_{fi}}{v_{fi}} p_{fi} \tag{3.2}$$

where m_{fi} is matches in island i of firm f and the ratio m_{fi}/v_{fi} is the matching probability. I assume perfect matching within each island for now such that $m_{fi} = \min\{u_{fi}, v_{fi}\}$ for both taxi and Uber. In the full model, I allow matching frictions within a location, especially for taxis. In (3.2) the demand elasticity of supply is not parameterized for simplicity. Assume all passengers and drivers make simultaneous

[7] I use these equation as linear approximation for discrete choice model. Like the discrete choice model, market share of a particular product depends on characteristics of other products. For completeness, one could add the term of opponent's demand which will make the solution complex.

decision and Nash equilibrium satisfies the conditions E1-E4:

$$\frac{m^*_{f1}}{v^*_{f1}} p_{f1} = \frac{m^*_{f2}}{v^*_{f2}} p_{f2} \qquad \forall f = y, x \tag{E1}$$

$$u^*_{fi} = \frac{\beta}{1-\alpha} v^*_{fi} + \frac{D^*_{fi}}{1-\alpha} \qquad \forall i = 1, 2,\ f = y, x \tag{E2}$$

$$m^*_{fi} = \min\{u^*_{fi}, v^*_{fi}\} \qquad \forall i = 1, 2,\ f = y, x \tag{E3}$$

$$N_f = v^*_{f1} + v^*_{f2} \qquad \forall f = y, x \tag{E4}$$

where $D^*_{fi} = \theta v_{-fi} + d_{fi}$. Condition E1 means the expected profits of the two islands are equal and drivers have no incentive to deviate. Condition E2 is obtained from the demand function. E3 follows perfect matching assumption. Finally, E4 means the total number of drivers is fixed at N_f. The equilibrium demands and supplies can be solved from the conditions above.

Given the equilibrium supplies and demands, the mismatches across islands from an aggregate perspective can be calculated as:

$$\begin{aligned} mismatch &= \min\{\Sigma_i u_{fi}, \Sigma_i v_{fi}\} - \Sigma_i \min\{u_{fi}, v_{fi}\} \\ &= \min\{\underbrace{\Sigma_i \max\{0, u_{fi} - v_{fi}\}}_{\text{aggregate excess demand}}, \underbrace{\Sigma_i \max\{0, v_{fi} - u_{fi}\}}_{\text{aggregate excess supply}}\} \end{aligned} \tag{3.3}$$

The mismatch depends on the parameters of the model including $\alpha, \beta, \theta, d_{fi}$ and exogenous variables N_f, p_{fi}. To illustrate the economic forces, I consider the equilibrium with excess supply in island 1 and excess demand in island 2 for taxis. In other words, focus on the equilibrium with mismatches for taxis and study how changing network effects (β) and competition (N_x) affects equilibrium mismatches of taxis.

3.2 Network Effects Calibration

To study how network effects change matching efficiency of taxis, competition from Uber is shut down by letting $N_x = 0$. Under condition E4, the equilibrium supplies of Uber are $v^*_{xi} = 0$. Though the demands for Uber may not be zeros, the equilibrium demands for taxis are independent of Uber. Without loss of generality, we consider the equilibrium with excess supply in island 1 and excess demand in island 2 as an example of mismatches. In such equilibrium, it requires $p_{y1} > p_{y2}$ and the solutions satisfy:

$$v_{y1}^* = \frac{p_{y1}}{p_{y2}} u_{y1}^* \qquad \text{(supply in island 1)}$$

$$u_{y1}^* = \frac{d_{y1}}{1 - \alpha - \beta \dfrac{p_{y1}}{p_{y2}}} \qquad \text{(demand in island 1)}$$

$$v_{y2}^* = N_y - v_{y1}^* \qquad \text{(supply in island 2)}$$

$$u_{y2}^* = \frac{\beta}{1-\alpha} v_{y2}^* + \frac{d_{y2}}{1-\alpha} \qquad \text{(demand in island 1)}$$

As long as $1 - \alpha - \beta \frac{p_{y1}}{p_{y2}} > 0$ and $N_y > v_{y1}^*$, there is excess supply in island 1. There will be excess demand in island 2 as long as d_{y2} is large enough. To study the impact of network effects, I calibrate the parameters and solve the equilibrium with various β. The parameter values are $p_{y1} = 3, p_{y2} = 2, N_y = 20, N_x = 0, d_{y1} = 10, d_{y2} = 100, \alpha = -3$ and β ranges from 1.5 to 2.5.

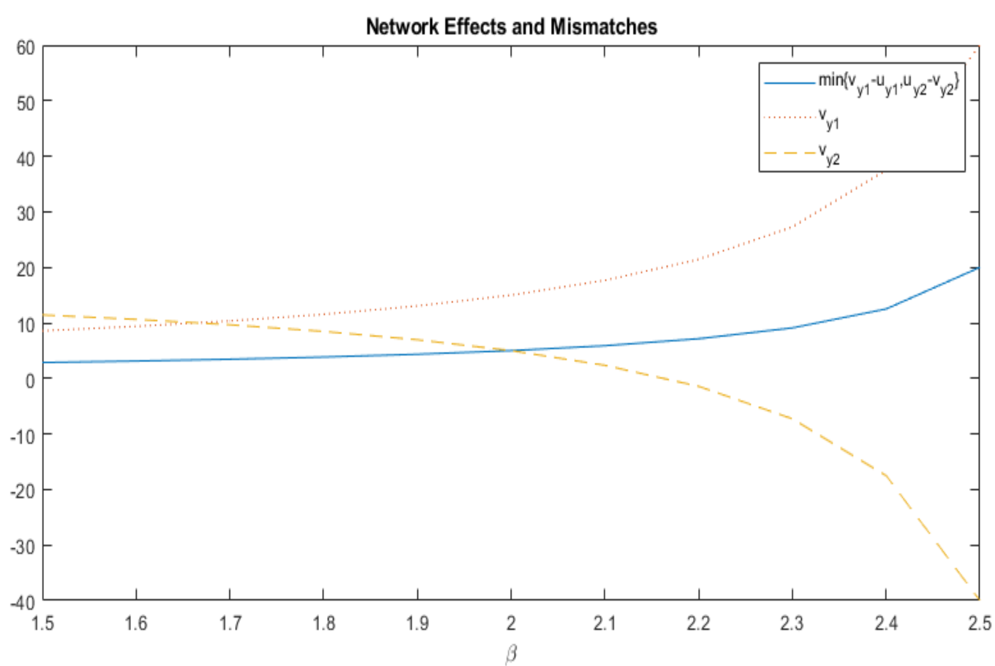


Figure 1: Comparative Statics of Network Effects and Taxi Mismatches

The result is illustrated in figure 1. As the indirect network effect β increases, the demands for taxis increase in both islands. Increased demand in island 1 further attracts more drivers choosing island 1 and more supply generates more demand. Whether it ends up with a new equilibrium with more excess supply or less depends on the relative elasticities. In this model, since demand and supply have a linear relationship in equilibrium, increased demand will make more excess supply in island 1. As for island 2, since the matching probability is already one, increased demand will not attract more supply and on the contrary more drivers will choose island 1 than 2. Whether excess demand in island 2 increases or decreases depends on whether the increase of β dominates the decrease of supply v_{y2}. In this calibration, a stronger network effect will make the matching less efficient for taxis in terms of aggregate mismatches.

3.3 Competition Calibration

To analyze how competition affect the matching efficiency, I calibrate the parameters with various N_x, because variation in N_x can simulate changes in regulatory policy on Uber. One can also vary the parameter θ to study the competition effect. This calibration still focuses on the particular equilibrium with mismatches for taxis, but with excess supply for Uber in both islands. The reasons are twofold: 1, it simplifies the discussion of matching efficiency by only considering one firm's mismatches. 2, the equilibrium is unique if both islands have excess supply of Uber drivers in comparison to the case that both islands have excess demand for Uber. The equilibrium demands and supplies satisfy:

$$v_{y1}^* = \frac{p_{y2}}{p_{y1}} u_{y1}^* = \frac{p_{y2}}{p_{y1}} \frac{\theta v_{x1}^* + d_{y1}}{1 - \alpha - \beta p_{y1}/p_{y2}} \tag{3.4}$$

$$u_{y2}^* = \frac{\beta}{1-\alpha} v_{y2}^* + \frac{\theta v_{x2}^* + d_{y2}}{1-\alpha} \tag{3.5}$$

$$v_{x1}^* = \frac{\theta v_{y1}^* + d_{x1}}{\theta N_y + d_{x1} + d_{x2}} N_x \tag{3.6}$$

$$v_{x2}^* = \frac{\theta v_{y2}^* + d_{x2}}{\theta N_y + d_{x1} + d_{x2}} N_x \tag{3.7}$$

To construct the equilibrium, set $p_{y1} > p_{y2}$ and make d_{y2} large enough such that island 2 has excess demand for taxi. I set $p_{x1} = p_{x2}$ such that they cancel out in condition (E1) for simplicity. To obtain excess supply for Uber, I adjust the values of N_x to guarantee it. The parameter values are $p_{y1} = 3, p_{y2} = 2, p_{x1} = p_{x2}, N_y = 3, d_{y1} = d_{x1} = d_{x2} = 10, d_{y2} = 100, \alpha = -3, \beta = 2, \theta = -2$ and N_x ranges from 12 to 24. Figure 2 shows the results. As N_x increases, demand for taxis decreases in both islands due to competition. Decreased demand for taxi in island 1 makes island 1 less profitable for taxi drivers and therefore supply of taxi in island 1 decreases. More taxi drivers choose to search island 2 and make the mismatches between islands smaller. In the opposite, if the number of Uber drivers is controlled by government and N_x decreases, the mismatches of taxi will increase as this particular calibration shows.

To summarize, the spatial search model illustrates the impacts of network effects and competition on aggregate matching efficiency. These non-price factors are as important as prices to understanding matching outcomes and social welfare. Especially, when evaluating policy outcomes, one could draw imprecise conclusion by ignoring these effects. The next section introduces the background of NYC taxi industry which will be used to empirically study matching efficiencies in a fully developed dynamic version of the search model.

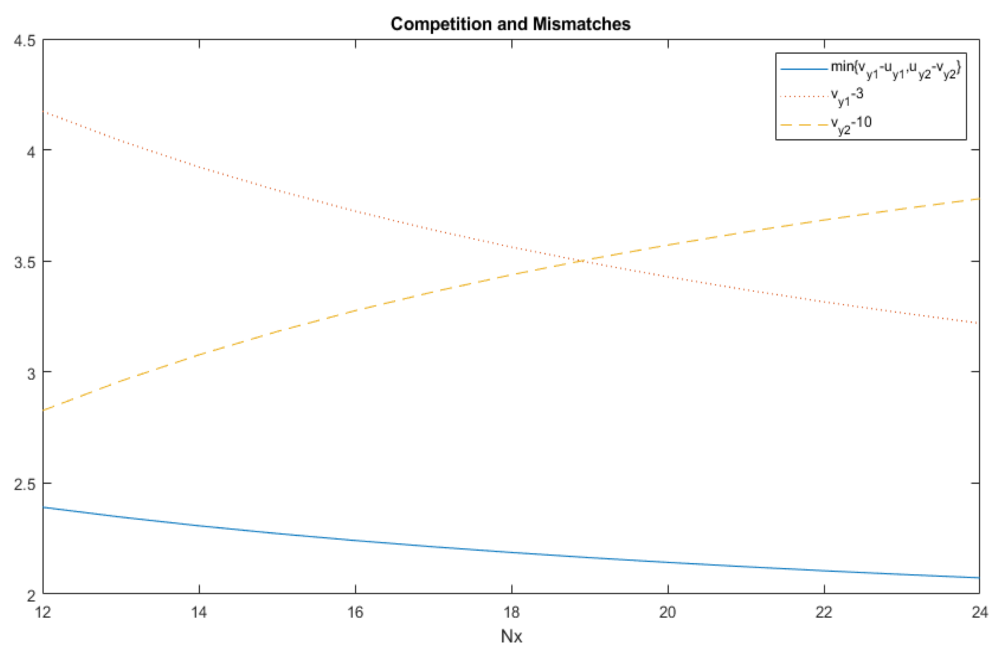


Figure 2: Comparative Statics of Competition and Taxi Mismatches

4 Background on New York City Taxi and Ridesharing Industry

In NYC, there are mainly two ways other than public transportation to get a ride, taxi or for-hire vehicle(FHV). Taxis can only pick up street hails and FHV can only take pre-arranged ride requests. These two markets are strictly separated under the regulation of NYC government. Running a taxi requires a medallion. The total number of medallions available is fixed by regulation at 13,587 in 2016. Medallion owners can trade medallions through auction and the price has dropped due to rideshare service in recent years. Along with yellow taxis, there are up to 18,000 boro taxis which were introduced gradually to the city since 2013. Boro taxis follow similar operation and pricing rules as yellow taxis except for some restrictions[8]. In this paper, I do not model Boro taxis since their trips are far less than yellow taxis and most trips of the city originate from Southern Manhattan. The taxi fare follows a fixed nonlinear pricing structure and it rarely changes under regulation. Most trips charge either metered fare (base price and price/mile) or flat fare (i.e. airport trip is $52). Taxi price is fixed in the sense that it only depends on the origin-destination of the trip not on the demand and supply levels. Most taxi drivers lease the vehicle from medallion owners. There are two shifts in a day: day shift starting from 6 a.m. and night shift starting from 4 p.m. with one rush hour in each shift.

The FHV provides pre-arranged transportation. It has different classes of service: livery, black car and luxury limousine of which black cars account for most of the trips. Unlike taxis, there is no restriction on entry of FHV. Since the operation of Uber in 2011, the number of FHV has increased by more than 60 percent to 63,000 vehicles by 2016. In early 2016, the total number of Uber cars was more than 25,000, far bigger than other ridesharing companies such as Lyft and Via[9]. Though Uber cars are mostly registered as black cars which provide service through a mobile app, its drivers are not dispatched from bases. Active Uber drivers search for passengers on the street. When there is a ride request, Uber sends it to the drivers nearby. In practice, Uber drivers can decline a request, but it is not observed in the data and it is not modelled. Unlike the

[8] Boro taxis can only pick up passengers in Northern Manhattan and Outer Boroughs. Moreover, the boro taxis can only pick up passengers at the airport by prearrangement.

[9] In this paper, I only model Uber as competitor to taxis without Lyft. One reason is that Lyft is not big enough during my sample period. Out of all black car trips, Uber accounts for 72.6% and Lyft accounts for 11.6% in April 2016. The other reason is that trip records of Lyft are not good.

fixed price of taxi, Uber uses surge pricing which charges a higher than normal price when demand is higher than supply in an area. On the one hand, high price attracts more drivers to increase supply, and on the other hand it excludes passengers with low willingness-to-pay to decrease demand. Uber's flexible price helps the market to clear and improve cross-location matching efficiency. Finally, unlike taxi drivers' fixed labor supply in a shift, Uber drivers work more flexible hours during the day which makes it difficult to tract real-time supply of Uber.

5 Data

The data used in this paper comes from three sources. The main data about taxis and Uber comes from trip records provided by the New York City Taxi and Limousine Commission (TLC). The taxis trip records include all trips completed by yellow taxis since 2009 and by boro taxis since 2013. Each observed trip in the data includes pick-up and drop-off timestamp, geographic location, trip distance and fare. Trip time can be calculated from the gap between pick-up and drop-off time. One useful variable not available in the trip records is vehicle's identifier for each trip. This variable can help to understand drivers' search patterns by comparing drop-off location and next pick-up location. Moreover, I cannot tell the real-time location of vacant taxis and therefore the potential supply of drivers. Neither can I tell how many potential passengers who want rides but fail to match with this data.

The TLC data also contains FHV trip records which include trips of Uber[10]. Due to different ways of data collection, Uber's trip records have less information than taxis. First, it only provides zone area instead of geographic coordinates for each trip[11]. Second, the zone and timestamp is only available for pickups but not dropoffs. Uber refuses to submit drop-off information of trips because of privacy concerns. Since I do not know the trip time and distance without observing dropoffs, I assume it takes the same travel distance and time for Uber as taxi for the same trip. Under this assumption, I can predict Uber' trip fares with the next data source.

The second source of data that supplements Uber's trip records is Uber's surge multiplier. Uber's fare during a normal time is calculated based on base price, trip

[10]The trips of Uber can be identified by the base id affiliated with each trip. TLC provides separate list for bases of black, livery and luxury cars and companies they are affiliated to.

[11]The taxi zones are not accurate as geographic locations which are areas defined by the TLC. There are about 263 taxi zones in the NYC.

distance and time. When Uber's supply is less than demand, Uber applies surge pricing which charges a higher price as the product of a surge multiplier and the normal price. In order to predict Uber's trip fare, I use Uber's API to collect the real time surge multiplier every 10 minutes at 79 selected location spots across the city during the sample period. Each request returns the surge multiplier of the time-location which can be matched with the trip data to construct the price of each Uber trip in the sample.

The third dataset is subway riderships obtained from Metropolitan Transportation Authority (MTA) of the city. This data is used to calculate the number of potential travellers of a given location-time as a measure of market size. The ridership data includes information on weekly aggregate entrances to each station of the NYC subway. For a given station, the riderships are sorted by various types of MetroCards used by customers such as day pass, student, and full fare. I only count travellers paying full fare as potential passengers of taxi&Uber since they are more likely to have the same travelling patterns as taxi&Uber passengers compared to routine metro commuters. Thus, the market size of a location at a given time is defined as the sum of taxi and Uber pickups and full fare metro riderships.

5.1 Sample Construction

In the empirical part of this paper, I model taxi&Uber drivers' dynamic search decisions and passengers' demands across locations in NYC in a representative day shift (6 a.m.-4 p.m.) of a weekday in April 2016. Thus, the model focuses on equilibrium evolution of supplies, demands and matches over this representative time interval. There are two reasons for doing this. First, in the data, pickups and dropoffs across locations and times for the day shift follow similar pattern among weekdays. Solving a dynamic game for every day is not only a computational burden but also a redundant exercise. Second, the number of active Uber drivers is not observed in the data. Given the flexible working hours of Uber drivers, the number of on-street Uber drivers is more stable during a weekday than the night shift and weekend when part-time or occasional drivers are more likely to enter[12]. In this paper, I need to assume at least the total number of on-street vehicles in order to estimate the unobserved supply. For taxis, the number of total taxi cars is fixed at 13,587 which is the number of medallions. For

[12] There are studies on labor supply by Chen et al.(2017) and Hall and Krueger (2016) for Uber and by Farber(2008) and Frechette et al. (2016) for taxi

Uber, I have to assume the total number of cars is also fixed, which is 3,000, for the purposely picked day shift of weekday[13].

Time and space for the representative shift is discretized in the following ways. A time period is defined every 10 minutes and there are 60 periods in total. The city is divided into 40 geographic locations as shown in figure 3. I define the area of each location by joining taxi zones to make locations comparable in number of pickups. For example, the area of locations in Queens and Brooklyn are large compared to those in Manhattan because the pickups in outer boroughs are quite sparse. I exclude central park from this map since pickups within it are on the boundaries and I assign those pickups to locations nearby. The $40 \times 60 = 2400$ location-period pairs are each defined as a market. Within each market, passengers and drivers randomly meet only once and successful matches become pickups in that market[14].

One set of variables constructed from the data for estimation are distance, travelling time of 96,000 ($40 \times 40 \times 60$) possible types of trip and fares of taxi and Uber for these trips. For any given type of trip, the average trip distance and time of all taxi trips of this type are calculated. Since the same data for Uber is not observed, I assume same type of trip costs the same distance and time for Uber as taxi. Having the trip distance and time, the average fares are calculated according to the pricing rule of taxi ($\$2.5 + 2.5 * \text{distance} + 0.8$) and Uber ($\$2.55 + 0.35 * \text{time} + 1.75 * \text{distance}$). The fares are adjusted for airport trip and Uber's $7 mimimum fare and surge multiplier. Another variable analogous to the travelling time (in minutes) is the travelling periods. The model does not distinguish drivers arriving at different minutes of a period. For example, a 25 minutes trip takes 3 periods for drivers to complete.

Another set of variables are the travelling patterns of passengers. For any given market, the distribution of destinations is obtained by calculating the proportions of dropoffs. For example, when a passenger shows up at Time Square, the distribution of destinations tells the probability that she will go to JFK airport or Brooklyn. This variable is useful to compute expected profit of drivers in a market. The issue is that only the transition probability of taxi passengers can be calculated according to

[13]In Chen et al. (2019), they calculate the transition matrix of Uber driver's types including evening driver, morning driver, late-night driver, weekend driver and infrequent driver. I calculate the stationary distribution of the markov chain which has 10% morning drivers. Thus, I assume there are 10% of 30,000 registered Uber drivers working during the day shift.

[14]Under this assumption that a driver only searches once in a 10-minute period, the model could underestimate supply. For example, a driver could finish a trip and search for the next in 10 minutes. The problem is negligible if intervals get thinner.

observed taxis' dropoffs. Considering the sorting of passengers between taxi and Uber, the transition of taxi passengers is not necessarily the transition of population. The difference between these two reflects the endogenous choice decisions of passengers which is a good identification strategy. Thus, the transition of taxi passengers in 2010, before the entry of Uber, is used as a proxy for the population transition in the sample period. This proxy is valid if travellers with different travelling patterns to those in 2010 do not enter the market as Uber enters.

Finally, for each market, the average pickups of taxi and Uber over days are calculated. I also construct market size as sum of taxi pickups, Uber pickups and subway riders paying full fare with subway being the outside option of the discrete choice model. Subway riders are assumed to include those who choose to request a ride but fail to match and those who choose outside option at the beginning. All people are implicitly assumed to leave the market at the end of period. In this way, the market size may be underestimated by ignoring the people who wait until future periods to leave. Another issue to notice is that the transition of taxi passengers in 2010 may not represent the transition of population, considering that choosing subway is also an endogenous outcome[15].

5.2 Sample Overview

Table 1 shows monthly aggregate statistics of weekday pickups in November 2010 (22 days) and April 2016 (21 days). It summarizes the distribution of pickups by location, shift and firm. In April 2016, taxis' monthly aggregate pickups during day shifts are 3.7 million. Almost 93% of the total pickups are in Manhattan, 4.59% are in JFK and Laguardia airports. My sample of 40 locations cover 99.38% of all taxis' pickups during day shifts. Outer boroughs has 2.24% pickups in total. The pickups of taxis drop largely from 4.6 million to 3.7 million for monthly aggregate day shifts of 2010-11 and 2016-04. The distribution of taxi pickups also changes such that share of airport increases from 3.42% to 4.59% and share of Manhattan drops from 93.67% to 93.17%. Though the difference between 2010 and 2016 is not obvious. More changes of pick-up distribution can be investigated if Manhattan is collapsed into smaller locations. This change of pick-up distribution reflects change of taxis' supply and demand distribution after Uber's entry. Table 1 also includes statistics on Uber's pickups. Uber has different

[15]For example, if those choosing subway at the beginning are long-distance travellers, then the 2010 transition of taxi passengers underestimate the probability of long-distance trips in the population.

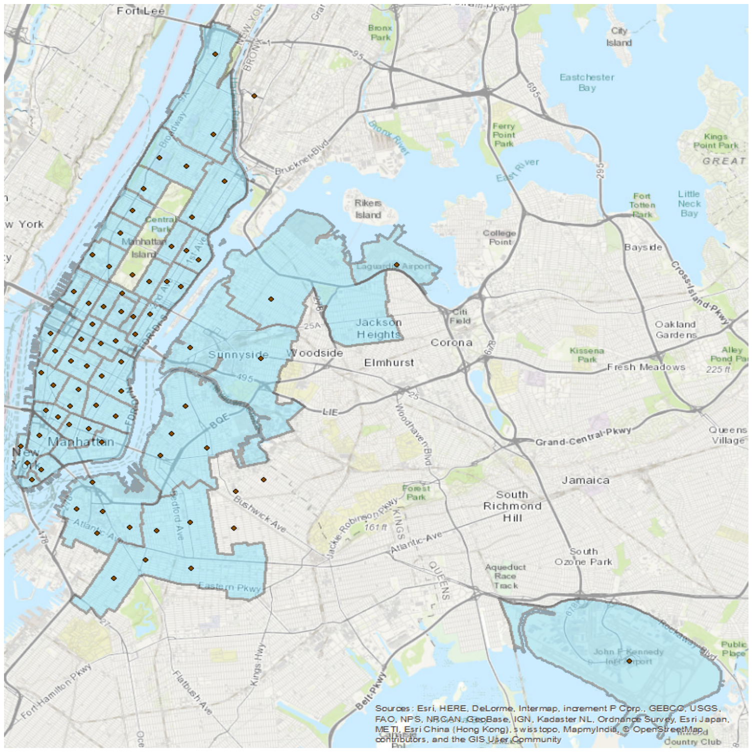


Figure 3: 40 select markets and 79 spots to collect Uber surge multiplier

distribution in comparison to taxis that 59.5% pickups are in Manhattan. The share of Uber's pickups in outer boroughs is 36.58% quite larger than taxis. The share of Uber's pickups covered by my 40 locations is 77.63%. The variation of pickups between taxi and Uber across markets helps to estimate the demand model. The complexity in the model is that shares of pickups are not exactly shares of demand due to the matching friction within a market.

Similar to demographics of a market, the exogenous travelling pattern of passengers represents the demographics in the market. As mentioned above, I use taxis' trip records of 2010 to approximate exogenous travelling patterns of market population. Taxis' trip records of 2016 are endogenous outcome of passengers' choices. Table 2 provides a rough overview of conditional distribution of dropoffs. The distribution is calculated for day shift of weekdays. The first panel shows that 95.11% of pickups in Manhattan are delivered within Manhattan and 3% to airports. Trips originating from airports have 73.45% ending up in Manhattan and 6.37% of them are inter-airport. Comparing 2010 and 2016, the dropoff distributions are slightly different that trips originating from airports to Manhattan decrease from 73.45% to 71%. Table 2 only shows travelling patterns among three highly aggregate areas, Manhattan, Airport and Other. More variations of travelling patterns can be discovered at the market level. I use this difference as an identification strategy. The variation of travelling patterns across market also helps identify price coefficient. For instance, given two markets with the same market size and supply level, but with different travelling patterns of population, a high price elasticity will generate quite different demand than a low price elasticity.

Table 1: Trip and share by firm, shift, and area

Firm&Shift	*Total*	*Manhattan*	*Airports*	*Other*	*40 mkt*
Yellow Taxi 2010.11(22)					
Day shift	4,627,258	93.67%	3.42%	2.91%	98.93%
Night shift	5,139,146	92.91%	3.46%	3.63%	98.99%
Yellow Taxi 2016.04(21)					
Day shift	3,730,326	93.17%	4.59%	2.24%	99.38%
Night shift	4,279,262	92.04%	4.94%	3.02%	99.27%
Uber 2016.04					
Day shift	1,322,507	59.5%	3.92%	36.58%	77.63%
Night shift	1,989,054	64.15%	4.1%	31.75%	82.43%

Table 2: Distribution of dropoffs in day shift by firm

	Obs.	*Manhattan*	*Airports*	*Queen&Brooklyn*	*not in 40*
Yellow Taxi 2010.11(22)					
Manhattan	4,334,266	95.11%	3%	0.98%	0.89%
Airports	158,410	73.45%	6.37%	8.15%	12.01%
Queen&Brooklyn	84,925	47.53%	4.43%	42.13%	5.88%
not in 40	49,657	41.50%	3.2%	9.04%	46.24%
Yellow Taxi 2016.04(21)					
Manhattan	3,475,467	94%	3%	0.97%	1.17%
Airports	171,407	71%	3.2%	10.73%	14.69%
Queen&Brooklyn	60,274	42.6%	4.26%	44.89%	8.2%
not in 40	23,178	45%	3.9%	17.3%	32.8%

Table 3 shows the statistics of the rest of variables in the sample. Given 40 locations and 60 periods, there are 2,400 markets. Uber's surge multiplier varies from 1 to 1.37 at 10% and 90% quantiles. Market pickups range from 8.8 to 163 for taxis and 8.33 to 34.95 for Uber. Taxi and Uber prices, distance, and time are calculated at the origin-destination-period level. Uber's fare on average is higher than taxi because Uber charges both trip time and distance, there is a surge multiplier and Uber charges a minimum fare of $7 which is higher than taxis for short trips.

Table 3: Summary statistics of key variables

variable	*Obs*	*mean*	*10%ile*	*90%ile*	*S.D.*
surge	2,400	1.14	1	1.37	0.18
taxi matches	2,400	72.28	8.80	163.02	57.34
Uber matches	2,400	20.37	8.33	34.95	10.25
taxi fare	96,000	17.68	7.92	28.4	10
Uber fare	96,000	22.39	9.12	37.95	12.48
trip distance	96,000	5.07	1.19	9.88	4.10
trip time	96,000	23.78	8.97	39.11	12.35

In order to show the heterogeneity of conditional expected profits across markets, I calculate the conditional expected profit for 2,400 markets. The distributions of expected profits for taxi and Uber are shown in figure 4. For taxis, the average expected profit is 15.27 and it ranges from less than 10 dollors to almost 60 dollors. The distribution for Uber has mean of 13.94 dollars without surge multiplier and 18.21 dollars with surge multiplier. It implies that Uber driver's expected profit is higher than taxi in general. This high heterogeneity in profitability as shown in figure 4 illustrates the incentive of drivers' search decisions and why some markets are oversupplied. However, the expected profit calculated here is static flow profit. Drivers' search decisions are made based on dynamic search values in the full model. Using the dynamic model developed in the next section, search values can be computed with the estimates and compared across markets.

6 Dynamic Search and Matching Model

The structural model fully extends the search and matching model discussed in section 3. Taxi and Uber drivers make dynamic spatial search decisions among I locations over T periods in a day shift. Potential passengers make a static discrete choice decision among Uber, taxi and subway. Drivers and passengers have perfect information about the size of either side in a given market. When making supply/demand decision, the agent accounts for both the indirect network effect from the other side of the market and direct network effect from the same side. This supply sensitive demand specification is one contribution of this paper to Buchholz (2022). The model allows two types of frictions that prevent the market from clearing. First, within a market taxis and passengers do not fully contact with each other due to coordination failure as in Burdett, Shi and Wright (2001). However, perfect matching is assumed for Uber within the market. Matching is perfect in airports for both taxis and Uber. In other words, the matching process within a market is modelled with an explicit functional form. Second, because of drivers' endogenous search decisions there are locations exhibiting excess supply along with other locations with excess demand. Both frictions result in inefficient matching at the city aggregate level.

The timeline within a market is as follows. At the beginning of each period, part of taxis and Uber cars will arrive at their destinations. If the car has a passenger on board (employed), it arrives at the dropoff location. If the car is vacant (unemployed),

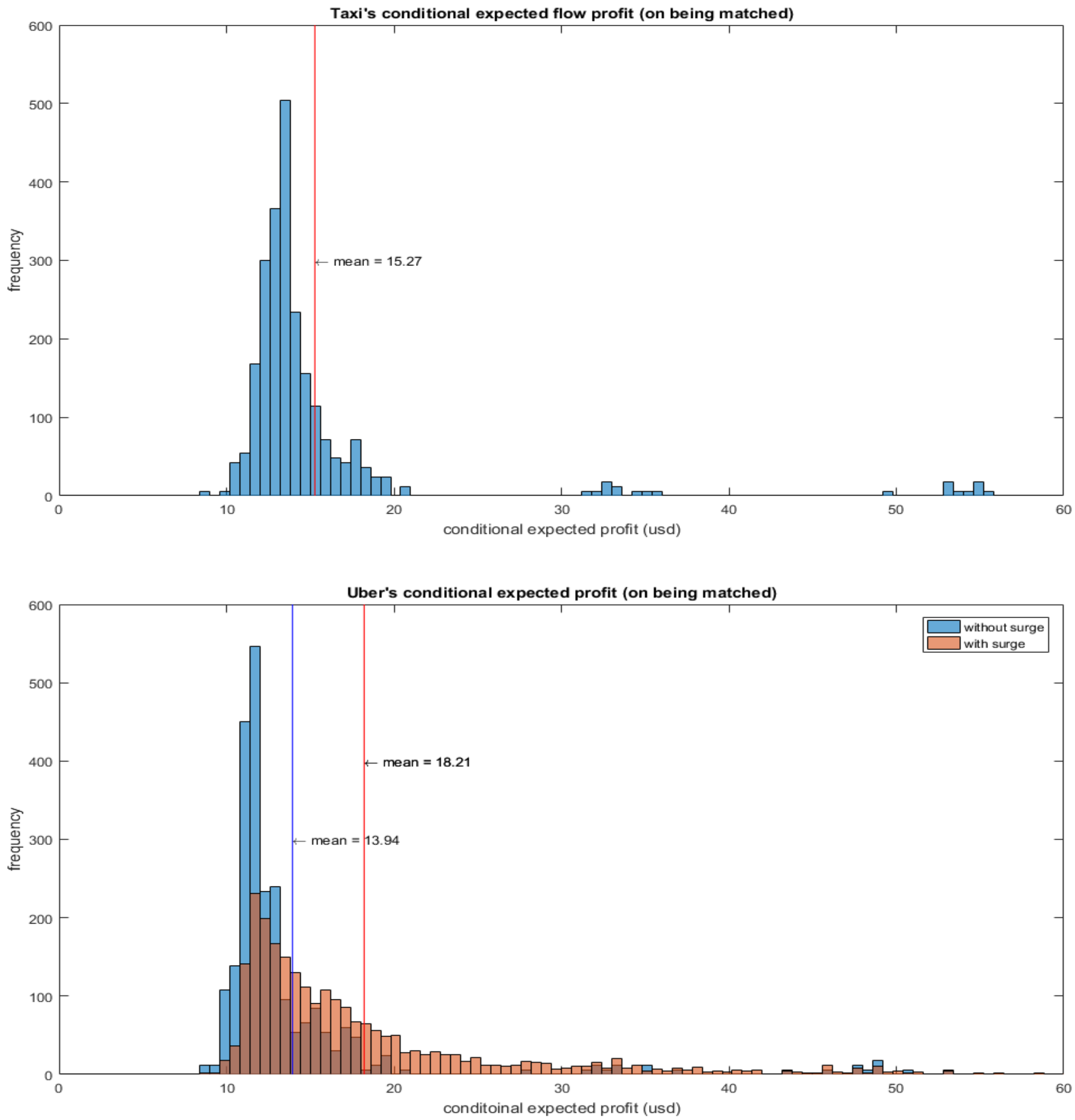
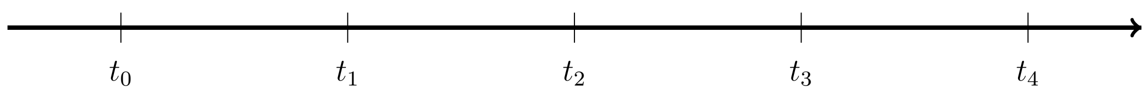


Figure 4: Distribution of expected profits of taxi and Uber

it arrives at the location based on the driver's search decision in the last decision period. Some of the cars either employed or unemployed are still on their way to the destinations and will not necessarily arrive at a location in this period. All arriving cars become supply to that market in this period. A passenger in this market has perfect information about fares and rational beliefs on demand/supply, how likely he will find a taxi or Uber car, and how long it takes to match[16]. Passengers make a static discrete choice decision. Aggregating all passengers' decisions returns demand for each firm in this market. Then matches are made within each firm. Unmatched passengers either due to excess demand or matching friction leave with the outside option (subway). Employed drivers deliver passengers to their destinations and unemployed drivers search next locations.

timing of supply, demand and match of a market



- t_0: Drivers arrive and become supply.
- t_1: Passengers make discrete choices.
- t_2: Drivers and passengers of the same firm are matched.
- t_3: Passengers either unmatched or choosing subway in t_1 leave with subway.
- t_4: Employed drivers deliver passengers and unemployed drivers make search decisions.

6.1 Passengers' Choice Problem

In each market, a group of potential travellers make a discrete choice among taxis, Uber and subway conditional on their exogenous destination with knowledge of prices, product qualities, supply and demand. In this model, Uber is denoted as x (UberX),

[16]Uber's supply can be perfectly learned by app which shows how many cars are around and how long to wait. Taxis' supply is hard to directly observe. However, the model characterizes demand and supply only in equilibrium such that passengers are fully experienced and know how likely to get a car without necessarily knowing how many cars nearby.

taxi as y (yellow taxi) and outside option as o. The utility of a passenger c in location i at period t choosing firm $f = x, y, o$ to travel to j prior to matching process is:

$$\begin{aligned} U^{ij}_{cft,pre} &= G(\tau_f(u^i_{ft}, v^i_{ft}), u^i_{ft}, v^i_{ft}, p^{ij}_{ft}, X^i_{ft}, \varepsilon^i_{cft}) \\ &= \tau^i_{ft} U^{ij}_{cft,post} + (1 - \tau^i_{ft}) U^{ij}_{cot,post} \end{aligned} \tag{6.1}$$

where u^i_{ft} is market demand for firm f, v^i_{ft} is firm f's supply and p^{ij}_{ft} is the price from i to j. The function $\tau^i_{ft} = \tau_f(u^i_{ft}, v^i_{ft})$ is the probability of being matched by choosing firm f and is determined by the firm specific demand and supply level in the market. The matching probability does not differ for different destinations j under the fact that it is illegal for drivers to discriminate and decline a ride[17]. The probability $\tau(u^i_{ft}, v^i_{ft})$ can be written as $m(v^i_{ft}, u^i_{ft})/u^i_{ft}$ with matching function $m(v^i_{ft}, u^i_{ft})$. The functional form of m will be discussed later. In addition to effects of u^i_{ft}, v^i_{ft} on the matching probability, they could also affect the utility through classical direct network effects. For example, the demand u^i_{ft} in the market could affect passenger's choice decision as outcome of consumers learning from each other or congestion effect (Goolsbee and Klenow (2002)). The supply v^i_{ft} affects choices through indirect network effect. Higher v^i_{ft} could increase utility by decreasing waiting time. The utility of the outside option is normalized to zeor, $U_{cot,post} = 0$, which implies that unmatched passengers end up with zero utility by taking subway. Furthermore, $U^{ij}_{cft,post}$ is specified as log-linear such that (6.1) can be rewritten as :

$$\begin{aligned} \ln(U^{ij}_{cft,pre}) &= \ln(\tau^i_{ft}) + \ln(U^{ij}_{cft,post}) \\ &= \underbrace{\theta_1 \ln(v^i_{ft}) + \theta_2 \ln(u^i_{ft}) + d_x + d_i + t + \xi^i_{ft}}_{\delta^i_{ft}} + \alpha^{ij} \ln(p^{ij}_{ft}) + \varepsilon^{ij}_{cft} \end{aligned} \tag{6.2}$$

Equation (6.2) is obtained by: (1), transforming τ^i_{ft} as a linear combination of $\ln u^i_{ft}, \ln v^i_{ft}$; (2), assuming $\ln(U^{ij}_{cft,post})$ is linear in $\ln u^i_{ft}, \ln v^i_{ft}$ and other characteristics including price p^{ij}_{ft}, Uber fixed effect d_x, market fixed effect d_i, time fixed effect d_t, unobserved firm-market demand shock ξ^i_{ft} and idiosyncratic consumer shock ε^{ij}_{cft}. The benefit of

[17] Uber drivers can decline a ride according to its destination which causes issues to the company. Uber tries to use destination filter for drivers to add their order preference in order to decrease the decline rate. In this paper, I do not model and allow such discrimination for Uber drivers.

these assumptions is that all endogenous variables of the model u^i_{ft}, v^i_{ft} are contained in parameter δ^i_{ft} such that demand is simple to solve. In other words, unobserved endogenous demand u and supply v , which need to be solved through the structure, are separated from estimating price coefficients[18]. The drawback of the log-linearity assumption of τ is that coefficients θ in (6.2) measures the joint effect of v^i_{ft} or u^i_{ft} without distinguishing channels through matching probability or through classic network effects (i.e. product variety, word-of-mouth.). The price coefficients α^{ij} depend on travel distance and trip type which are parameterized as:

$$\alpha^{ij} = \sum_{k=1,2,3} \alpha_k \mathbb{1}\{\text{dist}^{ij} \in \mathcal{I}_k\} + \alpha_4 \mathbb{1}\{\text{dist}^{ij} \in \mathcal{I}_{JFK}\} \tag{6.3}$$

where $\mathcal{I}_1$ is for trip distance less than 3 miles, $\mathcal{I}_2$ is for distance between 3 and 6 miles, $\mathcal{I}_3$ is for distance greater than 6 miles and $\mathcal{I}_{JFK}$ is trips between JFK airport and Manhattan which charges flat rate.

Finally, I assume the utility of choosing outside option before the matching process as:

$$\ln(U^{ij}_{cot,pre}) = \delta^i_{ot} + \varepsilon^{ij}_{c0t} \tag{6.4}$$

where δ^i_{ot} is normalized to zero. Since the subway fare is fixed for single trip regardless of trip length, there is not price in 6.4.

I allow substitution between taxi and Uber by assuming a nested logit demand model such that:

$$\varepsilon^{ij}_{cft} = \zeta^{ij}_{cgt} + (1-\beta)\nu^{ij}_{cft} \tag{6.5}$$

where ζ^{ij}_{cgt} is common to taxi and Uber which are categorized as one group, and subway alone as the other group. Variable ν^{ij}_{cft} is assumed to follow type I extreme value distribution. The distribution of ζ^{ij}_{cgt} satisfies that ε^{ij}_{cft} is also an extreme value random variable. The parameter β measures substitution between taxi and Uber. When $\beta = 0$, it is equivalent to the simple logit demand model. Larger β implies stronger substitution pattern between taxi and Uber. Then, the choice probability condition-

[18] In details, given values δ^i_{ft} and α^{ij}, demands are fixed when I iteratively solve equilibrium supply. If matching probabilities interacts with price, update of supply requires update of demand as well. More details are in section 7.

ing on route ij at time t becomes the product of choice probability within group and probability across group. Within each group of taxi and Uber, the choice probability is:

$$s^{ij}_{y|gt} = \frac{\exp((\delta^i_{ft} + \alpha^{ij}\ln(p^{ij}_{ft}))/(1-\beta))}{D_g} \tag{6.6}$$

where:

$$D_g = \exp((\delta^i_{yt} + \alpha^{ij}\ln(p^{ij}_{yt}))/(1-\beta)) + \exp((\delta^i_{xt} + \alpha^{ij}\ln(p^{ij}_{xt}))/(1-\beta)) \tag{6.7}$$

The probability of choosing the group with taxi and Uber is:

$$s^{ij}_{gt} = \frac{D_g^{1-\beta}}{1 + D_g^{1-\beta}} \tag{6.8}$$

Then the choice probability becomes $s^{ij}_{ft} = s^{ij}_{f|gt} * s^{ij}_{gt}$.

In the traditional way, demand is estimated by matching choice probabilities to the market shares obtained by dividing demand u^{ij}_{ft} by the number of people travelling from i to j as Berry (1994). However, there are two obstacles to this. First, only pickups m^{ij}_{ft} rather than demand u^{ij}_{ft} is observed. Thus, market share of demand can not be calculated from the data. Second, even though assuming matches equal to demand such that $u^i_{ft} = m^i_{ft}$, I cannot calculate u^{ij}_{xt} for any j of Uber without knowing the destination distribution of Uber trips. In other words, equation (6.2) cannot be directly estimated at the trip type level $\{i, j, t\}$.

Instead, I treat the choice probability as the model prediction for the conditional market share on destination and aggregate s^{ij}_{ft} over destinations j to calculate the unconditional market shares of firms at location i. The exogenous distribution of passengers' destination in a market $\{i, t\}$ is denoted as $A^i_t = \{a^{ij}_t\}_{\forall j}$ where a^{ij}_t is the probability that a passenger from this market travels to j. The unconditional market share predicted by the demand model is:

$$s^i_{ft} = \sum_j a^{ij}_t s^{ij}_{ft} \tag{6.9}$$

If the market size is λ^i_t. Then the potential demand before the matching process is:

$$u_{ft}^{i} = \lambda_t^i s_{ft}^i \tag{6.10}$$

In comparison to the exogenous distribution of travellers' destinations A_t^i, the dropoffs distribution of each firm $\tilde{A}_{ft}^i$ can be calculated as outcomes of passengers' discrete choice using Bayes' rule. Thus, the model predicted firm-specific destination distribution becomes:

$$\tilde{a}_{ft}^{ij} = \frac{a_t^{ij} s_{ft}^{ij}}{s_{ft}^i} \tag{6.11}$$

where $\tilde{a}_{ft}^{ij}$ is firm specific and distinguished from a_t^{ij} for the population.

To summarize the demand side, I assume passengers make demand decisions before the matching process but with rational belief of the demand and supply level as proxy for matching probability, waiting time, and network effects. Given a set of demand parameter values, the demand model can predict two main things. First, the model predicts market shares s_{ft}^i and demand u_{ft}^i. Second, it predicts the endogenous distribution of firm's dropoffs $\tilde{\boldsymbol{A}}_{\boldsymbol{f}} = \{\tilde{a}_{ft}^{ij}\}$. Though I do not directly observe demand and supply in the data, the estimation section will discuss how to solve supply/demand by fitting model predicted pickups to pickups observed in the data.

6.2 Drivers' Choice Problem

At the end of each period, if the driver is employed, he will travel to the destination requested by the passengers. Drivers cannot refuse to deliver a passenger once matched. The probability of an employed car of firm f in location i at time t travelling to destination j is $\tilde{a}_{ft}^{ij}$ which is obtained from equation (6.11). Search decisions are made only by unmatched drivers at the end of each period.

If the driver is unmatched after the current period, he makes a decision on which location to search for passengers in the next period. Drivers are identical within firm and make individual decisions without coordination by the firm. Similar to passengers, when drivers consider a location to search in the next period, they know the matching probability, expected profit conditional on being matched and continuation value if not matched in that location. In order to know the matching probability, drivers need to have rational expectation of the demand and supply distribution across locations in the future. In equilibrium, a driver's belief is consistent with the realized supply and demand distributions.

At the end of a period, an unmatched driver of firm f in location i makes a search decision after observing supply shocks $\{\epsilon^j\}_{\forall j}$ by choosing the location with maximum value:

$$j^* = \arg\max_j \{\underbrace{V^j_{ft+\chi^{ij}_t} - c^{ij}_t + \rho_f(V^j_{ft+\chi^{ij}_t} - \min_l\{V^l_{ft+\chi^{il}_t}\})\mathbb{1}_{\chi^{ij}_t=1}}_{\Delta^{ij}_{ft}} + \epsilon^j_f\} \tag{6.12}$$

where c^{ij}_t is the cost of travelling from i to j calculated as $c^{ij}_t = 0.75 * \text{distance}^{ij}_t$. The cost per mile is set to be 0.75 dollars. $V^j_{ft+\chi^{ij}_t}$ is the driver's ex-ante value of searching location j in period $t + \chi^{ij}_t$ before the matching process in period $t + \chi^{ij}_t$. The number of periods travelling from i to j at t is χ^{ij}_t which is time cost compared to c^{ij}_t. Drivers are assumed not to pick up passengers along his way to the search location. When the driver chooses j which is far from i, he has to account for the loss of not searching for passengers until the next χ^{ij}_t periods. This time costs plays two important roles in the model. First, it contributes to mismatches across locations due to mobility. For example, suppose that location i has many vacant cars at the end of period t and there are many passengers in another location j far from i in the next period. Drivers cannot arrive at j in one period and result in excess demand in j and excess supply in locations near i in $t + 1$. Second, I can study benefits of traffic improvement by changing χ^{ij}_t. Both c^{ij}_t and χ^{ij}_t are allowed to vary over periods t and route i, j. These two variables can be directly calculated from the data. Finally, the parameters ρ_f measure the extra benefits of searching locations that are close to current location i, $\chi^{ij}_t = 1$. The difference of $V^j_{ft+\chi^{ij}_t}$ minus the minimal values guarantee non-negative benefit. The ex-ante value is defined as:

$$V^j_{ft} = \phi^j_{ft}\underbrace{\Big(\sum_l \tilde{a}^{jl}_{ft}(p^{jl}_{ft} - c^{jl}_t + V^l_{ft+\chi^{jl}_t})\Big)}_{VS^j_{ft}} + (1 - \phi^j_{ft})\underbrace{\mathbb{E}_\epsilon\big[\max_l\{\Delta^{jl}_{ft} + \epsilon^l_f\}\big]}_{VF^j_{ft}}. \tag{6.13}$$

In equation (6.13), ϕ^j_{ft} denotes the matching probability of drivers, $\phi^j_{ft} = m^j_{ft}/v^j_{ft}$. The conditional expected value on being matched is denoted as VS^j_{ft}. Conditional on being matched, the expected profit is obtained by averaging over all possible destinations l with weights $\tilde{a}^{jl}_{ft}$. Recall that $\tilde{a}^{jl}_{ft}$ measures the firm-specific destination distribution obtained in (6.11). It is important to note that the competition across firms affects both ϕ^{j}_{ft} and $\tilde{a}^{jl}_{ft}$. However, the competition within a firm only affects ϕ^{j}_{ft}. The profit conditional on trip jl includes the fare of the trip, cost of travelling, and continuation value in location l after dropoff in $t + \chi^{jl}_{t}$ period.

The second part of (6.13) is the continuation value of not being matched in j. The interpretation of each variable is the same as (6.12). The conditional expected value on not being matched is denoted as VF^{j}_{ft}. Since drivers do not observe the realized supply shocks ϵ's until the end of period, the continuation value takes an expectation over all possible supply shocks[19]. The supply shock ϵ_f follows i.i.d T1EV distribution with scale parameter σ_f for each firm f such that the continuation value of being unmatched has an explicit form:

$$\mathbb{E}_{\varepsilon} \max_{l}\{\Delta^{jl}_{ft} + \epsilon^{l}\} = \sigma \log\left(\Sigma_l \exp(\Delta^{jl}_{ft}/\sigma_f)\right) \tag{6.14}$$

Given this feature of supply shock's distribution, the deterministic transition probability of unemployed drivers of firm f in location i searching j in the next period is:

$$\pi^{ij}_{ft} = \frac{\exp(\Delta^{ij}_{ft}/\sigma_f)}{\Sigma_l \exp(\Delta^{il}_{ft}/\sigma_f)} \tag{6.15}$$

The scale parameter σ_f controls for incentives of drivers searching certain locations captured by values Δ^{il}_{ft} other than shocks ϵ^{l}. For example, large σ_f implies that drivers' search decisions are largely driven by random supply shocks which leads to an even allocation of drivers' searches across locations.

Combining the transition of employed cars $\tilde{\boldsymbol{A}}_{\boldsymbol{f}} = \{\tilde{a}^{ij}_{ft}\}$ calculated in (6.11) and the policy function of unemployed cars $\boldsymbol{\Pi}_{\boldsymbol{f}}$ of equation (6.15) gives the law of motion for the state transition. The state includes the status of all in-transit cars. The state at the beginning of period t is a collection of $\{S^{i}_{t}\}_{\forall i}$ where S^{i}_{t} is a collection of $\{\tilde{v}^{i}_{ft,k}\}_{f,k}$

[19]It is important to note that equation 6.13 does not necessarily imply a positive correlation between search value and matching probability of a market. This positive correlation is crucial for drivers to prefer high demand, therefore providing high supply, given all else equal. The positive correlation requires $VS^{j}_{ft} > VF^{j}_{ft}$ such that drivers prefer to be matched. While this inequality is not imposed in the estimation, the result shows it is satisfied for all markets.

with $\tilde{v}^i_{ft,k}$ indicating the number of cars for firm f that will arrive at location i in the next k periods. When $k=1$, it implies that the supply at period t satisfies $v^i_{ft} = \tilde{v}^i_{ft,k=1}$. At the end of each period, the transition of employed and unemployed cars update the state such that:

$$\tilde{v}^i_{ft+1,k} = \tilde{v}^i_{ft,k+1} + \Sigma_j m^j_{ft}\tilde{a}^{ji}_{ft}\mathbb{1}_{\chi^{ji}_t=k} + \Sigma_j(v^j_{ft} - m^j_{ft})\pi^{ji}_{ft}\mathbb{1}_{\chi^{ji}_t=k}, \forall f, i, k \tag{6.16}$$

To interpret (6.16), at beginning of period $t+1$, the number of firm f drivers that will arrive at location i in k periods is composed of three parts: (1) those who will arrive at i in $k+1$ periods at the beginning of period t; (2) those who pickup passengers at time t and will arrive at i in k periods; (3) those unemployed drivers of period t who decide to search location i next but will arrive in k periods. The next section introduces the matching functions used to calculate matches m^i_{ft}.

6.3 Matching Function

During the matching process in each period within a location, I use an explicit functional form to predict the matching outcomes. Buchholz (2022) and Frechette et al.(2019) assume a matching process with friction for taxis within a location. Frechette et al.(2019) simulate the process of taxis searching over grids within a location for passengers. Buchholz (2022) assumes an urn-ball random matching process and a corresponding explicit functional form is derived by Burdett, Shi and Wright (2001). I use the same functional form as Buchholz (2022) and Burdett, Shi and Wright (2001) with a modification to reflect heterogeneity in frictions across locations. This matching process is only applied to taxis within locations outside the two airports. For taxi at airports and Uber in all locations, there is perfect matching within the location.

The matching function for a taxi is obtained in the following way. Given taxis' demand u^i_{yt} and supply v^i_{yt} in location i at period t, passengers are assumed to randomly visit the taxis and of those visiting the same car only one can be successfully matched. Other unmatched passengers will leave with the subway. I do not distinguish where passengers and cars are located within the market such that all cars are identical to the passengers. That means a passenger has equal probability of visiting any car. Each car receives a passenger's visit with probability $1/v^i_{yt}$. The probability of a taxi not receiving a visit is $(1 - 1/v^i_{yt})^{u^i_{yt}}$ and the probability of a taxi being matched is

$1-(1-1/v_{yt}^i)^{u_{yt}^i}$. However, in locations with larger size of area, it is more difficult to know the exact location of cars. Thus, I add an location-specific parameter γ_i such that the probability of a car being matched becomes $1-(1-1/(\gamma_i v_{yt}^i))^{u_{yt}^i}$. A higher value of γ_i will decrease the probability of being matched. To be specific, I define $\gamma_i = \gamma_1\{i \in \text{Manhattan}\} + \gamma_2\{i \in \text{Outer Borough}\}$. Since taxis within i have the same probability of being matched, the number of matches is:

$$
\begin{aligned}
m(u_{yt}^i, v_{yt}^i) &= v_{yt}^i\Big(1-(1-\frac{1}{\gamma_i v_{yt}^i})^{u_{yt}^i}\Big) \\
&\approx v_{yt}^i(1-\exp(-\frac{u_{yt}^i}{\gamma_i v_{yt}^i}))
\end{aligned}
\tag{6.17}
$$

Function (6.17) itself allows matching friction due to coordination failures such that there is possibility that some cars receive no visits and some passengers are not matched.

As for Uber and aiports, the matching between drivers and passengers is assumed to be frictionless. Uber use mobile technology to assign passengers and drivers into a one-to-one pair without coordination failure as above. At the airports, drivers are waiting in a queue to match passengers one by one without coordination failure as well. The explicit functional form of perfect matching is $m_{xt}^i = \min\{u_{xt}^i, v_{xt}^i\}$. However, one drawback of this function is that given $m_{xt}^i \approx v_{xt}^i$, the inverted demand satisfies $u_{xt}^i = m_{xt}^i \approx v_{xt}^i$. In other words, this function does not capture excess demand. Instead, I use a hyperbolic function to approximate perfect matching as shown in figure 5. When the demand-to-supply ratio is greater than 1, the matching probability of drivers approaches 1. When the ratio is less than 1, the matching probability approaches the 45 degree line that is $m_{xt}^i \approx u_{xt}^i$. In the case $m_{xt}^i \approx v_{xt}^i$, it returns an u_{xt}^i greater than v_{xt}^i as excess demand, though its value depends on the curvature rather than any economic assumption. The explicit matching function form of figure 4 is obtained by solving m_{xt}^i from (with small value of ϵ):

$$
(\frac{m_{xt}^i}{v_{xt}^i} - 1)(\frac{m_{xt}^i}{v_{xt}^i} - \frac{u_{xt}^i}{v_{xt}^i}) = \epsilon \tag{6.18}
$$

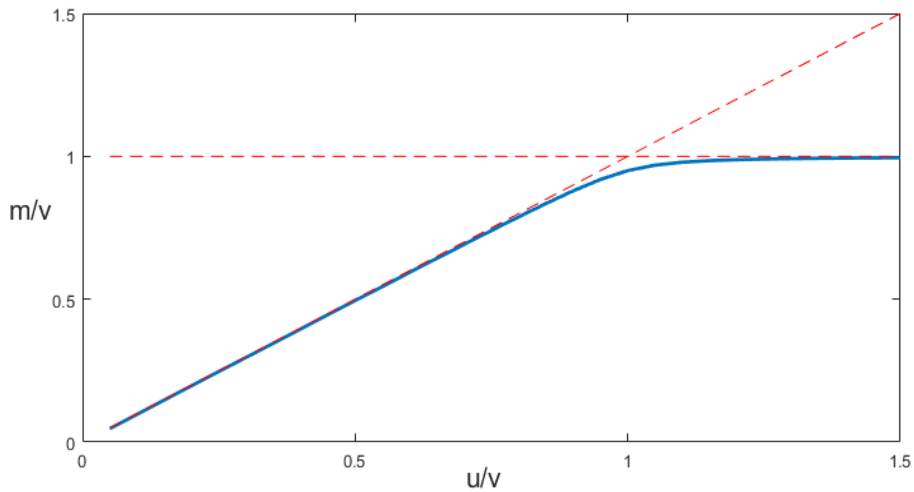


Figure 5: Hyperbolic function for perfect matching

6.4 Equilibrium

To solve the equilibrium of the model, a standard dynamic oligopoly model is inappropriate for this game due to the large number of drivers in the game. For example, the number of possible states of allocating N drivers into I locations will be C_{N+I-1}^{I-1} which is large when N is large[20]. The model would be intractable and computationally infeasible if drivers' expected profits are taken over all possible market states. Instead, I assume drivers make their search decision only based on their own state and knowledge of the deterministic market evolution of demand and supply distributions. This concept comes from oblivious equilibrium (Weintraub et al.(2008)) when players are atomistic and individual decisions do not measurably impact the aggregate market state. The key information about the market state is the distribution of supply and destination distribution of in-transit cars. A driver's own state is denoted as s_t which includes his location at time t. The state of the city at time t is denoted as the collection $\{\mathcal{S}_t^i\}_{i\in I}$. For any i, $\mathcal{S}_t^i$ includes information about arrival of cars in next K periods, hence collection $\{\tilde{v}_{ft,k}^i\}_{f,k\in K}$. In OE, drivers make optimal search decisions according to $\{s_t, \{\mathcal{S}_t^i\}_{\forall i,t}\}$. Drivers' belief on the evolution of market state (i.e. supply distri-

[20]This number is obtained by counting the number of outcomes of putting N balls in I different urns $C_{N+I-1}^{I-1} = \dfrac{(N+I-1)!}{(I-1)!(N)!}$ allowing for empty urns.

bution) is consistent with the realized state in equilibrium[21]. Given the deterministic evolution of supplies, equilibrium demand can be calculated for each market from the discrete choice model. The definition of equilibrium is summarized as follows:

Definition Equilibrium is a sequence of supply $\{v^i_{ft}\}$, beliefs of state transition $\{\hat{v}^i_{ft,k}\}$, policy function of unemployed cars $\{\pi^{ij}_{ft}\}$, transition of employed cars $\{\tilde{a}^{ij}_{ft}\}$ for $\forall i, t, f$ and given initial distribution of supply $\{v^i_{ft=1}\}$ such that:

1. At the beginning of period t, passengers make discrete choice between firms based on (6.1)-(6.8). Market demand is calculated from (6.10).

2. Matches are made between supply and demand for each firm within the market. The matching process follows (6.17) for taxi and (6.18) for Uber and airport.

3. Transition of employed cars follows $\{\tilde{a}^{ij}_{ft}\}$ obtained by Bayes' rule (6.11).

4. At the end of each period, unemployed drivers follow policy function π^{ij}_{ft} calculated in (6.15) based on beliefs of state transition $\{\hat{v}^i_{ft+1,k}\}$.

5. Realized state transition is obtained by combining both employed $\{\tilde{a}^{ij}_{ft}\}$ and unemployed cars $\{\pi^{ij}_{ft}\}$. State of next period is updated to $\tilde{v}^i_{ft+1,k}$ by (6.16).

6. At the beginning of next period, both employed and unemployed cars arrive and form the new supply $\tilde{v}^i_{ft,k=1} = v^i_{ft}$.

7. Drivers' belief is consistent such that $\hat{v}^i_{ft+1,k} = \tilde{v}^i_{ft+1,k}$ for all i, t, f, k

This model's equilibrium is quite similar to Buchholz(2022) and still satisfies finite horizon and finite action-space for existence of equilibrium.

7 Estimation

This section discusses the estimation process in detail. The key feature of estimation is that supply v^i_{ft} and demand u^i_{ft} in any market are not directly observed in the data. Instead, the data only has observation on pickups m^i_{ft} as the outcome of matching process. Estimation of the model is searching for parameter values, equilibrium demand

[21]Another way to understand the OE in this model is that instead of knowing the evolution of the supply distribution, drivers know the evolution of ex-ante search values $\{V^i_{ft}\}_{\forall i,t}$. Knowing the supply distribution or search values are interchangeable given one step calculation of (6.8).

and supply, that generate matches and transition of cars that fit the data. The demand side parameters include mean utility $\{\delta_{ft}^i\}_{\forall f,t,i}$, price coefficients $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$, and substitution parameter β within group of taxi and Uber. The supply side parameters include supply shock parameter σ_f, and bonus of searching nearby locations ρ_f. Finally, there are parameters defining the matching function of taxis $\{\gamma_1, \gamma_2\}$. These parameters are called structural parameters in the rest of this paper. Given estimates of the structural parameters, equilibrium supply $\boldsymbol{v_f}$ and demand $\boldsymbol{u_f}$, $f = 1, 2$ can be solved. The next step is to estimate reduced form parameters $\boldsymbol{\theta}$ in the mean utility $\boldsymbol{\delta}$. These parameters are important since they help build the feedback loop between demand and supply, the network effects. With $\boldsymbol{\theta} \neq 0$, demand responds to the change of equilibrium supply in the counterfactuals. Parameters in $\boldsymbol{\delta}$ can be estimated using linear regression.

7.1 Estimating Structural Parameters

The estimation algorithm is diagrammed in figure 6. The first step is to estimate demand and supply side parameters $\{\boldsymbol{\delta}, \boldsymbol{\alpha}, \beta, \boldsymbol{\gamma}, \boldsymbol{\sigma}, \boldsymbol{\rho}\}$. Since the mean utilities $\boldsymbol{\delta}$ has a larger dimension than other parameters, I separate the structural parameters into two subsets and estimate the subsets separately in two procedures. First, given the parameter values of $\{\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma}, \boldsymbol{\sigma}, \boldsymbol{\rho}\}$, equilibrium demand and supply are solved such that model generates the same pickups in markets as observed in the data. Second, $\{\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma}, \boldsymbol{\sigma}, \boldsymbol{\rho}\}$ are updated to minimize deviation of the model predicted distribution of dropoffs for taxi trips to the distribution in data. The first procedure is similar to Buchholz (2022) with an extended computation due to the introduction of the discrete choice demand to the model. In Buchholz (2022), he searches for equilibrium demand $\boldsymbol{u}$ which generates pickups in the sample. Instead of solving for demand $\boldsymbol{u}$, I am solving for mean utilities $\boldsymbol{\delta}$ which have a one-to-one mapping to demand according to BLP contraction mapping. Solving mean utilities is crucial to predict destinations of passengers for each firm $\boldsymbol{\tilde{A}_f}$ which forms objective function for optimization in the second procedure.

In details, beginning in the upper left corner of figure 6. Given initial values of $\{\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma}, \boldsymbol{\sigma}, \boldsymbol{\rho}\}$, initial values of mean utilities $\{\delta_{ft}^i\}_{\forall f,t,i}$ are calculated by using pickups as demands to calculate market shares for inverting. The procedure one to solve equilibrium demand and supply is further broken into two loops. In the inner loop, given the initial values of mean utilities and demands, I solve the corresponding equilibrium supplies $\{v^i_{ft}\}_{\forall f,t,i}$ using backward induction. In detail, from the last period T, given the demand distribution and zero search values $V^i_{ft} = 0, \forall t > T$, calculate continuation values $\{V^i_{fT}\}$ according to (6.13) for an arbitrary supply distribution $\{v^i_{fT}\}$[22]. For period $T-1$, given the search values $\{V^i_{fT}\}$ and demands $\{u^i_{fT-1}\}$, calculate search values $\{V^i_{fT-1}\}$ and updated supplies of next period $\{v^i_{fT}\}$ for an arbitrary supply distribution $\{v^i_{fT-1}\}$. Repeat this process backward until period $t = 1$. The supply distribution in period 1 is assumed to be proportional to pickups distribution of period 1. This process only provides the starting values of supplies $\{v^i_{ft}\}_{\forall f,t,i}$, because the supplies and search values are not consistent. Iterate this process updating supplies and search values from $t = 1$ forward to $t = T$ until equilibrium supplies are obtained[23]. The whole procedure is summarized in algorithm 1 in the appendix. This procedure generates a mapping from mean utilities to demands $\boldsymbol{u} = u(\boldsymbol{\delta})$ and supplies denoted as $\boldsymbol{v} = \Gamma(\boldsymbol{\delta})$. The pickups are $\boldsymbol{m} = m(u(\boldsymbol{\delta}), \Gamma(\boldsymbol{\delta}))$.

In order to fit pickups to the monthly average pickups in the data, denoted as $\bar{\boldsymbol{m}}$, I update mean utilities in the outer loop. To be specific, given the mean utilities $\boldsymbol{\delta}^k$ and solved equilibrium supply $\boldsymbol{v}^k$ of the kth iteration, I invert the matching function using pickups in the data to obtain the updated mean utilities such that $\boldsymbol{\delta}^{k+1} = u^{-1}(m^{-1}(\bar{m}, \Gamma(\boldsymbol{\delta}^k))$. Plug the updated $\boldsymbol{\delta}^{k+1}$ into the inner loop to update equilibrium supplies $\boldsymbol{v}^{k+1} = \Gamma(\boldsymbol{\delta}^{k+1})$ until the model predicted matches fit the data. The outer loop is summarized in algorithm 2 in the appendix.

The second procedure, in the bottom right corner of figure 6, estimates $\{\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma}, \boldsymbol{\sigma}, \boldsymbol{\rho}\}$. Given the estimated mean utilities conditional on the parameter values of $\{\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma}, \boldsymbol{\sigma}, \boldsymbol{\rho}\}$, the transition matrix of taxi passengers $\{\tilde{A}^i_{yt}\}$ can be calculated for all the markets in order to match the observations in the data. For each period, the transition probabilities form a 40 by 40 matrix and there are totally 1600×60 data point to match. Instead of matching the probabilities point-to-point to the data, I aggregate the pickups and dropoffs over locations and periods and recalculate the transition probabilities in larger areas over 20 half hours. This process applies to both data and model generated transitions. Then, the sum of weighted squared differences between model generated transition of taxi passengers and that of the data can be calculated as the objective

[22]The day ends for taxi drivers in period T due to shifts. However, Uber drivers have no shifts. Uber drivers may continue to work after period T with positive search values. In this estimation, I assume search values are equal over locations for $t > T$ and normalized to 0. Normalization won't affect transition probabilities since constants are cancelled out by equation 6.15.

[23]This process does not satisfy contraction mapping. In order to find the fixed point solution, I applies iterative method of average damping.

function of nonlinear least squares estimation. Estimators are defined as 7.1 :

$$\{\hat{\alpha}, \hat{\sigma}, \hat{\gamma}, \hat{\rho}, \hat{\beta}\} = \underset{\alpha,\sigma,\gamma,\rho,\beta}{\arg\min} \Sigma_{t,i} \big(\bar{m}^i_{yt}(\bar{\tilde{A}}^{ij}_{yt} - \tilde{A}^{ij}_{yt}(\alpha, \sigma, \gamma, \rho, \beta))\big)^2 \tag{7.1}$$

7.2 Estimating Parameters in Mean Utility

In the second step, given the estimates of $\{\Theta, \boldsymbol{m}, \boldsymbol{\delta}, \boldsymbol{u}, \boldsymbol{v}\}$ with $\Theta = \{\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma}, \boldsymbol{\sigma}, \boldsymbol{\rho}\}$ in step 1, variables in mean utility equation that are not directly obtained in data such as market demand and supply in equation (6.2) can be calculated. In the specification of mean utility, the coefficient on demand measures direct network effect and coefficient on supply measures indirect network effect. An OLS regression of equation (6.2) has endogeneity problem since demand and supply are all correlated with unobserved demand shock ξ^i_{ft}. For supply v^i_{ft}, I use arrival of employed cars of firm f at the beginning of period as instrument. The argument is that these cars visit location i because of their passengers' destination. It is reasonable to assume that demand shock of current period is not correlated with the destination of passengers picked up from other locations in previous periods. One exception to this assumption could be that passengers visit location i and leave i in the same period after arrivals. This instrument is correlated with supply as it constitutes supply together with arrival of unemployed cars. To solve the endogenous problem of demand, two instruments are used including market size λ^i_t and arrival of opponent's cars. Market size is exogenous and correlated with demand as in (6.10). Arrivals of opponent's cars are uncorrelated with demand shocks following the same argument above. It correlates with demand through the discrete choices of passengers.

To summarize, given demand parameter values and market size, market demand can be calculated. Then, given supply side parameter values, solve drivers' optimal decisions and calculate supply using backward induction in the inner loop of procedure one. In the outer loop of procedure one, the matching function can generate pickups given supply and demand to fit pickups in the data. The model is estimated in procedure two such that deviation of model generated transitions of employed taxis to the transitions in the data is minimized. Finally, mean utility equation is estimated using IV regression.

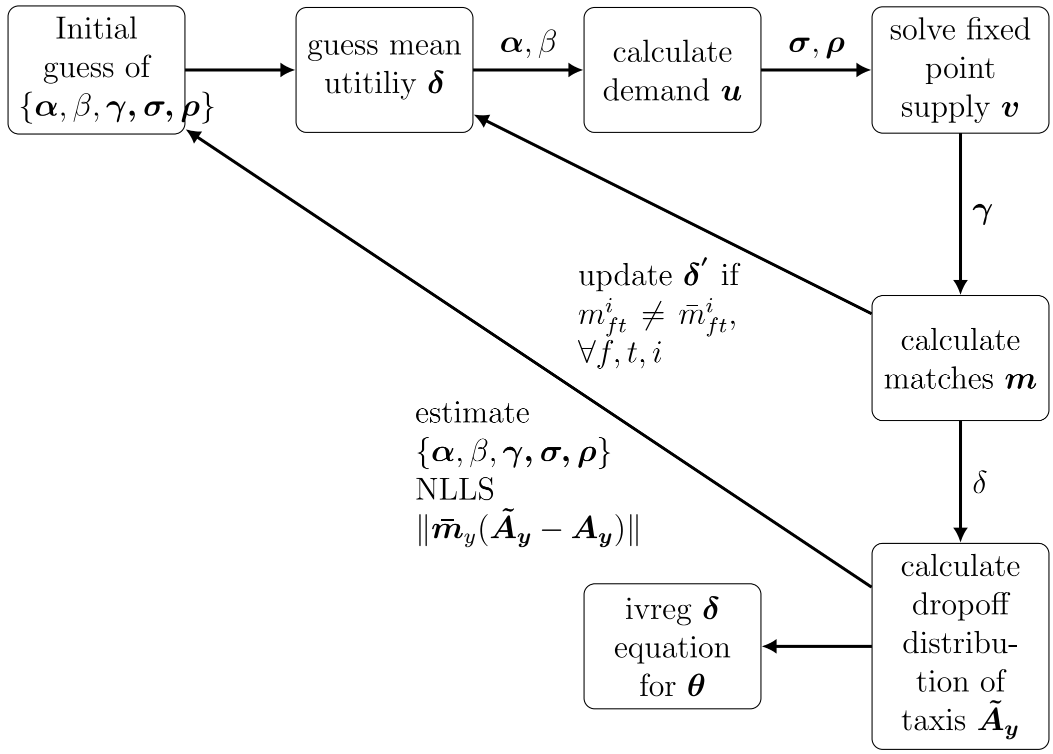


Figure 6: Overview of the estimation process

7.3 Identification

The parameters are identified by variation of pickups, prices, the travelling pattern of population and the pattern of taxi passengers over markets in the data. Given a set of nonlinear parameter values, mean utilities $\{\delta^i_{ft}\}$ are identified by the variation of pickups across firms and markets. The mapping from mean utilities to pickups follows algorithm 1&2 in which I firstly map mean utilities to demands and corresponding dynamic supply distributions followed by calculating matches given supply and demand. Consider two identical markets (i.e. same market size, travelling pattern and prices) with different pickups $\bar{m}^1 > \bar{m}^2$. In a static model, market 1 with higher pickups implies a higher demand and supply than market 2 and therefore $\delta^1 > \delta^2$. However, in a dynamic model, the corresponding supplies for given mean utilities are more complicated than in a static model. In the dynamic model, supply may not fully respond to demand variation across locations due to mobility restriction of cars conditional on their locations in the previous period. However, given the one-to-one mapping from inter-period demands to dynamic supplies, the dynamic pickup patterns in the data help to identify mean utilities.

Identification of the price coefficient α comes from variation in population travelling patterns over markets. For example, two markets with same mean utility $\delta^i_{ft} = \delta^j_{ft}$ and market size λ^i_t but with different destination distributions of passengers $A^i_{ft} \neq A^j_{ft}$ will have different demands $u^i_{ft} \neq u^j_{ft}$. The demand level relative to subway riderships also helps identifying price coefficient. As for supply shock parameter σ, it controls for transition of unemployed drivers. Given the search values over locations $V^i_{ft}, \forall i$, high σ implies equal probability of search in each location i.

Identification of the matching function parameters γ is less intuitive. They are crucial to connect the mapping from δ to pickups. Different γ do not affect equilibrium supply as much as equilibrium demand. The reason is that drivers' matching probability depends directly on successful pickups m^i_{ft} rather than potential demand u^i_{ft}. Given the pickups generated by model equal those of data in estimation, the matching function parameter γ does not change the probability much. However, given a fixed supply level, inefficient matching of γ affects estimated mean utilities δ. For example, given fixed pickup and supply, a large γ (less efficient) generates high potential demand and corresponding high δ. The mean utilities further affect destinations of passengers in equation (6.11) and drivers' profits. Since ex-ante search values V^i_{ft} depend not only on the matching probability but also on profits, equilibrium supply also reacts to changes of γ. Thus, in order to identify the $\{\delta^i_{ft}\}$ with restriction to γ, I use the dropoffs of taxis in the data such that model predicted dropoffs of taxis match the data. The reason is that different magnitudes of δ, which is shifted by γ, not only affect market shares relative to outside option but also affect the distribution of firm-specific dropoffs as in (6.11). High δ could dominate price effects on sorting passengers between taxi and Uber. As a result, it makes the distribution of taxi passengers' dropoffs close to the distribution of population[24]. Hence, I use taxis' dropoffs to identify the matching function parameters.

[24]For a given market i, t, the mean utility δ^i_{ft} is common for all destinations j. It shifts the conditional shares on routes uniformly. In the extreme case of taxis' mean utility large enough, destination of taxis' passengers is exactly same to population's.

8 Results

8.1 Estimates of Structural Parameters

The estimation results are listed in table 4. The estimates of price coefficients $\boldsymbol{\alpha} = \{\hat{\alpha}_k\}_{k=1,2,3,4}$ for different trip distances are specified in (6.3). The price coefficient for trip distance less than 3 miles is $\hat{\alpha}_1 = -0.81$ and becomes less sensitive for long distance trip $\hat{\alpha}_3 = -0.41$. The estimate $\hat{\alpha}_4 = -0.26$ is for trips between JFK and Manhattan which charges flat rate \$52 by taxi. The estimate of another demand parameter $\beta \in [0, 1]$ in nested logit demand defined in (6.5) is equal to 0.38. When $\beta \to 1$, demand shocks for taxi and Uber are highly correlated and when $\beta \to 0$ they are independent as in a simple logit model. The estimate of β indicates the substitution between taxi and Uber but the substitution is not quite strong. In the demand side, there are a large set of mean utilities $\{\delta^i_{ft}\}_{\forall f,t,i}$ to estimate and the statistics for taxi and Uber are listed below the demand parameters. The mean of taxis' $\{\delta^i_{yt}\}$ is 1.21 with maximum value at 5.05 and minimum value at -1.79. Uber's mean utilities $\{\delta^i_{xt}\}$ are less than taxi. The difference in mean utility can be inferred from the different market shares. If the supply is positively correlated with the mean utility, it may indicate existence of network effects between demand and supply as specified in (6.2).

There are two sets of parameters from the supply side. First, the estimates of supply shocks' scales $\sigma_f, f = y, x$ for taxi and Uber are 7.67 and 12.65. These two parameters affect the transition probability of unemployed cars in (6.15). Larger σ_f implies less effect of profit difference across locations on transition probability π^{ij}_{ft}. Conversely, a smaller σ_f will enlarge the difference in profits among locations such that drivers have a higher likelihood of searching high profit location. The estimate means that taxi drivers have greater incentive to search locations with higher search values than Uber drivers. Furthermore, controlling for profit difference across locations, these two estimates imply a higher chance that taxi drivers will overcrowd high profit locations and leave low profit locations undersupplied. In other word, σ_f enhances the role of profit gap caused by fixed price on matching frictions. The second set of supply parameters are ρ_f which measure the incentive of unmatched drivers to search locations nearby in the next period. This extra bonus from current location i to search location j is measured proportionally to $V^j_{ft+\chi^{ij}_t} - \min_l\{V^l_{ft+\chi^{il}_t}\}$ if $\chi^{ij}_t = 1$ (see 6.12). The estimate for taxi is 0.38 and for Uber is 0.27. Higher $\hat{\rho}_y$ means that taxi drivers have a greater incentive to search locations nearby than to visit a location far away.

The bottom of table 4 reports the estimates for random matching function 6.17. The $\gamma_i = \gamma_1\{i \in \text{Manhattan}\} + \gamma_2\{i \in \text{Outer Borough}\}$ measures within-market matching efficiency for taxis. I distinguish the efficiency in Manhattan and outer Boroughs. A high value of γ means inefficient matching within the market. For instance, given a fixed number of supply and demand, higher γ generates less successful matches. The estimates of this parameter in Manhattan area is 1.11 in comparison to 3.67 in Outer Boroughs indicating that the within-market matching is less efficient in Outer Boroughs.

Table 4: Estimates of structural parameters

parameter	*estimates*	s.e.
demand side parameters		
$\hat{\alpha}_1$	-0.81	(0.042)**
$\hat{\alpha}_2$	-0.58	(0.029)**
$\hat{\alpha}_3$	-0.41	(0.040)**
$\hat{\alpha}_4$	-0.26	(0.031)**
$\hat{\beta}$	0.38	(0.026)**
mean utilities	*mean*	*min/max*
$\hat{\delta}^i_{yt}$	1.21	-1.79/5.05
$\hat{\delta}^i_{xt}$	0.30	-1.62/3.94
supply side parameters		
$\hat{\sigma}_y$	7.67	(0.245)**
$\hat{\sigma}_x$	12.65	(0.513)**
$\hat{\rho}_y$	0.38	(0.023)**
$\hat{\rho}_x$	0.27	(0.025)**
matching function parameters		
$\hat{\gamma}_1$	1.11	(0.015)**
$\hat{\gamma}_2$	3.67	(0.345)**

** 1-percent or * 5-percent level significant

Along with the parameter estimates, summary statistics of several variables across the 2400 location-time markets are reported in table 5 for Uber and Taxi. The variables are the equilibrium values of demand, supply, matching probability, search values and difference in conditional expected values on being matched and unmatched. The mean of taxis' demand is 118.14 with 10th percentile at 19.63 and 90th at 268.19. In comparison, Uber's potential demand is much less, on average 20.92. The average supply of taxis across markets is 174.21 with percentiles 41.56 to 313.05. The average supply of Uber is 34.46, not much less than taxi considering the ratio of cars. The matching probabilities of taxi and Uber drivers are also different. Uber has a higher matching probability than taxis for two reasons. First, Uber has a smaller number of cars than taxi which indicates less cannibalization. Second, Uber has perfect matching as opposed to taxi's random matching within market which, ceteris paribus, increases the driver's matching probability. The ex-ante search value for taxi is, on average, 98.31 compared to 128.78 of Uber. The 10th and 90th percentile of search values for taxi is 21.97 and 174.92, the dispersion of which is mainly driven by the time of day in comparison to the market location. In general, Uber drivers have higher expected profit than taxis. This high profitability of Uber could be the result of surge pricing, matching efficiency and less competition within firm. The search value is a proxy for the revenue in dollars of an individual driver in a day shift. To interpret these numbers, one needs to realize that search values decrease over time because of the finite time horizon in dynamic game. For example, an Uber driver expects to earn a total revenue above 233.39 dollars at the beginning of the day. Finally, the table shows the difference between continuation value conditioning on being matched and not being matched. As mentioned in the model, the difference is not restricted but calculated and a positive sign implies positive correlation between search value and matching probability such that demand positively affects supply given all else equal.

Table 5: Statistics in equilibrium

	mean	*10th/90th percentile*
demand		
u^i_{yt}	118.14	19.63/268.19
u^i_{xt}	20.92	8.47/35.92
supply		
v^i_{yt}	174.21	41.56/313.05
v^i_{xt}	34.46	18.72/53.08
matching probability		
ϕ^i_{yt}	0.37	0.15/0.57
ϕ^i_{xt}	0.59	0.31/0.88
search value		
V^i_{yt}	98.31	21.97/174.92
V^i_{xt}	128.78	34.29/233.39
difference in conditional expected values		
$VS^i_{yt} - VF^i_{yt}$	15.10	7.13/18.82
$VS^i_{xt} - VF^i_{xt}$	15.56	10.25/21.96

I demonstrate the heterogeneous continuation values Δ_{ft}^{ij} across destination j for any origin location i in figure 7. The horizontal axis is time periods and vertical axis is the difference. Each line represents a location i over t about the variation of continuation values $\{\Delta_{ft}^{ij}\}_j$ over j. The differences of Δ_{ft}^{ij} directly determine the search choices of drivers based on equation (6.15). I use two standard deviations, $2{*}std_j\{\Delta_{ft}^{ij}\}$, to represent the differences in incentives of drivers to search among locations. A small value of deviation means drivers have equal incentive to search any of the 40 locations. For example, at the beginning of the day, an unemployed taxi driver in a location faces a heterogeneous continuation value across $j's$ with two standard deviations equal to 7. At the end of the day shift, since all continuation values go to zero, the lines converge to the horizontal axis. The variation of continuation values is stable across time for taxi, while it increases at 8:30 a.m.($t = 15$) for Uber and become stable afterwards. It means that during 8:30 a.m., no matter where the unemployed Uber drivers are located, some locations are much more profitable to search than some other locations.

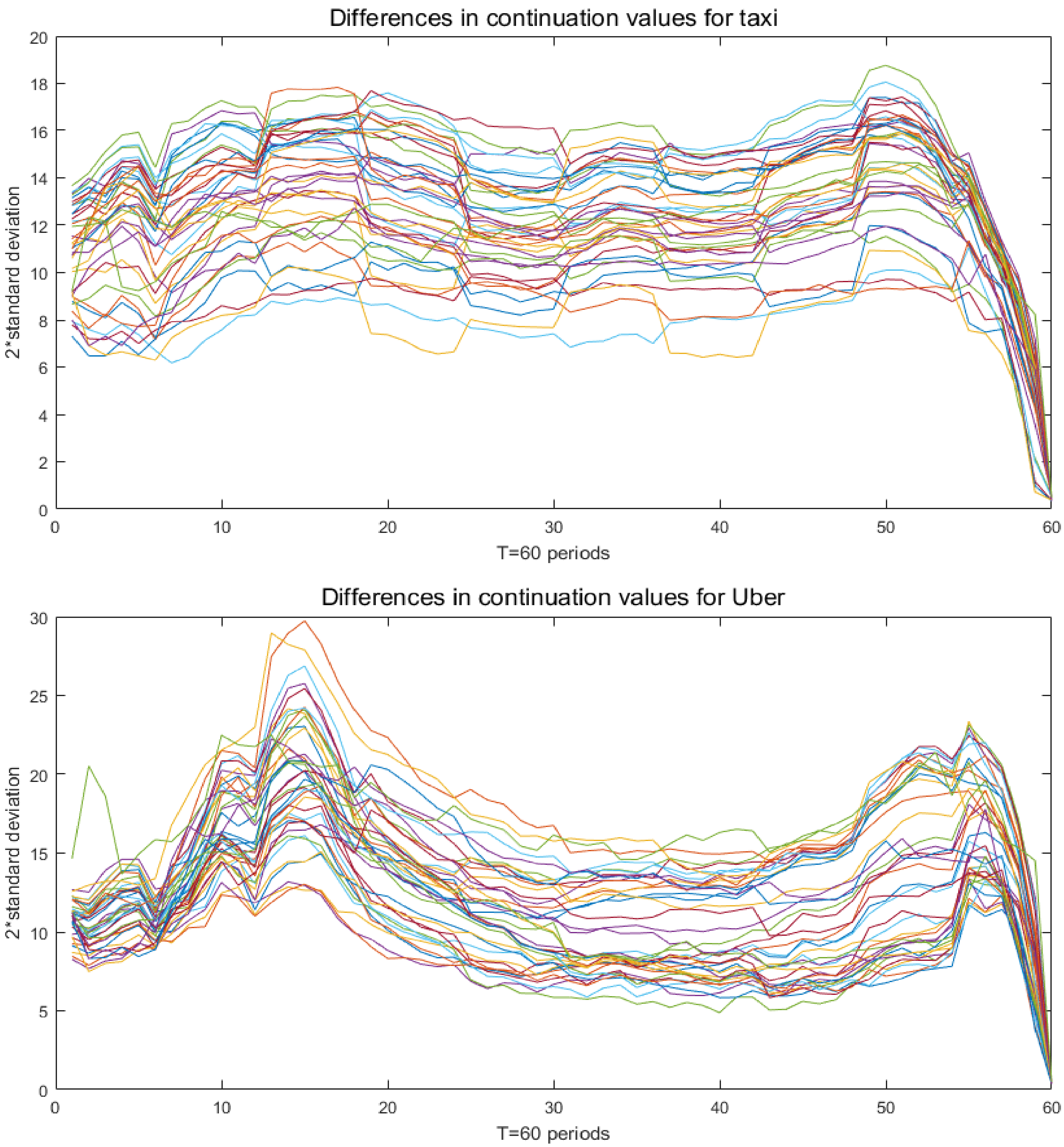


Figure 7: Heterogeneous expected continuation values conditional on being unmatched

8.2 Estimates of Reduced form parameters

Given previous results, I estimate a linear regression of mean utilities $\{\delta^i_{ft}\}$ on variables in equation (6.2). The coefficients of interest are θ_1 and θ_2, the coefficient on supply and demand in the utility function. The supply coefficient θ_1 measures indirect network effect from the other side of the market. It is the net effect of supply on the utility of choosing the product including, but not limited to, the impact via matching probability and waiting time. Likewise, the coefficient on demand θ_2 captures the net effect of demand on utility. One possible channel is that the demand decreases the likelihood of being matched and therefore negatively affects utility. Conversely, it can also positively affect utility of choices via consumers learning from each other and herding. The cause of this joint effect, as discussed in section 6.1, is approximation of matching probability and log linear assumption of ex-post utility, see 6.2.

I use market size, arrival of drivers from the same and opponent firm as instruments for $\ln u, \ln v$. The OLS and 2SLS regression results are shown in table 6. In the regression, I add an interaction term between Uber dummy and logarithm of demand and supply levels. The estimates show positive effects of both demand and supply on utility. Moreover, the effects are larger for utility of choosing taxis than for Uber. The positive sign of supply coefficient means that higher supply level increases the demand. Since demand will positively affect supply in the model, these two effects form the positive feedback loop between drivers and passengers. It is a little surprised that the coefficient on demand is also positive. Demand is expected to negatively affect matching probability after controlling for supply due to congestion effect. One way to explain this positive sign is that there exists strong positive direct network effect among passengers which offset the congestion effect on probability. For example, if my friend uses taxi (or Uber), I would also like to choose taxi (or Uber). Comparing coefficients on demand between OLS and 2SLSIV shows that OLS overestimates θ_2 due to endogeneity as expected. However, the IV regression has a higher estimate of θ_1 than OLS. One possible explanation is that it is easy to understand how demand correlates with demand shock controlling for supply but hard to understand how supply correlates with demand shock controlling for demand. These estimates are used in the counterfactual to compare the scenarios with or without network effect by allowing the mean utility to react to changes of demand and supply or not. This will generate different equilibria in order to understand the consequences of ignoring network effects.

Table 6: Estimates of network effects

Dependent variable δ^i_{ft}		
	OLS	*2SLSIV*
$\ln v$	0.054	0.496
	(0.019)**	(0.07)**
$\ln u$	0.53	0.249
	(0.021)**	(0.06)**
$\ln v \times d_x$	-0.053	-0.224
	(0.039)	(0.11)*
$\ln u \times d_x$	-0.014	-0.086
	(0.038)	(0.08)
Uber dummy d_x	0.19	1.21
	(0.054)**	(0.18)**
constant	-2.23	-3.71
location fixed effects	YES	YES
time fixed effects	YES	YES

** 1-percent or * 5-percent level significant

8.3 Benchmark Welfare

In this section, I discuss and analyze the matching efficiency of this industry given the model's implied demand, supply and matches. The key factors of interest are the two types of frictions in the model, within-location friction (type I) and cross-location friction (type II). Within-location friction is measured by the mismatches between drivers and passengers of the same firm within the market. This type of friction is mainly driven by the matching function (6.17). Cross-location mismatches occur in the same period, because some locations have more drivers than demand whereas other locations have more demand than supply. Excess supply in a location can be counted as $\max\{v^i_{ft} - u^i_{ft}, 0\}$ and likewise excess demand is counted as $\max\{u^i_{ft} - v^i_{ft}, 0\}$. These expressions do not account for the mismatches due to random matching within market so that I can distinguish these two types. In period t, the city level aggregate demand is $\Sigma_i u^i_{ft}$ and aggregate supply is $\Sigma_i v^i_{ft}$. The maximum aggregate matches that can be made without type I friction are $\Sigma_i \min\{u^i_{ft}, v^i_{ft}\}$. Given the aggregate demand and supply level fixed, the efficient matches should be $\min\{\Sigma_i u^i_{ft}, \Sigma_i v^i_{ft}\}$ from the city aggregate perspective[25]. The difference is:

$$\begin{aligned} &\min\{\Sigma_i u^i_{ft}, \Sigma_i v^i_{ft}\} - \Sigma_i \min\{u^i_{ft}, v^i_{ft}\} \\ &= \min\{\underbrace{\Sigma_i \max\{0, u^i_{ft} - v^i_{ft}\}}_{\text{aggregate excess demand}}, \underbrace{\Sigma_i \max\{v^i_{ft} - u^i_{ft}, 0\}}_{\text{aggregate excess supply}}\} \end{aligned} \tag{8.1}$$

Expression 8.1 counts the minimum of aggregate excess supply and demand. Unlike type I friction, type II friction is mainly driven by the endogenous decisions of drivers and passengers. There are three limitations for 8.1 to be a good measure of friction. First, it only counts the static mismatches in a given period. Less efficient matches in the current period could result in better matches in the next period considering the mobility of drivers across locations. Second, it counts the trips equally, but passengers and trips are not identical. I improve this measure by calculating trips' values in dollars. Third, this measure does not necessarily correlate with the number of matches, especially when excess supply is greater than excess demand. For example, an equilibrium may have both high type II friction and large number of matches.

The welfare statistics are listed in table 7. The first panel displays the type I

[25] Lagos (2000) treats this expression as efficient aggregate matches and an aggregate matching function generating fewer matches has friction. However, in his paper, the demand is exogenous and there is no demand-supply feedback loop.

friction in quantity and dollars. For example, there are a total of 95,547 within-location mismatches for all markets in a representative day shift. These mismatches are worth 1.3 million trip fares for taxi compared to total profit of 2.5 million. In other words, without coordination failure within market, drivers could make 50% more profits. The type I friction for Uber is negligible[26]. The type II friction as measured by 8.1 are shown in second panel. There are a total of 14,738 cross-location mismatches for taxis which are worth $203,530 trip fares. Uber has a fewer cross-location mismatches than taxi, 1,850.

[26]Though I assume perfect matching for Uber within market, the hyperbolic function still generates slightly more friction than perfect matching.

Table 7: Baseline welfare statistics

within-location mismatches	
Taxi	95,547
	$1,286,400
Uber	635
	$10,517
cross-location mismatches	
Taxi	14,738
	$203,530
Uber	1,850
	$34,450
Profits and welfare	
Taxi profit	$ 2,510,400
Uber profit	$ 779,380
Consumer welfare	505,210

The dynamics of frictions over time are provided in figure 8. It shows how the aggregate frictions and matches over locations for any given period evolve. Taxi pickups increase sharply after the first hour in the morning and decrease until 11 a.m. ($t = 30$). Uber's aggregate pickups are flatter than taxis which only increases slightly during the morning rush hours. Notably, the cross-location mismatches reach the highest daily level during the morning rush hours when taxi drivers are more likely to overcrowd some locations and leave others undersupplied. In the appendix, I show figures for demand, supply and pickups over time of day for several locations.

To summarize, I estimate a dynamic search and matching model of taxi & Uber drivers and passengers. The results show that: (1), there exists a feedback loop (indirect network effect) between demand and supply in a market and direct network effect within the same side; (2), when drivers make search decisions, they face a very heterogeneous search values among locations; (3), drivers are more likely to oversupply high profitability locations and to leave other locations undersupplied such that cross location mismatches exist; (4), the low matching probability due to oversupply counters the high conditional expected continuation value such that ex-ante search values V_{ft}^i are less variable across locations.

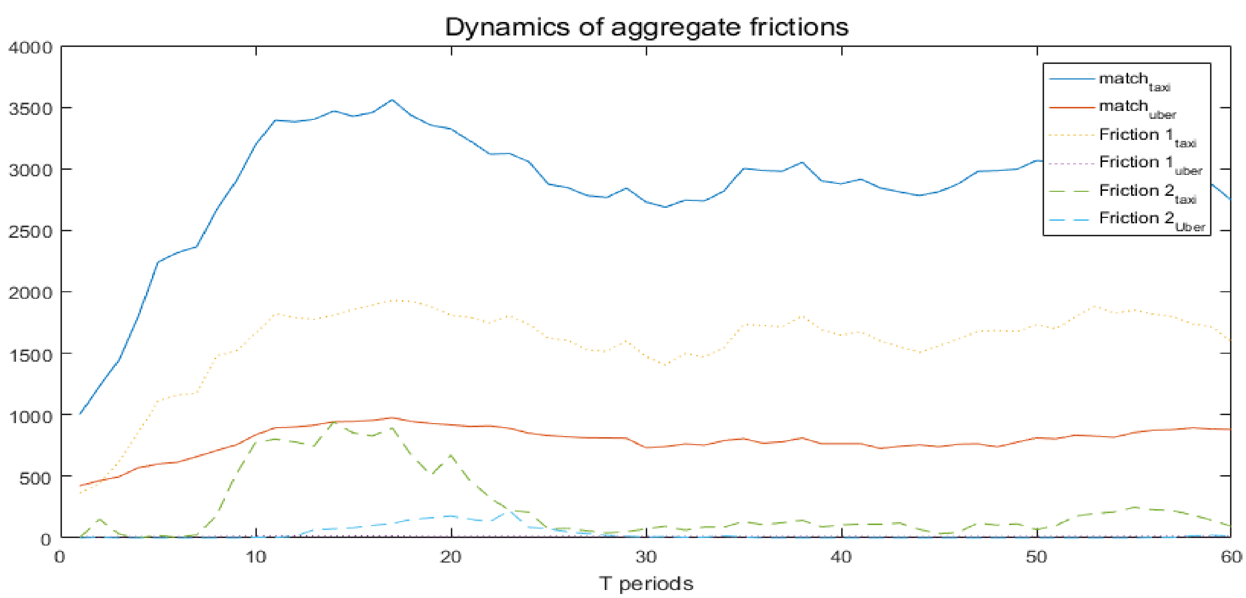


Figure 8: Aggregate frictions and matches of taxis& Uber over time

9 Counterfactuals

This paper characterize the "mis-allocation" of drivers through across-market friction. In order to understand the factors that affect the matching efficiency and social welfare, I simulate three counterfactuals. First, is a government proposal to cap the growth of Uber and its effects. The second, asks to what extent traffic conditions matter for matching efficiency. Third, analyzes how Uber surge pricing affects the efficiency of matching.

The simulation process that applies to all counterfactual scenarios is in the appendix. The difference between with or without network effect is whether or not to update mean utility $\boldsymbol{\delta_f}$ for the new equilibrium demand and supply. Not updating mean utility means that passengers will not respond to the change of demand and supply levels so that the feedback loop between two sides is shut down. This case is similar Buchholz (2022) which does not incorporate network effects. With network effects, the mean utilities adjust to any change of supply and demand. Furthermore, the supply will update to the change of demand and so forth until a new equilibrium is reached. As most structural models do with counterfactuals, I assume the demand shocks $\hat{\boldsymbol{\xi}}_f$ are fixed.

9.1 Regulating Uber's supply

In this counterfactual, I study the proposed regulatory policy of the NYC government regulating Uber. During the sample period of April 2016, Uber is growing rapidly and the total Uber licensed drivers outnumber the taxis[27]. There are two complaints about the growth of Uber. First, people argue that it contributes to traffic congestion. Second, it causes the taxis' profits to drop[28]. In 2015, the city mayor proposed to solve these problems by capping the growth of Uber, but it was dropped and not implemented after a protracted battle with the company. In early 2018, city mayor thought of regulating Uber again and a law was passed that the TLC will not issue new licenses for FHVs for one year[29]. To simulate the impact of this policy, I decrease

[27] There are $26,000$ Uber's licensed vehicles in comparison to $13,000$ taxi medallions. In the meantime, the number of Uber's affiliated vehicles is growing at a monthly rate of 3%.

[28] For example, the auction price of an independent unrestricted medallion dropped from \$ 0.7 million in 2011 to \$0.5 million in 2016.

[29] https://www.timeout.com/newyork/news/the-city-council-finally-remembered-that-uber-needs-to-be-regulated-in-nyc-030218

the assumed total number of Uber vehicles by 30% and simulate the new equilibria with and without network effect. The results are in table 8.

First, compare the third (without network effect) and the second (benchmark) column. After dropping 30% of Uber vehicles, the total supply of Uber cars decrease by 17,377 (21%). The total number of taxis' supply does not change . The reason is that demand for taxis in this case does not change due to fixed mean utility and price. Given the unchanged distribution of demand for taxis, the equilibrium supply of taxis does not change as well. Similarly, demand for Uber is also unchanged. But due to the decline of Uber's supply, total pickups of Uber decrease by 3,387 (7.1%). As for the frictions, I compare Uber's type II friction. Uber's cross-location mismatches increases by 3,020 which is worth \$ 56,702 fares. The increase is because of unchanged Uber demand and the decreased supply of Uber. However, demand of passengers should respond to change of supply as it affects the matching probability or waiting time. The importance of accounting for network effect is reflected in this example. Finally, taxis are not affected in this case without network effect. The total profit of Uber decrease by \$ 63,250 (8.11%).

Next, compare the last column that allows network effect with the previous two columns. Total supply of taxis does not change much but its demand increases by 3,790 (1.34%). However, the decline of Uber's demand is 6,353 (12.64%) which is greater than increased taxi's demand. The total pickups of taxi increase by 1,660 and pickups of Uber decrease by 7,004 more than without network effect. The difference is because the utility of choosing Uber declines due to less supply of Uber. As for frictions, taxi's type I friction increases a bit due to increased demand for taxi. Moreover, its type II friction also increases by 857 (5.81%). Though the difference is small, it reflects the direction that taxis' matching friction may go if there is less competition from Uber. Uber's cross-location mismatches is smaller than without network effect as expected, but it is still greater than the benchmark level by 684 (36.97%). In terms of profits, taxis make \$ 34,300 more money in a day shift after regulating Uber compared to \$ 119,320 profit loss of Uber. Finally, passengers are worse off by \$ 98,829 after this regulation[30].

[30]Consumer welfare is evaluated by inclusive value of log utility. The welfare loss in dollars is computed by compensating variation.

Table 8: Restricting Uber's supply

supply, demand, match	*Benchmark*	*change w/o network*	*change with network*
Taxi supply ($\Sigma_{i,t}v^i_{yt}$)	418,100	20(0.01%)	-490(-0.12%)
Uber supply ($\Sigma_{i,t}v^i_{xt}$)	82,707	-17,377(-21.01%)	-18,085(-21.86%)
Taxi demand ($\Sigma_{i,t}u^i_{yt}$)	283,510	-480(-0.17%)	3,790(1.34%)
Uber demand ($\Sigma_{i,t}u^i_{xt}$)	50,224	-42(-0.08%)	-6,353(-12.64%)
Taxi pickups ($\Sigma_{i,t}m^i_{yt}$)	173,230	-160(-0.09%)	1,660(0.96%)
Uber pickups ($\Sigma_{i,t}m^i_{xt}$)	47,738	-3,387(-7.09%)	-7,004(-14.67%)
two type friction			
Taxi type I friction	95,547	49(0.05%)	1,265(1.32%)
	\$ 1,286,400	\$ 700(0.05%)	\$ 18,400(1.43%)
Uber type I friction	635	65(10.23%)	-33(-5.19%)
	\$ 10,517	\$ 1,153(10.96%)	\$ -523(-4.97%)
Taxi type II friction	14,738	-371(-2.52%)	857(5.81%)
	\$ 203,530	\$ -4,910(-2.41%)	\$ 12,220(6%)
Uber type II friction	1,850	3,020(163.24%)	684(36.97%)
	\$ 34,450	\$56,702(164.59%)	\$ 12,314(35.74%)
welfare			
Taxi profit	\$2,510,400	\$-2,600(-0.10%)	\$34,300(1.36%)
Uber profit	\$779,350	\$-63,250(-8.11%)	\$-119,320(-15.31%)
Consumer welfare	505,210	0	-3,670(-0.72%)
Δ Consumer welfare	NA	\$0	\$ -98,829
Δ Social welfare	NA	\$-65,850	\$-183,849

Note: Type I friction is $\Sigma_{i,t}\min\{u^i_{ft}, v^i_{ft}\} - \bar{m}^i_{ft}$.
Type II friction is $\Sigma_t \min\{\Sigma_i \max\{u^i_{ft} - v^i_{ft}, 0\}, \Sigma_i \max\{v^i_{ft} - u^i_{ft}, 0\}\}$.

9.2 Improving traffic conditions

The second counterfactual simulates equilibrium after improving traffic conditions. As government blames Uber for traffic congestion, fluid traffic movement is welfare improving for all the citizens. Instead of regulating the number of cars on the street, government can also improve the infrastructure as the McKinsey report on assessing Uber impact suggests. Since I do not model the relationship between traffic conditions and Uber cars, I only simulate the traffic improvement as an exogenous shock. To do so, I replace the trip time $\{\chi_t^{ij}\}_{\forall i,j,t}$ calculated using 2016 data by the corresponding time in 2010. Figure 9 displays the comparison of trip time in 2010 and 2016. The difference in distributions implies worse traffic condition in 2016. For example, the median of trip time in 2010 is 18.73 minutes. The median of trip time in 2016 is 22.62 minutes. The new equilibrium suggests the importance of traffic conditions on matching efficiency. The new equilibrium and welfare in comparison to the benchmark is provided in table 9.

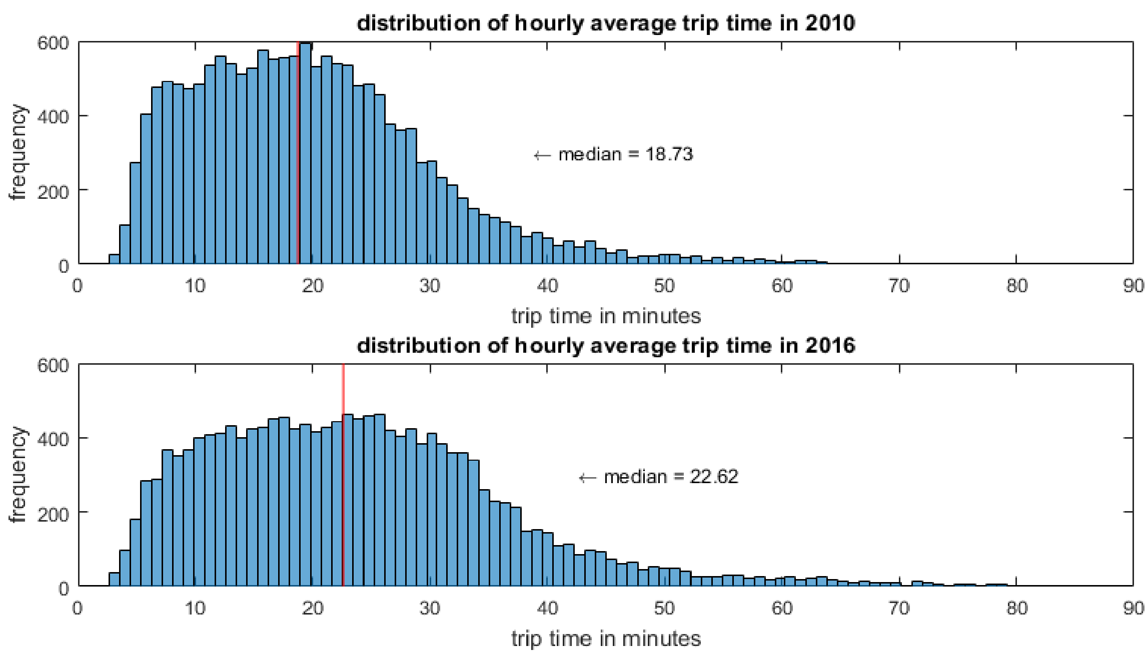


Figure 9: Comparison of traffic speed in 2010 and 2016

First, compare the efficiency change without network effects. Aggregate demands for taxi and Uber do not change in the new equilibrium. Daily aggregate supply of both taxi and Uber increase. For example, total supply/searches of taxi drivers increase by

68,610 (16.41%) and Uber's increase by 10,963 (13.25%). As consequence, the total pickups of taxis increase by 8,110 (4.68%) and of Uber increase by 971 (2.03%). The two types of friction also change. Both frictions for taxi decrease, especially for the cross-location mismatches. The total number of type II friction of taxis decreases by 6,142 (41.67%) which is worth \$ 80,620 trip fares. As for Uber, the within-location friction is negligible due to perfect matching assumption. Uber's cross-location mismatches also decrease by 1,014 and the loss of fares decrease by \$ 18,618. The total revenue of taxis increases by \$ 107,400 (4.28%) and of Uber increases by \$ 17,830 (2.29%). The welfare gain to consumers measured by compensating variation (CV) is zero because inclusive value of their expected utility prior to matching does not change without network effect.

The last column of table 9 shows the equilibrium with network effect which allows demand to respond to the change of supply. The total supply of taxi increases more than without the network effect. Uber's supply also increases compared to the benchmark, however, less than without network effect. There could be two possible explanations. One reason is that the choice decisions of passengers change after responding to the network effect. As consequence, not only the market share/demand changes, but also the destinations of Uber's passengers change such that Uber's passengers tend to travel longer distance. The other reason is that the search values of Uber change due to the demand change and Uber drivers are more likely to search a location far away. Both reasons make Uber drivers spend more time on travelling than searching. Taxi's demand increases more than without network effect but Uber's demand declines. One reason is that taxis have stronger network effect than Uber and increased mean utility of choosing taxis is higher than Uber. As a result, taxi's pickups increase by 18,470 (10.66%) which is larger than the 4.68% without network effect. It is interesting that Uber's pickups decreases in the new equilibrium rather than increase in the previous case which implies that we could even have opposite conclusions with or without network effect. As for frictions, the type I friction of taxis increases both in quantity and dollars. It is due to both increased demand and supply of taxi, and the random matching assumption. The type II friction of taxis also decreases compared to the benchmark but is slightly larger than the case without the network effect. The same finding applies to Uber's type II friction. This finding implies that network effects make the matching across locations less efficient which coincides with the calibration in section 3..

Table 9: Traffic improvement

supply, demand, match	*Benchmark*	*change w/o network*	*change with network*
Taxi supply ($\Sigma_{i,t} v_{yt}^i$)	418,100	68,610(16.41%)	80,820(19.33%)
Uber supply ($\Sigma_{i,t} v_{xt}^i$)	82,707	10,963(13.25%)	9,150(11.06%)
Taxi demand ($\Sigma_{i,t} u_{yt}^i$)	283,510	-480(-0.17%)	16,610(5.86%)
Uber demand ($\Sigma_{i,t} u_{xt}^i$)	50,224	-42(-0.08%)	-1,833(-3.65%)
Taxi pickups ($\Sigma_{i,t} m_{yt}^i$)	173,230	8,110(4.68%)	18,470(10.66%)
Uber pickups ($\Sigma_{i,t} m_{xt}^i$)	47,738	971(2.03%)	-1,057(-2.21%)
two type friction			
Taxi type I friction	95,547	-2,457(-2.57%)	3,376(3.53%)
	\$ 1,286,400	\$ -33,500(-2.60%)	\$ 41,800(3.25%)
Uber type I friction	635	0	-24(3.78%)
	\$ 10,517	\$ 14(0.13%)	\$-111(-1.05%)
Taxi type II friction	14,738	-6,142(-41.67%)	-5,237(-35.53%)
	\$ 203,530	\$-80,620(-39.61%)	\$ -68,610(-33.71%)
Uber type II friction	1,850	-1,014(-54.81%)	-753(-40.70%)
	\$ 34,450	\$-18,618(-54.04%)	\$ -12,643(-36.69%)
welfare			
Taxi profit	\$2,510,400	\$107,400(4.28%)	\$231,100(9.21%)
Uber profit	\$779,350	\$17,830(2.29%)	\$5,110(0.65%)
Consumer welfare	505,210	0	30,860(6.11%)
ΔConsumer welfare	NA	\$0	\$ 748,900
ΔSocial welfare	NA	\$ 125,230	\$ 985,110

9.3 Eliminating surge multiplier

The final counterfactual studies whether Uber's surge pricing improves matching efficiency across locations. Unlike fixed fare of taxis, Uber uses surge multiplier to efficiently adjust drivers' search incentives among locations. When some locations have higher demand than supply, Uber tends to charge a higher price than regular one to motivate more drivers to come. This higher price is the product of a surge multiplier and regular price. To investigate the effect of flexible pricing of Uber on matching efficiency, I eliminate Uber's surge multiplier such that all Uber's trips are calculated using the normal pricing structure. By comparing the new equilibrium with the benchmark, it helps to understand the effect of flexible pricing on matching efficiency. The results are listed in table 10.

Column 2 and 3 compare the new equilibrium without network effects to the baseline. After the decline of Uber's prices, the aggregate searches of drivers do not change much for both taxi and Uber. Demand is more sensitive to the lower price such that Uber's demand increases by 8,150 (16.23%). Taxis' total demand decreases by 5,510 (1.94%) due to the price competition. As a result of demand change, taxis' pickups also decline slightly by 1.39% whereas Uber's pickups increase by 9.11%. Comparing the frictions in the second panel, the type I friction of taxi decreases due to the decreased demand and supply. Most interesting findings are the cross-location mismatches. Taxi's cross-location mismatches decrease by 1,894 (12.85%) whereas Uber's mismatches increase by 2,811 (152%) trips. The increased type II friction of Uber is worth \$ 30,455 fares. The decreased cross-location mismatches of taxis may result from the competition effect, since Uber's product becomes more competitive with the lower price. Without the help of surge pricing, Uber's mis-allocation of drivers makes its matching less efficient. Then, taxi drivers search less efficiently due to less pressure from Uber's competition. As for the last panel, taxis' profits decrease due to price competition by \$ 30,300. Though Uber's demand increases due to lower price, its total profit decreases by \$ 57,990. The welfare gain of passengers is \$ 120,400.

The last column reports equilibrium with network effects such that demand and supply change the mean utilities of passengers. I find that both supply and demand of taxi and Uber change more than without the network effect in the same direction. For example, the total supply of taxis decreases by 3,240 which is more than previous without network effects. The demand of Uber increases by 23.58% compared to the 16.23% without network effects. This finding reflects the positive feedback loop in the two sided market. After Uber's price decreases, demand for Uber increases which further increases utility of choosing Uber through direct network effect. As a result, the total pickups of taxis decrease more and Uber's pickups increase more than without network effect. To conclude, existence of network effects exaggerates the effect of price drop on market share in this counterfactual. As for frictions, taxis' within-location friction decreases by 2.57%. This is mainly due to the decreased demand and supply of taxis. Cross-location mismatches of taxis and Uber have opposite results as well. Taxi's type II friction decreases by 2,573(17.46%) and Uber's increases by 3,152. Finally, in the last panel, taxis' profits decreases slightly more than the case without network effect. Uber drivers make more money after allowing network effect because it make more demand for Uber to compensate the price drop. However, consumer welfare gain decreases from \$ 120,400 to \$ 96,977 due to network effects.

Table 10: Eliminating surge multiplier

supply, demand, match	*Benchmark*	*w/o network*	*with network*
Taxi supply ($\Sigma_{i,t} v_{yt}^i$)	418,100	-870(-0.21%)	-3,240(-0.77%)
Uber supply ($\Sigma_{i,t} v_{xt}^i$)	82,707	668(0.81%)	1,512(1.83%)
Taxi demand ($\Sigma_{i,t} u_{yt}^i$)	283,510	-5,510(-1.94%)	-9,780(-3.45%)
Uber demand ($\Sigma_{i,t} u_{xt}^i$)	50,224	8,150(16.23%)	11,841(23.58%)
Taxi pickups ($\Sigma_{i,t} m_{yt}^i$)	173,230	-2,420(-1.39%)	-4,760(-2.74%)
Uber pickups ($\Sigma_{i,t} m_{xt}^i$)	47,738	4,350(9.11%)	6,633(13.89%)
two type friction			
Taxi type I friction	95,547	-1,204(-1.26%)	-2,455(-2.57%)
	$ 1,286,400	$ -14,100(-1.09%)	$ -29,600(-2.30%)
Uber type I friction	635	123(19.37%)	151(23.78%)
	$ 10,517	$ -1(-0.01%)	$ 287(2.73%)
Taxi type II friction	14,738	-1,894(-12.85%)	-2,573(-17.46%)
	$ 203,530	$-25,410(-12.48%)	$ -34,070(-16.74%)
Uber type II friction	1,850	2,811(151.95%)	3,152(170.38%)
	$ 34,450	$ 30,455(88.40%)	$ 32,153(93.33%)
welfare			
Taxi profit	$2,510,400	$-30,300(-1.21%)	$-57,500(-2.29%)
Uber profit	$779,350	$-57,990(-7.44%)	$-35,020(-4.49%)
Consumer welfare	505,210	5,930(1.17%)	4,810(0.95%)
ΔConsumer welfare	NA	$120,400	$ 96,977
ΔSocial welfare	NA	$32,110	$4,457

10 Conclusion

This paper studies the factors that determine supply and demand, and their effects on matching friction with an application to taxi and Uber drivers searching for passengers in New York City. In this industry, due to the fixed pricing structure of taxi, the market is not cleared in prices, leaving mismatches across locations. There are also non-price factors that could affect matching efficiency. The mostly important two are network effects and competition. To analyze their impacts, I model Uber and taxi drivers' dynamic search decisions among 40 defined locations in NYC and passengers' discrete choice decisions. I estimate the model using data of a representative weekday shift in April 2016.

Estimates of the model show the existence of network effects and a positive feedback loop between drivers and passengers. There is also a positive externality among passengers. How network effects affect spatial mismatches of drivers and passengers are studied in three counterfactuals. The first is a regulatory policy capping Uber's number of vehicles. I reduces the number of Uber cars by 30% and find that taxis' pickups increase by 1,660, far less than the decrease of Uber's pickups 7,004. Interestingly, due to less competition, taxis' mismatches increases. Uber's matching efficiency decreases due to the smaller number of cars. The second studies to what extent traffic conditions matter for matching efficiency by using the travel time in 2010. In the new equilibrium with better traffic, both Uber and taxis have fewer mismatches, 35.53% and 40,7% reduction, respectively. Their profits also increase. However, without considering network effect, Uber's pickups are predicted to increase whereas it decreases with network effects. Finally, I study whether Uber's surge pricing improves matching efficiency. By eliminating the surge pricing, there is a significant increase in Uber's mismatches. However, lower price without surge pricing results in more pickups for Uber, 6,633 trips. All counterfactuals show different results either in magnitude or sign depending on whether the network effects are recognized in the analysis. Sometimes the magnitude of change won't be large but sometimes they are. Moreover, the sign of change is flipped with network effects. Thus, ignoring network effects will lead to incorrect conclusions or inferences.

References

[1] Ackerberg, D.A. and G. Gowrisankaran (2006): "Quantifying equilibrium network externalities in the ACH banking industry," *RAND journal of economics*, 37(3), 738-761.

[2] Amstrong, M. (2006): "Competition in two-sided markets," *RAND Journal of Economics*, 37(3), 668-691.

[3] Berry, S. (1994): "Estimating discrete-choice models of product differentiation", *Rand Journal of Economics*, 25(2), 242-262.

[4] Blanchard, O. and P. Dianmond (1989): "The beveridge curve," *Brookings Papers on Economic Activity*, 1989, 1-76.

[5] Buchholz, N. (2022):"Spatial equilibrium, search frictions and dynamic efficiency in the taxi industry," *Revew of Economic Studies*, 556-591.

[6] Burdett, K., S. Shi, and R. Wright (2001): "Pricing and matching with frictions," *Review of Economic Studies*, 465-491.

[7] Butters, G. R. (1977): "Equilibrium distributions of sales and advertising prices," *Review of Economic Studies*, 465-491.

[8] Castillo, J.C. (2022): "Who benefits from surge pricing?", working paper.

[9] Castillo, J.C., D.Knoepfle, and G. Weyl (2022): "Matching in ride hailing: wild goose chases and how to solve them.", working paper.

[10] Chen, M. K. and M. Sheldon (2015): "Dynamic pricing in a labor market: surge pricing and flexible work on the Uber platform" working paper.

[11] Chen, M. K., J. A., Chevalier, P. E. Rossi and E. Oehlsen (2017): "The value of flexible work: evidence from Uber drivers," *Journal of Political Economy*, 127(6), 2735-2794.

[12] Chou, C. and O. Shy (1990): "Network effects without network externalities," *International Journal of Industrial Organization*,8, 259-270.

[13] Chu, J. and P .Manchanda (2016): " Quantifying cross and direct network effects in online consumer-to-consumer platforms," *Marketing Science*, 35(6), 870-893.

[14] Church, J, and N. Gandal (1992): “Network effects, software provision, and standardization,” *Journal of Industrial Economics*, 40, 85-103.

[15] Clements, M. T. and H. Ohashi(2005): “Indirect network effects and the product cycle: U.S. video games, 1994-2002.” *Journal of Industrial Economics*, 53(4), 515-542.

[16] Corts, K. S. and M. Lederman (2009): “Software exclusivity and the scope of indirect network effects in the US home video game market.” *International Journal of Industrial Organization*, 27(2), 121-136.

[17] Crawford, V.P. and J. Meng (2011): “New York City cab drivers’ labor supply revisisted: reference-dependent preferences with rational expectations targets for hours and income” *American Economic Review*, 101, 1912-1932.

[18] Dubé, J.-P. H., G. J. Hitsch, and P.K. Chintagunta (2010): “Tipping and concentration in markets with indirect network effects”, *Marketing Science*, 29(2), 216-249.

[19] Farber, H. S.(2005): “Is tomorrow another day? The labor supply of New York City cab drivers”,*Journal of Political Economy*, 113, 46-82.

[20] Farber, H. S. (2008): “Reference dependent preferences and labor supply: the case of New York City taxi drivers,” *American Economic Review*, Juen 2008, 1069-1082.

[21] Farrell, J. and G. Saloner (1986): “Installed base and compatibility: innovation, product preannouncements, and predation,” *American Economic Review*, 940-955.

[22] Frechette, G., A. Lizzeri and T. Salz (2019): “Frictions in a comeptitive, regulated market: evidence from taxis,” *American Economic Review*, 109(8),2954-2992.

[23] Gandal, N. (1994): “Hedonic price indexes for spreadsheets and an empirical test of network externalities,” *RAND Journal of Economics*, 25, 161-170.

[24] Gandal, N., Kende, M. and R. Rob (2000): “The dynamics of technological adoption in hardware/software systems: the case of compact disk players,” *RAND Journal of Economics*, 31, 43-62.

[25] Goolsbee, A. and P. Klenow (2002): " Diffusion of home computers", *Journal of Law and Economics*, 45, 317-343.

[26] Katz, M. L. and C. Shapiro (1985): "Network externalities, competition, and compatibility," *American Economic Review*, 424-440.

[27] Kreindler, G. (2018): "The welfare effect of road congestion pricing: experimental evidence and equilibrium implications," working paper.

[28] Lagos, R.(2000): "An Alternative approach to search frictions," *The Journal of Political Economy*, 108,851-873.

[29] Leccese, M. (2022): "Asymmetric taxation, pass-through and market competition: evidence from ride-sharing and taxis," working paper.

[30] Lee, R. S.(2013): " Vertical integration and exclusivity in platform and two-sided markets," *American Economic Review*, 103(7), 2960-3000.

[31] Liu, Y. and R. Luo (2022): "Network Effect and Multi-Network Sellers' Dynamic Pricing in the U.S. Smartphone Market," *Management Science*, 68(9),1-22.

[32] Nair, H., P. Chintagunta, and J.-P. Dubé (2004): "Empirical analysis of indirect network effects in the market for personal digital assistants," *Quantitative Marketing and Economics*, 2(1),23-58.

[33] Mortensen, D. T. and C.A. Pissarides (1999): "New developments in models of search in the labor market," *Handbook of Labor Economics*, 3, 2567-2627.

[34] Nevo, A. (2000): "A practitioner's guide to estimation of random-coefficients logit models of demand," *Journal of Economics and Management Strategy*, 9, 513-558.

[35] Ohashi, H. (2003): "The role of network effects in the US VCR market, 1978-1986," *Journal of Economics and Management Strategy*, 12(4),447-494.

[36] Park, S. (2004): "Quantitative analysis of network externalities in competing technologies: the VCR case," *Review of Economics and Statistics*, 86(4),937-945.

[37] Pissarides, C. A. (1984): "Search intensity, job advertising, and efficiency," *Journal of Labor Economics*, 128-243.

[38] Qi, S. (2013):"The impact of advertising regulation on industry: the cigarette ddvertising ban of 1971," *RAND Journal of Economics*, 44, 215-248.

[39] Rochet, J.-C. and J. Tirole (2003): "Platform competition in two-sided markets," *Journal of the European Economic Association*, 1(4).

[40] Rysman, M. (2004): "Competition between networks: a study of the market for yellow pages," *Review of Economic Studies*, 71, 483-512.

[41] Saeedi, M. (2014): "Reputation and adverse selection: theory and evidence from eBay," working paper.

[42] Saloner, G. and A. Shepard (1995): "Adoption of technologies with network effects: an empirical examination of the adoption of automated teller machines," *RAND Journal of Economics*, 26, 479-501.

[43] Shapiro, M. H.(2018): "Density of demand and the benefit of Uber," working paper.

[44] Weintraub, G., C. Benkard, P. Jeziorski and B. Van Roy (2008): "Markov perfect industry dynamics with many firms," *Econometrica*, 76, 1375-1411.

[45] Weintraub, G., C. Benkard, P. Jeziorski and B. Van Roy (2008): "Nonstationary oblivious equilibrium," Manuscript.

[46] Xu, Y. (2008): "A structural empirical model of R&D, firm heterogeneity, and industry evolution," *Journal of Industrial Economics*, forthcoming.

Appendix

The inner and outer loops of estimating equilibrium demand and supply, which corresponds to the small triangle in the upper right of figure 6, are described in the algorithm 1 and 2 below.

Algorithm 1 Solve Equilibrium Supply

1: Set parameter values for $\{\sigma, \alpha, \gamma, \rho, \beta\}$ and $\{\delta^i_{ft}\}_{\forall f,i,t}$
2: Guess supply $\{v^i_{fT}\}_{f,i}$, calculate $\{u^i_{fT}\}_{f,i}$, $\{V^i_{fT}\}_{f,i}$
3: for $\tau = T - 1$ to 1
4: Guess $\{v^i_{f\tau}\}_{f,i}$, reset state of in-transit cars $\{\tilde{v}^i_{ft,k}\}_{t>\tau} = 0$
5: for $t = \tau$ to $t = T$
6: Compute market share $\{s^i_{ft}\}$ and demand $\{u^i_{ft}\}$
7: Compute matches $m^i_{ft} = m(u^i_{ft}, v^i_{ft})$
8: Compute transition of employed cars $\{\tilde{a}^i_{ft}\}$
9: Compute transition policy of unemployed cars $\{\pi^i_{ft}\}$
10: Update $\{\tilde{v}^i_{ft+1,k}\}$ based on $\{\pi^i_{ft}\}, \{\tilde{a}^i_{ft}\}$
11: Update value $\{V^i_{ft}\}$ of current period
12: Update supply of next period $v^i_{ft+1} = \tilde{v}^i_{ft+1,k=1}$
13: end
14:end
15:Fix $\tau = 1$
16: Iterate step 5 to 13
17: Stop until step 11 and 12 won't update under certain tolerance level.

Algorithm 2 Solve Fixed Points of Mean Utility

1: Guess demand $\{u^i_{ft}\}_0$ based on observed pickups $\{\bar{m}^i_{ft}\}$
2: Calculate market share s^i_{ft} and guess initial $\{\delta^i_{ft}\}_0$
3: Iterate BLP contraction mapping to solve for $\{\delta^i_{ft}\}_1$ to match market shares
4: Plug $\{\delta^i_{ft}\}_1$ into algorithm 1 to solve for $\{v^i_{ft}\}$
5: Invert matching function with $\{v^i_{ft}, \bar{m}^i_{ft}\}$ for $\{u^i_{ft}\}$
6: a: $v^i_{ft} > \bar{m}^i_{ft}$, update u^i_{ft}
7: b: $v^i_{ft} \leq \bar{m}^i_{ft}$, don't update u^i_{ft}
8: Given updated $\{u^i_{ft}\}_1$, solve BLP contraction mapping for $\{\delta^i_{ft}\}_2$ and $\{\tilde{A}^i_{yt}\}$
9: a: $\Sigma_f u^i_{ft} < \lambda^i_t$, update δ^i_{ft}
10: b: $\Sigma_f u^i_{ft} \geq \lambda^i_t$, set $u^i_{xt} = u^i_{ft}\bar{m}^i_{xt}/\Sigma_f u^i_{ft}$ and $u^i_{yt} = \lambda^i_t - u^i_{xt}$, update δ^i_{ft}
11: Repeat step 4 to 10
12: Until $|\boldsymbol{\delta}^{k+1} - \boldsymbol{\delta}^k| < \epsilon$
13: Report $\{\tilde{A}^i_{yt}\}$, transition of employed taxis

The steps of solving new equilibria with and without network effects in the counterfactuals are described in algorithms below.

Simulation Algorithm Without Network Effects

1: Fix parameter values as estimates $\{\hat{\sigma}, \hat{\alpha}, \hat{\gamma}, \hat{\rho}, \hat{\beta}, \hat{\theta}\}$ and $\hat{\boldsymbol{\delta}}$
2: Given $\{\hat{\boldsymbol{\delta}}, \hat{\alpha}, \hat{\beta}\}$, calculate new eq demand $\boldsymbol{u'}$ and transition of passengers $\tilde{\boldsymbol{A}}_{\boldsymbol{f}}$
3: Run the iteration process in Algorithm 1 to solve for new eq supplies $\boldsymbol{v'}$

Simulation Algorithm With Network Effects

1: Fix parameter values as estimates $\{\hat{\sigma}, \hat{\alpha}, \hat{\gamma}, \hat{\rho}, \hat{\beta}, \hat{\theta}\}$
2: Set initial guess of $\boldsymbol{\delta^0} = \hat{\boldsymbol{\delta}}$
3: Iterate from $k = 0$
3: Given $\{\boldsymbol{\delta^k}, \hat{\alpha}, \hat{\beta}\}$, calculate new eq demand $\boldsymbol{u^k}$ and transition of passengers $\tilde{\boldsymbol{A}}_{\boldsymbol{f}}^{\boldsymbol{k}}$
4: Run the iteration process in Algorithm 1 to solve for $\boldsymbol{v_f^k}$
5: Plug $\{\boldsymbol{v_f^k}, \boldsymbol{u_f^k}, \hat{\boldsymbol{\xi}}, \hat{\theta}\}$ in to mean utility 6.2 and update $\boldsymbol{\delta^{k+1}}$
6: Stop until $\|\boldsymbol{\delta^{k+1}} - \boldsymbol{\delta^k}\| < \epsilon$
7: New equilibrium $\boldsymbol{v_f^*}, \boldsymbol{u_f^*}$

Figures below show demand, supply and pickups of taxi and Uber over time for several locations in figure 10- 13. One interesting finding by comparing Queens or Brooklyn with central Manhattan areas is that there are excess taxis' supply in Manhattan almost all the time during the daytime. In comparison, Queens and Brooklyn are more likely to have excess demand. Especially during the morning hours from 7 am to 9 am. It is harder for passengers to hail a taxi. Another finding is that Uber serves outer boroughs more intensively than taxis. For example, the number of Uber drivers in selected location of Brooklyn is close to that in Times Square.

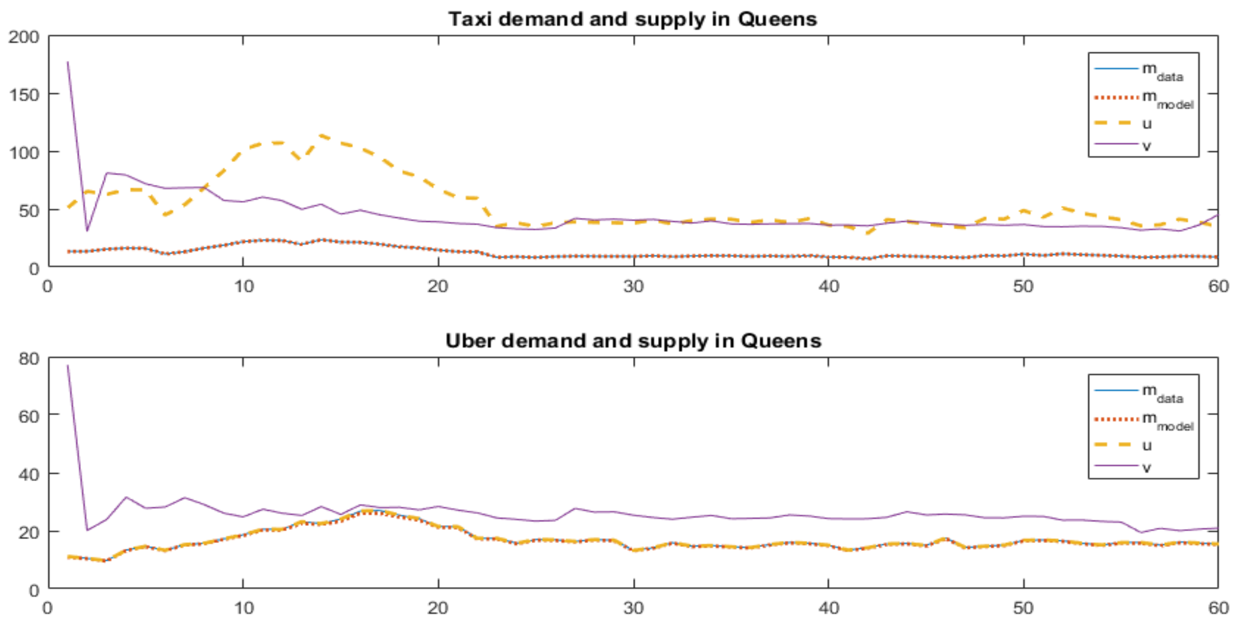


Figure 10: Demand, supply and matches in a location of Queens

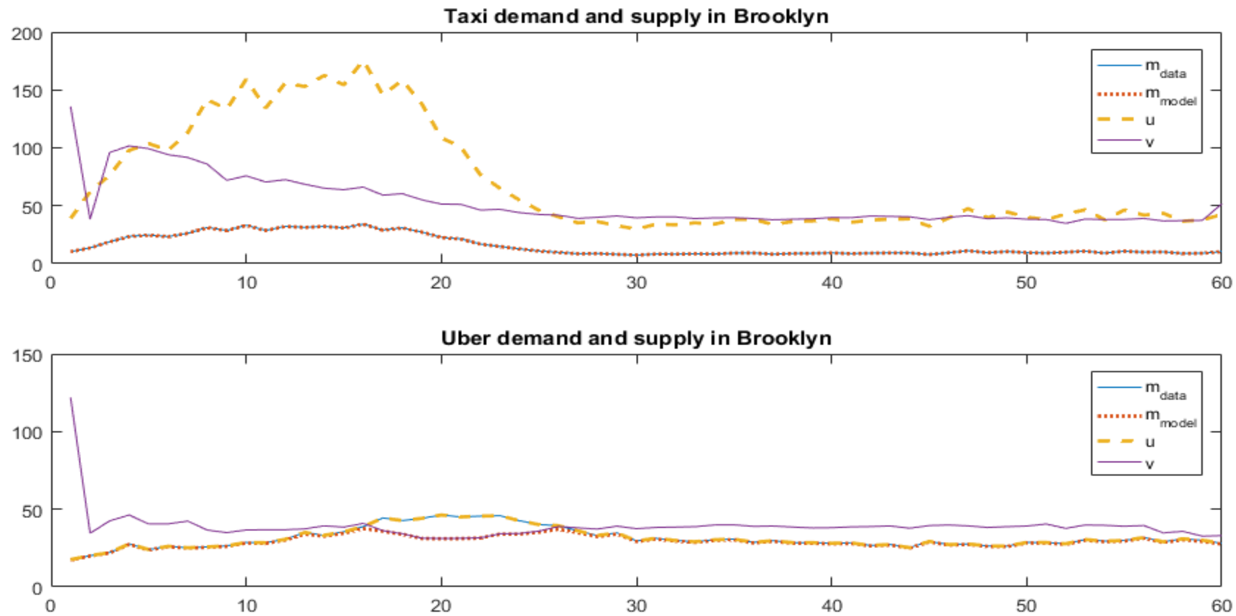


Figure 11: Demand, supply and matches in a location of Brooklyn

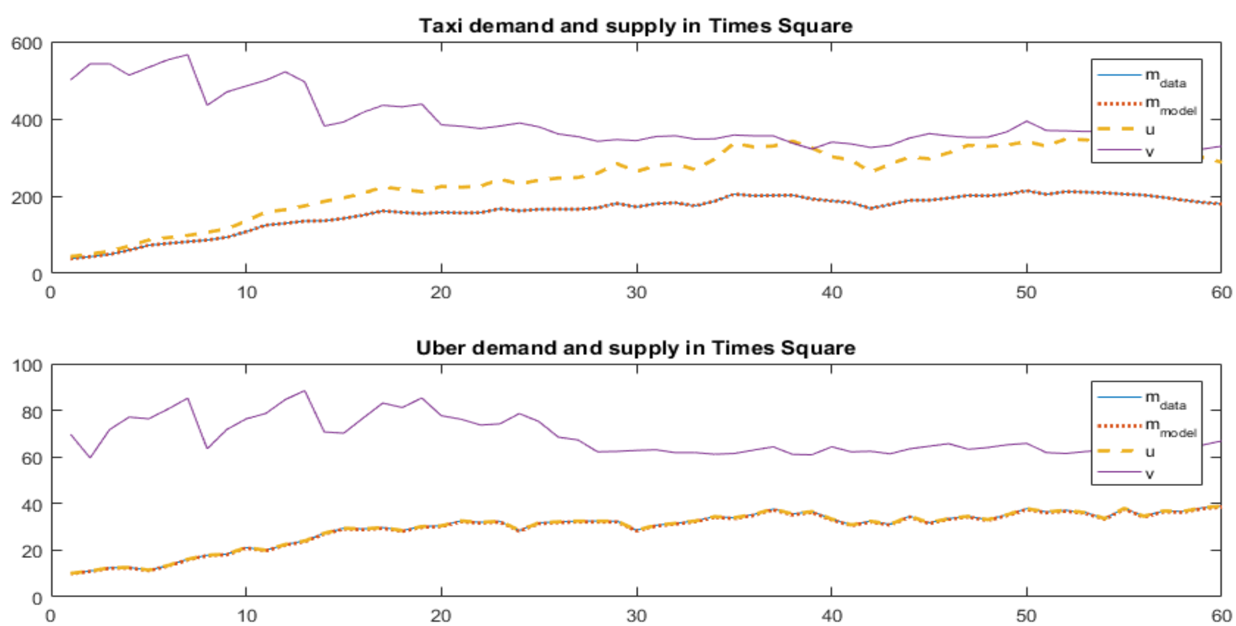


Figure 12: Demand, supply and matches in Times Square

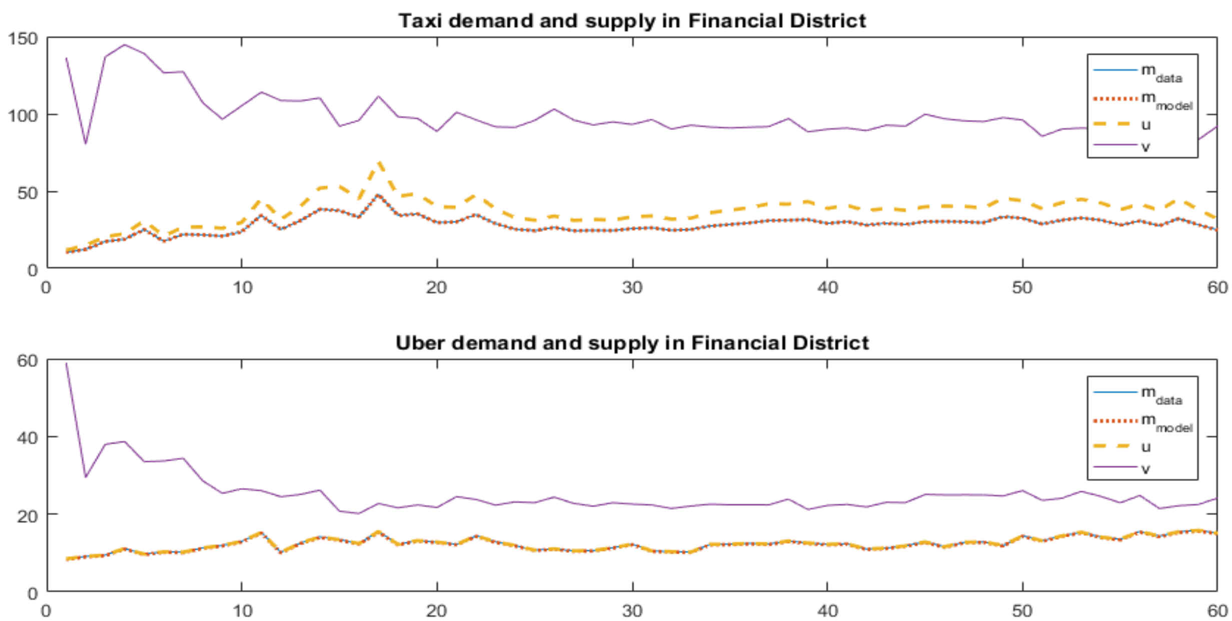


Figure 13: Demand, supply and matches in Financial District

Equilibrium Solution of the Exposition Model

3.2 Solution to Network Effects Calibration

Since we only consider equilibrium with excess supply in island 1 and excess demand in island 2, we have $m^*_{y1} = u^*_{y1}$ and $m^*_{y2} = u^*_{y2}$. By plugging these equations in to (E1), we have $v^*_{y1} = \frac{p_{y1}}{p_{y2}} u^*_{y1}$.

Substitute $v^*_{y1} = \frac{p_{y1}}{p_{y2}} u^*_{y1}$ in to the first equation in (3.1), we obtain the expression of solution u^*_{y1} as function of exgenous variables, which is $u^*_{y1} = \frac{d_{y1}}{1 - \alpha - \beta \frac{p_{y1}}{p_{y2}}}$.

Since the total number of drivers is N_y, it comes immediately that $v^*_{y2} = N_y - v^*_{y1}$.

Substitute $v^*_{y2} = N_y - v^*_{y1}$ into the second equation in (3.1), we can solve $u^*_{y2} = \frac{\beta}{1-\alpha} v^*_{y2} + \frac{d^*_{y2}}{1-\alpha}$. One can simply solve all the endogenous variables as function of exogenous variables by replacing $u^*_{y1} = \frac{d_{y1}}{1 - \alpha - \beta \frac{p_{y1}}{p_{y2}}}$.

3.3 Solution to Competition Calibration

We consider the equilibrium with excess supply of Uber in both islands. Therefore, we have $m^*_{x1} = u^*_{x1}$ and $m^*_{x2} = u^*_{x2}$. By substituting these equations in to (E1), we obtain $\frac{u^*_{x1}}{v^*_{x1}} p_{x1} = \frac{u^*_{x2}}{v^*_{x2}} p_{x2}$. Since we assume $p_{x1} = p_{x2}$, it turns out $\frac{u^*_{x1}}{v^*_{x1}} = \frac{u^*_{x2}}{v^*_{x2}} \equiv w^*$. Replace u^*_{xi} in (3.1) for Uber, we have:

$$((1-\alpha)w^* - \beta)v^*_{xi} = \theta v^*_{yi} + d_{xi}, \forall i = 1, 2 \qquad (1)$$

Summing equation (1) over i, we have:

$$((1-\alpha)w^* - \beta)N_x = \theta N_y + d_{x1} + d_{x2}, \forall i = 1, 2 \qquad (2)$$

Next, we solve equilibrium demand and supply for yellow taxis. Given the equilibrium with excess supply of taxi in island 1 and excess demand in island 2, we still have $m^*_{y1} = u^*_{y1}, m^*_{y2} = v^*_{y2}$ and $v^*_{y1} = \frac{p_{y1}}{p_{y2}} u^*_{y1}$. Replace v^*_{y1} in equation (3.1), we obtain:

$$(1 - \alpha - \beta \frac{p_{y1}}{p_{y2}}) u^*_{y1} = \theta v^*_{x1} + d_{y1} \rightarrow u^*_{y1} = \frac{\theta v^*_{x1} + d_{y1}}{1 - \alpha - \beta \frac{p_{y1}}{p_{y2}}}$$

which is equivalent to (3.4).

Equation (3.5) can be obtained simply from (3.1) when i=2. To solve endogenous variables as function of exogenous ones, one can use (3.4) and (3.7) to solve either v^*_{y1} or v^*_{x1} and plug the solution into (3.4)-(3.7).