Spatial Equilibrium, Search Frictions and Dynamic Efficiency in the Taxi Industry

Nicholas Buchholz *

May 8, 2019

Abstract

This paper analyzes the dynamic spatial equilibrium of taxicabs and shows how common taxi regulations lead to substantial inefficiencies. Taxis compete for passengers by driving to different locations around the city. Search costs ensure that optimal search behavior will still result in equilibrium frictions in the form of waiting times and spatial mismatch. Medallion limit regulations and fixed fare structures exacerbate these frictions by preventing markets from clearing on prices, leaving empty taxis in some areas, and leaving excess demand in other areas.

To analyze the role of regulation on frictions and efficiency, I pose a dynamic model of spatial search and matching between taxis and passengers. Using a comprehensive dataset of New York City yellow medallion taxis, I use this model to compute the equilibrium spatial distribution of vacant taxis and estimate intraday demand given price and medallion regulations. My estimates show that the weekday New York market achieves about $4.7 million in daily welfare split across 182 thousand trips, but an additional 45 thousand customers fail to find cabs due to search frictions. Counterfactual analysis shows that implementing pricing rules that vary by time, location or distance can enhance allocative efficiency and expand the market, offering daily net surplus gains of up to $232 thousand and 93 thousand additional daily taxi-passenger matches, substantially better than the gains due to a perfect static matching technology.

Key Words: dynamic games, dynamic pricing, spatial equilibrium, search frictions, regulation, taxi industry

JEL classification: C73; D83; L90; R12

*Department of Economics, Princeton University. Email: nbuchholz@princeton.edu. This is a revised version of my job market paper, previously circulated under the title Spatial Equilibrium, Search Frictions and Efficient Regulation in the Taxi Industry. Special thanks to Allan Collard-Wexler, Eugenio Miravete and Stephen Ryan for their guidance and support. I also benefitted from discussions with Hassan Afrouzi, John Asker, Austin Bean, Lanier Benkard, Laura Doval, Neal Ghosh, Andrew Glover, Jean-François Houde, Jakub Kastl, Ariel Pakes, Rob Shimer, Emily Weisburst, Daniel Xu, Haiqing Xu, anonymous referees and numerous seminar participants.
1 Introduction

It has been well documented that search frictions lead to less efficient outcomes.\(^1\) One particularly salient reason for the existence of search frictions is that buyers and sellers are spatially distributed, for example within a city, and meeting to trade requires costly transportation of one or both sides of the market. When the search literature incorporates buyer and seller heterogeneity, allocative efficiency is typically derived from an equilibrium distribution of match quality. In some spatial search markets, however, every trade involves a future re-allocation of buyers or sellers. This is particularly true in the transportation and shipping sector, where every individual transited or every load of cargo shipped entails a vehicle’s movement from one place to another. How buyers and sellers match with each other therefore impacts how the market evolves over time, as each trip affects the search friction faced by future buyers and sellers at each destination. Importantly, market participants will not internalize the impact of their trades on future participants, leading to potentially distorted allocations of supply and demand. An important question, therefore, is how much spatial misallocation is induced by this search externality and to what extent can pricing regulations serve to mitigate it? In this paper, I study the dynamic spatial search process, equilibrium, and the ability of prices to clear spatial markets in the presence of frictions. The setting is the regulated taxicab industry, where a decentralized geographic search process and tariff-based pricing leads to stark inefficiencies in the equilibrium spatial patterns of supply and demand.

The taxicab industry is a critical component of the transportation infrastructure in large urban areas, generating about $23 billion in annual revenues. New York City has long been the largest taxicab market in the United States, with this single market accounting for about 25% of national industry revenues in 2013. In New York, as well as many other cities, the taxi market is distinguished from other public transit options by a lack of centralized control; taxi drivers do not service established routes or coordinate search behavior. Instead, drivers search for passengers and, once matched, move them to destinations. How and where the search for passengers is conducted directly impacts the availability of taxi service across the city. This is determined by the distribution of customer demand in terms of both origin and destination, as well as the locations chosen for search among vacant taxis. These movements of capacity give rise to equilibrium search patterns that can leave some areas with little to no service while in other areas empty taxis will wait in long queues for passengers. Efficiency of these allocations depends on the transaction prices as well as the total externalities induced by cab-passenger matches.

\(^1\)Models of search conduct and equilibrium have been widely studied. Since the pioneering work of Diamond (1981, 1982a,b), Mortensen (1982a,b) and Pissarides (1984, 1985), the search and matching literature has focused on understanding the role of search frictions in impeding the efficient clearing of markets. The search and matching literature examines many markets where central or standardized exchange is not possible, including labor markets (e.g., Rogerson, Shimer, and Wright (2005)), marriage markets (e.g., Mortensen (1988)), monetary exchange (e.g., Kiyotaki and Wright (1989, 1993)), and financial markets (e.g., Duffie, Gärleanu, and Pedersen (2002, 2005)).
In this paper, I model taxi drivers’ location choices as a dynamic spatial oligopoly game in which vacant drivers choose where to provide service given the location and search behavior of their competition and the presence of search frictions. I show how price regulations impact the spatial distribution of demand and search behavior of taxi drivers, which in turn influence the equilibrium spatial distribution of taxi service. To empirically analyze this model, I use data from the New York City Taxi and Limousine Commission (TLC), which provides trip details including the time, location, and fare paid for all 27 million taxi rides in New York between August and October of 2012.\(^2\) Using TLC data together with a model of taxi search and matching, I estimate the spatial and intra-daily distribution of supply and demand in equilibrium. Importantly, the data only reveal matches made between taxis and customers as a consequence of search activity, but do not show underlying supply or demand; I do not observe how many taxis are vacant or the number of customers who want a ride in different areas of the city. Because these objects are necessary to measure search frictions and welfare in the market, I present an estimation strategy using the dynamic spatial equilibrium model together with a static location-specific matching function. Identification is obtained by first mapping the observed spatial distribution of matches into optimal policy functions of vacant taxis across the day. I then invert the matching function to recover an implied distribution of customer demand, and finally solve for matching efficiency using moments related to the inter-temporal variance of matches across days of the month. The most common methods for estimating dynamic oligopoly games involve solving for equilibrium policies for each type of agent at each point in time. These methods are infeasible here due to an extremely high-dimensional state space. Instead, I leverage the large number of agents in the model to solve for an equilibrium where each taxi driver plays against the distribution of his competitors.

I use this model to evaluate welfare and search frictions in this market. Model estimates further allow for estimating the effects of several alternative tariff pricing regimes as well as the effect of improved matching technology. Baseline estimates of welfare indicate that the New York taxi industry generates $1.58 million in consumer surplus and $3.1 million in taxi profits in each 9-hour day-shift, reflecting the aggregate surplus across 182 thousand taxi-passenger matches. Commensurate with these surpluses, however, is an average of 45 thousand failed customer searches per day and 5,405 vacant drivers at any point in the day. To what extent can a more sophisticated pricing policy mitigate these costs by better allocating available supply to demand? By simulating market equilibrium over nearly one million potential pricing rules, I am able to solve for a dynamically optimal flexible fare structure and show that a flexible tariff that scales with trip distance can provide up to a 12% increase consumer welfare, a 4% increase in taxi profits and a 37% increase in trips. Importantly, these gains come with the added benefit of reduced passenger

\(^2\)New York’s taxi industry is approximately 25% of the U.S. taxi market, the largest in the United States. Source: my own calculation based on 2013 data from the NYC Taxi and Limousine Commission and Brennan (2014)
waiting times. Alternative policies offering flexible tariffs by location and time yield slightly smaller benefits, but all of the counterfactual policies tested offer unambiguous benefits to the market even after accounting for search and matching frictions. I contrast these results with a counterfactual simulation of ride-sharing technology that implements perfect local-neighborhood matching, and show that optimal pricing policies can produce 30% more trips than the matching technology can alone.

Related Literature

This paper integrates ideas from the search and matching literature, spatial economics, and empirical studies of industry dynamics. My model is built around the aggregate matching function concept of canonical search-theoretic models.\(^3\) Traditionally, matching functions incorporate frictions as a functional form assumption, but more recent literature has derived matching functions under explicit micro-foundations. Notably, Lagos (2000) studies endogenous search frictions using a stylized environment of taxi search and competition. The model predicts how meeting probabilities adjust to clear the market and how misallocation can occur as an equilibrium outcome. Lagos (2003) uses the Lagos (2000) model to empirically analyze the effect of fares and medallion counts on matching rates and medallion prices in Manhattan. I draw elements from the Lagos search model, but make several changes to reflect the real-world search and matching process. Specifically, I add non-stationary dynamics, a more realistic and flexible spatial structure, stochastic and price-sensitive demand, fuel costs, and heterogeneity of the matching process in different locations. Further, I build a tractable framework for the empirical analysis of dynamic spatial equilibrium by providing tools for the estimation and identification of the model. Examples include specifying flexible location-specific matching functions and accounting for unobservables in drivers’ location choices.\(^4\) As with the latter study, my specification generates endogenous frictions within each search area as a function of the number of buyers and sellers. Finally, I explicitly estimate the unobservable supply and demand in each neighborhood and time, as well as demand functions in different locations and times-of-day, to quantify the spatial and inter-temporal patterns of welfare achieved in the market and the ability of prices to enhance dynamic efficiency.


\(^3\)A starting point for the implementation of these models in studying labor markets can be found in surveys of search-theoretic literature by Mortensen (1986), Mortensen and Pissarides (1999) and Rogerson, Shimer, and Wright (2005). Hall (1979) introduces the aggregate matching function concept, and notably uses the urn-ball framework adapted in this paper.

\(^4\)See also recent work of Brancaccio, Kalouptsidi, and Papageorgiou (2017) for an application of related search models to study endogenous trade costs in the global shipping industry.
and their competitors. They key difference is that the firm choices in this model are spatial, and the collective policies of drivers lead to a dynamic spatial equilibrium. There are may spatial areas that are relevant for driver choices, and drivers also value spaces differently at different times of day, so the dynamic problem is high-dimensional and non-stationary. To mitigate computational burden, my approach exploits the large number of agents in the market by assuming that state transitions are perfectly forecastable. This assumption not only permits solving for equilibrium search policies of taxis, but also reflects a more realistic behavioral model: here, an agent’s strategic location choices involve playing against the expected distribution of competitors throughout the day, rather than all possible realizations of competitors’ states. This approach closely relates to oblivious equilibrium (Weintraub, Benkard, and Van Roy (2008b)), the literature on non-stationary firm dynamics (e.g., Weintraub, Benkard, Jezierski, and Van Roy (2008a), Melitz and James (2007)) as well as in auction models with many bidders (e.g., Hong and Shum (2010)).

To my knowledge, this is the first empirical analysis of pricing and welfare in this market and the first to study how tariffs impact spatial allocations. A related study is Frechette, Lizzeri, and Salz (2017), which models the labor supply dynamics of taxis to ask how customer waiting times and welfare are impacted by medallion regulations and matching technology. As with my paper, Frechette, Lizzeri, and Salz (2017) study the effect of regulations on search frictions and welfare. The key difference is that they focus on the labor supply decision rather than the spatial location decision. As a result they model the dynamic equilibrium effects of regulation on aggregate (i.e., city-wide) levels of service and demand, whereas I model the dynamic equilibrium effects of regulation on spatial distributions of supply and demand. Though these approaches differ substantially, they lead to similar predictions when comparing similar counterfactuals.

There is a recent literature on the benefits of dynamic pricing used in ride-hail services (e.g., Hall, Kendrick, and Nosko (2015), Castillo, Knoepfle, and Weyl (2017)). This paper similarly highlights the impact of pricing on efficiency, but with two distinct differences. First, I focus on posted tariffs as opposed to real-time price adjustment. Posted tariffs are a feature of both traditional taxis and ride-hail services that affect the search behavior of taxi drivers. Second, I explicitly model the influence of prices on the dynamic path of supply and demand. I leverage this model to show how prices can be configured to induce efficient allocations of supply and demand.

There is an additional body of literature on taxi drivers’ labor supply choices, including Camerer, Babcock, Loewenstein, and Thaler (1997), Farber (2005, 2008), Crawford and Meng (2011), and Thakral and To (2017). These studies investigate the labor-leisure tradeoff for drivers. They ask how taxi drivers’ labor supply is determined and to what extent it is driven by daily wage targets and other factors. Buchholz, Shum, and Xu (2017) estimate a dynamic labor supply model of taxi drivers to show that behavior consistent with dynamic optimization may appear as a behavioral bias when viewed in a static setting.

There is also a literature in empirical industrial organization which studies the allocative distortions induced by search frictions in different industries. This includes work on airline parts Gavazza (2011) and mortgages Allen, Clark, and Houde (2014).
while accounting for the flow of reallocated of cabs due to passenger trips.

Finally, a diverse set of literature exists to address whether taxi regulation is necessary at all. Among this literature, both the theoretical and empirical findings offer mixed evidence. These studies point to successful regulation’s function to reduce transaction costs (Gallick and Sisk (1987)), prevent localized monopolies (Cairns and Liston-Heyes (1996)), correct for negative externalities (Schrieber (1975)), and establish efficient quantities of vacant cabs (Flath (2006)). Other authors assert that regulations have lead to restricted quantities and higher prices (Winston and Shirley (1998)) and that low sunk and fixed costs in this industry are sufficient to support competition (Häckner and Nyberg (1995)). My paper shows how existing regulatory levers are inefficiently configured, and that a better implementation of posted tariffs leads to more efficient spatial allocations and higher rates of utilization.

This paper is organized as follows. Section 2 details taxi industry characteristics relating to search, regulation, and spatial sorting, as well as a description of the data. Section 3 presents the dynamic model of taxi search and matching. Section 4 discusses the empirical strategy for computing equilibrium and estimating model parameters. Results are presented in section 5, with an analysis of counterfactual policies in section 6. Section 7 concludes.

2 Regulation and Search Frictions in The Taxi Industry

2.1 Industry Characteristics

Fragmented Firms

In the U.S., taxi service is highly fragmented. The market share of the largest firm is less than 1%, and the largest four firms make up less than 3% of the overall market. There are many individual firms, some consisting of a single owner-operator. While there is increased concentration at the local level in which taxis operate, most non-owner drivers lease taxis from owners. The typical lease arrangement has drivers paying a fixed leasing cost, paying for their own gas and insurance, and collecting all residual revenue as profit. Given these arrangements, drivers do not centrally coordinate their search behavior. Instead, each driver independently searches for passengers, competing with other drivers for the same demand. This is especially apparent in New York City, where yellow medallion taxis are only permitted to operate a street-hail service and may not pre-arrange rides with customers.

Regulation

Taxi markets are typically highly regulated by local municipalities. There are two predominant taxi regulations. The first is a fixed two-part tariff fare pricing structure, where fares are based on
a one-time flag-drop fee and a distance-based fee. The second is the imposition of entry restrictions as a limit on the number of legal taxis that may operate. This is typically implemented by requiring drivers to hold a permit or medallion, the supply of which are capped. Entry restrictions are often controversial; critics argue that they are a product of regulatory capture, and serve to enrich medallion owners by limiting competition. Proponents of regulation highlight several market failures that arise in an unregulated environment: congestion externalities, localized market power in remote locations, and potentially high bargaining costs.

In recent years, several firms including Uber, Lyft, Curb, and Sidecar have entered the taxi industry, all of which take advantage of mobile technology to match customers with cabs, thereby greatly reducing frictions associated with taxi search and availability. In addition, these firms implement pricing schemes which, instead of being uniform across the day as with traditional taxis, tend to adjust to supply and demand conditions. The precipitous expansion and success of these firms is suggestive of the enormous benefits associated with both the reduced search costs and more flexible pricing compared with traditional taxi markets.

Another potential reason for this expansion is that taxi regulations are often at odds with this new wave of technology-centered entrants. These firms tend to enjoy much less stringent entry restrictions than the more regulated incumbents, leading to a variety of legal disputes as stakeholders in the traditional taxi business absorb losses. The stakes of this debate are large, and highlight the need for analysis surrounding the effects of these new entrants. This is a central question of this paper: how regulation and matching technology impacts market equilibrium and welfare, with particular attention to spatial availability. This paper aims to characterize the location incentives faced by taxis, the equilibrium spatial distribution of supply and demand, and to analyze the role of regulations on market outcomes.

---

7These licenses are also tradable, and the mere fact that they tend to have positive value, sometimes in excess of one million dollars, implies that this quantity cap is binding and below that of an unrestricted equilibrium.
8See, e.g., forbes.com/sites/ellenhuet/2015/06/19/could-a-legal-ruling-instantly-wipe-out-uber-not-so-fast/.
9The spatial availability of taxis is of evident concern to municipal regulators around the country: policies of various types have been enacted in different cities to control the spatial dimension of service. For example, in the wake of criticism over the availability of taxis in certain areas, New York City issued licenses for 6,000 additional medallion taxis in 2013 with special restrictions on the spatial areas they may service (See, e.g., cityroom.blogs.nytimes.com/2013/11/14/new-york-today-cabs-of-a-different-color/). Specifically, these green-painted “Boro Taxis” are only permitted to pick up passengers in the boroughs outside of Manhattan. Though the city’s traditional yellow taxis have always been able to operate in these areas, it’s apparent that service was scarce enough relative to demand that city regulators intervened by creating the Boro Taxi service. This intervention highlights the potential discord between regulated prices and the location choices made by taxi drivers.
2.2 Data Overview

New York City is the largest taxi market in the United States, with 236 million passenger trips in 2014, or about 25% of all U.S. service. In 2009, The Taxi and Limousine Commission of New York City (TLC) initiated the Taxi Passenger Enhancement Project, which mandated the use of upgraded metering and information technology in all New York medallion cabs. The technology includes the automated data collection of taxi trip and fare information. I use TLC trip data from all New York City medallion cab rides given from August 1, 2012 to September 30, 2012. An observation consists of information related to a single cab ride. Data include the exact time, date and GPS coordinates of pickup and drop-off, trip distance, and trip time length for approximately 27 million rides.\(^\text{11}\) New York cabs typically operate in two separate shifts of 9-12 hours each, with a mandatory shift change between 4-5pm. I focus on the day-shift period of 7am until 4pm, after which I assume all drivers stop working.

A unique feature of New York taxi regulation is that medallion cabs may only be hailed from the street and are not authorized to conduct pre-arranged pick-ups, which are the exclusive domain of licensed livery cars. As a result, the TLC data only record rides originating from street-hails. This provides an ideal setting for analyzing taxi search behavior since all observed rides are obtained through search. Table 1 provides summary facts for this data set. Additional monthly-level statistics are in Appendix A.3.

Most of the time, New York taxis operate in Manhattan. When not providing rides within Manhattan, the most common origins and destinations are to New York’s two city airports, LaGuardia (LGA) and John F. Kennedy (JFK). Instead of conducting a search for passengers, taxis form queues and wait in line for next available passengers. The costly waiting times and travel distances is offset by larger fares, however. Table 2 below provides statistics related to the frequency and revenue share of trips between Manhattan, the two city airports, and elsewhere.

Lastly, Uber began operating in New York City beginning in 2011, but service remained minimal. In an October 2012 interview, the CEO reported that 160 drivers had participated in providing trips in the city since opening.\(^\text{12}\) This represents about 1% of licensed yellow cab drivers, and likely much less in trip volume as these drivers were not necessarily operating consistently throughout the prior year, nor full time.

\(^{11}\) Using this information together with geocoded coordinates, we might learn for example that cab medallion 1602 (a sample cab medallion, as the TLC data are anonymized) picks up a passenger at the corner of Bowery and Canal at 2:17pm of August 3rd, 2012, and then drives that passenger for 2.9 miles and drops her off at Park Ave and W. 42nd St. at 2:39pm, with a fare of $9.63, flat tax of $0.50, and no time-of-day surcharge or tolls, for a total cost of $10.13. Cab 1602 does not show up again in the data until his next passenger is contacted.

### Table 1: Taxi Trip and Fare Summary Statistics

Taxi trip and fare data come from New York Taxi and Limousine Commission (TLC). This table provides statistics related to individual taxi trips taken in New York City between August 1, 2012 and September 30, 2012 for two fare types. The first is the standard metered fare (TLC rate code 1), in which standard fares apply, representing 98.1% of the data. The second is a trip to or from JFK airport (TLC rate code 2). Total Fare and Distance data are reported for each ride in the dataset. The two main fare components are a distance-based fare and a flag-drop fare. I predict these constituent parts of total fare using the prevailing fare structure on the day of travel and the distance travelled, though they are not separately reported from each other or from waiting costs. Flag fare calculations include the presence of time-of-day surcharges. Any remaining fare is due to a charge for idling time. The first set of statistics relate to the full sample of all New York taxis rides across the two months, and the second set relates to the smaller sample used in this analysis: weekdays, day-shift trips occurring within the space described in Figure 4.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Rate Type</th>
<th>Variable</th>
<th>Obs.</th>
<th>10%ile</th>
<th>Mean</th>
<th>90%ile</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Data</td>
<td>Standard Fares</td>
<td>Total Fare ($)</td>
<td>27,475,614</td>
<td>4.50</td>
<td>9.51</td>
<td>16.00</td>
<td>5.57</td>
</tr>
<tr>
<td></td>
<td>Standard</td>
<td>Dist. Fare ($)</td>
<td>27,475,621</td>
<td>1.50</td>
<td>5.59</td>
<td>12.00</td>
<td>6.14</td>
</tr>
<tr>
<td></td>
<td>Standard</td>
<td>Flag Fare ($)</td>
<td>27,475,621</td>
<td>2.50</td>
<td>2.83</td>
<td>3.50</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td>Standard</td>
<td>Distance (mi.)</td>
<td>27,475,621</td>
<td>0.82</td>
<td>2.70</td>
<td>6.00</td>
<td>2.74</td>
</tr>
<tr>
<td></td>
<td>Standard</td>
<td>Trip Time (min.)</td>
<td>27,475,621</td>
<td>4.00</td>
<td>12.04</td>
<td>22.52</td>
<td>8.23</td>
</tr>
<tr>
<td>JFK Fares</td>
<td>Standard</td>
<td>Total Fare ($)</td>
<td>491,689</td>
<td>45</td>
<td>48.32</td>
<td>52</td>
<td>3.58</td>
</tr>
<tr>
<td></td>
<td>Standard</td>
<td>Distance (mi.)</td>
<td>491,689</td>
<td>3.02</td>
<td>16.25</td>
<td>20.58</td>
<td>5.95</td>
</tr>
<tr>
<td></td>
<td>Standard</td>
<td>Trip Time (min.)</td>
<td>491,689</td>
<td>22.75</td>
<td>39.49</td>
<td>60.00</td>
<td>17.33</td>
</tr>
<tr>
<td>Weekdays, Day-Shift, Manhattan &amp; Boro.</td>
<td>Standard</td>
<td>Total Fare ($)</td>
<td>8,164,678</td>
<td>4.50</td>
<td>10.17</td>
<td>17.70</td>
<td>6.42</td>
</tr>
<tr>
<td></td>
<td>Standard</td>
<td>Dist. Fare ($)</td>
<td>8,122,794</td>
<td>1.20</td>
<td>4.66</td>
<td>9.60</td>
<td>5.33</td>
</tr>
<tr>
<td></td>
<td>Standard</td>
<td>Flag Fare ($)</td>
<td>8,122,794</td>
<td>2.50</td>
<td>2.5</td>
<td>2.5</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Standard</td>
<td>Distance (mi.)</td>
<td>8,122,794</td>
<td>0.71</td>
<td>2.28</td>
<td>4.67</td>
<td>2.37</td>
</tr>
<tr>
<td></td>
<td>Standard</td>
<td>Trip Time (min.)</td>
<td>8,122,794</td>
<td>4.00</td>
<td>12.74</td>
<td>23.8</td>
<td>8.49</td>
</tr>
<tr>
<td>JFK Fares</td>
<td>Standard</td>
<td>Total Fare ($)</td>
<td>171,223</td>
<td>45.00</td>
<td>48.28</td>
<td>52.00</td>
<td>3.60</td>
</tr>
<tr>
<td></td>
<td>Standard</td>
<td>Distance (mi.)</td>
<td>171,223</td>
<td>2.00</td>
<td>16.14</td>
<td>20.91</td>
<td>6.09</td>
</tr>
<tr>
<td></td>
<td>Standard</td>
<td>Trip Time (min.)</td>
<td>171,223</td>
<td>26.18</td>
<td>45.65</td>
<td>67.00</td>
<td>19.16</td>
</tr>
</tbody>
</table>

2.3 Evidence of Frictions

Search frictions occur when drivers cannot locate passengers even though supply and demand coexist at some point in time. Frictions in this market manifest as waiting time experienced by drivers to match with a passenger.\(^{13}\) The TLC data provides evidence of the presence of search frictions for drivers that vary across space and time of day. Using driver ID together with the time of pick-up and drop-off, I compute the waiting time between trips. The mean waiting time for different trips is displayed in Figure 1. Panel 1 divides the day into 5-minute segments and shows the probability that a driver will find a passenger in each period, as well as the expected waiting time to find a passenger, in units of 10-minutes (i.e., a value of 0.5 equals 5 minutes). It shows

\(^{13}\)Even if all passengers are matched as soon as they start looking for a ride, frictions exist because taxis compete with one another: each driver wants to find the scarce passengers before their rivals.
Table 2: Taxi Trips and Revenues by Area

<table>
<thead>
<tr>
<th>Time</th>
<th>Place</th>
<th>Obs.</th>
<th>Mean Fare</th>
<th>Trip Share</th>
<th>Rev. Share</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Times</td>
<td>Intra-Manhattan Trips</td>
<td>24,835,103</td>
<td>$9.28</td>
<td>89%</td>
<td>73%</td>
</tr>
<tr>
<td></td>
<td>Airport Trips</td>
<td>1,568,699</td>
<td>$33.77</td>
<td>6%</td>
<td>17%</td>
</tr>
<tr>
<td></td>
<td>Other Trips</td>
<td>1,563,501</td>
<td>$19.83</td>
<td>6%</td>
<td>10%</td>
</tr>
<tr>
<td>Weekdays,</td>
<td>Intra-Manhattan Trips</td>
<td>7,813,226</td>
<td>$9.33</td>
<td>91%</td>
<td>76%</td>
</tr>
<tr>
<td>Day-shift</td>
<td>Airport Trips</td>
<td>503,711</td>
<td>$34.80</td>
<td>6%</td>
<td>18%</td>
</tr>
<tr>
<td></td>
<td>Other Trips</td>
<td>270,883</td>
<td>$19.62</td>
<td>3%</td>
<td>6%</td>
</tr>
</tbody>
</table>

Taxi trip and fare data come from New York Taxi and Limousine Commission (TLC). This table provides statistics related to the locations of taxi trips taken in New York City between August 1, 2012 to September 30, 2012. Intra-Manhattan denotes trips which begin and end within Manhattan, Airport Trips are trips with either an origin or destination at either LaGuardia or JFK airport, and Other Trips captures all other origins and destinations within New York City. Statistics are reported for all times as well as the day-shift period of a weekday, from 6am until 4pm. The latter category is the focus of my analysis.

substantial intra-day variation in search times, with the best times of day for finding passengers around 9am and closer to 4pm, with average wait times around 6 minutes and 5-minute finding rates around 50%. The worst times are in early morning and mid-day, where average wait times are nearly 10 minutes and finding rates fall as low as 25%. Panel 2 divides New York into 37 locations (see Section 4 for details) and shows the same driver match probabilities and waiting times by region averaged over time of day and all weekdays of the month. Again there is heterogeneity across space, with relatively higher match probabilities and lower waiting times in lower-Manhattan (1-8) and Midtown (9-18), declining into the uptown neighborhoods (19-34) and even worse in Brooklyn (35-37). In aggregate, drivers spend about 47% of their time vacant during the sample period of weekdays during the day-shift. This suggests that among 11,500 active drivers, an average of 5,405 are vacant at any one time.

Figure 1 provides a snapshot of the frictions faced by drivers by time-of-day and neighborhood. The data do not reveal the frictions faced by customers; it is impossible to tell how long a customer has been waiting before pick-up, nor is it possible to tell if a customer arrived to search for a taxi and gave up. To analyze frictions, I will exploit observed moments in the data and estimate a model of dynamic spatial search and matching to recover the distribution of customer demand as well as the elasticities of demand across space and time. The estimated model will permit measuring the impact of frictions and the extent of misallocation due to inefficient pricing regulation.

There is additional evidence that drivers often relocate to find passengers: 61.3% of trips begin in a different neighborhood than the neighborhood where drivers last dropped off a passenger. This suggests that some spatial search frictions are present for drivers, as finding a customer requires relocation.
3 Model

A city is a network of $L$ nodes called “locations”, connected by a set of routes. A location can be thought of as a spatial area within the city. Time within a day is divided into discrete intervals with a finite horizon, where $t = \{1, ..., T\}$. At time $t = 1$ the work day begins; at $t = T$ it ends. Model agents are vacant taxi drivers who search for customers within a location $i \in \{1, ..., L\}$. When taxis find passengers, they drive them from origin location $i$ to a destination location $j \in \{1, ..., L\}$. Denote $v_t^i \in \mathbb{R}$ as a measure of vacant taxis and denote $u_t^j \in \mathbb{R}$ as a measure of customers looking for a taxi in each location at each time. The distance between each location is given by $\delta_{ij}$ and the travel time between each location is given by $\tau_{ij}$.

The model has four basic ingredients. First, there is a demand system that describes, for every possible route in the city, how many customers will arrive to the market in search of a taxi as a function of the price of service along that route. Second, there is a payoff vector associated with every route that taxis service. Payoffs include the revenues from each ride as well as a service cost due to fuel expenses. Third, there is a model of period-by-period market clearing. Here I use a search and matching model to map supply and demand into matching probabilities, which clear the

---

15e.g., a series of blocks bounded by busy thoroughfares, different neighborhoods, etc.
market given that prices are fixed.\textsuperscript{16} Finally, these three components are combined in a dynamic model of location choice. The dynamic model captures the decision making process of vacant drivers, who make period-by-period location choices accounting for future search opportunities. These four components are presented in more detail below.

3.1 Demand

In each location $i$ at time $t$, the number of customers that wish to move to a new location, $u_{it}^l$, is drawn from a Poisson distribution with parameter $\lambda_{it}^l$. Moreover, $\lambda_{it}^l$ is a sum of independent Poisson parameters $\lambda_{it}^l = \sum_j \lambda_{ij}^l(P_{ij})$, where $\lambda_{ij}^l(P_{ij})$ represents the destination-$j$-specific Poisson arrival of customers in location $i$ at time $t$. The parameters $\lambda_{ij}^l$ are functions of the price of a taxi ride between $i$ and $j$, $P_{ij}$. Conditional on a set of prices, the probability that a customer in $i$ wants to travel to location $j \in \{1, \ldots, L\}$ at time $t$ is given by $M_{ij}^t$, so that $\lambda_{ij}^l(P_{ij}) = M_{ij}^t \cdot \lambda_{i}^l(P)$, where $P$ denotes a vector of prices between all locations.

I assume that taxi drivers face a constant-elasticity demand curve. Demand depends on the origin and destination of the trip, its price, and the time of day. Price elasticities depend on whether the trip involves an airport (a binary index denoted $a$) and the distance of the trip (indexed by discrete categories $s$). Taxi demand takes the form:

$$\ln(\lambda_{ij}^l(P_{ij})) = \alpha_{0,i,t,s,a} + \alpha_{1,s,a} \ln(P_{ij}) + \eta_{is}. \quad (1)$$

In addition, I assume that customers demand taxi services for one period. After this period, consumers use a different method of transit.\textsuperscript{17}

3.2 Revenue and Costs

Taxis earn revenue from giving rides and pay a cost of fuel. Earning revenue is assumed to be the single objective faced by working drivers. At the end of each ride, the taxi driver is paid according to the fare structure. The fare structure is defined as follows: $b$ is the one-time flag-drop fare and $\pi$ is the distance-based fare, with the distance $\delta_{ij}$ denoting the distance between $i$ and $j$. Thus the fare revenue earned by providing a ride from $i$ to $j$ is $b + \pi \delta_{ij}$.

Drivers have two sources of costs. First, there is a fixed daily fee for leasing the taxi and medallion license (or a financing cost for drivers who also own their own medallion). Second there are intra-daily fuel costs, which I denote as $c_{ij}$. On any particular day a driver is working, the sunk

\textsuperscript{16}Note that in a setting of ride-hail, in which prices adjust to neighborhood market conditions, we might instead recast this model as one of localized price formation instead of search frictions.

\textsuperscript{17}In the empirical analysis to follow, I define one period as five minutes. In the context of New York City, there are plenty of alternative transport options, and this assumption suggests that customers will choose to travel via one of these alternatives upon failing to find a taxi.
cost of leasing a medallion for that day is immaterial to the driver’s search choices, and so I ignore these costs in the model.

The net revenue of any passenger ride is given by

\[ \Pi_{ij} = b + \pi \delta_{ij} - c_{ij}. \]  

(2)

This profit function sums the total fare revenue earned net of fuel costs in providing a trip from location \( i \) to \( j \).

3.3 Searching and Matching

There are two types of locations, search locations and airports. Search locations comprise most a city; they are locations in which cabs drive around in search of passengers. In airport locations, however, taxis wait in queue for a guaranteed passenger ride upon reaching the front of the queue.\(^{18}\)

The next two subsections detail how matches are formed in these two location types.

At the start of each period, taxis search for passengers. The number of taxis in each location at the start of the period is given by the sum of previously vacant taxis who have chosen location \( i \) to search, plus the previously employed taxis who have dropped off a passenger in location \( i \). This sum is denoted as \( v_t^i \). I make the following assumptions about matching: (1) matches can only occur among cabs and customers within the same location, (2), matches are randomly assigned between taxis and customers, and (3) once a driver finds a customer, a match is made and the driver may not refuse a ride.\(^{19}\)

The expected number of matches made in location \( i \) and time \( t \) is given by an aggregate matching function \( m_i(\lambda_t^i, v_t^i) \). The ex-ante probability that a driver will find a customer is then given by \( p_t^i = \frac{m_i(\lambda_t^i, v_t^i)}{v_t^i} \). Figure 2 illustrates the within-period search and matching process.

3.3.1 A Model of Search Locations

In the Lagos (2000) model, space is defined as a discrete set of points. At a particular point, a frictionless matching technology is specified by \( m = \min\{u, v\} \). In the context of continuous space, a sufficiently small area (for example a single street corner where all customers and taxi drivers see each other) could be reasonably specified by frictionless matching. A larger area, such as a whole neighborhood, is made up of many such small points and will exhibit frictions when drivers are on different blocks from customers.

To model location-level search frictions, I use an aggregate matching function given by equation 3.

---

\(^{18}\)Airport rides comprise roughly 6% of all taxi trips, and 16% of revenues.


\(^{20}\)This function is derived from an urn-ball matching problem first formulated in Butters (1977) and Hall (1979). Where the original model characterizes matches from discrete (i.e., integer) inputs, I have reformulated this function.
passengers arrive at Poisson rate $\lambda_i^t$

taxis w/ passengers

$\# \text{ vacant cabs} = v_i^t$

$\# \text{ customers} = u_i^t \sim F(\lambda_i^t)$

matches: $m_i^t = m_i(\lambda_i^t, v_i^t)$

$p_i^t = E\left[\frac{m_i(\cdot, \cdot)}{v_i^t}\right]$

Prob($\text{match}; \lambda_i^t, v_i^t$))

vacant taxis + taxis dropping off passengers

vacant taxis

Figure 2: Flow of demand, matches, and vacancies within a location

This illustration depicts the sources of taxi arrivals and departures in location $i$ and time $t$. At the beginning of a period, all taxis conducting search in location $i$ are either dropping off passengers or vacant and searching from previous periods. Matches are then made according to the matching function $m_i(u_i^t, v_i^t)$. At the end of the period, newly employed taxis leave for various destinations and vacant taxis continue searching.

$\alpha_r > 0$ is an efficiency parameter; all else equal, larger values of alpha generate fewer matches. $r$ denotes a region, or a subset of locations as described in section [4.1.1] This parameter depends on location as it reflects the difficulty of search within a region, such as the complexity of the street grid or limitations of visibility. These are physical characteristics which are assumed to be fixed with the region and across the day. This function serves as a reduced-form characterization of within-neighborhood search and matching, and via offers a flexible parameterization via the efficiency parameter $\alpha_r$. An illustration of the aggregate matching function and the role of $\alpha_r$ is depicted in Figure [3]

\[ m(\lambda_i^t, v_i^t, \alpha_r) = v_i^t \cdot \left(1 - e^{-\frac{\lambda_i^t}{\alpha_r v_i^t}}\right) \] (3)

3.3.2 A Model of Airport (Queueing) Locations

At airports, taxis pull into one of multiple queues and wait for passengers to match with cabs at the front of the queue. Demand in each location-time $i, t$ is a Poisson random variable with parameter $\lambda_i^t = \sum_j \lambda_{ij}^t (P_{ij})$. I assume there is a measure congestion in the taxi lane, $\omega_i^t$ which reflects how, once a passenger has arrived, it takes a very short time until that passenger is in the front taxi and to characterize continuous inputs. See more detail in appendix XXX.

In principle $\alpha$ could depend on both time and location, say if time were thought to influence the matching technology at the expense of imposing greater requirements on data to get credible identification of additional parameters.

---

21In principle $\alpha$ could depend on both time and location, say if time were thought to influence the matching technology at the expense of imposing greater requirements on data to get credible identification of additional parameters.
the taxi has exited the cab-stand to make way for the next taxi. The total number of cabs at the airport in each period is $v^t_i$. Thus the total number of matches made is given by $\min\{u^t_i, w^t_i\}$, where $w^t_i(\omega^t_i, v^t_i)$ is the number of taxis in the queue that can pick up a customer within the time of each period, given the presence of some congestion which prevents instantaneous clearing of the queue. Finally, I assume that airport consumers are captive in that they don’t have an outside option once arrived to the taxi stand, and will wait until the line clears.

3.4 Dynamics: Taxi Drivers’ States, Actions and Payoffs

A taxi driver’s behavior is dependent on his own private state $(\ell^t_a, e^t_a)$ and the market state, $S^t$. Specifically a driver $a$’s own location at time $t$ is given by $\ell^t_a \in \{1, \ldots, L\}$, and employment status (vacant or employed) is $e^t_a \in \{0, 1\}$. The market state $S^t$ at time $t$ is a measure of vacant taxis $v^t_i$ in each location $i$, as well as a measure of employed taxis $v^t_{e,k}$ actively in-transit between locations, where all in-transit states are indexed by $k \in \{1, \ldots, T\}$, where $k$ records how many periods until arrival at each drop-off location when the taxi becomes vacant once again.\footnote{At any moment, employed taxis are not directly competing with vacant taxis for passengers. Accounting for the number and location of employed taxis is an important component of the state variable because the eventual arrival of employed taxis and subsequent transition to vacancy is payoff-relevant for the decision-making problem of forward-looking vacant taxis.} Thus the market state at time $t$ is given by

$$S^t = \{\{v^t_i\}_{i \in \{1, \ldots, L\}}, \{v^t_{e,k}\}_{k \in \{1, \ldots, K\}}\}. \quad (4)$$
Denote \( \mathcal{S} = \{ \mathcal{S}^t \} \) \( \forall t \) so that \( \mathcal{S} \) reflects the entire spatial and intertemporal distribution of vacant and employed taxis. At the beginning of each period, taxi drivers make a conjecture about the current-period state and transition probabilities into the next period. Given this conjecture, they assign value \( V_t^i \) to each \( i, t \)-pair.

I define the drivers’ ex-ante (i.e., before observing any shocks and before any uncertainty in passenger arrivals is resolved) value as

\[
V_t^i(\mathcal{S}^t) = \mathbb{E}_{\mathcal{S}^t|\mathcal{S}^t}\left[ p(\lambda_t^i, v_t^i|\alpha_i) \left( \sum_j M_t^{ij} \cdot (\Pi_{ij} + V_{j}^{t+\tau_{ij}}(\mathcal{S}^{t+\tau_{ij}})) \right) + (1 - p(\lambda_t^i, v_t^i|\alpha_i)) \cdot \mathbb{E}_{\varepsilon_j,a} \left[ \max_{j \in \mathcal{A}(i)} \{ \tilde{V}^{t+\tau_{ij}}(\mathcal{S}^{t+\tau_{ij}}, \gamma) - c_{ij} + \varepsilon_{j,a} \} \right] \right]. \tag{5}
\]

This expression is decomposed as follows. Drivers in \( i \) at time \( t \) expect to contact a passenger with probability \( p(\lambda_t^i, v_t^i|\alpha_i) \). Drivers’ payoff for providing a trip is equal to the net profit of a trip \( \Pi_{ij} \) plus the associated continuation values \( V_{j}^{t+\tau_{ij}} \) of being in location \( j \tau_{ij} \) periods later. Therefore the expected value of a trip is simply the value of a trip to each \( j \) weighted by the probability that a passenger picked-up in \( i \) chooses destination \( j \), which is given by \( M_t^{ij} \).\(^{23}\)

At the end of the period, any cabs which remain vacant can choose to relocate or stay put to begin a search for passengers in the next period. Relocation over longer distances requires more time. Vacant drivers choose to search next period in the location that maximizes total expected payoff as the sum of continuation values \( \tilde{V}_{j}^{t+\tau_{ij}}(\mathcal{S}, \gamma) \), fuel costs \( c_{ij} \) and a contemporaneous agent-location-specific shock \( \varepsilon_{j,a} \), where

\[
\tilde{V}_{j}^{t+\tau_{ij}}(\mathcal{S}, \gamma) = \begin{cases} 
V_{j}^{t+\tau_{ij}}(\mathcal{S}) & i \neq j \\
\gamma V_{j}^{t+\tau_{ij}-1}(\mathcal{S}) + (1 - \gamma)V_{j}^{t+\tau_{ij}}(\mathcal{S}) & i = j
\end{cases} \tag{6}
\]

The parameter \( \gamma \) represents an extra time-payoff associated with choosing to stay put: when taxis continue to search in the current location, they will receive extra value associated with time and effort savings.

\( \varepsilon_{j,a} \) is a driver \( a \)-specific i.i.d. shock to the perceived value of search in each alternative location \( j \), which I assume to be drawn from a Type-I extreme value distribution. This shock accounts for unobservable reasons that individual drivers may assign a slightly greater value to one location over another. For example, traffic conditions and a taxi’s direction of travel within a location may make

\(^{23}\)Note that \( M_t^{ij} \) has superscript \( t \) because preferences of passengers change throughout the day.
it inconvenient to search anywhere but further along the road in the same direction.\textsuperscript{24}

The set $A(i)$ reflects the set of locations available to vacant taxis. I assume that $A(i)$ is the entire set of locations in the city. If a driver chooses an $j \in A(i)$ which is not adjacent to $i$, this means the driver opts out of search along the path. For example, vacant cabs might move from lower Manhattan to Midtown via the west-side highway, which would have less travel-time (via $\tau$) compared with a central path involving adjacent locations.

Time ends at period $T$. Continuation values beyond $t = T$ are set to zero: $V_{t}^{i} = 0 \ \forall t > T, \forall i$.

At the end of each period, vacant drivers decide where to search for passengers in the next period by choosing the location with the highest present value of search net of transportation costs. Vacant drivers in location $i$ move to location $j^*$ by solving the last term in equation $5$:

$$j^* = \arg \max_{j} \{ \tilde{V}^{t + \tau_{ij}} (S^{t + \tau_{ij}}, \gamma) - c_{ij} + \varepsilon_{ja} \}.$$  \hfill (7)

To compute the driver’s strategies, I define the ex-ante choice-specific value function as $W_{t}^{i}(j_{a}, S^{t}, \gamma)$, which represents the net present value of payoffs conditional on taking action $j_{a}$ while in location $i$, before $\varepsilon_{ja}$ is observed:

$$W_{t}^{i}(j_{a}, S^{t}, \gamma) = \mathbb{E}_{S^{t + \tau_{ij}} \leftarrow S^{t}} \left[ \tilde{V}^{t + \tau_{ij}} (S^{t + \tau_{ij}}, \gamma) - c_{ij} \right].$$  \hfill (8)

Defining $W_{t}^{i}$ separately from $V_{t}^{i}$ permits a simple expression of taxi drivers’ conditional choice probabilities: the probability that a driver in $i$ will choose $j \in A(i)$ conditional on observing state $S^{t}$, but before observing $\varepsilon_{ja}$, is given by the multinomial logit formula.

$$P_{i}^{t}[j_{a} \mid S^{t}] = \frac{\exp(W_{t}^{i}(j_{a}, S^{t}, \gamma)/\sigma_{\varepsilon})}{\sum_{k \in A(i)} \exp(W_{t}^{i}(j_{k}, S^{t}, \gamma)/\sigma_{\varepsilon})}.$$  \hfill (9)

This expression defines drivers’ policy functions $\sigma_{t}^{i}$ in each location $i$ and time $t$ as the probability of optimal transition from an origin $i$ to all destinations $j$ conditional on future-period continuation values.

### 3.5 Intraday timing

At time $t = 1$, taxis have an initial spatial distribution, which I denote as $S^{1}$. At this time taxis conduct search and some match with passengers while others are left vacant. Newly employed taxis disappear from the stock of vacant cabs, earning period profits associated with each trip. Vacant taxis realize zero period profits but face continuation values associated with each possible move.

\textsuperscript{24}The terms $\varepsilon_{ja}$ also ensure that vacant taxis leaving one location will mix among several alternative locations rather than moving to the same location, a feature broadly corroborated by data.
Equation [7] defines which locations they will opt to search next. In period $t = 2$, the locations of vacant taxis are updated, some taxis drop off passengers from period $t = 1$, and the locations of employed taxis who are still in-transit are updated (in $v^t_{e,k}$). Any vacant taxis then search for passengers, repeating the same process as above. The full inter-period timing is as follows:

1. Taxis are distributed according to $S^1$. Aggregate demand per location is given by $\lambda^t_i$.
2. $m_i(\lambda^1_i, v^1_i | \alpha_i)$ random matches occur in each location.
3. $m_i(\lambda^1_i, v^1_i | \alpha_i)$ employed taxis leave for destinations.
4. The remaining $\lambda^t_i - m_i(\lambda^1_i, v^1_i | \alpha_i)$ customers leave the market.
5. The remaining $v^t_i - m_i(\lambda^1_i, v^1_i | \alpha_i)$ vacant taxis choose a location to search in next period according to policy functions.25
6. Previously vacant and hired taxis arrive in new locations, forming distribution $S^2$.26
7. The process repeats from $S^2$, $S^3$, etc. until reaching $S^T$.

### 3.6 Transitions

Policy functions $\sigma^t_i$ are a vector of transition probabilities from origin $i$ to destinations $j \in L$. Note that only vacant taxis transition according to these policies. Employed taxis will transition according to passenger transition probabilities given by $M^t_i$, also a vector of transition probabilities from origin $i$ to destinations $j \in L$. Together, these two processes combine to generate a law of motion for the state variable $S$, comprising the match probabilities, vacant transitions, and employed transitions.

The transition kernel of employed taxis is given by $\nu(v^{t+1}_e | v^t_e, M^t, m^t)$ where $v^t_e$ is the distribution of employed taxis across locations in period $t$, $M^t = \{M^t_{ij}\}$ for $i, j \in \{1, ..., L\}$ is the set of transition probabilities of each matched passenger at time $t$ and $m^t = \{m_i(u^t_i, v^t_i)\}$ for $i = \{1, ..., L\}$ is the distribution of matches in each time $t$. $\nu$ specifies for all employed taxis in location $i$ at time $t$, $v_e, i^t$, their expected distribution across locations in period $t + 1$.

25Regarding item 6, if a vacant taxi perceives some far-away location $j_0$ as best, he may either choose to move directly to that far-away location over the course of several periods, in which case he is not available to give rides until arriving in that location, or else he may move in the direction of that far away location by moving to, say, an adjacent location $j_1$ and continuing to search along the way to $j_0$. Which choice is made depends on which destination $j_0$, or $j_1$, solves equation [7].

26Regarding item 7, “dropping-off taxis arrive”: many hired taxis are in-transit for more than one period. Suppose hired taxis providing service from location $i$ to $j$ will take 3 periods to complete the trip. Then only the taxis who were 1 period away at time $t - 1$ will arrive in $j$ in period $t$. 
Likewise, the transition kernel of vacant taxis is given by $\mu(v^{t+1}_t | v^t_e, \sigma^t)$. As with $\nu$, $\mu$ specifies the expected $t+1$ spatial distribution of period $t$ vacant taxis, given the transitions generated from policies $\sigma^t = \{\sigma^t_i\}$ for $i = \{1, ..., L\}$. The combined set of transitions forms an aggregate transition kernel that defines the law-of-motion, given by $Q(S^{t+1} | S^t) = \nu(v^{t+1}_e | v^t_e, M^t, m^t) + \mu(v^{t+1}_e | v^t_e, \sigma^t)$. Explicit formulas for the state transitions are provided in Appendix A.5.

### 3.7 Equilibrium

Taxi drivers’ policy functions depend on the current state, beliefs about the policies of competitors, as well as an information set which includes the fixed price schedule, the arrival parameter of demand, and the geography of routes and distances. The current state is unobservable; taxi drivers do not see where other taxis are, but rather have beliefs about the distribution and policy functions of their competitors.\textsuperscript{27} Beliefs over competitors’ policies, conditional on the distribution of all vacant cabs, allow taxis to infer how the state will update in future periods. This implied transition of the current state as a function of the policies of competitors is denoted as $\tilde{Q}^t_i$. A driver is assumed to have already learned the arrival parameters of demand so that any observed deviation from the expected number of people hailing a taxi in a given location is taken as a draw from the known Poisson distribution. The optimization problem facing taxis is the choice of where to locate when vacant. Since the time $t$ state and transition beliefs summarize all relevant information about the competition, taxis condition only on the current-period so that an optimal location choice at time $t$ can be made using time $t-1$ information. This Markovian structure permits a definition of equilibrium as follows:

**Definition** Equilibrium is a sequence of state vectors $\{S^t_i\}$, transition beliefs $\{\tilde{Q}^t_i\}$ and policy functions $\{\sigma^t_i\}$ over each location $i = \{1, ..., L\}$, and an initial state $\{S^0_i\}_{vi}$ such that:

(a) In each location $i \in \{1, ..., L\}$, at the start of each period, matches are made according to equation 3 and are routed to new locations according to transition matrix $M^t$. The aggregate movement generates the employed taxi transition kernel $\nu(v^{t+1}_e | v^t_e, M^t, m^t)$ where $v^t_e$ is the distribution of employed taxis across locations in period $t$ and $m^t$ is the distribution of matches across locations.

(b) In each location $i \in \{1, ..., L\}$, at the end of each period, vacant taxi drivers (indexed by $a$) follow a policy function $\sigma^t_{i,a}(S^t_i, \tilde{Q}^t_i)$ that (a) solves equation 7 and (b) derives expectations under the assumption that the state transition is determined by transition kernel $\tilde{Q}^t_i$. The

\textsuperscript{27}Of course drivers will see other taxis while driving around, but since other taxis may be vacant or employed, and on- or off-duty, I assume drivers do not update beliefs based on this noisy measure of competition within a neighborhood.
aggregate movement generates the vacant taxi transition kernel \( \mu(v^{t+1}_v | \sigma^t, \tilde{S}) \) where \( v^t_v \) is the distribution of vacant taxis in period \( t \).

(c) State transitions are defined by the combined movement of vacant taxis and employed taxis, defined by \( Q(S^{t+1}|\tilde{S}^t) = \nu(v^{t+1}_e | v^t_e, M^t, m^t) + \mu(v^{t+1}_v | v^t_v, \tilde{S}^t) \).

(d) Agent’s expectations are rational, so that transition beliefs are self-fulfilling given optimizing behavior: \( \tilde{Q}^t_i = Q^t_i \) for all \( i \) and \( t \).

Existence of this equilibrium is a direct consequence of the finite horizon and finite action-space (e.g., Maskin and Tirole (2001)). The following proposition establishes that equilibrium is uniquely determined by the initial condition and a set of observed data elements. This result is leveraged to identify demand parameters, as discussed in the next section.

**Proposition 3.1.** Suppose that there exists at least one equilibrium such that if \( V^{t_0}_i > V^{t_0}_j \) for any distinct locations \( i \) and \( j \) and for some time period \( t_0 \), then \( V^t_i > V^t_j \) for all \( t \in \{t_0, \ldots, T\} \). Then given initial condition \( S_0 \) and demand functions \( \lambda_{ij}^t(P) \), \( S^* \) is uniquely determined by expected matches, \( m^t_i \), expected transitions \( \{M^t_{ij}\} \), and the costs and benefits associated with travel \( \{\tau_{ij}, \delta_{ij}, \Pi_{ij}\} \).

**Proof.** See Appendix A.7

Equilibrium delivers a distribution of vacant taxi drivers such that no one driver can systematically profit from an alternative policy: there is no spatial arbitrage-opportunity that would make search more valuable (ex-ante) in any location other than the optimum one. Vacant taxis are therefore more clustered in locations with highly-profitable rides, but these profits are offset by higher search frictions and vice-versa for low-profitability locations. Equilibrium conditions give rise to tightly correlated value functions in each location over time.\(^{28}\) See additional discussion and illustration in Appendix Section A.11.

4 Empirical Strategy

4.1 Discretizing time and space

I use trip-level New York taxi data from 6am-4pm on Weekdays from August 1st to September 30th, 2012.\(^{29}\) This time range corresponds to a typical day-shift among New York taxi drivers

---

\(^{28}\)In a setting with no time or transportation costs, equilibrium value functions would equate across locations in each period.

\(^{29}\)The August-September period of 2012 was selected because it straddles a change in regulated tariff prices, which I will use in part to estimate customers’ demand elasticities.
and it represents a period in which nearly all medallions are utilized. To prepare the data for estimation, I establish a period length of five minutes.\textsuperscript{30} I also define 39 spatial areas and link these with observed GPS points of origin and destination for each taxi trip.\textsuperscript{31} These locations represent 98\% of all taxi ride originations, and are depicted in Figure 4. For scale with respect to the time units, the average empirically observed travel time between one location to a neighboring location is 2 minutes, 45 seconds, or about one-half of a five-minute period. This suggests that the 5-minute period is reasonably well-suited to this geographic partitioning. For additional details on location selection and construction see Appendix A.2.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure4.png}
\caption{39-location Map of New York City}
\end{figure}

Each divided section of Manhattan depicts a location \(i\). Locations are created by aggregating census-tract boundaries, which broadly follow major thoroughfare divisions. The expected travel time and distance between these locations is computed separately for each origin and destination pair as the average of all observations within each \(ij\) cell.

\subsection*{4.1.1 Regions}
I further denote five regions, or distinct subsets of all 39 locations, according to Figure 5. Each region is characterized by unique mixes of geographical features and transit infrastructure. I will estimate the efficiency of search across each of these five regions. Region I is lower Manhattan, an older part of the city where streets follow irregular patterns, and where numerous bridges, tunnels

\textsuperscript{30}Note in Table 1 that standard trips average 11 minutes, with a 10\textsuperscript{th} percentile of 4 minutes.

\textsuperscript{31}This association is achieved via the point-in-polygon matching procedure outlined in Brophy (2013). Thanks to Tim Brophy for the code.
and ferries connect to nearby boroughs and New Jersey. Region II is midtown Manhattan, with fewer traffic connections away from the island, but denser centers of activity including the major transit hubs Penn Station and Grand Central Station. Region III is uptown Manhattan, where streets follow a regular grid pattern, but at the same time are longer and more spread out. Few bridges, tunnels or stations offer direct connection outside to other boroughs. Region IV is the large area encompassing Brooklyn and Queens. This area is much more residential and less congested than Manhattan. Region V consists of the two airports JFK and LaGuardia.

![Five-region Map of New York City](image)

Figure 5: Five-region Map of New York City

Each divided section of Manhattan depicts a region $r$, indicated with Roman numerals I-V. Regions are characterized by similarities in transit infrastructure, road layouts, and zoning.

4.2 Expectations and Matching Probabilities

A key object of interest in the empirical application is the expectation of matches, which is used to form matching probabilities. Taking a measure of taxis as $v_i^t$ and expected demand as $\lambda_i^t$, the expected matching function is given by Equation 3. I index the efficiency parameter as $\alpha_r$, so that each region is characterized by a unique match efficiency.

Because taxis search on different blocks, and because they cannot easily observe if others taxis are occupied or not, I assume that taxi drivers cannot observe the total number of cabs searching in each location. This assumption implies that drivers form beliefs about matching probabilities based on learned experience alone. I leverage this assumption to equate observable moments in the data (the average matches that occur in each location and time period) with the expected matches implied by the matching function, a component of drivers’ optimization problem.
4.3 Equilibrium Computation

Solving the full dynamic programming problem with large state spaces is burdened by the curse of dimensionality. Under the assumption that taxi drivers are symmetric and atomistic agents whose actions do not measurably impact the payoffs of competitors, model estimation reduces to a single-agent problem; there is only one policy function to solve for at any location and time period, which is a function of the market state. However, the total number of taxis still contributes to the computational burden because it affects number of market states over which continuation values must be computed.\textsuperscript{32}

I implement a method that takes advantage of the large number of agents in my empirical application: I take the 11,500 active taxis in the model to approximate a continuum of agents. When there is a continuum of agents facing state transitions, any probabilistic transition matrices become deterministic.\textsuperscript{33} In the taxi model, state transitions are composed of the combined transitions of vacant and employed taxis. Under the continuum approximation, these transitions are therefore also deterministic. The advantage of this approach, then, is that instead of computing policy and value functions for all possible states, I only need to compute a single, deterministic equilibrium path for the state \( \{S_t^i\} \) for \( t = \{1, \ldots, T\} \).\textsuperscript{34}

Moreover, drivers’ policies are a function of expectations that are formed through repeated observation of the profitability of search in each location over many days, which implies that the expected transitions are the appropriate basis for computing continuation values. This is similar to the experience-based equilibrium concept of Fershtman and Pakes (2012). Because drivers cannot observe competition directly (as described in Section 4.2), their strategies coincide with an oblivious strategy of Weintraub, Benkard, Jeziorski, and Van Roy (2008a), in which agents form strategies on the basis of expected future paths of the state.

One tradeoff associated with this method is that it requires an initial exogenous state \( S_0 \) in the first period. I develop an algorithm to compute equilibrium conditional on \( S_0 \). To do this, I first guess the entire state vector \( S_t^i \) for \( t = 1, \ldots, T \). I then solve for continuation values backwards from the last period \( T \) and subsequently use forward simulation to find optimal transition paths.

\textsuperscript{32}Specifically, when there are \( N \) taxis, \( L \) locations, and \( T \) time periods, continuation values and (symmetric) policy functions must be computed for every point in the state space, which numbers \( T \) times \( \frac{N+L-1}{L-1} P_N \), or: number of states \( (T, N, L) = T \cdot \frac{N+L-1}{L-1} P_N \). Storing and solving for policy functions and value functions in this setting would be infeasible.

\textsuperscript{33}This insight appears in firm dynamics literature where continuums of firms are modeled, such as Hopenhayn (1992), as well as in literature which views this approach as an approximation, as in Weintraub et al. (2008a).

\textsuperscript{34}Using traditional computational methods in a non-stationary, finite-time environment, continuation values can be computed by evaluating the probability of transitioning into each possible state in the next period multiplied by the value function computed at each state. Policy functions are also computed from every possible state to determine transitions. By implementing the continuum assumption, continuation values in every period are computed by evaluating only one point in the state space: next period’s known, deterministic state. Similarly, if the current period’s state is known, the policy function need only be computed from that state.
given these continuation values. The job of the algorithm is to find equilibrium transition paths and value functions which are mutually consistent in equilibrium. This method reduces the total number of equilibrium continuation values and policy functions that I need to search for to $T \cdot L$, vastly fewer than if the state transitions were stochastic, as in a granular version of the model. Finally, by solving equilibrium for the hour before 7am (i.e., $t = -12, -11, ..., 0$), I can mitigate the initial condition problem. Hereafter, I refer to this algorithm as the Taxi Equilibrium Algorithm (TEA). The details of TEA and tests for robustness to the initial condition are in Appendix (A.6).

### 4.4 Model Estimation and Identification

Five parameters of the model are identified directly off the data. Each is a set of time and location averages. The first four parameters are expected quantities related to time, distance, and transitions. The fifth parameter relates to fuel costs, where average values of the taxi fleet’s fuel economy are the only available data.

1. $M_{ij}^t$ is the transition probability of employed taxis in each period and location. In each period, I record the probability of transition from each origin to each destination conditional on a taxi matching with a passenger. The mean of these probabilities over each weekday of the month, computed for each origin $i$, destination $j$ and hour $t$, generates expected transition probabilities $M_{ij}^t$.

2. $\tau_{ij}$ is the travel time between each origin and destination. As above, I record the average of all travel times between each $i$ and $j$, for each hour $t$, over all weekdays of the month. I set $\min(\tau_{ii}) = 1$, so that within-location trips must take at least 1 period.

3. $\delta_{ij}$ is the distance between each origin and destination. With the trip distance variable in TLC data, I record the mean distance between each $i$ and $j$ across all weekdays of the month. Note that $\delta_{ii} > 0$ as trips can occur within a location.

4. $\omega_i^t$ is the congestion measure at each airport ($i = \{38, 39\}$). I treat the mean number of pick-ups in each airport location as direct observations of $\lambda_{38}^t$ and $\lambda_{39}^t$. However, there are sequences of large and small queues at each airport and a logistical system to allocate arriving taxis to these queues. I also compute the mean rate of pickups at each airport as the per-taxi expected wait time. With this estimate, taxis’ expected wait times upon arrival at the airport queue each period are computed as the product of all taxis at the airport $v_i^t$ (having arrived in previous periods) multiplied by this mean wait time.
Table 3: Parameter List

<table>
<thead>
<tr>
<th>Parameter</th>
<th>No. Elements</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_t^i$</td>
<td>4,212</td>
<td>Demand arrival parameters</td>
</tr>
<tr>
<td>$\alpha_r$</td>
<td>4</td>
<td>Matching efficiencies</td>
</tr>
<tr>
<td>$\sigma_\varepsilon$</td>
<td>1</td>
<td>Variance of $\varepsilon_t^i$ shocks</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>1</td>
<td>Extra value to staying-put</td>
</tr>
</tbody>
</table>

This table describes the full set of parameters to be estimated. Refer to section 3 for model details.

5. Finally the cost of fuel per mile $c$, taken as the average fuel price in New York City in 2012, divided by the average fuel economy in the New York taxi fleet, 29 mpg.\textsuperscript{35} Using $c$, I compute the cost of traveling between any origin and destination as $c_{ij} = c \cdot \delta_{ij}$. Note that $\delta_{ii} > 0$ implies $c_{ii} > 0$.

After I record the distances between each origin and destination, I can derive $\Pi_{ij}$, the expected profit associated with each possible trip. Recall from equation 2 that $\Pi_{ij} = b + \pi \delta_{ij} - c_{ij}$, where the regulated fare structure is given by the set $\{b, \pi\}$. With these parameters, and given data on expected matches, the spatial equilibrium model is configured to identify the equilibrium spatial distribution of vacant taxis.

4.4.1 Estimation and Identification of Demand and Efficiency Parameters

In the following, I show that the parameters $\theta_1 = \{\{\lambda_t^i\}, \{\alpha_r\}\}$ and $\theta_2 = \{\sigma_\varepsilon, \gamma\}$ can be identified given the available data. Table 3 outlines the set of parameters to be estimated. A summary of this process is as follows.

1. Given the estimated objects that are derived directly from the data, $\{\tau_{ij}, \delta_{ij}, \Pi_{ij}\}$, and given some value of $\sigma_\varepsilon$ and $\gamma$, all that is necessary to solve for the equilibrium state $\{v_t^i\}$ are the expected number of matches across locations and times, $m_t^i$.

2. Given the expected number of matches and the equilibrium $\{v_t^i\}$, I can identify $\lambda$ and $\alpha$.

3. I resolve step 1 over a grid of different values of $\sigma_\varepsilon$ and $\gamma$, and choose the values of each that best match a set of simulated moment conditions to data.

The next two subsections describe this process in more detail.

\textsuperscript{35}Data come from the New York City Taxi and Limousine Commission 2012 Fact Book. The taxi fleet is approximately 60% hybrid vehicles. Volatility of fuel prices is low in this period: cost-per-mile fluctuates within a range of $\$0.01$ during the sample period.
First, I discuss identification of $\theta_1$. Note that the parameters of interest, as well as the equilibrium state variables, all reflect average patterns: the expected locations of vacant taxis and the mean customers arrivals. Likewise, the key empirical moments of interest are also averages. Let $\tilde{m}^t_i$ denote the expected number matches in each location $i$ and time $t$. Because matches are observed over many days, the $\tilde{m}^t_i$ are therefore obtained in the data. As discussed above, $\tilde{m}^t_i$ together with observed data moments $\tau_{ij}$, $\delta_{ij}$, and $\Pi_{ij}$ give rise to a unique equilibrium state $S^* = \{v^*_i\}$. Because $\{\tau_{ij}, \delta_{ij}, \Pi_{ij}\}$ are estimated directly from data and internalized by the solution to the dynamic problem, I will write for simplicity $\{v^*_i(\tilde{m})\}$ where $\tilde{m} = \{m^t_i\}$.

In a first step, I identify the set of ratios $\{\frac{\lambda^i_t}{\alpha_r}\}$ using $\{v^*_i(\tilde{m})\}$ and properties of the matching function.

**Proposition 4.1.** Suppose a vector of expected matches by location and time, $\tilde{m}$, is observed. Further, suppose $\{v^*_i(\tilde{m})\}$ is unique and $v^*_i(\tilde{m}) \neq 0$ for all $i, t$. Then the ratio $\frac{\lambda^i_t}{\alpha_r}$ is identified.

**Proof.** Appendix (A.7) shows in detail that the expected number of matches $E[m(v^t_i, \lambda^i_t, \alpha_r)|v^*_i]$ is one-to-one in $\frac{\lambda^i_t}{\alpha_r}$. Thus as long as $v^*_i(\tilde{m})$ is unique and strictly positive, the expected matching function can be uniquely inverted for $\frac{\lambda^i_t}{\alpha_r}$ given $v^*_i = v^*_i(\tilde{m})$ and $m^*_i = \tilde{m}^t_i$:

$$\frac{\lambda^i_t}{\alpha_r} = -v^*_i(\tilde{m}) \cdot \ln \left(1 - \frac{\tilde{m}^t_i}{v^*_i(\tilde{m})}\right).$$

This ratio scales the actual number of arriving customers by the matching efficiency parameter. To separately identify these objects, I will exploit an additional moment in the data: the variance of matches in each $i, t$ cell, across days of the sample. Specifically, using the density function of the Poisson distribution, an analytic expression for the variance of matches can be derived. Equation (10) is derived under the assumption that any variance in $v^*_i$ across days is negligible.

$$Var(m^t_i) = (v^*_i)^2 e^{-2\frac{\lambda^t_i}{\alpha_r} \frac{1}{\alpha_r}} \left(1 - e^{\frac{\lambda^t_i}{\alpha_r}} - 1\right).$$

**Proposition 4.2.** Suppose all assumptions of Proposition 4.1 hold, and suppose a vector $\tilde{\sigma}^2_m$ of the variance of matches by time and location across days is observed. Further suppose that the $v^*_i$ are constant across days. Then $\{\lambda^t_i\}$ and $\{\alpha_r\}$ are identified.

---

36This assumption is motivated by two facts. First, the daily variance in $v^*_i$ induced by variance in prior matches is low: $E_{i,t} \left[\frac{Var(v^*_i)}{Var(m^t_i)}\right] = 0.087$. Second, on average, supply exceeds demand in equilibrium: $E_{i,t} \left[\frac{v^*_i}{m^t_i}\right] = 5.87$. Thus the contribution of the variance in taxis to the variance of all matches will be proportionally smaller. Additional details and derivations are in Appendix A.8.
Proof. \( \lambda_t^i \) and \( v^*_i \) are previously obtained as in Proposition \[.\] Denote \( \hat{\lambda} = \frac{\lambda_t^i}{\alpha_r} \) and \( \hat{\sigma}^2_{m, it} = \{Var(\hat{m}_t^i)\} \), where the variance of matches in each location and time is taken across days. Then inverting equation 11 for \( \alpha_r \) gives:

\[
\alpha_r = \frac{\hat{\lambda}_{it}}{v^*_it} \left( \ln \left( e^{2\hat{\lambda}_{it}} \left( \frac{\hat{\sigma}^2_m}{v^*_it} \right) + 1 \right) \right)^{-1}.
\]

(12)

With estimates of \( \alpha_r \) and \( \frac{\lambda_t^i}{\alpha_r} \), the demand parameters \( \{\lambda_t^i\} \) may be recovered directly.

The intuition behind this result is that, if matching were completely efficient, then the variance of matches across days would be driven entirely by the variance in customer arrivals. Since matching is not perfectly efficient, the variance of realized matches will reflect a dampening of the Poisson variance. Equation 12 shows that \( \alpha_r \) is overidentified as each region \( r \) is made up of several locations \( i \) and times-of-day \( t \). While \( \alpha \) could be treated as \( i \)- or \( t \)-specific, a choice to model frictions on the basis of broader regions will help obtain more credible results for each \( \alpha_r \), as there will be error in the measurement and estimation of the right-hand-side parameters and moments. From here, \( \alpha_r \) can be estimated via NLLS. Additional details are in Appendix A.10.

4.4.2 Estimation of \( \gamma \) and \( \sigma_\epsilon \)

The other two parameters to be estimated, denoted above as \( \theta_2 \), consist of the scale parameter \( \sigma_\epsilon \) and the same-location value \( \gamma \). To identify \( \theta_2 \), I use a Method of Simulated Moments (MSM) estimator designed to rationalize the following moments in the data: drivers’ vacant waiting times between trips, total distance traveled with passengers, the probability of the next ride being given from location \( i \) conditional on the last drop-off being in location \( i \), and average matching probabilities within sections I-IV depicted in Figure 5.

These moments relate to each parameter. The greater \( \sigma_\epsilon \), the less taxis’ behavior will correspond to the model and instead be driven by unobservables. A high-\( \sigma_\epsilon \) equilibrium will lead to drivers that are more spread out spatially, serving different customers with, on average, longer distance trips and longer trip times. Note that this scale parameter is identifiable as the profit function is observed directly and carries no parameters to be estimated. The value of \( \gamma \) will influence drivers’ tendencies to continue searching within the same location after dropping off a passenger. A higher \( \gamma \) will increase the chances of pick-up in the same location as the last drop-off.

To implement the MSM estimator in conjunction with the estimation of \( \theta_1 \), I solve for \( v^*_k(\hat{m}) \) and estimate \( \theta_1,k \) for each point \( k \) in a two-dimensional grid of \( \theta_2 \) parameter values. I then identify the point \( k^* \) and the resulting \( \theta_2,k^* = (\sigma_\epsilon^*, \gamma^*) \) which minimizes the GMM criterion function com-
paring empirical moments with their simulation counterparts. The resulting values $v_k^*(\hat{m})$ and \{${\theta}_{1,k^*}, {\theta}_{2,k^*}$\} are then recorded as the solution and estimates, respectively.$^{37}$

### 4.4.3 Estimation of Demand Elasticities

To compute market welfare, I estimate the demand elasticities in equation 1. I exploit a broader panel of data, spanning an additional month, during which there was a change in regulated prices. On September 4, 2012, the distance fee increased by $0.50 per-mile, and the JFK airport flat-fee increased by $7. Using September 2012, data, I re-estimate the model for \{${v}^*_i$\}, \{${\lambda}^*_i$\}, and \{${\alpha}_r$\}.$^{38}$ In the analysis that follows, I use price variation across months and across space to estimate demand elasticities, where the demanded quantity is the average customer arrivals in each $i$ with destination $j$ at time $t$ given prices $P_{ij}$.

For distance categories $r$, I use bins \{${s}_0, {s}_1, {s}_2, {s}_3$\} = \{0-2 mi., 2-4 mi., 4-6 mi., 6+ mi.\}, roughly corresponding to trip-distance quartiles. Price elasticities $\alpha_{1,s,a}$ are different for each ride-length category. Origin and time fixed effects captures the heterogeneity of locations: some have more or less public transit stations, bus stops, or walkability.

In this demand system, all customers of a given type $s, a$ have the same price elasticities. Identifying variation comes from two sources: differences in prices for trips from a given origin to all other destinations (within category $s, a$), and differences in prices before and after the September 2012 fare change.$^{39}$ To estimate parameters, I estimate an empirical analogue of equation 1 for each $s$ category using OLS. Since prices are fixed within a location and time period, this specification does not suffer from simultaneity bias as would traditional non-instrumented demand models.

### 5 Empirical Results

This section presents estimation results of the spatial equilibrium model. Table 4 shows a summary of the estimated parameters and the corresponding equilibrium spatial supply of taxis. Panel A shows estimation results for the per-period Poisson parameters for customer arrivals \{${\lambda}^*_i$\}, as well as point estimates for three additional parameters: the variance of unobservable shocks, $\sigma_\epsilon$, the matching efficiency parameters $\alpha_r$, and the extra value to cabs for choosing to remain in the same location to search next period. Note that the parameter $\gamma$ is computed and expressed as a fraction of value functions between period $t$ and period $t+1$. Thus the value 0.1 means there is an extra

---

$^{37}$Details on the MSM estimator are given in Appendix (A.9).

$^{38}$Since road conditions, traffic patterns, average weather patterns, etc. may change, I allow $\alpha_r$ to change by month. I assume that $\sigma_\epsilon$ and $\gamma$ are unchanged, however.

$^{39}$For example, a non-airport trip of distance 2.2 miles from location $i$ is compared with another non-airport trip of distance 2.4 miles, also from $i$. These trips have slightly different prices as well as different customer arrivals. In addition, there is a change in prices and arrival rates for all trips between August and September.
Table 4: Results Summary

This table presents a summary of estimation results and equilibrium solutions using the baseline August 2012 data. Point estimates for matching efficiency parameters $\alpha_r$ correspond to Table 5 sections I-IV, respectively.

value equal to $0.1 \cdot (V^t_i - V^{t+1}_i)$, implying the choice is worth an additional 1/10 of a period of search value over moving to any adjacent location.

While I provide several selected results below, the full set of result figures is provided in Appendix Figure A.11.

5.1 Spatial Distributions and Intra-day dynamics

Figure 6 depicts supply and demand for taxi rides across all locations, averaged across all periods of the day. It depicts how both taxi supply and passenger demand are most highly concentrated in the central part of Manhattan. We see that, in general, the number of vacant taxis is sufficient to meet demand in the absence of search frictions. The notable exception is in two central regions with very high demand, where average demand exceeds average supply. Because this view aggregates across time, I also present time-of-day results below for two selected locations to illustrate how matches are formed from supply and demand dynamics within a location.

Figure 8 shows results for two busy locations, representing the Wall Street district (Panel 1) and Brooklyn (Panel 2). Both graphs depict the equilibrium supply of vacant taxis, estimated arrival of customers looking for a taxi, the equilibrium number of matches, and the model’s fit against the observed number of matches in the data. Each series is shown from 7a-4p, in 5-minute increments. Panel 1 shows that there are periods of relative oversupply and undersupply (compared
This figure shows the average supply and demand estimates for August 2012, where the average is taken over time-of-day. The left panel shows the average customer arrivals in each location from 7a-4p, and the right panel shows the average number of vacant taxis in each location from 7a-4p.

to demand) of taxis at different times of day. Panel 2 shows an oversupply of taxis at the same moment there is an undersupply shown in Panel 1. These illustrate evidence of spatial misallocation as an equilibrium outcome: there is mismatch across locations, as across Panels 1 and 2, and there is the more granular friction captured in the estimated matching function. This is pictured as the vertical space between supply and demand at any point (i.e., min\{v_t^i, \lambda_t^i\}) compared with matches (i.e., m_t^i).\textsuperscript{40}

When taxis and customers match each period, taxis are redistributed across space. How much are the spatial patterns of taxi supply driven by customer preferences versus taxis’ search behavior? Figure 9 aggregates taxi supply across all 39 locations into five regions (identical to those in Figure 5) and depicts the net flow of matches by region (defined as the sum of drop-offs minus pick-ups in each location, summed across all locations per region) as well as net flow of driver location choices by region. It shows that in the first half of the day, many more taxis are dropping off passengers into the Midtown region compared with pickups. Conversely, vacant drivers are leaving this region much more than entering it. The situation is reversed for Uptown, as there is a disproportionate number of pick-ups compared with drop-offs. In equilibrium, vacant taxis choose this region more often precisely to make up for the supply glut. This figure shows how the equilibrium search patterns of

\textsuperscript{40}Results for more locations are available in Appendix [A.11].
This figure shows intra-day results for two example locations. Panel 1 is Location 2 in far south-east Manhattan. Panel 2 is Brooklyn, just across the East River from Location 2. Each figure depicts the equilibrium supply of taxis (red, dashed line) and the estimated arrival of passengers (blue, dot-dash line) from 7am to 4pm, compared with the expected number of matches. Matches are shown in two forms: the purple (dotted) line shows the expected matches in each minute, for each location, where the expectation is taken over days of the month. The yellow (solid) line shows a smoothed-over-time version of the former, with smoothing implemented by fitting a sixth-order polynomial. Each point depicts the over- or under-supply of taxis relative to demand in each 5-minute interval.

Vacant taxi drivers will compensate for the redistributions created by the movement of employed taxis.

The dynamic spatial equilibrium described above is solved by holding trip prices $P_{ij}$ fixed, as is true of New York City’s two-part tariff that is constant through the day. Holding prices fixed, the equilibrium allocation of taxis is obtained above. Note that in each time period $t$ and location $i$, the spatial model estimation recovers $\lambda^t_i$. The destination-specific demand parameters are then recovered via $\lambda^t_{ij} = M^t_{ij} \cdot \lambda^t_i$, the demand for taxis at price $P_{ij}$.

5.2 Demand Elasticities

Table 5 provides estimation results for the demand model in equation 1. As outlined in Section 4.4.3, I re-estimate the model for September 2012, following a change in the regulated tariff, and use this price variation to aid in identifying demand elasticities with respect to prices. An observation is an arrival-rate of customers within an origin-location, destination-location, and five-minute period during a weekday from 7a-4p. Table 5 reports price elasticities of passenger arrivals between -0.64 for short trips under 2 miles and -3.41 for longer trips of 4-6 miles.
This figure shows the net flow of matches and vacant taxis for August 2012. The top panel shows the net flow of matches, defined as the sum of matches with destinations into each region minus the sum of pick-ups headed out of each region. The bottom panel similarly shows the net movement of vacant cabs into and out of locations within each region. Positive values therefore reflect a net inflow of vacant cabs in each location due to taxis dropping off customers (in Panel I) and previously vacant cabs (in Panel II).

Table 5: Estimation results: Demand Elasticities for Non-airport Origins

Demand data come from model estimates. Dependent variable is log($\lambda_{ij}^e$). Observations are mean customer arrivals and prices for an origin-destination pair, time-of-day, month and year. For example, one observation is the estimated mean customer arrivals in location 1 (Battery Park) with destination location 25 (Central Park), from 2:00p-2:05p on any weekday in August 2012. Each row is a separate regression by trip distance with separate elasticities for airport trips (a trip with an airport as origin or destination) and non-airport trips. Specification includes hour-of-day and origin-location fixed effects. Standard errors are clustered at the level of origin-location.

5.3 Frictions and Welfare

Given demand functions estimates, I compute estimates of consumer and producer surplus by integrating under the demand function in each origin-destination-hour and summing across all locations and times. For airport trips under 2 miles, which represent 0.28% of all trips, welfare
functions are not integrable because demand elasticities fall below unity. In these cases I use a choke-pricing approach with threshold $150. Welfare is obtained by the (randomly-assigned) fraction of customers served, $E[m_{ij}^t/Q^d_{i,j,t}]$, where $Q^d_{i,j,t} = \lambda_{ij}^t$ is the expected total demand for trips in each $i,j,t$ market. Explicit formulas for these integrals are in Appendix A.12.

These calculations are illustrated in Figure 10. Panel I shows the welfare measure for demand functions with price elasticities greater than 1. The full area $A \cup B$ reflects the entire available surplus in this market at price $p^t_{ij}$. The area $B$ is the lost surplus due to not all demand being matched and the area $A$ is the realized welfare for each sub-market $(i,j,t)$. Note that $A$ is an endogenous object due to $p^t_{ij}$ and $m^t_{ij}$. Panel II depicts realized $(C)$ and lost $(D)$ welfare for demand functions with price elasticities below 1, corresponding to short airport trips (Table 5). Customers are matched randomly, so the lost surplus due to matching frictions is represented as area $D$. $\bar{p}$ is the choke price. Aggregate welfare is then computed as $\sum_{i,j,t} W_{ijt}$.

Total estimated welfare for each weekday, day-shift is shown in Table 6 which shows that there is a lower bound consumer welfare of $1.58M per day for New York taxi service during weekdays from 7a-4p in September 2012. Taxi profits in each shift are $3.09M, or $269 per driver.\textsuperscript{41} Beyond

\textsuperscript{41}The discrepancy between taxi profits and consumer welfare in part reflects the rents accruing to medallion holders as a consequence of medallion limits and high trip prices.
<table>
<thead>
<tr>
<th>Grouping Category</th>
<th>Grouping</th>
<th>Cons. Surplus ($, thousand)</th>
<th>Taxi Profits ($, thousand)</th>
<th>Matches</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggregate</td>
<td>-</td>
<td>1,576.7</td>
<td>3,094.8</td>
<td>182,309</td>
</tr>
<tr>
<td></td>
<td>7a-9a</td>
<td>226.4</td>
<td>525.9</td>
<td>29,535</td>
</tr>
<tr>
<td></td>
<td>9a-11a</td>
<td>394.8</td>
<td>766.4</td>
<td>46,249</td>
</tr>
<tr>
<td></td>
<td>11a-1p</td>
<td>373.0</td>
<td>695.9</td>
<td>42,082</td>
</tr>
<tr>
<td></td>
<td>1p-4p</td>
<td>584.0</td>
<td>1,106.5</td>
<td>64,443</td>
</tr>
<tr>
<td>Time-of-day</td>
<td>Sec. I</td>
<td>402.3</td>
<td>550.5</td>
<td>32,738</td>
</tr>
<tr>
<td></td>
<td>Sec. II</td>
<td>543.2</td>
<td>1,336.1</td>
<td>87,563</td>
</tr>
<tr>
<td></td>
<td>Sec. III</td>
<td>471.5</td>
<td>825.9</td>
<td>52,753</td>
</tr>
<tr>
<td></td>
<td>Sec. IV</td>
<td>53.5</td>
<td>79.1</td>
<td>3,614</td>
</tr>
<tr>
<td></td>
<td>Sec. V</td>
<td>106.2</td>
<td>303.0</td>
<td>5,641</td>
</tr>
<tr>
<td>Origin Region</td>
<td>0-2 mi.</td>
<td>713.0</td>
<td>1,378.6</td>
<td>114,472</td>
</tr>
<tr>
<td></td>
<td>2-4 mi.</td>
<td>566.6</td>
<td>778.9</td>
<td>43,355</td>
</tr>
<tr>
<td></td>
<td>4-6 mi.</td>
<td>90.5</td>
<td>292.6</td>
<td>11,405</td>
</tr>
<tr>
<td></td>
<td>6+ mi.</td>
<td>206.1</td>
<td>643.1</td>
<td>13,014</td>
</tr>
</tbody>
</table>

Table 6: Estimated Results: Daily, Single-Shift Welfare Measures

This table depicts welfare measures decomposed by category. Consumer welfare is summed across all $i, j, t$ in each category. Taxi profits derive from total matches multiplied by prices for each origin, destination, and time-of-day. Profits reflect daily, single-shift revenues net of fuel costs. Sections I-V refer to those in Fig. 5.

the baseline welfare estimates, Table 6 also shows that afternoon trips, central Manhattan trips, and short trips generate the largest surplus for both supply and demand.

Another measure of demand-side frictions is given by counting up the raw levels of excess demand. Excess demand is given by $\sum_{i,t} (\lambda_i^t - m_i(\lambda_i^t, v_i^t,^*, \alpha_r))$. Total daily demand is 227,343 whereas total matches is 182,309, implying that 45,034 demanded customer trips are unmet each day, or an average of 417 each period. This contrasts with the 5,405 taxis that are vacant at any one time, suggesting the presence of substantial search frictions on both sides of the market.

This can be further decomposed by the unmet demand that is attributable to within-location versus across-location frictions. Within-location frictions are due to matching function frictions. These can be measured via $\sum_{i,t} \left( \min(\lambda_i^t, v_i^t,^*) - m_i(\lambda_i^t, v_i^t,^*, \alpha_r) \right)$, where the left term captures a frictionless matching function. These frictions can potentially be mitigated with better matching technologies such as the platforms used by ridesharing firms like Uber and Lyft. Within-location frictions amount to 35,149 unmet passengers, or 77%. The remaining 10,575 lost trips are due to spatial mismatch between vacant drivers and passengers. Even with more efficient matching, these frictions exist when supply and demand are farther apart and thus not readily matched. Since equilibrium demand and supply in each location and each period are functions of the trip tariffs, the presence of such across-location frictions highlights the possibility of distortions in the way trips are priced. The next section studies how more granular pricing can be used to induce dynamically efficient spatial allocations of supply and demand.
6 Counterfactuals: Pricing for Dynamic Efficiency

Modern taxi alternatives, including ride-sharing services such as Uber and Lyft, have implemented platform features to enable market-driven pricing and a reduction in search costs. However, these technologies often fail to address two potential sources of intra-day dynamic inefficiency. The first is due to customers: since providing trips will redistribute the taxi supply, and because customers do not internalize the effects of this redistribution on future customers, the future distribution of taxis will be inefficient. This is true even in real-time pricing settings which serve to clear markets at a highly local level. The second inefficiency is due to taxi drivers’ behavior: real-time pricing incentivizes drivers to re-allocate search efforts towards higher-priced locations, but if those prices fail to capture their endogenous impact on supply later in the day (via customers’ location preferences), drivers will likewise fail to internalize the costs of such redistribution. In both cases, taxi relocation to different areas is time consuming, and so inefficient allocations will take time to mitigate.

In this section, I use the estimated model to investigate how introducing different types of price flexibility can improve the dynamic spatial allocations of supply and the equilibrium number of taxi-passenger matches within neighborhoods and across time, with the goal of finding the optimal pricing policy given a set of adjustment margins. Specifically, I permit prices to adjust along different dimensions across each of three counterfactual experiments: time, space, and trip distance. A price adjustment means introducing a change in the overall price for every trip in a category (for example, a 10% increase in prices of all trips originating in Midtown). In each experiment, I conduct thousands of distinct counterfactual simulations to search over sets of prices, with the following sequence of steps:

1. For each counterfactual \( k \in K \):
   (a) Change the set of prices faced by customers (and earned by drivers)
   (b) Recompute demand for each origin-destination-hour according to demand estimates
   (c) Recompute the matrix of customer transitions given the new distribution of demand
   (d) Recompute the profit matrix faced by taxis
   (e) Input updated objects from steps (b)-(d) into the taxi’s dynamic problem and resolve for equilibrium policies, valuations, and spatial allocations of vacancies and matches
   (f) Given the above, recompute consumer welfare by integrating the demand curve over all realized matches in each origin, destination, and time.

2. For each \( k \in K \) find the set of outcomes from step 1 in which consumer welfare is increased over the baseline measurements, subject to the constraints that (1) taxi profits are no worse
and (2) customer wait times are no worse. Choose the highest realization of consumer welfare from this set.

I compute each counterfactual by resolving equilibrium across \( K = 2401 \) points (7^4 points between 0.5 and 2.0), and linearly interpolate values for points in between. One point is a set of four prices, as outlined in the next paragraph. This produces a total of \( K = 923,521 \) points (31^4 points of increment 0.05 between 0.5 and 2.0). While I am interested in understanding how better pricing rules can induce improved spatial allocations of supply and demand, I note that each location-time market will clear not only on prices, but also on the search friction which manifests as matching probabilities. The extent of the friction experienced by taxis is fully internalized by taxi drivers, but limited demand-side data makes it difficult to estimate the elasticity of demand with respect to the friction. To account for this, I compute a measure of customer waiting time in each location and time as the probability of not being matched times the period length of 5 minutes. All pricing counterfactuals will consider price changes that only lead to outcomes in which average equilibrium waiting times and aggregate taxi profits are at least as good as in the current regime.

The three experiments in price flexibility are described below.

**Time-based Pricing:** In the first experiment, I allow prices to vary in segments of 2-3 hours across the day (the segments are 7a-9a, 9a-11a, 11a-1p, 1p-4p). This will explore whether pricing to intraday demand patterns (e.g., commute times) can improve overall efficiency.

**Location-based Pricing:** In the second experiment, I allow prices to vary in each of 4 areas, depicted as zones I-IV in Figure 5. This experiment tests the impact of pricing by pick-up location, where demand may be very different. While the areas are large enough to limit concerns about endogenous customer movement, I assume that customers near the border will not relocate.

**Distance-based Pricing:** In the third experiment, I allow prices to vary by trip distance, grouping trip distances into four bins corresponding to the demand elasticity bins: 0-2 miles, 2-4 miles, 4-6 miles, and 6+ miles. The tariffs price short rides differently from long and medium rides, inducing a shift in both the number and distribution of customers by type (destination-preference) within each location, potentially generating very different equilibrium spatial patterns of taxi supply.

The aim is to test whether an optimal implementation of price flexibility given an exogenous set of pricing categories can improve welfare by generating different equilibrium allocations of supply and demand in each time segment compared with uniform pricing. The results will highlight the value of more granular pricing in dynamic spatial markets, suggesting that new technologies can
aid in market clearing by setting flexible prices and transparently communicating the prices that are in effect.

Price changes in each experiment and each dimension of flexibility are measured as percentage increases or reductions in the existing tariff. I also compare results with the welfare induced by a frictionless matching technology. Throughout, I hold fixed the basic institutional details of the New York City taxi market: the network geography, medallion limits, and medallion rental prices.

6.1 Results

To isolate the impact of price changes on market efficiency due solely to spatial re-allocation across the day, I search for new sets of prices across each of the three experiments such that (a) total daily taxi profits are at least as high as before and (b) total daily consumer waiting time frictions are at least as low as before. Any consumer welfare gains generated from re-pricing are thus not at the expense of profits or additional search frictions.

Table 7 displays, for each set of counterfactual pricing adjustments, the highest gain to consumer welfare achievable under the above conditions, where location-specific waiting times are restricted to be no more than two times in the worst location-time combination. The table shows that gains to consumer surplus and profits are simultaneously achievable. In particular, optimal pricing by distance yields the largest equilibrium welfare gains of about 13% in consumer surplus and 4% in profits, while mean waiting times fall. This result corroborates the theoretical insights of Schmalensee (1981) and Varian (1985) which find that a necessary condition for price discrimination to enhance social welfare is that it accompanies an increase in output.42

The final row of Table 7 contrasts these results with a counterfactual in which the within-location matching technology is frictionless, given by $m(u^i_t, v^i_t) = \min(u^i_t, v^i_t)$. This model approximates that of a modern ride-share platform, where local supply and demand are guaranteed to find one another, but taxis must still choose locations to search. While I find substantial improvements to matches and waiting times, consumer welfare actually falls as taxis’ spatial distribution is shifted away from the highest-value locations around Penn Station and Grand Central Station. This is because in the baseline case these locations already resemble perfect matching, as taxis are almost fully utilized (see Figure A6). With perfect matching there are fewer vacant taxis nearby to enter these high-demand locations. Thus, matching the two sides more efficiently within-location only produces modest overall gains to trip volumes compared with policies that explicitly target changes to the spatial distribution of supply.43

42 These ideas can be directly traced to Robinson (1933), representing perhaps some of the earliest literature in industrial organization.

43 Note that this counterfactual does not simulate gains from other attributes of ride-sharing services such as the value of less waiting, the certainty of a match, app-based payments, etc.
### New Policy Price

<table>
<thead>
<tr>
<th>Description</th>
<th>Price</th>
<th>Consumer Surplus</th>
<th>Profits</th>
<th>Mean Waiting Time</th>
<th>Matches</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline, Aug. 2012</td>
<td>$p = {$2.50 + $2.00/\text{mi.}}$</td>
<td>$1.52$ Million</td>
<td>$1.83$ Million</td>
<td>$2.03$ minutes</td>
<td>$247.9$ thous.</td>
</tr>
<tr>
<td><strong>New Policy</strong></td>
<td><strong>Price</strong></td>
<td><strong>△ Consumer Surplus</strong></td>
<td><strong>△ Profits</strong></td>
<td><strong>△ Mean Waiting</strong></td>
<td><strong>△ Matches</strong></td>
</tr>
<tr>
<td><strong>Location Pricing</strong></td>
<td>Sec. 1: $1.40 \cdot p$</td>
<td>$20.0K$ (1.66%)</td>
<td>$76.7K$ (4.18%)</td>
<td>$-0.34$ seconds (&lt; 0.01%)</td>
<td>$+26,068$ (10.5%)</td>
</tr>
<tr>
<td></td>
<td>Sec. 2: $0.80 \cdot p$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sec. 3: $1.05 \cdot p$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sec. 4: $1.45 \cdot p$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Time Pricing</strong></td>
<td>7a-9a: $0.75 \cdot p$</td>
<td>$13.1K$ (1.10%)</td>
<td>$41.5K$ (2.27%)</td>
<td>$-0.27$ seconds (&lt; 0.01%)</td>
<td>$+6,909$ (2.8%)</td>
</tr>
<tr>
<td></td>
<td>9a-11a: $1.00 \cdot p$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>11a-1p: $1.10 \cdot p$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1p-4p: $1.10 \cdot p$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Distance Pricing</strong></td>
<td>0-2 mi.: $0.80 \cdot p$</td>
<td>$153.3K$ (12.7%)</td>
<td>$79.0K$ (4.31%)</td>
<td>$-0.48$ seconds (&lt; 0.01%)</td>
<td>$+92,667$ (37.4%)</td>
</tr>
<tr>
<td></td>
<td>2-4 mi.: $1.50 \cdot p$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4-6 mi.: $1.45 \cdot p$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>&gt; 6 mi.: $0.50 \cdot p$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Platform Matching (within-location)</strong></td>
<td>$p$</td>
<td>$-31.2K$ (-2.10%)</td>
<td>$12.7K$ (&lt; 0.01%)</td>
<td>$-18.0$ seconds (-14.8%)</td>
<td>$+17,563$ (7.1%)</td>
</tr>
</tbody>
</table>

**Table 7: Counterfactual Results: Optimal Flexible Prices by Policy**

This table shows, for each weekday day-shift (7a-4p), the estimated change in total consumer surplus, total profits, average waiting times and matches, across each counterfactual policy of price flexibility. Each pricing rule shown is a rule that applies four multipliers on the base prices $p$. The indicated price multipliers are chosen, in increments of 0.05 and across a support of [0.5, 2.0], to maximize consumer surplus subject to (1) weakly positive change in profits and (2) weakly negative change in average wait times. The final row simulates equilibrium under a matching function of the form $m_i^t = \min(u_i^t, v_i^t)$ at baseline prices.

Figure 11 shows the change in supply and demand under the distance pricing counterfactual, with average changes (across each period) plotted on a map of New York. It shows that under the optimal configuration, relative to the baseline outcomes, demand changes in patterns that reflect the fraction of trips demanded by distance; areas with the most demand for 2-6 mile trips see the largest declines, and areas with high fractions of short and very long trips see gains. The vacant taxi supply improves nearly everywhere, however. This is because the equilibrium redistribution taxi trips to short-distances (especially central Manhattan in which the popular short trips are now cheaper) leads to substantially increased rates of utilization.\(^{44}\) Even though very long trips are also cheaper, these trips are in relatively low demand and tend to also improve utilization rates by increasing the number of taxis who are able to return from the airport or distant locations employed with passengers instead of empty. Figure 13 shows the hourly effect of time-of-day pricing, in Panel (a), and distance-based pricing, in Panel (b). Panel (a) shows how time-of-day pricing necessarily

\(^{44}\)Although the vacant taxi supply increases in the model, it is due to high utilization rates of short trips, where taxis drop off passengers at the end of each period.

38
6.2 Discussion: Comparing Tariff Changes with Real-time Pricing

In pursuit of enhanced intra-day dynamic efficiency, the above pricing counterfactuals are configured to expected patterns of supply and demand. The optimal policies do not have any direct relation to real-time pricing, however; real-time pricing would re-adjust prices in each period to accommodate unexpected shifts in supply or demand. Note that some degree of real-time pricing, by both location and time, is implemented, for example, in Uber’s “Surge Pricing” or Lyft’s “Prime Time”, whereas implementing better dynamic pricing across hours of the day is not the goal of these pricing rules. Nevertheless, better average pricing of the form indicated in Table 7 may enable the regulator to alleviate some of the need for additional real-time pricing, as supply and demand will be more efficiently allocated across space and time in a way that is consistent with dynamic evolution of supply from one period to the next.
Figure 13: Detailed Counterfactual Results: Time and Distance Pricing

This figure shows the equilibrium changes in supply and demand under two counterfactual policies. Panel (a) shows the changes due to optimal time-of-day pricing. Panel (b) shows the changes due to optimal distance-based pricing. Changes are computed with respect to the equilibrium allocations under the August 2012 uniform pricing.

7 Conclusion

Supply and demand in the taxi market are uniquely shaped by space. Regulation influences how taxis and their customers search for one another and how often they find each other. This paper models a dynamic spatial equilibrium in the search and matching process between taxis and passengers, showing how both supply and demand can be recovered using data on intra-daily spatial matches. Using such data from New York yellow taxis, I estimate this model to recover the expected spatial and inter-temporal distribution of taxis and mean customer arrivals. By using variation in prices, I further specify and estimate demand curves for each time-of-day and across 39 locations within New York. Demand elasticities allow for welfare estimates and counterfactual analysis of demand. Finally I recompute the equilibrium taxi supply, spatial matches, search frictions and welfare outcomes under alternative price schedules and a better matching technology.

I show that consumer welfare attained in the New York market is $1.6 million per day-shift on a typical weekday, but a more flexible tariff pricing system could provide up to 12% more welfare and 25% more trips. Implementing an optimal distance-based tariff alone leads to an aggregate welfare gain of at least $153 thousand per-shift by better allocating vacant taxis with respect to current and future customer demand through the end of the day. A more sophisticated tariff might offer different prices by location, time and distance, and these results suggest additional benefits to increasing the complexity and granularity of the price, provided they are optimally configured.45 The gains due to optimal dynamic pricing are substantially better than gains due to better inter-location

45In other words, offering both additional dimensions on which to set price as well as more categories within each dimension could further boost efficiency.
matching, suggesting the importance of effective tariff pricing implementation. Such a mechanism is well within scope of the modern app-based ride-hailing platforms that permit transparent price communication to both drivers and customers.
Bibliography


A Appendix

A.1 Data Cleaning

Taxi trip and fare data are subject to some errors from usage or technology flaws. A quick analysis of GPS points reveals that some taxi trips appear to originate or conclude in highly unlikely locations (e.g., the state of Maine) or even impossible locations (e.g., the ocean). I first drop any apparently erroneous observations. Next, I drop observations outside of the locations of interest, Manhattan and the two airports. This section describes how data are cleaned and provides some related statistics.

Data Cleaning Routine

1. Begin with merged trip and fare data from August 2012 to September 2012.
2. Drop observations outside of USA boundaries.
3. Drop observations outside of the New York area.
4. Drop duplicates in terms of taxi driver ID and date-time of pickup.
5. Drop observations outside of Manhattan (bounded above by 125th st.) or either airport.
6. Drop observations that cannot be mapped to any of the 39 locations in Figure 4.

Table A1 shows the incidence of each cleaning criterion.

<table>
<thead>
<tr>
<th>Procedure</th>
<th>Criterion Applied</th>
<th>Obs. Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drop Errors</td>
<td>1. Initial Data</td>
<td>28,927,944</td>
</tr>
<tr>
<td></td>
<td>2. obs. outside USA</td>
<td>-749,623</td>
</tr>
<tr>
<td></td>
<td>3. obs. outside NYC</td>
<td>-5,298</td>
</tr>
<tr>
<td></td>
<td>4. drop duplicates</td>
<td>-57</td>
</tr>
<tr>
<td></td>
<td>5. keep manhattan + airports</td>
<td>-3,622,803</td>
</tr>
<tr>
<td></td>
<td>6. un-mapped data</td>
<td>-117,249</td>
</tr>
</tbody>
</table>

Final Data Set: 24,432,914 observations

Table A1: Data Cleaning Summary

This table summarizes the data cleaning routine for TLC data from 8/1/2012-9/30/2012.
A.2 Map Preparation

The 39 spatial locations shown in Figure 4 are created by uniting census tracts, representing 98% of all taxi ride originations. While there is some arbitrariness involved in their exact specification, the number of locations used is a compromise between tradeoffs; more locations give a richer map of spatial choice behavior, but impose greater requirements on both data and computation. Because of the sparsity of data in the other boroughs, I focus on the set of locations falling within Manhattan below 125th street, three nearby areas within Brooklyn and Queens, and the two New York City airports, Laguardia and J.F.K.

The following graphics show how raw GPS data points are converted to locations. I begin with New York census tracts, 425 of which cover the locations of interest. From these, I examine taxi activity, and group census tracts into areas with clusters of activity. Figure A1 shows the origin of each trip in a 10-percent sample of TLC data. It can be seen that trip origins are most heavily concentrated around major streets, particularly north-south and diagonal thoroughfares in the north, with more scattered origin points in lower Manhattan and Midtown Manhattan. The densest neighborhoods are clearly those in Midtown. I have grouped census tracts to form locations in a way that attempts to minimize the number of location boundaries that overlap clusters of activity, for example the clusters around a busy transit station.

A.3 Summary Statistics by Month

Table A2 decomposes the trip and fare summary statistics by month, before and after the fare change.
Figure A1: Mapping GPS points to Locations

This figure shows TLC data for a 10 percent sample of taxi trips taken in August 2012. Each dot on the map is the GPS origin of a trip.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Rate Type</th>
<th>Variable</th>
<th>Obs.</th>
<th>10%ile</th>
<th>Mean</th>
<th>90%ile</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weekdays, Day-Shift, Manhattan &amp; Boro. (Aug. 2012)</td>
<td>Standard Fares</td>
<td>Total Fare ($)</td>
<td>4,299,644</td>
<td>4.50</td>
<td>9.20</td>
<td>15.7</td>
<td>5.46</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dist. Fare ($)</td>
<td>4,299,645</td>
<td>1.04</td>
<td>4.19</td>
<td>8.96</td>
<td>4.67</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Flag Fare ($)</td>
<td>4,299,645</td>
<td>2.50</td>
<td>2.5</td>
<td>2.5</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Distance (mi.)</td>
<td>4,299,645</td>
<td>0.72</td>
<td>2.29</td>
<td>4.68</td>
<td>2.34</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Trip Time (min.)</td>
<td>4,299,645</td>
<td>4.00</td>
<td>12.25</td>
<td>22.48</td>
<td>7.95</td>
</tr>
<tr>
<td></td>
<td>JFK Fares</td>
<td>Total Fare ($)</td>
<td>85,531</td>
<td>45</td>
<td>45</td>
<td>45</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Distance (mi.)</td>
<td>85,531</td>
<td>1.87</td>
<td>15.99</td>
<td>20.88</td>
<td>6.24</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Trip Time (min.)</td>
<td>85,531</td>
<td>26.00</td>
<td>45.27</td>
<td>65.83</td>
<td>19.27</td>
</tr>
<tr>
<td>Weekdays, Day-Shift, Manhattan &amp; Boro. (Sep. 2012)</td>
<td>Standard Fares</td>
<td>Total Fare ($)</td>
<td>3,823,147</td>
<td>5.00</td>
<td>11.23</td>
<td>20.00</td>
<td>7.16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dist. Fare ($)</td>
<td>3,823,149</td>
<td>1.20</td>
<td>5.17</td>
<td>11.00</td>
<td>5.89</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Flag Fare ($)</td>
<td>3,823,149</td>
<td>2.50</td>
<td>2.5</td>
<td>2.5</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Distance (mi.)</td>
<td>3,823,149</td>
<td>0.70</td>
<td>2.27</td>
<td>4.65</td>
<td>2.38</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Trip Time (min.)</td>
<td>3,823,149</td>
<td>4.12</td>
<td>13.30</td>
<td>25.0</td>
<td>9.01</td>
</tr>
<tr>
<td></td>
<td>JFK Fares</td>
<td>Total Fare ($)</td>
<td>85,692</td>
<td>52.00</td>
<td>51.56</td>
<td>52.00</td>
<td>3.60</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Distance (mi.)</td>
<td>85,692</td>
<td>3.42</td>
<td>16.29</td>
<td>20.93</td>
<td>5.94</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Trip Time (min.)</td>
<td>85,692</td>
<td>26.82</td>
<td>46.04</td>
<td>68.42</td>
<td>19.05</td>
</tr>
</tbody>
</table>

Table A2: Taxi Trip and Fare Summary Statistics by Month

Taxi trip and fare data come from New York Taxi and Limousine Commission (TLC). This table provides statistics related to individual taxi trips taken in New York City in the months of August 2012 and September 2012 for two fare types. The first is the standard metered fare (TLC rate code 1), in which standard fares apply. The second is a trip to or from JFK airport (TLC rate code 2). Total Fare and Distance data are reported for each ride. While not reported directly or separated from waiting costs, I predict distance and flag fares using the prevailing fare structure on the day of travel and the distance travelled.
A.4 Medallion Counts

Figure A2 shows the unique number of medallions observed each day of August and September 2012 in the TLC data during weekdays during the day shift. The mean across all days is 11,911.88. It should be noted however, that about 2% of trips occur outside of the 39 locations defined in this paper during this period. This implies that approximately 11,673 medallions are active within the locations, with some additional diminishment in reality due to breaks, refueling, etc. The second point of this figure is that the medallion counts seem fairly stable between price changes, lending support for the assumption that this overall level remains constant. The drop on September 3rd seems to reflect the extra servicing of metering equipment just prior to the tariff change on September 4th.

Figure A2: Medallions per day, Aug-Sep 2012

This figure depicts the unique number of medallions observed each day of August and September 2012 in the TLC data during weekdays during the day-shift.

A.5 Details on State Transitions

The combined set of transitions forms an aggregate transition kernel that defines the law-of-motion, given by

\[ Q(S_{t+1}^t | S_t^t) = \nu(v_{e, t+1}^t | v_{e, t}^t, M_t^t, m_t^t) + \mu(v_{v, t+1}^t | v_{v, t}^t, \sigma_t^t). \]

Let the following objects be defined:

- \( v_{e}^t \) be the \((L + K) \times 1\) vector of employed cabs at the start of period \( t \), where \( L \) is the total number of search locations and \( K \) is the total number of positions between locations (e.g., if a route takes 4 periods to travel, there is a pickup-location \( i \), 2 in-between positions, and a drop-off location \( j \)).
- \( m_t^t \) is the \((L + K) \times 1\) vector of matches in period \( t \), where the first \( L \) entries are the matches in each location and the next \( K \) entries are zeros (as no matches occur while cabs are
employed and in-transit. $M_t^e$ be the $(L + K) \times (L + K)$ vector of one-period transition probabilities of customers from all locations $\{1, \ldots, L\}$ and all in-between positions $\{1, \ldots, K\}$. The number of in-between positions is based on the mean number of periods it takes to travel from any locations $i$ to $j$, rounded to the nearest period (e.g., an average 16-minute trip would be considered 3.2 periods, and then rounded to be 3 periods, with a single in-between position). $m^{t-\tau_{ji}}$ describes how many drop-offs will occur in period $t$, which is the number of matches made in each pick-up location in $\tau_{ji}$ prior periods, and transition matrix $M_{e}^{t-\tau_{ji}}$ re-distributes those earlier matches to locations at time $t$.

Given these objects, we can write the state transitions of employed cabs as follows, reflecting the transitions of new matches and already-employed taxis at time $t$, minus the time $t$ drop-offs:

$$v_{e}^{t+1} = ((v_{e}^{t} + m^{t}) \times M_{e}^{t}) - (m^{t-\tau_{ji}} \times M_{e}^{t-\tau_{ji}}).$$  \hspace{1cm} (13)

Next, I define the state transitions of vacant taxis. Let $v_{v}^{t}$ be the $(L + K) \times 1$ vector of vacant taxis in all search locations and in-between locations $\{1, \ldots, (L + K)\}$. Note that there may be taxis in the in-between locations. For example, driving vacant to the airport may take more than one period. Let $v_{v}^{t}$ be the $(L + K) \times (L + K)$ vector of one-period transition probabilities of vacant taxis from all locations $\{1, \ldots, L\}$ and all in-between positions $\{1, \ldots, K\}$. Then the state transitions of vacant cabs is given by the vector of vacant cabs at the start of period $t$ minus the period $t$ matches, multiplied by the policy functions in each period:

$$v_{v}^{t+1} = (v_{v}^{t} - m^{t}) \times \sigma^{t}. \hspace{1cm} (14)$$

Summing these two transition formulas defines the state transitions from $t$ to $t+1$.

A.6 Taxi Equilibrium Algorithm Details

The algorithm that I implement takes as inputs all model primitives, parameters, and a time zero state, and returns the equilibrium state and policy functions for each location and each time period. Equilibrium states constitute a $L \times T$ matrix (i.e., how many taxis are in each location in each period), and equilibrium policy functions constitute a $L \times L \times T$ matrix (i.e., the probability of vacant taxi transition from any location $i \in \{1, \ldots, L\}$ to any location $j \in \{1, \ldots, L\}$ in each period). Broadly, the algorithm uses backwards iteration to solve for continuation values and forward simulation to generate transition paths. The algorithm moves in an alternating, asymmetric backwards and forwards sequence through the current time step $t \in \{1, \ldots, T\}$, where backwards moves update continuation values and forwards moves update transition paths. The algorithm terminates when all transition paths and continuation values are self-fulling and consistent with equilibrium. Below I provide an outline of the taxi equilibrium algorithm.
**Algorithm 1** Taxi Equilibrium Algorithm

1: Load empirical matches \( \{ \hat{m}_{ij}^t \} \) and \( \{ \hat{m}_{i}^t \} = \{ \sum_j \hat{m}_{ij}^t \} \)
2: Fix parameters \( \sigma, \gamma \) (solved outside of algorithm)
3: Set counter \( k = 0 \)
4: Guess \( S_0^T \) and compute \( V^T(S_0^T, \lambda^T) \)

5: for \( t = T - 1 \) to 1 do  \( \triangleright \) Backwards Iteration

6: Guess \( S_t^0 \) and compute \( V^t(S_t^0, \lambda^t) \)

7: for \( t = 1 \) to \( T - 1 \) do  \( \triangleright \) Fwd. Iteration to \( T \) for each step back

8: Derive choice-specific value functions \( W_t^i(j, S^t) \) for all \( t, i, j \).

9: Find policy fcts. \( \sigma_k^t(W_{k+1}^t) \) to determine vacant taxi transitions

10: \( \sigma_0^t \) and \( m_{ij}^t \) imply transition to \( S^{t+1} \)

11: Update next period state \( S_{k+1}^{t+1} \leftarrow \hat{S}^{t+1} \)

12: Update next period continuation values as \( V^{t+1}(S_{k+1}^{t+1}, \hat{m}^t) \)

13: \( k \leftarrow k + 1 \)

14: end for

15: end for

16: repeat

17: Iterate on steps 6 to 15

18: until \( |V_k^t - V_{k-1}^t| \leq \epsilon \) \( \forall t \)

The Taxi Equilibrium Algorithm begins with an initial guess of the market state \( S_0 \) (i.e., the number of all vacant taxis across locations and time-of-day).\(^{46}\) With \( S_0 \) as well as observations of the empirical distributions of taxi-passenger matches, \( \hat{m} \), I compute value functions \( V_t^i(S_t^0; \hat{m}) \) for each \( i \) and \( t \) via backwards induction, beginning at period \( T \) and stepping backwards to period 1, updating continuation values in each step. Next, using the value functions, I compute choice-specific value functions and optimal policies as in equation 9. Next, I use the computed policy functions and, starting at time \( t = 1 \) at \( S_0^t \), I forward simulate the optimal transition paths and update the initial state for \( t = 2, \ldots, T \), resulting in a new guess of the state, \( S_1 \). With \( S_1 \), I again combine the same observations \( \hat{m} \) to update value functions \( V_t^i(S_1; \hat{m}) \). This process repeats until value and policy functions converge.

### A.6.1 Initial conditions

Recall that \( S_0^t \) is a state vector of the number of vacant taxis in each location at time \( t \). The initial guess of the state in each period, \( S_0^0 \), is assigned by allocating the exogenous total number of taxis according to the empirical distribution of matches.\(^{47}\) As the algorithm runs, each vector \( S_0^t \) for

---

\(^{46}\) Note that the algorithm will take as given the parameters \( \gamma \) and \( \sigma \). These are solved for in a second stage, as an outer loop to the algorithm.

\(^{47}\) Recall the total number of taxis equals 11,500, as discussed in Section 4.1.
<table>
<thead>
<tr>
<th>Initial Condition</th>
<th>$\Delta v^t_i$ (mean)</th>
<th>$% \Delta v^t_i$ (mean)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Uniform</td>
<td>2.10</td>
<td>0.0175</td>
</tr>
<tr>
<td>Edge</td>
<td>4.52</td>
<td>0.0336</td>
</tr>
</tbody>
</table>

Table A3: Alternative Initial Conditions

This table shows the change in taxis’ spatial equilibrium distribution given changes in initial conditions. Baseline is the initial condition used throughout the paper, as described above. Uniform imposes an initial distribution that is uniform across all locations at 6am. Edge imposes an initial distribution that uniformly puts all vacant taxis across edge locations: all peripheral locations with adjacent access to the outer boroughs and New Jersey.

$t \geq 2$ is updated as $t - 1$ transitions are computed given the $t - 1$ initial state and value functions for $t, t + 1, ..., T$. Only one term, $S_0$ remains exogenously chosen.

To mitigate any issues related to this remaining first-period exogenous initial state, I define $t = 1$ as 6:00am. In this period, the assumption that all available cabs are actively searching or with customers is less credible. Nevertheless, by starting the equilibrium algorithm at 6:00am, a wide range of initial conditions quickly wash out within 5-6 periods. This is verified by setting alternative initial conditions and comparing equilibrium levels of taxi supply across locations. By the intended starting time for estimation, of 7:00am, then, the spatial distribution of taxis is in equilibrium. Table A3 shows the impact of initial conditions on the equilibrium supply of taxis under increasingly heterogeneous starting points. The baseline case, as described above, is compared with (1) a uniform initial distribution and (2) a distribution in which all initial vacant cabs are distributed at edge locations: those locations adjacent to the boundaries of the map. The latter edge distribution is meant to simulate the case in which taxis start the day by driving from garages where they are stored. The locations of these garages are not available in my data, so this condition serves as an extreme case in which all taxis are stored in outer boroughs.

### A.7 Proofs

#### Proof of Proposition 3.1

This proof demonstrates that, given some initial condition $S_0$ and demand functions of the form $\lambda_{ij}(P_{ij})$, the equilibrium distribution of taxis $S^*$ is uniquely determined by the expected matches $m_{ij}^t$, customer transitions $M_{ij}^t$ and travel cost features $\Pi_{ij}, \delta_{ij},$ and $\tau_{ij}$. The logic of the proof is

---

48 Recall from Figure ?? and earlier discussion that the data do not allow for distinguishing whether fewer matches in the morning are due to low supply or demand; and thus it is impossible to say how many cabs are actually on the road at any point.

49 Boundary locations are all peripheral locations with adjacent access to the outer boroughs and New Jersey. This includes all locations in Manhattan with bridges and those bordering 125th street, all Brooklyn and Queens locations, and each Airport.
to first posit the existence of two distinct equilibria $S^{\ast a}$ and $S^{\ast b}$, and then demonstrate that the optimization behavior of drivers (represented by the policies of the form $\sigma_{ij}^{\ast}(S^{\ast})$) is not consistent with this conjecture, as any two distinct equilibria imply value function differences at some fixed state, which is incompatible with driver optimization from the initial state.

**Proof.** Suppose there are two distinct equilibrium states, $S^{\ast a}$ and $S^{\ast b}$, such that $v_{i,t}^{\ast a} > v_{i,t}^{\ast b}$ for some $i$ in $\mathcal{L}$ and $t$ in $\{1, \ldots, T\}$. This difference implies differences in value functions $V_{i,t}^{\ast a}$ and $V_{i,t}^{\ast b}$ when one of two conditions hold: either $v_{i,t}^{\ast a} > v_{i,t'}^{\ast b}$ for all periods $t' > t$ and all locations $i$ in $\mathcal{L}$, or $t = T$, after which continuation values are everywhere set to zero. Since $t = T$ is always reached, this condition can be met for some $i$ and $t$. Then let $\hat{t}$ be a period in which this condition holds. Since $\frac{\partial p_i(u_{i,t}, v_{i,t})}{\partial v_{i,t}} > 0$, it must be the case that $V_{i,t}^{\ast a} < V_{i,t}^{\ast b}$, since period payoffs are different but continuation values are otherwise the same. Consider two distinct cases:

**CASE 1:** Relative match probabilities are the same across equilibrium $a$ and $b$ in period $\hat{t}$. i.e.,

$$\left\{ \frac{\mathbb{E}[p_i(u_{i,t}^{\hat{t}}, v_{i,t}^{\hat{a},\hat{t}}|\lambda_i^{\hat{t}})]}{\mathbb{E}[p_j(u_{j,t}^{\hat{t}}, v_{j,t}^{\hat{a},\hat{t}}|\lambda_j^{\hat{t}})]} \right\} \forall i,j,$$

Since there is an $i$ for which $v_{i,t}^{\ast a} > v_{i,t}^{\ast b}$, and since value functions in period $\hat{t}$ are strictly increasing in match probabilities, CASE 1 implies that value functions everywhere are lower in equilibrium $a$ than in $b$: $V_{i,t}^{\ast a} < V_{i,t}^{\ast b}$ $\forall j \in \mathcal{L}$. Since match probabilities are strictly decreasing in the taxi input $v_{i,t}$, CASE 1 also implies that $v_{i,t}^{\ast a} > v_{i,t}^{\ast b}$ for all $i$. Recall that the total number of cabs in the market is fixed in each period (e.g., so that $\sum_j (v_{j,t}^{\text{vacant}} + v_{j,t}^{\text{employed}}) = 11,500$). Then it must be the case that there are more employed taxis in transit under equilibrium $b$, so that $\sum_k v_{e,k}^{\ast b} > \sum_k v_{e,k}^{\ast a}$.

Since $\frac{\partial \sigma_{ij}^{\ast -1,a}}{\partial v_{i,t}^{\ast a}} > 0 \forall j \in \mathcal{L}$, more in-transit cabs in equilibrium $b$ implies that $\sigma_{ij}^{\hat{t}-s,a} > \sigma_{ij}^{\hat{t}-s,b}$ for some $j$ s.t. $\tau_{ij} > 1$ and some $s \geq 1$. However, since $V_{j,t}^{\ast a} < V_{j,t}^{\ast b}$ $\forall j \in \mathcal{L}$, we have $\sigma_{ij}^{\hat{t}-s,a} < \sigma_{ij}^{\hat{t}-s,b}$. Thus $V_{j,t}^{\ast a} < V_{j,t}^{\ast b}$ cannot be supported in two equilibria for any $j$. Since $\hat{t}$ is generically chosen, this implies a contradiction.

**CASE 2:** Relative match probabilities are *not* the same across equilibrium $a$ and $b$.

The same logic above now implies that there is a persistent difference in value functions between $V_{i,t}^{\ast a}$ and $V_{j,t}^{\ast a}$ for some $j$. This difference must be larger than the fuel cost $c_{ij}$, or else equilibrium policies in $\hat{t} - 1$ would have ensured that some taxis reallocate until the differences equal the fuel cost.\(^50\) If CASE 2 is true, then label $i$ and $j$ such that $V_{i,t}^{\ast a} < V_{j,t}^{\ast a}$. Thus, some taxis in $i$ would prefer to search in $j$ at time $\hat{t}$ but could not have feasibly relocated to $j$ in the prior period. This could happen for example if passenger demand took all available cabs out of locations adjacent to $j$.

\(^50\)Note that this could happen due to feasibility constraints on the equilibrium transition path of the state variable, which prevents vacant taxis from relocating instantly to higher-valued locations.
or else because of policies of vacant cabs prior to \( \hat{t} \) took all available cabs out of locations adjacent to \( j \). This implies that for some earlier period \( \tilde{t} < \hat{t} \), we have \( V_{\tilde{t}}^i > V_{\hat{t}}^j \). Iterating back, this therefore must also be true for some even earlier \( t < \tilde{t} \) and so on back to \( t = 0 \). At \( t = 0 \), we know that \( \{ V_{1,b} \} \) is feasible by assumption, which implies there is no set of optimal policies which lead to \( \{ V_{1,a} \} \). This implies a contradiction.

\[ \square \]

Note that the condition stated in the claim above refers to the particular equilibrium which is numerically recovered. The stated pattern in value functions can be observed in Appendix Figure A8.

**Proof of Proposition 4.1**

**Proof.** I first prove that the expected number of matches \( E[m(v_t^i, \lambda_t^i, \alpha_r)|v_t^i(\tilde{m})] \) is one-to-one in \( \lambda_t^i/\alpha_r \). In the following, I omit subscripts \( i, t, \) and \( r \) for clarity: Let \( m \) denote \( m_t^i \), \( v \) denote \( v_t^i(\tilde{m}) \), \( \lambda \) denote \( \lambda_t^i \) and \( \alpha \) denote \( \alpha_r \). Let \( f_\lambda(k) \) be the probability mass function of the Poisson distribution with parameter \( k \). We can write \( E[m|v, \lambda] \) as follows:

\[
E[m|v, \lambda] = v \sum_{k=0}^{\infty} \left( 1 - \left( 1 - \frac{1}{\alpha v} \right)^k \right) f_\lambda(k).
\]

Let \( \rho = \left( 1 - \frac{1}{\alpha v} \right) \). Then with some algebra we can write:

\[
E[m|v, \lambda] = v - v \sum_{k=0}^{\infty} \rho^k f_\lambda(k)
= v - v \sum_{k=0}^{\infty} \rho^k \lambda^k e^{-\lambda} / k!
= v - v \cdot e^{-\lambda} \sum_{k=0}^{\infty} (\rho \lambda)^k e^{-\rho \lambda} / k!
= v - v \cdot e^{-\lambda(1-\rho)}
= v(1 - e^{-\frac{\lambda}{\alpha v}}).
\]

It is straightforward to show from this step that \( E[m(v, \lambda)|v] \) is strictly increasing in \( \lambda/\alpha \) given \( v > 0 \). Since a strictly increasing function is one-to-one, \( \{m(\cdot, \cdot)|v\} \leftrightarrow \lambda/\alpha \) is one-to-one.

55
Recall that \( \hat{m} \) is observed. Including the model subscripts, it therefore follows that as long as \( v^*(\hat{m}) \) is unique and strictly positive, \( E[m(v^*_i, \lambda^*_t, \alpha^*_r)|v^*_t] \) can be uniquely inverted for \( \frac{\lambda^*_t}{\alpha^*_r} \) given \( v_{it} = v^*_i(\hat{m}) \) and \( m_{it} = \hat{m}_{it} \). \( \square \)

### A.8 Relative Variances of \( v^*_i \) and \( u^*_i \)

Equation 11 is derived under the assumption that the variance of \( v^*_i \) is negligible as a determinant of \( \alpha \). This subsection provides additional support for this assumption by highlighting two sets of equilibrium moments that are independent of \( \alpha \). First, across days, the variance in the equilibrium level of taxis in this model is solely driven by variance in the arrival of customers each day, which induces variance in taxi-customer matches as well as variance in the number of vacant taxis left to search. Below I provide evidence that this variance is small relative to that of demand. Second, I will show that generally large levels of supply relative to demand are present in equilibrium, implying that any variance in supply will have a comparatively small impact on the variance of matches.

The level of taxis in each location and time in each day is correlated with the levels in all other locations and previous times of that same day, and likewise correlated with the draws of \( u^*_i \) in all previous times of that same day. An exact result for taxi variance would require a large simulation and lack an interpretable analytical solution. Instead I will provide evidence that the variance in \( v^*_i \) is small relative to that of \( u^*_i \) by computing an upper-bound on the variance of incoming taxis in each location attributable to past draws of \( u^*_i \). This is done by fixing \( \text{Var}(m^*_i) = \text{Var}(u^*_i) \), an overestimate of match variance, and then computing the variance of incoming matches to each location and time, \( \text{Var}(v^*_{i, \text{matches}}) \) plus the variance of incoming searchers, \( \text{Var}(v^*_{i, \text{search}}) \).

Note that \( \text{Var}(u^*_{ij}) = (M^*_{ij})^2 \cdot \text{Var}(u^*_i) \) and \( \text{Var}(u^*_i) = \lambda^*_i. \) If matches have \( \text{Var}(u^*_{ij}) \), then

\[
\text{Var}(v^*_{i, \text{matches}}) = \sum_j \text{Var}(u^*_{ij}) = \sum_j (M^*_{ij} - \tau^*_{ij})^2 \lambda^*_{ij},
\]

and

\[
\text{Var}(v^*_{i, \text{search}}) = \sum_j (\sigma^*_{ij})^2 \cdot \text{Var}(1 - u^*_{ij}) = \sum_j (\sigma^*_{ij} - \tau^*_{ij})^2 \lambda^*_{ij}.
\]

Thus,

\[
\text{Var}(v^*_i) = \text{Var}(v^*_{i, \text{matches}}) + \text{Var}(v^*_{i, \text{search}}) - 2 \sum_j (M^*_{ij} - \tau^*_{ij})^2 (\sigma^*_{ij} - \tau^*_{ij})^2
\]

is an upper-bound measure of taxi variance attributable to demand. I compute this number and divide by the variance of demand \( \lambda^*_i \) in all non-airport locations from 7a-4p. Figure A3 Panel 1

56
Figure A3: Equilibrium Variance and Level Ratios of Supply and Demand

shows the mean variances for all locations within each region. The mean taxi variance as a fraction of demand variance is 0.0308, or 3.1%.

Figure A3, Panel 2 shows the equilibrium ratio of supply to demand across regions. While there are many reversals (see detailed estimates in Figure A6), the average ratio of taxis-to-customers across regions is between 3-10 throughout most of the day. Since the more constrained input to the matching function will exert a proportionally greater influence on matches (see, e.g., Figure 3), this further diminishes the role of supply variance on the realized variance of matches each day.

Together these facts suggest that the variance of taxis can be regarded as negligible with respect to their impact on the overall variance of matches used to identify the efficiency parameter $\alpha$.

A.9 Estimation and Simulated Moments

I identify $\sigma_\varepsilon$ and $\gamma$ by simulating individual taxi trip data and comparing simulation moments with their empirical counterparts. The moments are as follows: (1) Mean total vacancy times per taxi (2) Mean total distance travelled with passengers, (3) the probability that a driver’s next match is in the same location as his most recent drop-off and (4)-(7) the average probability of matching with a customer in each of sections I-IV, where sections are defined in Figure 5. Note that these moments will depend at least in part on these parameters; $\sigma_\varepsilon$ reflects how much of a taxi driver’s location choice depends on observable features within the model. A high value of $\sigma_\varepsilon$ should lead to behavior that appears random from the perspective of the model, including longer vacancy periods, whereas a low value implies that the model is capturing incentives well, and thus behavior should conform to the model’s valuation of locations. $\gamma$ reflects the value of saved time when drivers choose to stay in the same location next period as they ended up in last period. A low gamma means that drivers are, all else equal, indifferent between searching in the current location and adjacent ones. Thus, a lower probability of staying-put is expected. A high gamma likewise implies a higher probability of staying-put. For these reasons, both parameters will also impact match probabilities
<table>
<thead>
<tr>
<th>Moment</th>
<th>Data Average</th>
<th>Simulation Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Vacant Wait Time (min.)</td>
<td>155.21</td>
<td>184.95</td>
</tr>
<tr>
<td>Total Employed Distance Travelled (mi.)</td>
<td>38.37</td>
<td>35.74</td>
</tr>
<tr>
<td>( \Pr(\text{pickup in } i \mid \text{drop-off in } i) )</td>
<td>.3714</td>
<td>.3809</td>
</tr>
<tr>
<td>Match Probability (Section 1)</td>
<td>.4205</td>
<td>.3809</td>
</tr>
<tr>
<td>Match Probability (Section 2)</td>
<td>.4912</td>
<td>.4543</td>
</tr>
<tr>
<td>Match Probability (Section 3)</td>
<td>.3842</td>
<td>.2214</td>
</tr>
<tr>
<td>Match Probability (Section 4)</td>
<td>.2477</td>
<td>.4735</td>
</tr>
</tbody>
</table>

Table A4: Data and Simulation Moments

This table summarizes the data and simulation moments used to estimate remaining model parameters.

Across space. Empirical moments of these probabilities are recorded as the percentile associated with the distribution of waiting times in each location being equal to five minutes, the length of one period.\(^{51}\) Table A4 displays each simulation moment compared with its observed value.

A.10 Details on Estimating \( \alpha_r \)

Given parameter values \( \sigma \) and \( \gamma \), I use matches date \( \{ \tilde{m} \} \) and the TEA procedure to numerically solve for the equilibrium state, \( S = \{ v_{i,t}^* (\tilde{m}) \} \). Proposition 4.4 shows that, given a level of taxis \( v_i^t \), the matching function can be inverted to solve for \( \lambda_{i,t} \) in location \( i \) (within region \( r \)) and time \( t \).

Next, I use an analytic expression of the variance of matches, given by equation 11. This function depends on both \( \alpha_r \) as well as the ratio \( \frac{\lambda_{i,t}}{\alpha_r} \). From here I set up the following estimator:

\[
\alpha_r = \arg \min_{\alpha} \sum_{i \in R_r,t} \left( \text{Var}_d(\tilde{m}_{i,d}) - (v_{i,t}^*)^2 e^{-2 \frac{\lambda_{i,t}}{\alpha_r} \frac{1}{v_{i,t}}} \left( e^{\frac{\lambda_{i,t}}{\alpha_r} \frac{1}{v_{i,t}}} - 1 \right) \right),
\]

(15)

where \( R_r \) denotes the set of locations within region \( r \), and \( \tilde{m}_{i,d} \) refers to the number of observed taxi-passenger matches that take place in location \( i \), time \( t \), and day-of-month \( d \). The variance is then taken with respect to all observations within the weekdays in a given month, across days of the month.

A.11 Detailed Estimation Results

Figure A4 shows aggregate supply and demand results, summing all 39 locations into the five regions corresponding to Figure 5. The results above demonstrate that the while taxi supply maintains some coverage across all locations throughout the day, there are intra-day trends in

\(^{51}\)For example, if half of all taxis in a particular location matched with a passenger within 2.5 minutes, then the probability of matching within a period would equal 50%.
(a) Total Customer Arrivals ($\lambda$)  
(b) Total Vacant Taxis ($S^*$)

Figure A4: Equilibrium Vacant Taxis: Weekdays 7a-4p, 9/2012 (Five Region Aggregates)

This figure depicts the equilibrium spatial distribution of taxis and mean arrival of customers across the Five Regions shown in Figure 5. Results across all 39 locations are summed to these five areas. Results are depicted for the weekday taxi drivers’ day shift, from 7a-4p in September 2012.

spatial availability and demand. Spatial mismatch is evident, as the relative proportions of supply and demand are not the same across each region.

Figures A6 and A7 show detailed results of supply and demand in all locations. Note that location numbers 1-34 roughly track from South to North in Manhattan, locations 35-37 track South to North from Brooklyn to Queens, location 38 is LaGuardia airport and location 39 is JFK airport. We see that most locations have a surplus of taxis except for a few areas of very high demand. Lower Manhattan, parts of midtown Manhattan and far North-east Manhattan all demonstrate particularly large constraints in the ratio of vacant taxis to demand. All locations demonstrate some search frictions on both sides of the market, but we see here that the impact is felt more on the taxi side.

Figure A8 shows the evolution of Value functions by time of day. Each series is the value for a single location. The high correlation between each value function reflects the equilibrium result that drivers’ policy functions ensure that there is no spatial arbitrage possible. The remaining differences between each location’s value is due to the transportation cost that prevents perfect cross-location arbitrage. As the day reaches its 4pm end, the value of search in each location systematically drops to zero.

A.12 Welfare Calculation

Consumer welfare is computed by integrating under the the estimated CES demand curves in each origin, destination, time pair (i.e., each $i, j, t$). The integral can be computed analytically, and is
as follows:

\[
W_{ijt}(m_{ij}^t, \hat{\lambda}_{ij}^t, p_{ij}, \beta) = \frac{m_{ij}^t(\hat{\lambda}_{ij}^t, v_i^t(\lambda))}{\hat{\lambda}_{ij}^t(p_{ij})} \cdot \frac{\alpha_{1,r,a}}{\alpha_{1,r,a} + 1} \cdot \frac{1}{\alpha_{1,r,a} + 1 - \hat{\lambda}_{ijt}(p_{ij})} \cdot \hat{\lambda}_{ijt}(p_{ij}) \cdot \beta_{ijt}(p_{ij}) \cdot \beta_{ijt}(p_{ij})
\]

(16)

where \(\alpha_{0,r,a}, \alpha_{1,r,a}, \delta_{tr},\) and \(\xi_{ir}\) are the estimated parameters of the demand system, and where
This figure depicts the equilibrium value functions for all 39 locations, by time of day, estimated from August 2012 data. Each line depicts a separate location. The highest-valued function is that of LGA airport and the least-valued function is that of JFK airport. All other locations’ values fall in-between.

\( \hat{\lambda}_{ijt} \) is the predicted level of demand (the mean number of customer arrivals given price \( p_{ij} \)), and \( v^t_i(\lambda) \) is the equilibrium number of taxis in each location, a function of the entire distribution of demand across locations and time.

Taxi profits are computed as follows:

\[
W^\text{taxi}_{ijt}(m^t_{ij}, \hat{\lambda}^t_{ij}, p_{ij}, c_{ij}) = \frac{m^t_{ij}(\hat{\lambda}^t_{ij}, v^t_i(\lambda))}{\hat{\lambda}^t_{ij}(p_{ij})} \left( \frac{\hat{\lambda}^t_{ij}(p_{ij}) \cdot (p_{ij} - c_{ij})}{\text{frac. successful matches}} \right),
\]

where \( c_{ij} \) is the fuel cost for a trip from \( i \) to \( j \).