

Spatial Equilibrium, Search Frictions and Efficient Regulation in the Taxi Industry

Nicholas Buchholz *

December 15, 2017

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 search and matching between taxis and passengers under regulation. 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. My estimates show that the New York market, faced with price regulation and search frictions, achieves about \$4.2 million in daily welfare, about a third of which is realized as consumer surplus, split across 223 thousand trips. Counterfactual analysis reveals that allowing tariffs to vary by time, location or distance can enhance allocative efficiency given the presence of search frictions, offering daily net surplus gains of at least \$194 thousand and 31 thousand additional daily taxi-passenger matches.

Key Words: dynamic games, spatial equilibrium, search frictions, regulation, taxicabs

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. Special thanks to Allan Collard-Wexler, Eugenio Miravete and Stephen Ryan for their guidance and support. I also benefitted from discussions with Hassan Afrouzi, Fernando Alvarez, John Asker, Pat Bayer, Austin Bean, Lanier Benkard, Laura Doval, Neal Ghosh, Jean-François Houde, Andrew Glover, Ariel Pakes, Rob Shimer, Emily Weisburst, Daniel Xu, Haiqing Xu, anonymous referees and numerous seminar participants.

1 Introduction

How are supply and demand organized across space and time? Spatial markets are ubiquitous, encompassing labor markets, international trade, real estate, local retail and dining, and countless others. Understanding how trading partners locate relative to one another in these settings is critical, as the ability of markets to clear necessitates that supply and demand have to meet. This process is not always trivial; a long theory literature in search and matching tells us that in the presence of search frictions, markets will fail to fully clear, simultaneously leaving unmatched supply and demand.¹ Where buyers and sellers go following the search process, whether successful or not, impacts how the market evolves over time. Location incentives also depend on the prevailing prices across locations, but search frictions induce an important externality; market participants do not account for the impact of their pricing or search behavior on the frictions faced by other traders. How large are these misallocated populations and to what extent can pricing regulations serve to mitigate the search friction? In this paper, I study dynamic spatial search and equilibrium and the ability of prices to clear spatial markets in the presence of frictions. This paper pursues general insights about modeling and estimating dynamic spatial equilibrium in the presence of search frictions and price regulation. I study these topics in the context of the regulated taxicab industry, a setting where an uncoordinated geographic search process leads to equilibrium spatial patterns of supply and demand. Estimating this model allows me to measure the extent to which location- and time-specific pricing rules can better allocate available vacant taxis to neighborhoods with more customers.

The taxicab industry is a critical component of the transportation infrastructure in large urban areas. Over 800 million passengers are transported annually in the United States, generating \$23 billion in revenues.² Urban taxi markets are distinguished from most other public transit options by a lack of centralized control; taxis do not service established routes or coordinate search behavior. Instead, individual drivers decide where to provide service. The supply of taxis in any area is determined by the aggregation of these choices. Vacant taxis and customers do not know each others' locations, so in expectation, empty cabs and waiting customers will coexist while search is conducted. This paper analyzes these location-based frictions and the extent to which they are driven by the search behavior of vacant taxis. How and where drivers decide to search for passengers

¹Since 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. 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, Gârleanu, and Pedersen (2002, 2005)). In each case, market outcomes differ from the canonical, Walrasian ideal of perfectly frictionless trade and competition, as prices and quantities fail to reflect the intersection of supply and demand.

²See technical report Brennan (2014) for more information.

directly affects the availability of taxi service across the city. From a driver's perspective, location decisions have important economic tradeoffs. Taxis are drawn to search in the most concentrated pedestrian areas, but competitive pressure (i.e., the threat of losing fares to other taxis) incentivizes drivers to spread search among a broader set of locations. The confluence of these incentives gives rise to equilibrium search behavior that can leave some areas with little to no service while in other areas, empty taxis will wait in long queues for passengers.

A second defining characteristic of the U.S. taxi market is that both prices and quantities are typically regulated by municipal authorities. Although the scope and magnitude of regulations vary across cities, nearly all local regulators implement the same pair of instruments: two-part tariff pricing and entry restrictions. The two-part tariff price structure consists of a one-time fixed fare and a distance-based fare (e.g., \$2.50 per-ride plus \$2.50 per-mile, as in New York City). Entry restrictions strictly limit the number of taxis that are licensed to provide service (e.g., the 13,237 total permits to operate a taxi service in New York). Although local taxi regulations might be aimed at mitigating several potential externalities like congestion or pollution, fixed prices and quantities prevent traditional market mechanisms from efficiently rationing supply and demand within the city, so taxi drivers will not be fully incentivized to focus their search in the highest-demand areas. Further, in traditional taxi markets, prices are typically fixed within *and* across markets: a 2-mile trip to "point A" has a price that will not change in the face of different supply and demand conditions, nor will it change for an identical-distance trip to "point B", despite imposing very different opportunity costs on drivers. These uniform price rules inherently lead to distorted location incentives, and culminate into equilibrium patterns of supply that may be far from efficient.

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 locations of their competition and profitability of different search locations. I show how price regulations impact taxi drivers' search behavior, which in turn influences 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 2012.³ Using TLC data together with a model of taxi search and matching, I estimate the spatial, inter-temporal 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

³New 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)

an estimation strategy using the dynamic spatial equilibrium model together with a flexible 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 taxis play against the *distribution* of competitors instead of individuals.

I use this model to evaluate welfare in this market as well as the welfare effects associated with several potential policy alternatives. Baseline estimates of welfare indicate that the New York taxi industry generates \$1.48 million in consumer surplus and \$2.8 million in taxi profits in each 9-hour day-shift, reflecting the aggregate surplus across 223 thousand taxi-passenger matches. The impact of search frictions is substantial, however. By removing search frictions altogether, automatically matching available supply and demand could lead to daily efficiency gains of an additional 75 thousand matches, a 22% increase in taxi profits (from \$2.80M to \$3.43M) and a doubling of consumer welfare (from \$1.48M to \$2.97M).⁴ To what extent can a more sophisticated pricing policy mitigate these costs by better allocating available supply to demand? By simulating market equilibrium over approximately 10,000 potential pricing rules, I am able to solve for an optimal flexible fare structure and show that a flexible tariff which changes according to trip distance (i.e., the fixed price and price-per-mile are changing with distance) leads to an increase of \$76.3 thousand in daily consumer welfare, a gain of 6.3%, and an increase of \$117.4 thousand in daily taxi profits, a gain of 2.6%. Importantly, these gains come with the added benefit of reduced aggregate frictions due to 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.

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.⁵ Traditionally, matching functions incorporate fric-

⁴These calculations are in fact a lower bound: a partial equilibrium counterfactual that simply matches existing supply and demand, but does not account for any demand expansion due to the lack of frictions.

⁵A 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

tions 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.⁶ 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 welfare achieved in the market as well as the welfare impact of implementing more sophisticated pricing rules that are afforded by modern technological innovations in this industry.

I integrate the search and matching framework with a dynamic oligopoly model in the tradition of Hopenhayn (1992) and Ericson and Pakes (1995), which characterize Markov-perfect equilibria in entry, exit, and investment choices given some uncertainty in the evolution of the states of firms and their competitors. Instead of these choices, however, I model taxi drivers' problem as a series of location choices made throughout the day whenever cabs fail to find passengers. Given the non-stationary environment, this problem generates a particularly high-dimensional state space, so I use numerical approximate dynamic programming methods to mitigate computational burden. My approach exploits the large number of agents (taxi drivers) in the system by assuming that state transitions are perfectly forecastable by drivers. 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 game involves a play against the expected distribution of competitors throughout the day, rather than all possible realizations of competitors' states. This approach has precedent in literature on non-stationary firm dynamics (e.g., Weintraub, Benkard, Jeziorski, and Van Roy (2008), Melitz and James (2007)) as well in auction models with many bidders (e.g., Hong and Shum (2010)).

adapted in this paper.

⁶See also recent work of Brancaccio, Kalouptsidi, and Papageorgiou (2017) for an application of the Lagos model to study endogenous trade costs in the global shipping industry.

This paper also contributes to literature on spatial equilibrium by modeling the spatial dispersion of firms as an equilibrium outcome of the search process. While there is no unifying theory of spatial economics, there is a long history of literature tracing the interactions between space and economic activity.⁷ My model further underscores some common themes in the spatial equilibrium literature. Classic migration models of Rosen (1979), Roback (1988), and more recent work (e.g. Diamond (2012), Allen and Arkolakis (2014)), highlight equilibrium tradeoffs between city characteristics, wages, and migration. My spatial model reveals a conceptually similar no-arbitrage equilibrium that demonstrates tradeoffs between competition, profitability of locations, and the movement of vacant taxis over space.

Finally, this paper contributes to our understanding of the taxi industry by analyzing driver search behavior and the effects of regulation. To my knowledge, this is the first empirical analysis of the spatial search dynamics of taxis. The most closely related study is Frechette, Lizzeri, and Salz (2017) [hereafter, FLS], 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, FLS 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.⁸ 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.⁹

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 regulation has lead to restricted quantities and higher prices (Winston and Shirley (1998)) and

⁷For a history of spatial economics, see Ohta and Thisse (1993) For a review of modern spatial economics, see Fujita, Krugman, and Venables (1999). Important early literature capturing these effects appear in *The Isolated State* (von Thünen (1966)), Von Thünen's 1826 treatise showing how spatial sorting and agglomeration patterns develop purely through competition and market-determined land prices, as well as Hotelling (1929), a workhorse model that predicts the spatial locations of competitive firms as an equilibrium outcome.

⁸There is an additional body of literature on taxi drivers' labor supply choices, including Camerer, Babcock, Loewenstein, and Thaler (1997), Farber (2005, 2008), and Crawford and Meng (2011). These studies investigate the labor-leisure tradeoff for drivers. They ask how taxi drivers' labor supply is determined and to what extent is it 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.

⁹There is also a growing literature in empirical industrial organization which studies the welfare impacts and allocative distortions induced by search frictions in different industries. This includes work on airline parts (Gavazza (2011) and mortgages (Allen, Clark, and Houde (2014))).

that low sunk and fixed costs in this industry are sufficient to support competition (Häckner and Nyberg (1995)). My paper contributes to this debate by showing where regulation can help taxi markets by incentivizing more efficient spatial allocations of supply.

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 parameters with the data. Results are presented in section 5, with an analysis of counterfactual policies in section 6. Section 7 concludes.

2 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.

It is important to note that some cities permit taxis to operate both a street-hail service, where passengers are acquired via driving and searching, as well as a for-hire-vehicle (FHV) service, which pre-arranges rides with customers via telephone or internet. In other cities, including New York City, these two services are treated as separate markets by regulation. Yellow medallion taxis are only permitted to operate a street-hail service, and as such may not pre-arrange rides with customers. New York's FHVs are separately licensed for pre-arranged rides, and are likewise not permitted to operate a street-hail service.

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. The precipitous expansion and success of these firms is suggestive of the enormous benefits associated with reduced search costs compared with traditional taxi markets. At the moment, these companies also enjoy relatively little regulation, though this environment is changing rapidly as local regulators revise taxi laws to address the new platforms. My analysis, which focuses on the traditional, street-hail taxi markets

that still dominate service in the largest cities, will reveal the extent of spatial frictions induced by regulation as well as those induced by a random versus directed search mechanism (i.e., sorting customers by willingness-to-pay).

Regulation

The U.S. taxi market is highly regulated by local municipalities. The most common regulatory scheme is the combination of a fixed two-part tariff fare pricing structure and entry restrictions. The two-part tariff fare structure should be familiar to taxi customers; it consists of a one-time fixed fee and a distance-based fee. The other most common form of regulation is entry restriction. Most U.S. cities limit the number of legal taxis by requiring them to hold a permit or medallion, the supply of which are capped.^{10,11} 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.¹²

Across the U.S. and elsewhere around the world, taxi regulations are at odds with the new wave of technology-centered entrants in the taxi industry. These firms often implement different pricing regimes and much looser entry restrictions than those imposed by regulatory authorities, leading to a variety of legal disputes as stakeholders in the traditional taxi business suffer losses and as policy makers argue that safety and worker-rights regulations have been undermined.¹³ These ongoing policy questions highlight the need for an analysis of the effects of these new entrants, a central question of this paper.

Spatial Frictions

In the street-hail market, taxis and customers need to find one another through search. This is a consequence of spatial mis-allocation: often drivers are in one place and customers are in another, and where there are no dispatching services, a driver's process of finding customers is to drive around and search. The time it takes to match with a customer is the essential friction in this market, which could add substantially to the cost of service on both the demand- and supply-side. Search time imposes opportunity costs on drivers and customers as well as congestion, wasted fuel, pollution, etc.

¹⁰For a survey of entry restrictions across forty U.S. cities, see Schaller (2007).

¹¹These 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.

¹²See the discussion of regulation papers above.

¹³See, e.g.,

forbes.com/sites/ellenhuet/2015/06/19/could-a-legal-ruling-instantly-wipe-out-uber-not-so-fast/.

The extent and impact of frictions, however, is not primitive. The way that taxis search, the rate at which they find customers, and the rate at which customers arrive to the market are all shaped by regulation. Price regulations such as tariffs typically impact relative prices for taxis and customers across space (e.g., a high distance tariff makes longer rides relatively less attractive for customers). In particular, prices influence where cabs search and the transit patterns of customers. The transit patterns of customers further influence where taxis end up to conduct further search.¹⁴ This paper aims to characterize the location incentives faced by taxis, the equilibrium spatial distribution of supply, and to analyze the role of regulations on market outcomes.

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 34% 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 and date of pickup and drop-off, GPS coordinates of pickup and drop-off, trip distance, and trip time length for approximately 27 million rides.¹⁶ 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.

¹⁴The 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.¹⁵ 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.

¹⁶Using 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

Sample	Rate Type	Variable	Obs.	10%ile	Mean	90%ile	S.D.
All Data	Standard Fares	Total Fare (\$)	27,475,614	4.50	9.51	16.00	5.57
		Dist. Fare (\$)	27,475,621	1.50	5.59	12.00	6.14
		Flag Fare (\$)	27,475,621	2.50	2.83	3.50	0.36
		Distance (mi.)	27,475,621	0.82	2.70	6.00	2.74
		Trip Time (min.)	27,475,621	4.00	12.04	22.52	8.23
	JFK Fares	Total Fare (\$)	491,689	45	48.32	52	3.58
		Distance (mi.)	491,689	3.02	16.25	20.58	5.95
		Trip Time (min.)	491,689	22.75	39.49	60.00	17.33
	Weekdays, Day-Shift, Manhattan & Boro.	Total Fare (\$)	8,164,678	4.50	10.17	17.70	6.42
		Dist. Fare (\$)	8,164,681	1.20	4.66	9.60	5.33
		Flag Fare (\$)	8,164,681	2.50	2.5	2.5	0
		Distance (mi.)	8,164,681	0.71	2.29	4.71	2.37
		Trip Time (min.)	8,164,681	4.00	12.76	23.9	8.49
	JFK Fares	Total Fare (\$)	152,259	45.00	48.31	52.00	3.60
		Distance (mi.)	152,259	10.12	16.60	20.97	5.69
		Trip Time (min.)	152,259	28.98	46.78	67.67	18.20

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 5.

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.

2.3 Evidence of Frictions

The TLC data reveals every passenger trip that is taken linked with driver identifiers. Using driver ID together with the exact time stamps of pick-up and drop-off, I compute the “waiting time” drivers experience between trips. The mean waiting time for different trips is displayed in Figure

Table 2: Taxi Trips and Revenues by Area

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

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.

1. Panel 1 divides the day into 5-minute segments and computes 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 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 chapter 4 for details) and computes the same driver match probabilities and waiting times by region, this time averaged over time of day and all weekdays of the month. Again there is heterogeneity in frictions across space, with relative 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).¹⁷ In aggregate, the data suggests that 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.¹⁸

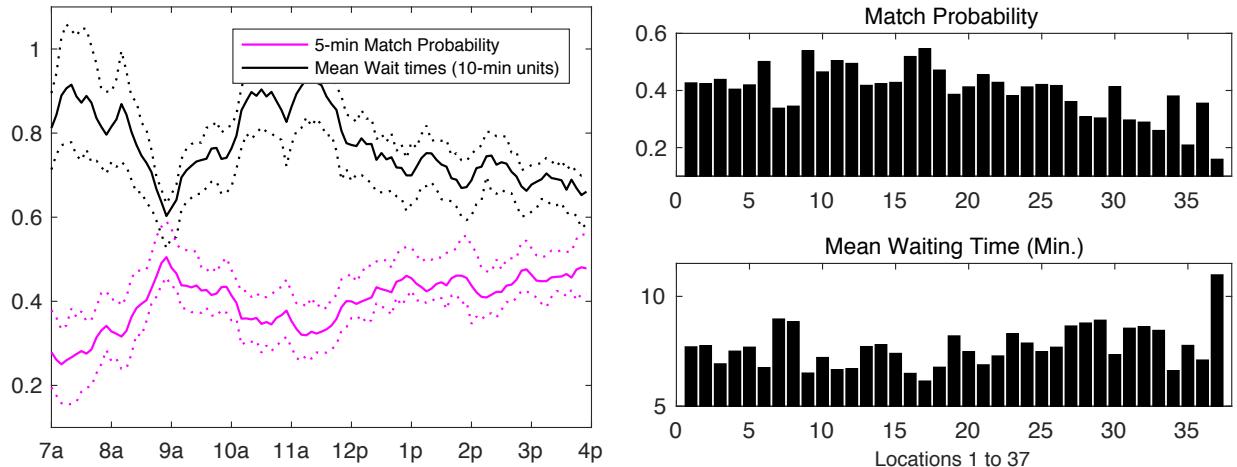
Figure 1 provides a snapshot of the frictions faced by drivers by time-of-day and neighborhood. The data do not reveal the frictions faced by customers; it is impossible to tell how long a customer has been waiting before pick-up, nor is it possible to tell if a customer arrived to search for a taxi and gave up.¹⁹ To analyze frictions, I will exploit observed moments in the data and estimate a

¹⁷There is additional evidence that drivers must 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 some spatial search frictions are present for drivers, as finding a customer requires relocation.

¹⁸See Section 4 for discussion regarding the total number of vacant drivers.

¹⁹The evidence that customers face frictions in the New York market is anecdotal, yet sufficiently widespread that

Figure 1: Taxi wait times and match probabilities by time-of-day and location



TLC Data from August 2012, Monday-Friday from 7am until 4pm, within regions indicated on Figure 5. Left Panel: Each series represents a taxi driver's 5-minute probability of finding a customer and mean waiting times, taken as the average across all drivers and all weekdays of the month for each 5-minute period. Dotted lines show the 25th and 75th percentiles of each series. Right Panel: This graphic divides the New York market into discrete locations and computes the driver waiting times and matching probabilities, averaged across all weekdays and times-of-day from 7am-4pm, but disaggregated by the location of each drivers' last customer drop-off. Location numbers follow a roughly South-to-North trajectory from 1-37.

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.

3 Model

A city is a network of L nodes, or “locations”, connected by a set of routes. Time is discrete, and the time horizon $t = \{1, \dots, T\}$ is finite, where t can be thought of as intra-day times such as minutes. At time $t = 1$ the work day begins; at $t = T$ it ends. A location can be thought of as a spatial area within the city.²⁰ In the empirical analysis below, I explicitly define locations by dividing a map into areas. Model agents are vacant taxi drivers who search for customers within a location $i \in \{1, \dots, L\}$. When taxis find passengers, they drive them from some origin location to a destination location $j \in \{1, \dots, L\}$. Denote $v_i^t \in \mathbb{N}$ as a count of vacant taxis and denote $u_i^t \in \mathbb{N}$ as a count of customers that will attempt to hail a taxi in each location at each time period. Each customer wants to travel to location $j \in \{1, \dots, L\}$ at time t with probability M_{ij}^t .

it is both a frequent complaint of New Yorkers and a common reference in pop-culture.

²⁰e.g., a series of blocks bounded by busy thoroughfares, different neighborhoods, etc.

The distance between each location is given by δ_{ij} where $i \in \{1, \dots, L\}$ represents the trip origin and $j \in \{1, \dots, L\}$ represents the destination. Similarly, the mean travel time between each location is given by τ_{ij} . In each location i at time t , the number of customers that wish to move to a new location is random and given by u_i^t where u_i^t is drawn from a Poisson distribution with parameter λ_i^t . I assume that taxi drivers know the distribution of demand at every location and time.

3.1 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 π is the distance-based fare, with the distance δ_{ij} denoting the distance between i and j . Thus the fare revenue earned by providing a ride from i to j is $b + \pi\delta_{ij}$.

At the same time, taxis spend time and money in pursuit of passengers. When a driver departs from location i to location j , with or without a passenger, he will arrive τ_{ij} periods later. Since there is a finite amount of time in a day, the cost of time is the opportunity cost of earning other revenue while in-transit; taxis may spend time to drive to more profitable locations. Taxis also pay a cost of fuel when traveling, given by c_{ij} .

Thus the net revenue of any passenger ride is given by

$$\Pi_{ij} = b + \pi\delta_{ij} - c_{ij}. \quad (1)$$

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

3.2 Searching and Matching

There are two types of locations, *searching locations* and *airports*. Searching locations comprise most a city; they are locations in which cabs drive around in search for passengers, with no guarantee that they will find one. Airport locations require taxi queueing. Taxis wait in queue for a guaranteed passenger ride once they reach the front of the queue.²¹ 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_i^t . Matches can only occur among cabs and customers within the same

²¹Airport rides comprise roughly 6% of all taxi trips, and 16% of revenues.

location. I assume that once a cab meets a customer, a match is made and the cab is obliged to go wherever the passenger wants within the city. A cab may not refuse a ride after contact is made.²² The meetings within each location-time are random and the number of matches made in location i and time t is given by an aggregate matching function $m_i^t(u_i^t, v_i^t)$. The exact specification of the matching function is explained in detail below.

A taxi driver's behavior is dependent on the probability that he will match with a customer. Within a location, $m_i^t(u_i^t, v_i^t)$ matches are randomly assigned between supply and demand. As such, the ex-ante probability that a cab will find a mover in location i at time t in a given period is $p_i^t = \frac{\mathbb{E}_{u_i^t}[m(u_i^t, v_i^t) | \lambda_i^t]}{v_i^t}$.

In addition, I make the following assumptions about demand: (1) The number of customers searching for a taxi ride (arriving at the beginning of each period and each day) is a Poisson random variable with parameter λ_i^t in location i and time t . (2) Customers will search for one period, after which they leave the market. These assumptions are necessary in the absence of empirical observation about the behavior of customers. The first specifies Poisson-distributed demand, a standard model for stochastic arrivals. The second says that customers will not wait for more than one period.²³ The matching between taxis and passengers within a location is illustrated in Figure 2.

3.2.1 Space, Time, and the Matching Function

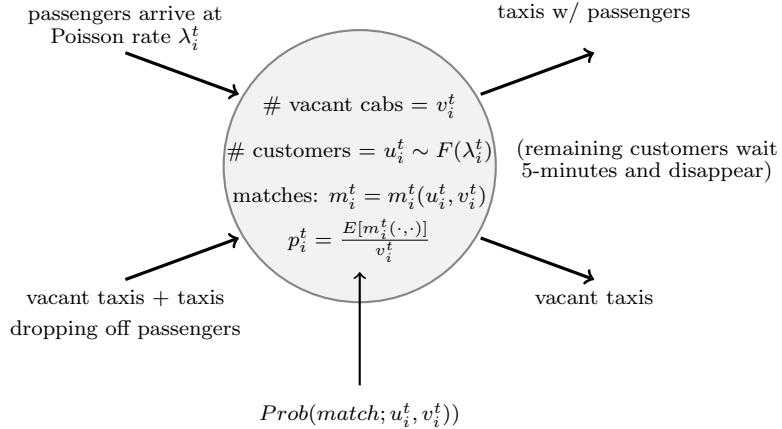
Figure 2 shows that matching probabilities depend on an aggregate matching function, $m_i^t(u_i^t, v_i^t)$. This function describes how many trades occur as a function of the number of trading partners on both sides of the market. If each trading partner in market i can make at most one trade per period t , then the most efficient matching function would be $\min\{u_i^t, v_i^t\}$. In a spatial search context, this frictionless matching function may aptly describe the matching technology governing a very small point in space, for example a single street corner where all buyers and sellers see each other. A larger area, such as a neighborhood, is made up of many such small points.

One problem with any empirical implementation of Leontief matching and an dynamic empirical search model, however, is that the time units must be consistent with the spatial search area. In other words, if it takes thirty seconds to search a street intersection, and two minutes to walk to the next intersection, then the underlying dynamic model should be consistent with the timing of these decisions. As there are several thousand street intersections in New York, an empirical

²²In New York, the TLC prohibits refusals, c.f. www.nyc.gov/html/tlc/html/rules/rules.shtml.

²³In 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.

Figure 2: Flow of demand, matches, and vacancies in search location i



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 $m(u_i, v_i)$. At the end of the period, newly employed taxis leave for various destinations and vacant taxis continue searching.

analysis of this market with this type of model would require, for reasons which will become clear below, infeasible amounts of data and colossal computing power.

Given these constraints, I propose another way to estimate search frictions and spatial mis-allocation by specifying a matching technology that governs an *area* as opposed to a *point*. As described above, an area, or neighborhood, can be thought of as a collection of points. *Within* a neighborhood, though, there will still be area-wide frictions arising from mis-match across the small “point” locations, as Lagos (2000) shows. For example, an empty taxi could be one street over from a person searching for a taxi, and so they fail to match. By defining a search area, many points may be collected into a location or zone of search and likewise, data and time can be aggregated to substantially ease the burden of data and computational requirements. There is a tradeoff, however. Moving from a frictionless matching function to one with frictions requires identifying additional parameters, those that determine the efficiency of the matching technology.²⁴

To model location-level search frictions, I use a simple aggregate matching function, given by equation 2. This function is derived from an urn-ball matching problem first formulated in Butters (1977) and Hall (1979), where u balls are randomly placed in v urns, and a “match” only occurs for the first ball placed in any urn.²⁵ To motivate its use in this study, I note its empirical tractability: it contains only one parameter, generates well-defined match probabilities (i.e., between 0 and 1),

²⁴I discuss identification in detail in section 4.4.

²⁵An equilibrium matching function based on this problem is derived in Burdett, Shi, and Wright (2001), where u represents buyers and v represents firms selling a single unit of some good. This basic premise has appealing analogs to the taxi market.

and is invertible in each argument. I make the following assumptions about search activity within a period: (1) customers and drivers can only match if they are in the same neighborhood and (2) each taxi can provide only one ride. It is therefore possible that some customers search on the same block and only the ones who first see cabs are matched, while elsewhere within the neighborhood some cabs encounter no passengers.

In location i at time t , with u_i^t customers and v_i^t taxis, aggregate matches are given by

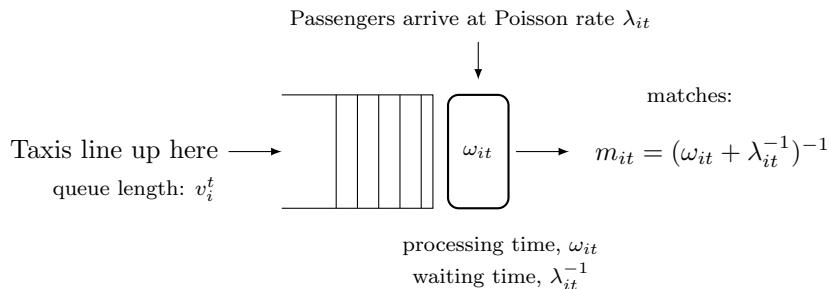
$$m_i^t(u_i^t, v_i^t; \alpha_r) = v_i^t \left(1 - \left(1 - \frac{1}{\alpha_r v_i^t} \right)^{u_i^t} \right). \quad (2)$$

I denote α_r as 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. I chose this parameter to depend 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.²⁶

3.3 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 i is governed by a Poisson arrival rate λ_{it}^t , and there is an additional expected waiting time ω_{it}^t which reflects how, once a passenger has arrived, it takes a very short time until that passenger is in the front taxi and the taxi has exited the cab-stand to make way for the next taxi. The expected total time for each taxi to be matched with a passenger once he is at the front of the line is therefore given by $\omega_{it}^t + \lambda_{it}^{-1}$. The expected waiting time experienced by a taxi at the back of a queue of length v_i^t is given by $v_i^t \cdot (\omega_{it}^t + \lambda_{it}^{-1})$.

Figure 3: Flow of demand, matches, and vacancies in airport location i



²⁶In principle α 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.

Figure 3 depicts the sources of taxi arrivals and departures at an airport location i at time t . All taxis searching for a customer at the airport are vacant and must arrive to a queue. The queue waiting time endogenously depends on how many taxis are in it and on the arrival rate of passengers at the airport. After waiting, taxis are matched with customers with probability 1, after which time they leave for various destinations. To understand how queues fit into observed data on matches, note that all taxi customers at the airport eventually get a ride and all queued taxi drivers eventually get a passenger. Thus matches are taken to be equal to the minimum between the number of customers and the number of taxis who can service the queue in one period, given by ω_{it}^{-1} for airport location i at time t .

3.4 Taxi Drivers' States, Actions and Payoffs

A taxi driver's behavior is dependent on his own private state (ℓ_a^t, e_a^t) and the industry state, \mathcal{S} . Specifically a driver a 's own location at time t is given by $\ell_a^t \in \{1, \dots, L\}$, and employment status (vacant or employed) is $e_a^t \in \{0, 1\}$. \mathcal{S} describes the expected number of taxis across all locations including those in transit between locations. The industry state at time t is a measure of vacant taxis v_i^t in each location i , as well as a measure of employed taxis $v_{e,k}^t$ actively *in-transit* between locations, where all “in-between” states are indexed by $k \in \{1, \dots, K\}$.²⁷ Thus the industry state at time t is given by

$$\mathcal{S}^t = \{\{v_i^t\}_{i \in \{1, \dots, L\}}, \{v_{e,k}^t\}_{k \in \{1, \dots, K\}}\}. \quad (3)$$

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 five-minute 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_i to each location $i \in \{1, \dots, L\}$.

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

²⁷At 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 dynamic decision-making problem of the currently-vacant taxis.

$$V_i^t(\mathcal{S}) = \mathbb{E}_{p_i|\lambda_i, \mathcal{S}^t} \left[p_i(u_i^t, v_i^t) \overbrace{\left(\sum_j M_{ij}^t \cdot (\Pi_{ij} + V_j^{t+\tau_{ij}}(\mathcal{S})) \right)}^{\text{Exp. Value of Fare}} + \right. \\ \left. (1 - p_i(u_i^t, v_i^t)) \cdot \mathbb{E}_{\varepsilon_{j,a}} \underbrace{\left[\max_{j \in A(i)} \left\{ V_j^{t+\tau_{ij} + \mathbb{I}_{[j=i]}\gamma}(\mathcal{S}) - c_{ij} + \varepsilon_{j,a} \right\} \right]}_{\text{Exp. Value of Vacancy}} \Big| \lambda_i^t \right]. \quad (4)$$

This expression can be decomposed as follows: Drivers in i expect to contact a passenger with probability $p_i(u_i^t, v_i^t)$. I assume matches are randomly determined within a location, so $p_i(u_i^t, v_i^t)$ is computed as the expected number of matches in location i divided by the total number of taxis searching in that same location, as detailed in sub-section 3.2.²⁸ Conditional on a passenger contact, drivers expect to receive mean profits, given by the probability of a trip to each location j (denoted by the transition probabilities given by $\{M_{ij}^t\}$), multiplied by the current profit associated with each possible ride Π_{ij} plus the continuation value $V_j^{t+\tau_{ij}}$ of being in location j τ_{ij} periods later.²⁹

A search for passengers occurs within the period. 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 adjacent location that maximizes the net present value of profits, given by continuation values $V_j^{t+\tau_{ij} + \mathbb{I}_{j=i}\gamma}$, fuel costs c_{ij} and a shock ε_{ja} . The parameter γ represents an extra time-payoff associated with choosing to stay put: when taxis continue to search where they are currently located, they will receive extra value associated with time and effort savings.³⁰

ε_{ja} 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 it inconvenient to search anywhere but further along the road in the same direction.³¹ 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, except when the day is almost over, then $A(i)$ reflects the locations that are attainable within the amount of time left in the day.

²⁸Note that the subscript i on p_i indicates that the matching function is not identical across locations. This is due to the different efficiency levels α_r for regions r .

²⁹Note that M_{ij}^t has superscript t because preferences of passengers change throughout the day.

³⁰In other words, the total continuation value of staying put in time t is neither equal to V_i^t nor is it equal to the same value a full period later, V_i^{t+1} but rather somewhere in between.

³¹The terms ε_{ja} further ensure that equilibrium will be in pure strategies, as drivers' will have best responses even among otherwise identical locations. Further, it ensures 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.

Thus V_i^t is the expected value of a search, evaluated prior to realizing a draw from the distribution of arriving customers (which affects the match probabilities p_i by influencing the number of matches experienced in each location), and before realizing the draw of valuation shocks ε_{ja} . Taxi drivers have beliefs over the policy function of vacant cabs governing state transitions in each period, denoted $\tilde{\sigma}^t$. Because time ends at period T , all continuation values beyond period T are set to zero: $V_i^t = 0 \forall t > T, \forall i$.

3.5 Drivers' Choice Problem

To form a strategy, taxi drivers compare the expected present values to search in each alternative location and account for the costs associated with traveling, both in terms of fuel expense and the opportunity costs of time. The strategy dictates where a taxi driver will search next for passengers, provided that the current period search is unsuccessful.

The last term of equation 4 gives the expected continuation value associated with ending up vacant. At the end of each period, vacant drivers must 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 4:

$$j^* = \arg \max_j \{V_j^{t+\tau_{ij} + \mathbb{I}_{[j=i]}\gamma} - c_{ij} + \varepsilon_{ja}\}. \quad (5)$$

To compute the firm's strategies, I define the ex-ante choice-specific value function as $W_i(j_a, \mathcal{S}^t, \lambda^t)$, which represents the net present value of payoffs conditional on taking action j_a while in location i , before ε_{ja} is observed:

$$W_i^t(j_a, \mathcal{S}^t) = \mathbb{E}_{S^{t+\tau_{ija}}} \left[V_{ja}^{t+\tau_{ija}}(\mathcal{S}^{t+\tau_{ija}}, \cdot, \cdot) - c_{ija} \right]. \quad (6)$$

Defining W_i^t separately from V_i^t 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 \mathcal{S}^t , but before observing ε_{ja} , is given by the multinomial logit formula:

$$P_i[j_a | S^t] = \frac{\exp(W(j_a, \mathcal{S}^t)/\sigma_\varepsilon)}{\sum_{k \in A(i)} \exp(W(j_k, \mathcal{S}^t)/\sigma_\varepsilon)}. \quad (7)$$

This expression defines drivers' policy functions σ_i^t 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.

3.6 Transitions

The collection of these policy functions form a transition probability matrix from any origin to any destination, which I denote as $\{\sigma_i^t\}$. Note that only vacant taxis transition according to these policies. Employed taxis will transition according to passenger transition probabilities given by M_i^t . Together, these two processes combine to generate a law of motion for the state variable \mathcal{S} .

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

Likewise, the transition kernel of vacant taxis is determined by policy functions. It is given by $\mu(\mathbf{v}_v^{t+1} | \mathbf{v}_v^t, \sigma^t)$. As with ν , μ specifies the expected $t + 1$ spatial distribution of period t vacant taxis, given the transitions generated from policies $\sigma^t = \{\sigma_i^t\}$ for $i = \{1, \dots, L\}$. The combined set of transitions forms an aggregate transition kernel that defines the law-of-motion, given by $Q(\mathcal{S}^{t+1} | \mathcal{S}^t) = \nu(\mathbf{v}_e^{t+1} | \mathbf{v}_e^t, \mathbf{M}^t, \mathbf{m}^t) + \mu(\mathbf{v}_v^{t+1} | \mathbf{v}_v^t, \sigma^t)$. Explicit formulas for the state transitions are provided in Appendix A.4.

3.7 Intraday timing

At time $t = 1$, taxis have an initial spatial distribution, which I denote as \mathcal{S}^1 and take as exogenous. Next, nature decides which taxis match with passengers and which ones are left vacant. The employed taxis disappear from the stock of vacant cabs while employed and accept a payoff $\Pi_{ij} + V_j^{t+\tau_{ij}+\mathbb{I}_{i=j}\gamma}$. The vacant taxis realize a profit of 0 in $t = 1$ and an average present value of his own state is given by last term in equation 4. Equation 5 defines which locations they will search in during the next period. In period $t = 2$, the locations of vacant taxis are updated, the employed taxis who are still in-transit are noted, and nature again decides which vacant taxis find passengers.

3.7.1 Timing

1. Taxis are distributed according to \mathcal{S}^1 .
2. Draws are taken from each Poisson distribution with parameters $\lambda_1^1, \dots, \lambda_L^1$, giving u_1, \dots, u_L .
3. Nature randomly assigns m_i^1 matches in each location according to the matching function.
4. Employed taxis leave for destinations.
5. Remaining customers disappear.

6. Remaining vacant taxis choose a location to search in during the next period.³²
7. Vacant and previously hired taxis arrive in new locations, forming new distribution \mathcal{S}^2 .³³
8. The process repeats from S^2 , S^3 , etc. until reaching \mathcal{S}^T .

3.8 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 rates 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.³⁴ 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}_i^t . A driver is assumed to have already "learned" the arrival rates 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 states $\{\mathcal{S}_i^t\}$, transition beliefs $\{\tilde{Q}_i^t\}$ and policy functions $\{\sigma_i^t\}$ over each location $i = \{1, \dots, L\}$, and an initial state $\{S_i^0\}_{\forall i}$ such that:*

- (a) *In each location $i \in \{1, \dots, L\}$, at the start of each period, matches are made according to equation 2 and are routed to new locations according to transition matrix \mathbf{M}^t . The aggregate movement generates the employed taxi transition kernel $\nu(\mathbf{v}_e^{t+1} | \mathbf{v}_e^t, \mathbf{M}^t, \mathbf{m}^t)$ where \mathbf{v}_e^t is the distribution of employed taxis across locations in period t and \mathbf{m}^t is the distribution of matches across locations.*

³²Regarding 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 5.

³³Regarding item 7, "some hired 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 .

³⁴Of 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.

- (b) In each location $i \in \{1, \dots, L\}$, at the end of each period, each vacant taxi driver (indexed by a) follows a policy function $\sigma_{i,a}^t(S^t, \tilde{Q}_i^t)$ that (a) solves equation 5 and (b) derives expectations under the assumption that the state transition is determined by transition kernel \tilde{Q}_i^t . The aggregate movement generates the vacant taxi transition kernel $\mu(\mathbf{v}_v^{t+1} | \mathbf{v}_v^t, \tilde{\sigma}^t, \mathcal{S}^t)$ where \mathbf{v}_v^t 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 $Q(S^{t+1} | \tilde{S}^t) = \nu(\mathbf{v}_e^{t+1} | \mathbf{v}_e^t, \mathbf{M}^t, \mathbf{m}^t) + \mu(\mathbf{v}_v^{t+1} | \mathbf{v}_v^t, \tilde{S}^t)$.
- (d) Agent's expectations are rational, so that transition beliefs are self-fulfilling given optimizing behavior: $\tilde{Q}_i^t = Q_i^t$ 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)). Uniqueness can be established by a simple backward induction argument, presented in Appendix A.6.

Proposition 3.1. *Given data on expected matches, m_i^t , expected transitions $\{M_{ij}^t\}$, and the costs and benefits associated with travel $\{\tau_{ij}, \delta_{ij}, \Pi_{ij}\}$, \mathcal{S}^* is uniquely identified.*

Proof. See Appendix A.6 □

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. The equilibrium conditions give rise to tightly correlated location-specific value functions over time, where lack of equality among the value function in each location is due to the transportation costs and time costs of moving around. See additional discussion and illustration in Appendix Section A.9.

4 Empirical Strategy

4.1 Discretizing time and space

The spatial equilibrium model is one of discrete times and locations. To connect these features with data, I define a discrete period length as well as a set of locations on a map of New York. The period length is set to five minutes.³⁵ Taxis form expectations of the state, location profitability, transition probabilities, demand arrivals, and policy functions for every five-minute period of a weekday and

³⁵Note in Table 1 that standard trips average 11 minutes, with a 10th percentile of 4 minutes.

month, so that, for example, on all weekdays in June from 1:00pm to 1:05pm, taxis will be playing the exact same strategy. In this sense, t is taken to be a {5-minute interval, Weekday Y/N, Month} combination.

I use data from 6am-4pm on Weekdays from August to September, 2012.³⁶ This time range corresponds to a typical day-shift among New York taxi drivers and it represents a period in which nearly all medallions are utilized. Figure 4 illustrates this by showing the number of active taxis by hour.³⁷ Note that levels shown will be smaller than the total medallion count of 13,237. This happens for a couple of reasons: medallions will not appear immediately when taxis enter the shift because it takes time to find passengers. Thus, there may be significant downward bias in the mornings. Some of this discrepancy is also attributable to out-of-service cabs (e.g., due to maintenance work), and lost observations from cleaning inaccurate data.³⁸ For these reasons, and because I also focus on a slightly smaller area than the entirety of New York City, I assume for estimation that 11,500 cabs are operating between Manhattan, Brooklyn, Queens, and the two airports between 7am-4pm.³⁹ While I make no assumptions about the number of taxis present between 6am and 7am, I will use the data in this range to establish the “initial” distribution of the 11,500 cabs at 7am.⁴⁰

To link data with spatial locations, I observe GPS points of origin for each taxi trip and then associate these points with one of 39 spatial areas.⁴¹ These locations, shown in Figure 5, are created by uniting census tracts. 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.⁴² Further, approximately 98% of all taxi rides originate within the area defined by these locations. 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. Figure 5 depicts the locations on the map of New York.⁴³

³⁶The 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.

³⁷To compute this, I first calculate the hours in which a taxi is on a shift by using the shift definitions from Frechette, Lizzeri, and Salz (2017), which specify that a driver is on a shift when he shows up in the data (i.e., a ride is given by the driver) and remains on shift until a gap of five or more hours between observed rides occurs. Next, I count how many taxis are working a shift in each hour and plot this count by hour.

³⁸See Appendix A.1 for more details on data cleaning procedures.

³⁹The only requirement for identification is that the total number of cabs at any point in time is exogenous.

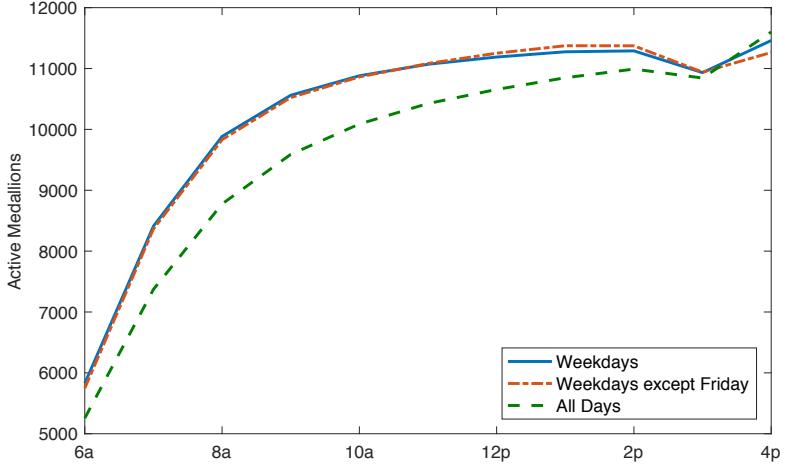
⁴⁰Appendix A.3 provides further discussion on the evolution of observed medallions across days of the month, lending additional support for the assumption of 11,500 active medallions.

⁴¹This association is achieved via the point-in-polygon matching procedure outlined in Brophy (2013). Thanks to Tim Brophy for the code.

⁴²See Appendix A.2 for additional details on creating these locations.

⁴³To provide a sense of 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.

Figure 4: Total medallion activity by hour

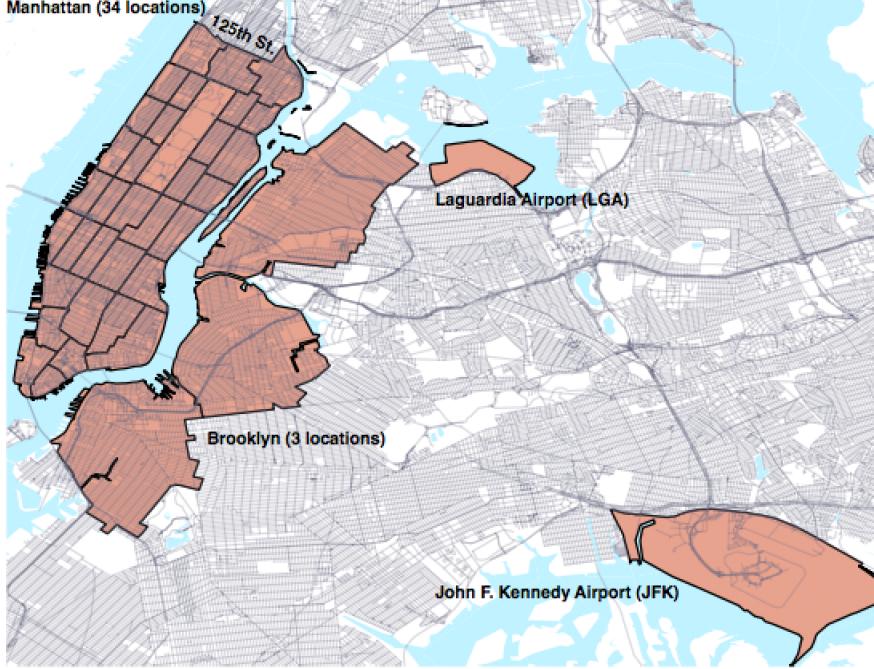


This figure is derived from August 2012 TLC trip data. It shows the number of unique medallions present in the data by hour-of-day, where presence is determined by finding the first- and last-appearance times within the day-shift, and counting each taxi as “active” in the hours between (inclusive of end-points). It approximates the total number of taxis working during the day shift, by hour, averaged across each day in the data set. Note that earlier hours will be systematically downward biased because drivers who begin a shift are not observed until finding their first passenger.

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. Though the realizations of these quantities may have variance, each parameter enters the value functions as a part of future expectations. As such, average values are the objects of interest. 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. τ_{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.
3. δ_{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.
4. ω_i^t is the expected per-passenger waiting time at each airport ($i = \{38, 39\}$). Because airport queues guarantee a match in the queue, I treat the mean number of pick-ups in each airport

Figure 5: 39-location Map of New York City



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.

location as direct observations of λ_{38}^t and λ_{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 only seek to identify the expected per-passenger waiting time faced by taxis, $\omega_i^t + \lambda_{it}^{-1}$. To compute this value, I compute the mean rate of pickups at each airport within all weekdays of the month. With this estimate, taxis' expected wait times at the start of the queue are computed as the product of all taxis at the airport v_i^t multiplied by the mean wait times.

- 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.⁴⁴ Using c , I compute the cost of traveling between any origin and destination as $c_{ij} = c \cdot \delta_{ij}$

After I record the distances between each origin and destination, I can derive Π_{ij} , the expected profit associated with each possible trip. Recall from equation 1 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

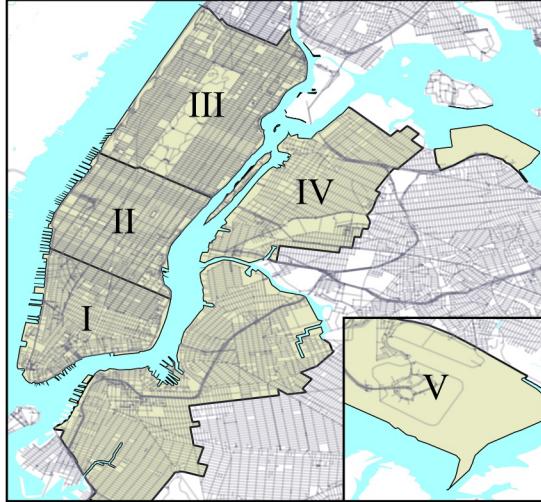
⁴⁴Data come from the New York City Taxi and Limousine Commission 2012 Fact Book. The high fuel economy rating is due to the medallion taxi fleet being made up of approximately 60% hybrid vehicles.

expected matches, the spatial equilibrium model is configured to identify the equilibrium spatial distribution of vacant taxis.

4.1.1 Regions

I further denote five “regions”, or distinct subsets of all 39 locations, according to Figure 6. Each region is characterized by unique mixes of geographical and transit infrastructure features. 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 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.

Figure 6: 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

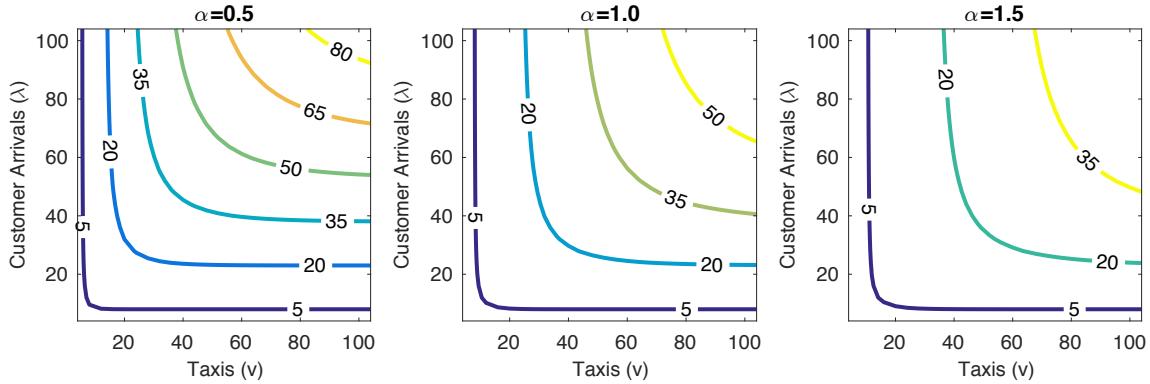
Equation 2 describes how many matches occur when u_i^t customers and v_i^t taxis arrive in a location, where u_i^t is drawn from a Poisson distribution with parameter λ_i^t . Because there is some underlying

variance in the realized number of matches each day, and because drivers make choices on the basis of expectations, the key moment of interest in the empirical application is the *expected matches*. The expected matching function, conditional on λ_i^t , can then be derived by integrating actual matches, given by Equation 2, over the probability density of each u_i^t given the Poisson parameter λ_i^t . Because the number of taxis v_i^t showing up at any moment is also stochastic due to the earlier draws of u_i^t each day, and due to the realizations of shocks ε_{ja} , let $v_i^t \in \mathbb{R}$ now denote the expected number of taxis in each location i and time t . This function is the following:

$$E[m_i^t(u_i^t, v_i^t) | \lambda_i^t, \alpha_r] = v_i^t \cdot \left(1 - e^{-\frac{\lambda_i^t}{\alpha_r v_i^t}}\right). \quad (8)$$

I assume that taxi drivers cannot observe the total number of cabs within the same location since they can neither observe competitors on nearby blocks nor can they easily see if a rival taxi is employed or not. Thus taxi drivers will base decisions on the probabilities of matching with respect to the average number of vacant taxis and the arrival rate of customers. Note that within a time t and location i , a new draw u_i^t occurs each day. Thus, if λ_i^t is the same across days (for example, take the set of all weekdays in September), repeated observation each day would reveal $E[m_i^t(u_i^t, v_i^t) | \lambda_i^t]$.⁴⁵ These expectations are one of the fundamental moments of the model that are also directly computable from observed data elements, a key ingredient to the identification of customer arrival rates. An illustration of the aggregate matching function, efficiency, and the role of the efficiency parameter α_r is depicted in Figure 7.

Figure 7: Matching Efficiency and α



This figure shows contour plots of the matching function over three values of α . Contour levels depict the expected number of matches produced in a given location when the level of taxis is v and the expected arrivals of customers is λ , for each shown level of α .

⁴⁵One could imagine weather, holidays, or other seasonality issues would also be conditions to ideally control for, though this would impose additional burden on data.

Equation 8 also highlights the way in which the demand parameters and state vector relate to expected matches. The ratio $\frac{\lambda_i^t}{\alpha_r v_i^t}$ determines the taxis' probability of matching p_i^t , and as shown in Figure 2, $p_i^t \cdot v_i^t = E[m_i^t]$.

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 industry state. However, the total number of taxis still contributes to the computational burden because it affects number of industry states over which continuation values must be computed.⁴⁶

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 probability transition matrices become deterministic transition matrices.⁴⁷ 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 $\{\mathcal{S}_i^t\}$ for $t = \{1, \dots, T\}$.⁴⁸

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 indeed the appropriate basis for computing continuation values. 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 (2008), in which agents form strategies on the basis of expected future paths of the state.

⁴⁶Specifically, 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 NP_L , or: number of states $(T, N, L) = T \cdot \frac{N!}{(N-L)!L!}$. Storing and solving for policy functions and value functions in this setting would be infeasible.

⁴⁷This 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. (2008).

⁴⁸Using 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.

Table 3: Parameter List

Parameter	No. Elements	Description
λ_i^t	4,212	Demand arrival rates
α_r	4	Matching efficiencies
σ_ε	1	Variance of ε_i^t shocks
γ	1	Extra value to staying-put

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

One tradeoff associated with this method is that it requires an initial, exogenous state \mathcal{S}^0 . I develop an algorithm to compute equilibrium conditional on \mathcal{S}^0 . I then solve for continuation values backwards from the last period T and use forward simulation to find optimal transition paths given continuation values. Given some state \mathcal{S}^t , remaining period continuation values $\{V^{t+1}, V^{t+2}, \dots, V^T\}$, I solve for optimal policies at time t which, given the deterministic transitions approximation, results in an equilibrium path from \mathcal{S}^t to \mathcal{S}^{t+1} . The job of the computational algorithm is to find equilibrium transition paths and value functions which are mutually consistent. This method reduces the total number of equilibrium continuation values and policy functions that I need to search for to $T \cdot L$, offering a drastic improvement in computational burden. Finally, by solving equilibrium for the hour *before* 7am (i.e., $t=-12, -11, \dots, 0$), I can mitigate the initial condition problem. Hereafter, I refer to this algorithm as the Taxi Equilibrium Algorithm (TEA). The details of TEA as well as tests for robustness to the initial condition can be found in Appendix (A.5).

4.4 Model Estimation and Identification

In this section I show that the parameters $\theta_1 = \{\{\lambda_i^t\}, \{\alpha_r\}\}$ and $\theta_2 = \{\sigma_\varepsilon, \gamma\}$ can be identified given the available data. Table 3 outlines the set of parameters to be estimated. The next two subsections describe this process in more detail.

4.4.1 Estimation and Identification of Demand and Efficiency Parameters

First, I discuss identification of θ_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 arrival rates of customers. Likewise, the key empirical moments of interest are also averages. Let \tilde{m}_i^t denote the *expected* matches in each location i and time t . Because matches are observed over many days, the \tilde{m}_i^t are therefore obtained in the data, computed as the average number of matches in each location and time-of-day cell across all weekdays of the month. As discussed above, \tilde{m}_i^t together with observed data moments τ_{ij} , δ_{ij} , and Π_{ij} give rise to a unique equilibrium state $\mathcal{S}^* = \{v_i^{t*}\}$.

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 $\{v_i^{t*}(\tilde{\mathbf{m}})\}$ where $\tilde{\mathbf{m}} = \{m_i^t\}$.

In a first step, I identify the set of ratios $\{\frac{\lambda_i^t}{\alpha_r}\}$ using $\{v_i^{t*}(\tilde{\mathbf{m}})\}$ and properties of the matching function. Then I exploit data on the variances of realized matches across days to separately identify the numerator and denominator.

Proposition 4.1. *Suppose a vector of expected matches by location and time, $\tilde{\mathbf{m}}$, is observed. Further, suppose Proposition 3.1 holds and $v_{it}^*(\tilde{\mathbf{m}}) \neq 0 \forall i, t$. Then the ratio $\frac{\lambda_i^t}{\alpha_r}$ is identified.*

Proof. Appendix (A.6) shows in detail that the expected number of matches $E[m(v_i^t, \lambda_i^t, \alpha_r)|v_i^t]$ is one-to-one in $\frac{\lambda_i^t}{\alpha_r}$. Thus as long as $\mathbf{v}^*(\tilde{\mathbf{m}})$ is unique and strictly positive, the matching function can be uniquely inverted for $\frac{\lambda_i^t}{\alpha_r}$ given $v_{it} = v_{it}^*(\tilde{\mathbf{m}})$ and $m_{it} = \tilde{m}_{it}$:

$$\frac{\lambda_i^t}{\alpha_r} = -v_{it}^*(\tilde{\mathbf{m}}) \cdot \ln \left(1 - \frac{\tilde{m}_{it}}{v_{it}^*(\tilde{\mathbf{m}})} \right). \quad (9)$$

□

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:

$$Var(m_i^t) = (v_i^{t*})^2 e^{-2\frac{\lambda_{it}}{\alpha_r} \frac{1}{v_{it}^*}} \left(e^{\frac{\lambda_{it}}{\alpha_r^2} \frac{1}{v_{it}^{*2}}} - 1 \right). \quad (10)$$

Proposition 4.2. *Suppose all assumptions of Proposition 4.1 hold, and suppose a vector $\tilde{\sigma}_{\mathbf{m}}^2$ of the variance of matches by time and location across days is observed. Then $\{\lambda_i^t\}$ and $\{\alpha_r\}$ are separately identified.*

Proof. $\frac{\lambda_i^t}{\alpha_r}$ and v_i^{t*} are previously obtained as in Proposition 4.1. Denote $\hat{\lambda} = \frac{\lambda_i^t}{\alpha_r}$, and $\tilde{\sigma}_{\mathbf{m}, it}^2 = \{Var(\tilde{m}_i^t)\}$, where the variance of matches in each location and time is taken across days. Then inverting equation 10 for α_r gives:

$$\alpha_r = \frac{\hat{\lambda}_{it}}{v_{it}^*(\tilde{\mathbf{m}})} \left(\ln \left(e^{2\hat{\lambda}_{it}} \left(\frac{\tilde{\sigma}_{\mathbf{m}}^2}{v_{it}^{*2}(\tilde{\mathbf{m}})} \right) + 1 \right) \right)^{-1}. \quad (11)$$

Equation 11 shows that α_r is overidentified as each region r is made up of several locations i and times-of-day t . While α could be specified as i or t specific, a choice to model frictions on the basis of broader regions will help obtain more credible results for each α_r , as there will be error in the measurement and estimation of the right-hand-side parameters and moments. From here, α_r

can be estimated via NLLS. Additional details and estimating equations can be found in Appendix A.8. With estimates of α_r and $\frac{\lambda_i^t}{\alpha_r}$, the demand parameters $\{\lambda_i^t\}$ may be recovered directly. \square

4.4.2 Estimation of Remaining Parameters

The other two parameters to be estimated, denoted above as θ_2 , consist of the scale parameter σ_ε and the same-location value γ . To identify θ_2 , I use a standard Method of Simulated Moments (MSM) estimator designed to rationalize a few key 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 each of sections I-IV depicted in Figure 6.

These moments should be impacted by each parameter. The greater σ_ε , the less taxis' behavior will correspond to the model and instead be driven by unobservables. Thus, a high- σ_ε equilibrium will have drivers more "spread out", serving different customers with, on average, longer distance trips and longer trip times. The value of γ will influence drivers' tendencies to continue searching within the same location after dropping off a passenger. A higher γ will thus tend to increase the chances of pick-up in the same location as the last drop-off. These "tuning parameters" help to ensure a better prediction of the spatial distribution of taxis, as this distribution is an input to identifying demand and search parameters. Both parameters will impact matching probabilities, as they are inputs to drivers' optimization problem.

To implement the MSM estimator in conjunction with the estimation of θ_1 , I solve for $\mathbf{v}_k^*(\tilde{\mathbf{m}})$ and estimate $\theta_{1,k}$ for each point k in a two-dimensional grid of θ_2 parameter values. I then identify the point k^* and the resulting $\theta_{2,k^*} = (\sigma_\varepsilon^*, \gamma^*)$ which minimizes the GMM criterion function comparing empirical moments with their simulation counterparts. The resulting values $\mathbf{v}_{k^*}^*(\tilde{\mathbf{m}})$ and $\{\theta_{1,k^*}, \theta_{2,k^*}\}$ are then recorded as the solution and estimates, respectively.⁴⁹

4.5 Demand System

The preceding sections outline the estimation of two quantities for *each* location and time period: the equilibrium supply of taxis, v_i^{t*} , and the demand parameters, λ_i^t . These are estimates of the average level of supply and demand on weekdays in August 2012, given the regulated prices in effect at the time. The role of prices in this estimation is limited only to computing revenues Π_{ij} . Next, I will exploit a broader panel of data, spanning additional months, during which there was a change in regulated prices. Price variation will permit computing elasticities of demand, which will further allow for estimating welfare as well as predicting counterfactual market outcomes.

⁴⁹Details on the MSM estimator are given in Appendix (A.7).

On September 4, 2012, the distance fee increased by \$0.50 per-mile, and the JFK airport flat-fee increases by \$7. Using September 2012, data, I re-estimate the model for $\{v_i^{t*}\}$, $\{\lambda_i^t\}$, and $\{\alpha_r\}$.⁵⁰ In the analysis that follows, I use price variation across months and across space to estimate demand elasticities, where “demand” is simply the average customer arrivals in each i with destination j at time t .

I assume that for each airport-category a , and distance-category r , taxi drivers face a constant-elasticity demand curve of the form:

$$\ln(\lambda_{ijt}) = \sum_a \mathbb{I}_{a=\{0,1\}} (\alpha_{0,r,a} + \alpha_{1,r,a}) \ln(P_{ij}) + \delta_{tr} + \xi_{ir}. \quad (12)$$

The left-hand side of this equation, λ_{ijtr} , is the destination-specific mean number of customers in location i going to location j at hour t , for a trip with distance range r . It is computed using estimated quantities as follows: $\lambda_{ij}^t = \lambda_i^t \cdot M_{ij}^t$. Recall that the travel preferences of customers are observed in the data and given by the transition probabilities $\{M_{ij}^t\}$. Since M_{ij}^t simply scales the total arrival rate, the destination-specific arrival rate will therefore also be Poisson.⁵¹

For distance categories r , I use bins $\{r_0, r_1, r_2, r_3\} = \{0, 2, 4, 6\}$ miles, roughly corresponding to trip-distance quartiles. Thus, there are four ride-length categories: {0-2 mi., 2-4 mi., 4-6 mi., 6+ mi.}. Trips are further classified as an *airport trip*, or $a = 1$, if either origin or destination involves an airport. The right-hand side has the following components: for each trip of origin-destination-time (i, j, t) in group a, r , there is a price P_{ij} , a time-of-day fixed effect δ_{tra} , and an origin-location fixed effect ξ_{ira} . Price elasticities $\alpha_{1,r,a}$ are different for each ride-length category; separating elasticities by ride-lengths reflects the idea that different types of customers demand different types of rides. Origin and time fixed effects captures the heterogeneity of locations: some have more or less public transit stations, bus stops, walkability, all of which impact demand.

In this demand system, all customers of a given type r, 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 r, a), and differences in prices before and after the September 2012 fare change.⁵² To estimate parameters, I estimate an empirical analogue of equation 12 for each r category using OLS. Recall that 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.

⁵⁰Since road conditions, traffic patterns, average weather patterns, etc. may change, I allow α_r to change by month. I assume that σ_ε and γ are unchanged, however.

⁵¹The estimates λ_{ij}^t can be regarded as the expectation of points on the demand curve, e.g. $E^\lambda(Q_{ij}^t|a, P_{ij}, \xi_{ir}, \delta_{tr})$.

⁵²For 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 arrival rates. In addition, there is a change in prices and arrival rates for all trips between August and September.

Table 4: Results Summary

Panel A: Parameter Estimates Summary			
Estimated Parameter	Number of Estimates	Point Estimate	Mean/Min/Max
$\{\lambda_i^t\}$	4,212 (108×39)	See Figs. (A5)-(A6)	69.4 / 1.45 / 1,999.7
α_r	4	{0.87, 0.84, 1.06, 0.99}	n.a.
σ_ε	1	2.0	n.a.
γ	1	0.1 periods	n.a.

Panel B: Equilibrium Summary			
Estimated Object	Number of Elements	Computed Value	Mean/Min/Max
$\mathcal{S} = \{v_i^t\}$	4,212 (108×39)	See Figs. (A5)-(A6)	219.10 / 3.17 / 475.0
$\{V_i^t\}$ (non-airport i)	3996 (108×37)	See Fig. (A7)	\$186.6 / \$0.45 / \$367.5
$\{V_i^t\}$ (airport i)	216 (108×2)	See Fig. (A7)	\$188.6 / \$0 / \$571.76

This table presents a summary of estimation results and equilibrium solutions using the baseline August 2012 data. Point estimates for matching efficiency parameters α_r correspond to Table 6 sections I-IV, respectively.

5 Empirical Results

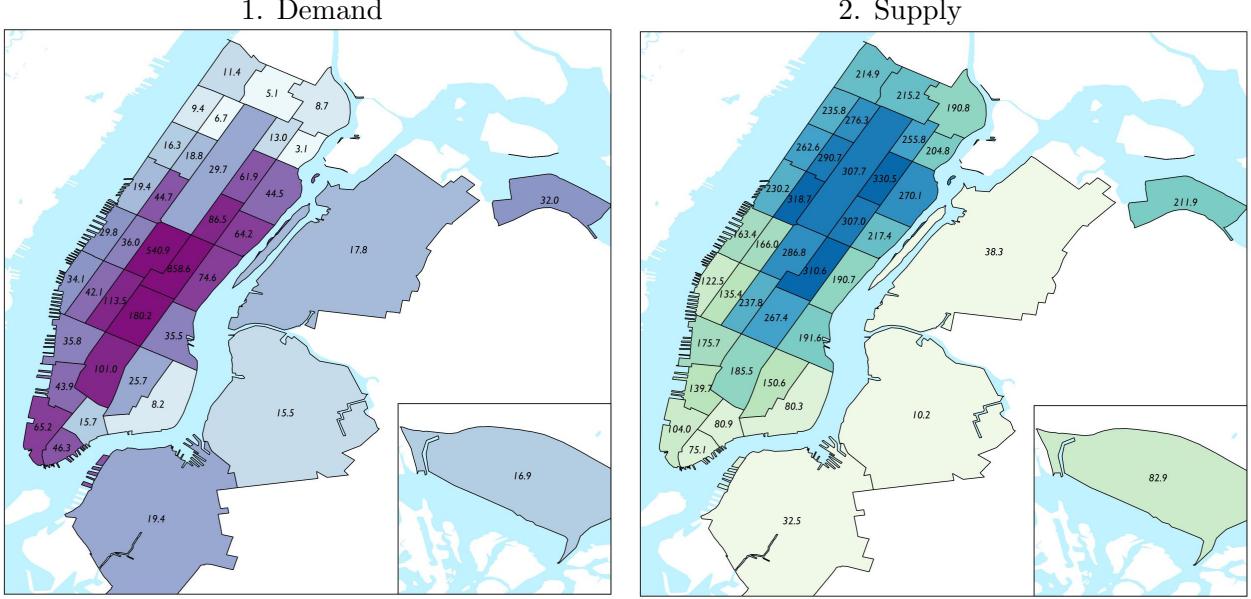
This section presents estimation results of the spatial equilibrium model. Table 4 shows a summary of the estimated parameters and the corresponding equilibrium spatial supply of taxis. Panel A shows estimation results for the per-period Poisson arrival rates of customers λ_i^t across time and locations, as well as point estimates for three scalar parameters: the variance of unobservable shocks, σ_ε , the matching efficiency parameter α , and the extra value to cabs for choosing to remain in the same location to search next period. Note that the parameter γ 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 a “bonus” equal to $0.1 \cdot (V_i^t - V_i^{t+1})$, 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 results figures is provided in Appendix Figure A.9.

5.1 Spatial Distributions and Intra-day dynamics

Figure 8 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

Figure 8: Map of Estimated Supply and Demand



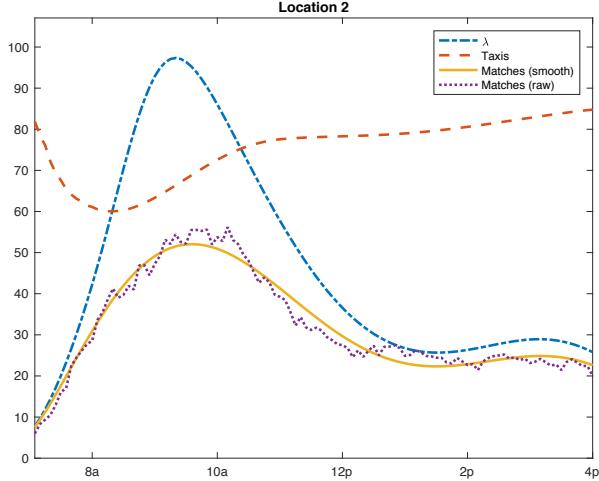
This figure shows the average supply and demand estimates for August, 2012, where the average is taken over time-of-day. The left panel shows the average customer arrivals in each location from 7a-4p, and the right panel shows the average number of vacant taxis in each location from 7a-4p.

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.

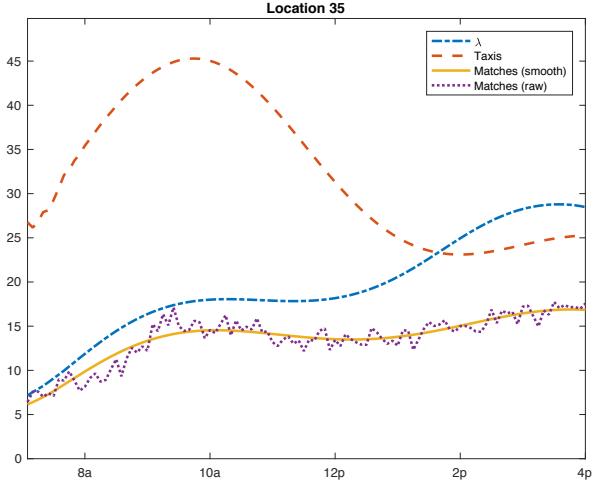
Figure 9 shows results for two busy locations, representing the Wall Street district (Panel 1) and Brooklyn (Panel 2). Both graphs depict the equilibrium supply of vacant taxis, estimated arrival rates of customers looking for a taxi, the equilibrium number of matches, and the model's fit against the 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 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, which gives rise to a difference between the levels supply and demand and the total number of matches. This is pictured as the vertical space between supply and demand at any point (i.e., $\min\{v_i^t, \lambda_i^t\}$)

Figure 9: Dynamic Spatial Misallocation Example

1. Undersupply: Lower Manh. (Loc. 2)



2. Oversupply: Brooklyn (Loc. 35)



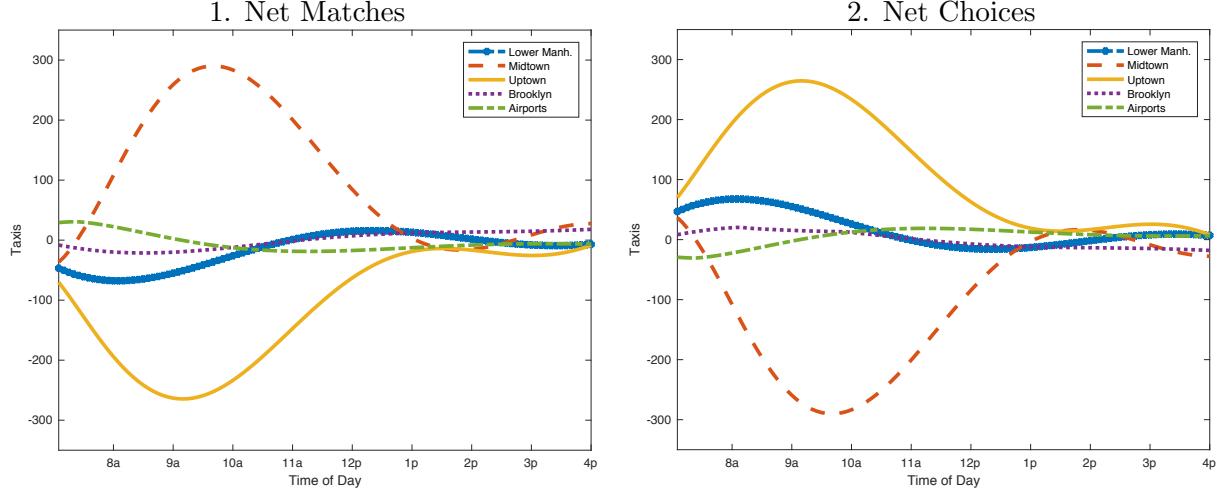
This figure shows intra-day results for two example locations. Panel 1 is Location 2 in far south-east Manhattan. Panel 2 is Brooklyn, just across the East River from Location 2. Each figure depicts the equilibrium supply of taxis (red, dashed line) and the estimated arrival of passengers (blue, dot-dash 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.

and the realized matches (i.e., m_i^t).⁵³

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 10 aggregates taxi supply across all 39 locations into five regions (identical to those in Figure 6) 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 vacant taxi drivers will “undo” the redistributions created by the movement of employed taxis.

⁵³Results for more locations are available in Appendix (A.9).

Figure 10: Equilibrium Flow of Matches and Vacancy Choices by Region



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 taxis 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).

5.1.1 Comment on Pricing

The dynamic spatial equilibrium described above is solved by holding prices $\pi_{ij} = b + \delta_{ij}$ as exogenous and fixed, as is true of the regulated prices in New York City: a two-part tariff that is constant through the day. Holding prices fixed, the dynamic spatial equilibrium is the result of search and matching across locations throughout the day. Thus, in each time period t and origin-destination pair i , the spatial model estimation recovers λ_{ij}^t , a *point* on the demand curve at price π_{ij}^t . Estimation of the demand system then recovers the price elasticities of demand. At any particular price, however, there will be unmet demand and/or supply (as in Figure 9) due to the presence of search frictions both within and across locations.

We might like to know if the market would be more efficient under a more flexible tariff pricing schedule. For example, in Figure 9, would a higher price in the morning in Lower Manhattan and a lower price in Brooklyn better ration supply and demand? This seems intuitive, but the answer is not obvious; the supply of taxis is determined by the equilibrium flow of supply and arrivals of demand throughout the entire network and over time. To analyze these kinds of counterfactuals, I need to understand how price changes will impact demand and the resulting equilibrium spatial allocations of supply. I proceed by estimating the price elasticities of demand. With the full set of taxi and arrival-rate estimates, I exploit price variation across different origin-destination pairs, as well as variation stemming from a regulated fare increase, to estimate demand elasticities and

welfare.

5.2 Demand Estimation

Table 5 provides results of the demand estimation of equation 12. As outlined in Section 4.5, 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 origin-location, destination-location, and five-minute period within a weekday from 7a-4p. The dependent variable is the expected passenger arrivals (λ_{ij}^t) in each 5-minute time interval. The level of price and quantity variation is origin-destination-time. In each pair of i, j, t I observe the average customer fares paid p_{ij}^t , and I estimate λ_{ij}^t , the mean customer arrivals.⁵⁴ 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.

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

Trip Distance	Price Elasticity		Specification Details	n	R^2
	City Trips	Airport Trips			
0-2 mi.	-1.7274 (.4474)	-.8311 (.0802)	Hour, Origin Loc.	87,186	.1233
2-4 mi.	-1.3543 (.4791)	-1.3844 (.4862)	Hour, Origin Loc.	73,646	.2680
4-6 mi.	-3.2173 (.4165)	-2.8815 (.4502)	Hour, Origin Loc.	43,630	.4368
6+ mi.	-2.2699 (.3918)	-1.4041 (.3507)	Hour, Origin Loc.	87,570	.3779

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

Demand data come from model estimates. Dependent variable is $\log(\lambda_{ij}^t)$. 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

With demand functions estimated, I directly 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.⁵⁵

There are two complicating factors associated with computing welfare in this context. The first is that there are some limitations to computing welfare from a single elasticity number in a constant-

⁵⁴I account for the actual prices paid after adding in fees for tax, tolls and idling.

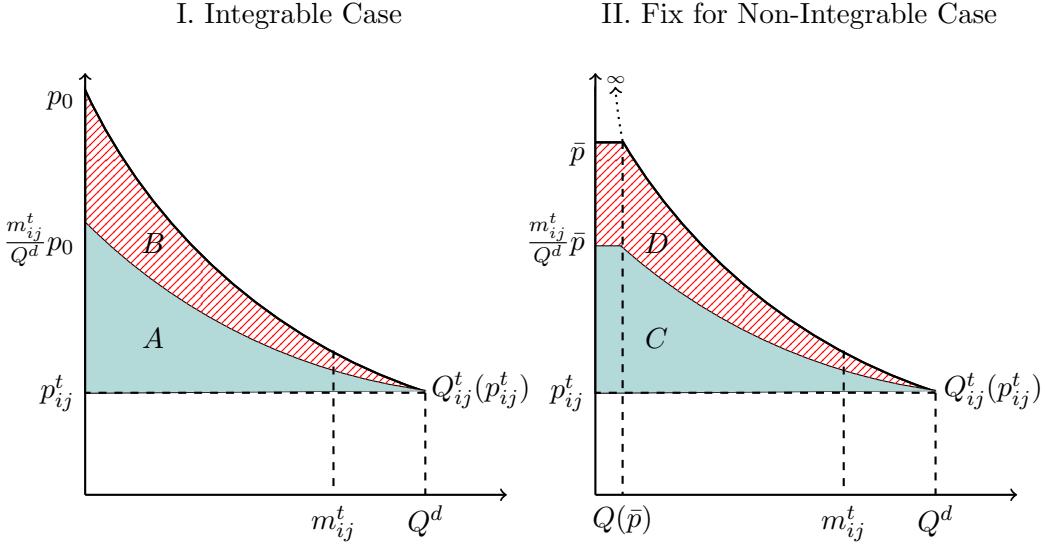
⁵⁵Explicit formulas used to compute these measures are provided in Appendix section A.10.

elasticity model. Any prediction of quantity based on these estimates will be increasingly sensitive to the functional form assumption and less reliable the further out-of-sample the price inputs are. Moreover, when elasticities are below unity, the inverse demand function has no intercept and thus the integral is infinite. With these caveats in mind, I will find an approximate welfare calculation for the few cases of below-one elasticities and, on the whole, argue that these measures may still be regarded as approximations as they generate “reasonable” predictions on the upper- and lower-bounds of willingness to pay. The approximation is simple: I winsorize consumer valuations above a certain dollar-valuation threshold, as in Figure 11 Panel II. While there is no perfect substitute for taxis, we might guess that even the highest-value consumers still have limits to their willingness to pay (unlike the limitless case produced by elasticities below one). This method attempts to correct for consumer valuations that are extraordinarily large, replacing them with this threshold. In the reported results below, I chose a threshold of \$150. Note that this will only apply to airport trips of 0-2 miles, representing only 0.28% of all trips, as shown in Table 5.

The second issue to consider in computing welfare is to properly account for search frictions. In this model there are two types of frictions: the within-location frictions generated by the matching function and modulated by the estimated efficiency parameters, and the across-location frictions stemming from equilibrium spatial mismatch between supply and demand. Because welfare is recorded at the level of the market (i.e., an origin-destination-time pairing), the across-location frictions are already accounted for in that there will be different levels of taxis and customers in each location and time, and different *types* of customers (based on destination) in each location and time. The second component is the within-location frictions, which I account for as follows: the demand functions provide a level of customers who wish to find a taxi for each market (i, j, t) , while solving spatial model solution produces the equilibrium number of matches m_{ij}^t generated in each market. Because I assume that taxis are rationed to passengers randomly as a result of the street-hail process in each period, this means that only a fraction of all customers demanding a ride will actually get one, which will shrink welfare by a factor of $m_{ij}^t / \lambda_{ij}^t$.

These calculations are illustrated in Figure 11. Panel I depicts welfare calculation 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_i^t , and the area B is the lost surplus due to not all demand being matched. The area A is therefore measured as the welfare for each sub-market (i, j, t) , and will be an endogenous equilibrium object through the influence of both p_{ijt} and m_{ij}^t . Panel II depicts the welfare calculation for demand functions with price elasticities below 1, corresponding to shorter-distance airport trips (Table 5). Again, $C \cup D$ reflects the total available surplus. Customers are matched randomly, so the lost surplus due to matching frictions is represented as area D . The difference from panel I is that I use “choke prices” at \bar{p} , in both the total available surplus as well as the match surplus. Thus, aggregate welfare numbers are a sum $\sum_{i,j,t} W_{ijt}$ across the recorded

Figure 11: Consumer Surplus with Frictions



This figure depicts how welfare is calculated in each sub-market (i, j, t) in the presence of search frictions. Panel I depicts the case in which inverse demand functions have finite intercepts, as when elasticities are greater than one. Panel II depicts this calculation when inverse demand functions grow infinitely large as $Q \rightarrow 0$. In this case, I truncate inverse demand at \bar{p} and record welfare as if willingness-to-pay is no greater than \bar{p} in the region $Q \in [0, Q(\bar{p})]$. The depicted scale of each curve is for illustration purposes, the scale and curvature of demand differs across markets.

equilibrium welfare in each (i, j, t) .

Total estimated welfare for each weekday, day-shift is shown in Table 6, which shows that there is a lower bound consumer welfare of \$1.48M per day for New York taxi service during weekdays from 7a-4p in September, 2012. Taxi profits in each shift are \$2.80M, or \$224 per driver.⁵⁶ A crude approximation of annual welfare, then, is obtained by multiplying this number by two shifts and again by 365 days. This calculation yields \$1.1 billion in consumer surplus and \$2.0 billion in taxi profits.⁵⁷ Beyond the baseline welfare estimates, Table 6 also depicts several decompositions: by time-of-day, trip origin location, and distance of trip. It shows that mid-morning trips, central Manhattan trips, and shorter distance trips are generally more profitable for taxis and customers alike, though not necessarily in proportion to one another. These relative differences reveal the wedge between taxi incentives versus the planners, highlighting the possibility of misaligned incentives in the way trips are priced.

⁵⁶I use the term “profits” to denote revenues net of fuel costs.

⁵⁷This discrepancy between taxi profits and consumer welfare in part reflects the rents accruing to medallion holders in this period as a consequence of high average prices together with medallion limits.

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

Grouping Category	Grouping	Cons. Surplus (\$, thousand)	Taxi Profits (\$, thousand)	Matches
Aggregate	-	1,477.2	2,796.7	223,278
Time-of-day	7a-9a	374.9	608.4	32,981
	9a-11a	514.2	768.7	46,415
	11a-1p	440.0	695.9	42,115
	1p-4p	543.9	723.7	64,479
Origin Region	Sec. I	327.3	499.2	32,745
	Sec. II	387.6	1,199.2	87,568
	Sec. III	371.4	755.2	52,807
	Sec. IV	391.0	781.0	36,191
	Sec. V	94.7	265.1	5,637
Trip-Length	0-2 mi.	664.5	1,230.0	139,623
	2-4 mi.	531.7	707.3	52,810
	4-6 mi.	84.9	269.1	13,775
	6+ mi.	195.6	588.9	16,992

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. 6.

6 Counterfactuals: Price Flexibility in Dynamic Spatial Markets

The recent rise of modern taxi alternatives, including Ride-sharing services such as Uber and Lyft, has demonstrated the potential value of several innovations, including a platform for on-demand labor, improved coordination between customers and drivers, increased competition, and the ability to adjust prices in real time depending on market conditions in a neighborhood. If these services are able to increase competition, lower prices, and increase output by dismantling the market power generated by a captured regulator, then it won't be surprising that welfare and efficiency are enhanced. An open question, however, is how much can price flexibility be used to enhance allocative efficiency in the dynamic spatial setting?

In this section, I use the estimated model to investigate the extent to which introducing different types of price flexibility can improve the equilibrium spatial allocations of supply and taxi-passenger matches. Specifically, I permit prices to adjust along different dimensions across each of three counterfactual experiments: time, space, and trip distance. “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 a multidimensional grid 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.

While I am interested in understanding how price flexibility 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 these margins, I compute the “waiting time” experienced by all customers as the probability of not being matched times the waiting time of 5 minutes. All 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 market, and in which any location-specific deterioration in customer waiting times

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 6. 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 (and approximate quartiles of the empirical distribution of trip distances): 0-2 miles, 2-4 miles, 4-6 miles, and 6+ miles. These new 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 can improve welfare by generating different equilibrium allocations of supply and demand in each time segment compared with uniform pricing.⁵⁸ The results will highlight the value of price flexibility 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. Throughout, I hold fixed the basic institutional details of the New York City taxi market: the network geography, two-part tariff pricing, 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, (b) total daily consumer waiting time frictions are at least as low as before, and (c) any *location-specific* consumer waiting time friction is limited (as specified below). 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 (even though, in aggregate, expected waiting times decrease). The table shows that gains to consumer surplus and profits are simultaneously achievable by permitting price flexibility, even while constraining that flexibility along only four dimensions within each category. In particular, flexible pricing by distance yields the largest equilibrium welfare gains of about 6.3% in consumer surplus and 2.6% in profits, while the worst location only experiences an additional 11 seconds of expected wait times. There are additional benefits to gain if we maintain that aggregate profits and average waiting times are at least as good as in the baseline case, but allow for increased heterogeneity in consumer search frictions: in other words, if we allow for worse waiting times in some locations.⁵⁹ Because these prices induce more efficient utilization of the taxi supply, there is an overall increase in matches. This result corroborates the theoretical insights of Schmalensee (1981) and Varian (1985) which find

⁵⁸ “Optimal” here is subject to the caveat that the margins of flexibility (i.e., how locations, times and trip distances are split) are exogenously determined in these counterfactuals.

⁵⁹ Appendix B.1 shows how efficiency gains scale when additional heterogeneity is allowed in the search friction.

Table 7: Counterfactual Results: Optimal Flexible Prices by Policy

Pricing Policy	Price	Consumer Surplus	Profits	Mean Waiting Time	Matches
Baseline, Aug. 2012	$p = \{\$2.50 + \$2.00/mi.\}$	\$1.21 Million	\$4.56 Million	1.86 minutes	257.9 thous.
Pricing Policy	Price	Δ Consumer Surplus	Δ Profits	Δ Mean Waiting	$\max(\Delta wait_{it})$
Location Pricing	Sec. 1: $1.10 \cdot p$ Sec. 2: $0.80 \cdot p$ Sec. 3: $1.00 \cdot p$ Sec. 4: $1.35 \cdot p$	\$29.2K (2.40%)	\$78.5K (1.72%)	-0.4 seconds	+109.2 seconds
Time Pricing	7a-9a: $0.70 \cdot p$ 9a-11a: $1.00 \cdot p$ 11a-1p: $1.05 \cdot p$ 1p-4p: $1.10 \cdot p$	\$17.9K (1.47%)	\$68.0K (1.49%)	-0.5 seconds	+ 21.6 seconds
Distance Pricing	0-2 mi.: $0.85 \cdot p$ 2-4 mi.: $1.10 \cdot p$ 4-6 mi.: $1.40 \cdot p$ > 6 mi.: $0.70 \cdot p$	\$76.3K (6.29%)	\$117.4K (2.57%)	-0.3 seconds	+11.1 seconds

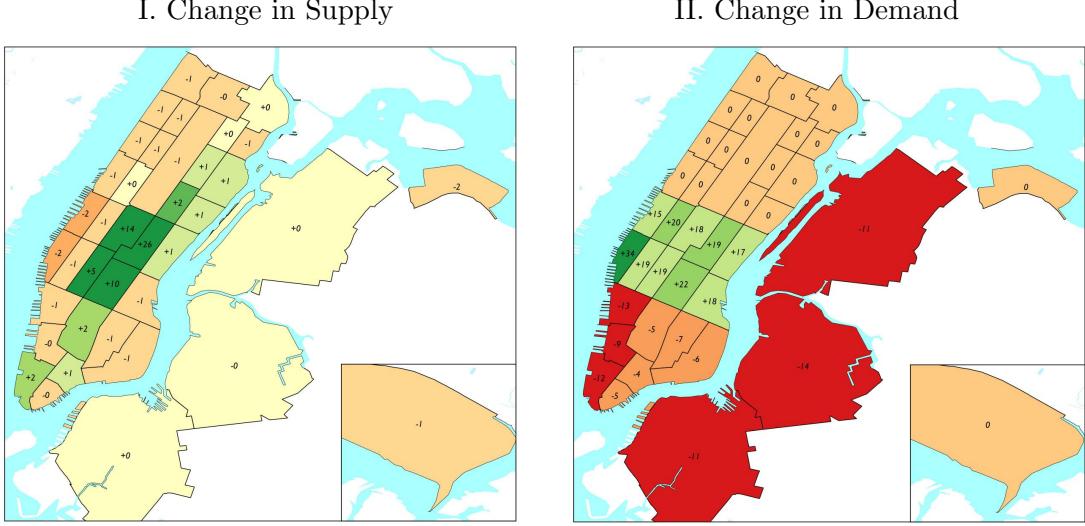
This table shows, for each weekday day-shift (7a-4p), the estimated change in total consumer surplus, total profits, average waiting times and matches, across each counterfactual policy of price flexibility. Each pricing rule shown is a rule that applies four multipliers on the base prices p . The indicated price multipliers are chosen, in increments of 0.05 and across a support of [0, 10], to maximize consumer surplus subject to (1) weakly positive change in profits, (2) weakly negative change in average wait times, and (3) maximum wait time change across *all locations and times* can be no more than 200%, or about 2 minutes.

that a necessary condition for price discrimination to enhance social welfare is that it accompanies an increase in output.⁶⁰

Figure 12 shows the change in supply and demand under location-based pricing, with average changes (across each period) plotted on a map of New York. It shows that under the optimal location pricing, demand changes predictably reflect the new pricing, while the changes in equilibrium vacant taxi supply leave taxis redistributed to the busiest part of Manhattan, in part because of the high propensity of passengers in midtown to take short trips within the same region. Gains are modest, however, as supply and demand (and thus, matches, welfare and profits) drop off in other regions. Figure 13 shows the effect of time-of-day pricing, in Panel I, and distance-based pricing, in Panel II. Panel I shows how time-of-day pricing necessarily trades off benefits to customers earlier in the day versus later in the day, inducing similar tradeoffs as in the spatial pricing policy. Panel II suggests how optimal distance pricing can completely alleviate the time-of-day tradeoff: by charging higher prices for longer trips, taxi supply is increased relative to the baseline case, and by charging lower prices for high-demand short rides, total overall demand is also increased. These changes create improvements in the utilization rates of taxis, leading to greater overall efficiency gains.

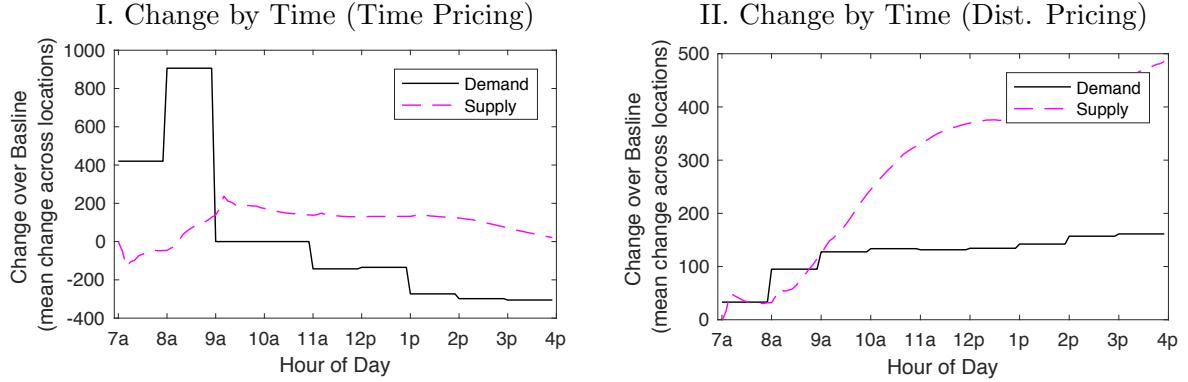
⁶⁰These ideas can be directly traced to Robinson (1933), representing perhaps some of the earliest literature in industrial organization.

Figure 12: Detailed Counterfactual Results: Optimal Location Pricing



This figure shows the equilibrium changes in supply (Panel I) and demand (Panel II) under the optimal location pricing rule (as in Table 7), where changes are computed with respect to the equilibrium allocations under the August 2012 uniform pricing.

Figure 13: Detailed Counterfactual Results: Time and Distance Pricing



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

6.2 Discussion: Comparing Tariff Changes with Real-time Pricing

Under a traditional uniform tariff, equilibrium spatial patterns of supply will reflect distortions stemming from the divergence between the profit incentives for taxis and the welfare maximizing incentives of the planner (or, in this case, the regulator). The results above suggest that there are significant gains to be generated by better aligning these incentives through flexible pricing.

These pricing counterfactuals specifically deal with the notion that there are expected average patterns of supply and demand that can be more efficiently cleared through better pricing policy. These counterfactuals do not speak 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”. Nevertheless, better average pricing of the form indicated in Table 7 will enable the regulator to alleviate some degree of the need for real-time pricing, as supply and demand will be more efficiently allocated across space and time, and in a way that is consistent with dynamic evolution of supply from one period to the next. These results then suggest that if a real-time pricing were to be adopted, the frequency and magnitude of adjustments could be reduced by simply pricing better to the average levels of supply and demand.

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 daily matching data. Using trip observations from New York yellow taxis, I estimate this model to recover the unobservable spatial and inter-temporal distribution of customer arrival rates. By using variation in prices, I further specify and estimate demand curves for each time-of-day and across 39 locations within New York. Having identified demand elasticities, a spatial equilibrium model of taxi supply permits estimates of welfare outcomes and the ability to simulate counterfactual equilibrium outcomes under alternative pricing rules.

I show that welfare attained in the New York market is \$1.48 million per day-shift on a typical weekday, but a more flexible tariff pricing system could enhance welfare on both sides of the market. Implementing an optimal distance-based tariff alone leads to an aggregate welfare gain of at least \$194 million per-shift by better allocating vacant taxis with respect to *current* customer demand as well as the entire dynamic path of supply and demand through 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.⁶¹ Such a mechanism is well within scope of the modern app-based matching platforms such as Uber and Lyft, which permit transparent price communication to both drivers and customers.

⁶¹In other words, offering both additional dimensions on which to set price as well as more categories within each dimension could further boost efficiency.

Bibliography

- ALLEN, J., R. CLARK, AND J.-F. HOUDE (2014): “Search frictions and market power in negotiated price markets,” Tech. rep., National Bureau of Economic Research.
- ALLEN, T. AND C. ARKOLAKIS (2014): “Trade and the Topography of the Spatial Economy,” *Quarterly Journal of Economics*, 1085–1140.
- BRANCACCIO, G., M. KALOUPTSIDI, AND T. PAPAGEORGIOU (2017): “Geography, Search Frictions and Endogenous Trade Costs,” Working Paper.
- BRENNAN, A. (2014): “IBISWorld Industry Report 48533 Taxi & Limousine Services in the US,” Tech. rep., IBIS World.
- BROPHY, T. (2013): “gpsmap: Routine for verifying and returning the attribute table of given decimal degree GPS coordinates,” in *2013 Stata Conference*, Stata Users Group, 14.
- BUCHHOLZ, N., M. SHUM, AND H. XU (2017): “Closed-Form Estimation of Dynamic Discrete Choice Models,” Working Paper.
- BURDETT, K., S. SHI, AND R. WRIGHT (2001): “Pricing and matching with frictions,” *Journal of Political Economy*, 109, 1060–1085.
- BUTTERS, G. R. (1977): “Equilibrium distributions of sales and advertising prices,” *The Review of Economic Studies*, 465–491.
- CAIRNS, R. D. AND C. LISTON-HEYES (1996): “Competition and regulation in the taxi industry,” *Journal of Public Economics*, 59, 1–15.
- CAMERER, C., L. BABCOCK, G. LOEWENSTEIN, AND R. THALER (1997): “Labor supply of New York City cabdrivers: One day at a time,” *The Quarterly Journal of Economics*, 112, 407–441.
- CRAWFORD, V. P. AND J. MENG (2011): “New york city cab drivers’ labor supply revisited: Reference-dependent preferences with rational expectations targets for hours and income,” *The American Economic Review*, 101, 1912–1932.
- DIAMOND, P. A. (1981): “Mobility costs, frictional unemployment, and efficiency,” *The Journal of Political Economy*, 798–812.
- (1982a): “Aggregate demand management in search equilibrium,” *The Journal of Political Economy*, 881–894.

- (1982b): “Wage determination and efficiency in search equilibrium,” *The Review of Economic Studies*, 49, 217–227.
- DIAMOND, R. (2012): “The determinants and welfare implications of us workers diverging location choices by skill: 1980-2000,” *Job Market Paper, Harvard University December*.
- DUFFIE, D., N. GÂRLEANU, AND L. H. PEDERSEN (2002): “Securities lending, shorting, and pricing,” *Journal of Financial Economics*, 66, 307–339.
- (2005): “Over-the-Counter Markets,” *Econometrica*, 73, 1815–1847.
- ERICSON, R. AND A. PAKES (1995): “Markov-perfect industry dynamics: A framework for empirical work,” *The Review of Economic Studies*, 62, 53–82.
- FARBER, H. S. (2005): “Is Tomorrow Another Day? The Labor Supply of New York City Cab-drivers,” *The Journal of Political Economy*, 113, 46–82.
- (2008): “Reference-dependent preferences and labor supply: The case of New York City taxi drivers,” *The American Economic Review*, 98, 1069–1082.
- FLATH, D. (2006): “Taxicab regulation in Japan,” *Journal of the Japanese and International Economies*, 20, 288–304.
- FRECHETTE, G., A. LIZZERI, AND T. SALZ (2017): “Frictions in a Competitive, Regulated Market: Evidence from Taxis,” Working Paper.
- FUJITA, M., P. KRUGMAN, AND A. VENABLES (1999): *The Spatial Economy*, MIT Press.
- GALLICK, E. C. AND D. E. SISK (1987): “Reconsideration of Taxi Regulation, A,” *JL Econ. & Org.*, 3, 117.
- GAVAZZA, A. (2011): “The role of trading frictions in real asset markets,” *The American Economic Review*, 101, 1106–1143.
- HÄCKNER, J. AND S. NYBERG (1995): “Deregulating taxi services: a word of caution,” *Journal of Transport Economics and Policy*, 195–207.
- HALL, R. E. (1979): “A theory of the natural unemployment rate and the duration of employment,” *Journal of monetary economics*, 5, 153–169.
- HONG, H. AND M. SHUM (2010): “Pairwise-difference estimation of a dynamic optimization model,” *The Review of Economic Studies*, 77, 273–304.

- HOPENHAYN, H. A. (1992): “Entry, exit, and firm dynamics in long run equilibrium,” *Econometrica: Journal of the Econometric Society*, 1127–1150.
- HOTELLING, H. (1929): “Stability in Competition,” *The Economic Journal*, 39, 41–57.
- KIYOTAKI, N. AND R. WRIGHT (1989): “On money as a medium of exchange,” *Journal of Political Economy*, 97, 927–954.
- (1993): “A search-theoretic approach to monetary economics,” *The American Economic Review*, 63–77.
- LAGOS, R. (2000): “An Alternative Approach to Search Frictions,” *The Journal of Political Economy*, 108, 851–873.
- (2003): “An analysis of the market for taxicab rides in New York City,” *International Economic Review*, 44, 423–434.
- MASKIN, E. AND J. TIROLE (2001): “Markov perfect equilibrium: I. Observable actions,” *Journal of Economic Theory*, 100, 191–219.
- MELITZ, M. AND C. JAMES (2007): “The Dynamics of Firm-Level Adjustment to Trade,” in *The Organization of Firms in a Global Economy*, ed. by ed. E. Helpman, D. Marin, and T. Verdier, Harvard University Press.
- MORTENSEN, D. T. (1982a): “The matching process as a noncooperative bargaining game,” in *The economics of information and uncertainty*, University of Chicago Press, 233–258.
- (1982b): “Property rights and efficiency in mating, racing, and related games,” *The American Economic Review*, 968–979.
- (1986): “Job search and labor market analysis,” *Handbook of labor economics*, 2, 849–919.
- (1988): “Matching: finding a partner for life or otherwise,” *American Journal of Sociology*, S215–S240.
- MORTENSEN, D. T. AND C. A. PISSARIDES (1999): “New developments in models of search in the labor market,” *Handbook of labor economics*, 3, 2567–2627.
- OHTA, H. AND J.-F. THISSE (1993): *Does Economic Space Matter?*, St. Martin’s Press.
- PISSARIDES, C. A. (1984): “Search intensity, job advertising, and efficiency,” *Journal of Labor Economics*, 128–143.

- (1985): “Short-run equilibrium dynamics of unemployment, vacancies, and real wages,” *The American Economic Review*, 676–690.
- ROBACK, J. (1988): “Wages, rents, and amenities: differences among workers and regions,” *Economic Inquiry*, 26, 23–41.
- ROBINSON, J. (1933): “The theory of money and the analysis of output,” *The Review of Economic Studies*, 1, 22–26.
- ROGERSON, R., R. SHIMER, AND R. WRIGHT (2005): “Search-theoretic models of the labor market: A survey,” *Journal of economic literature*, 43, 959–988.
- ROSEN, S. (1979): “Wages-based Indexes of Urban Quality of Life,” in *Current Issues in Urban Economics*, ed. by ed. P. Mieszkowski and M. Straszheim, John Hopkins Univ. Press.
- SCHALLER, B. (2007): “Entry controls in taxi regulation: Implications of US and Canadian experience for taxi regulation and deregulation,” *Transport Policy*, 14, 490–506.
- SCHMALENSEE, R. (1981): “Output and welfare implications of monopolistic third-degree price discrimination,” *The American Economic Review*, 71, 242–247.
- SCHRIEBER, C. (1975): “The Economic Reasons for Price and Entry Regulation of Taxicabs,” *Journal of Transport Economics and Policy*, 9, 268–79.
- VARIAN, H. R. (1985): “Price discrimination and social welfare,” *The American Economic Review*, 75, 870–875.
- VON THÜNEN, J. (1966): *Isolated state: an English edition of Der isolierte Staat*, Pergamon Press.
- WEINTRAUB, G. Y., C. L. BENKARD, P. JEZIORSKI, AND B. VAN ROY (2008): “Nonstationary Oblivious Equilibrium,” *Manuscript, Columbia University*.
- WINSTON, C. AND C. SHIRLEY (1998): *Alternate route: Toward efficient urban transportation*, Brookings Institution Press.

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.
 - Most of these appear to be erroneous.
5. Drop observations outside of Manhattan (bounded above by 125th st.), LaGuardia Airport and JFK Airport.
 - JFK and LGA airport areas are defined by their corresponding census tract.
6. Drop observations which cannot be mapped to any of the 48 locations summarized in Figure 5.
 - E.g., GPS point falls inside of Manhattan boundaries, but in impossible location such as the East River.

Table A1 shows the incidence of each cleaning criterion.

A.2 Map Preparation

The following graphics show how raw GPS data points are converted to locations. Figure A1 is a graphic showing raw census tracts. 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 A2 shows the origin of each trip in a 10-percent sample of TLC data.

Table A1: Data Cleaning Summary

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

Final Data Set: 24,432,914 observations

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

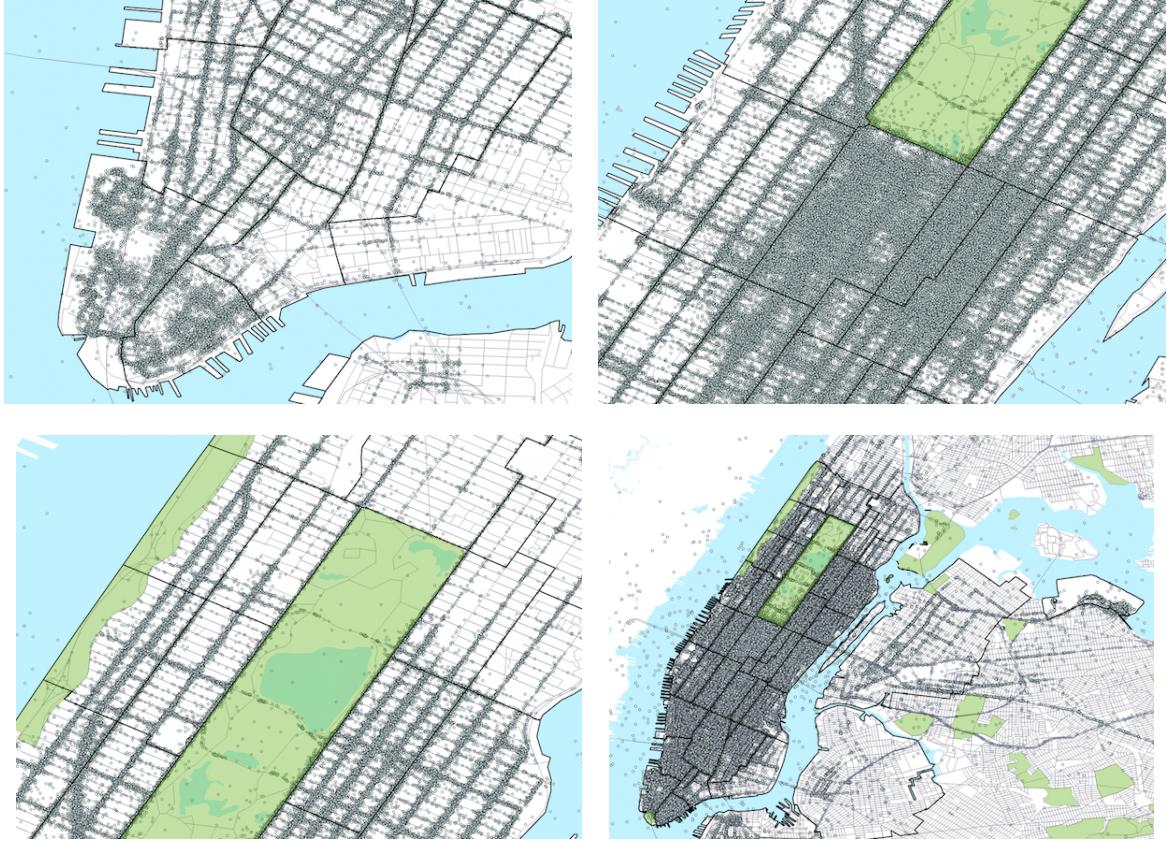
It can be seen that trip origins are most heavily concentrated around major streets, particularly north-south and diagonal thoroughfares in the north, with more scattered origin points in lower Manhattan and Midtown Manhattan. The densest neighborhoods are clearly those in Midtown. I have grouped census 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.

Figure A1: New York Census Tracts Overview



This figure shows a subset of New York census tracts around Manhattan.

Figure A2: Mapping GPS points to Locations

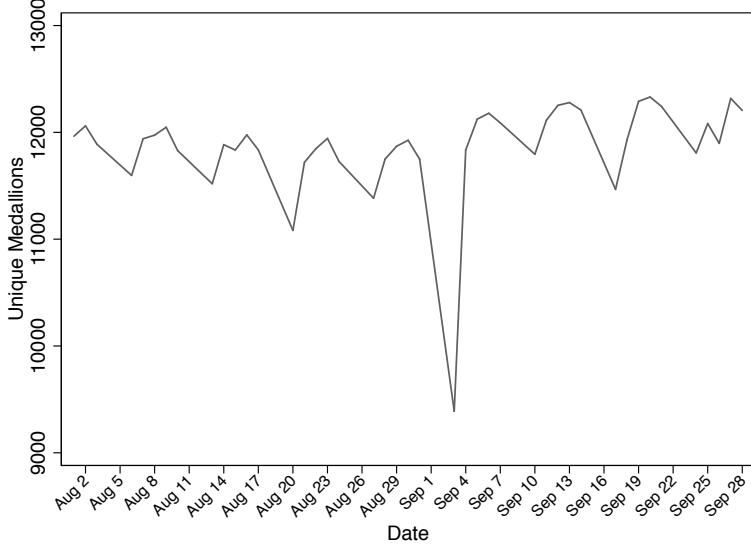


This figure shows TLC data for a 10 percent sample of taxi trips taken in August, 2012. Each dot on the map is the GPS origin of a trip.

A.3 Medallion Counts

Figure A3 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 A3: 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.4 Details on State Transitions

The combined set of transitions forms an aggregate transition kernel that defines the law-of-motion, given by $Q(\mathcal{S}^{t+1}|\mathcal{S}^t) = \nu(\mathbf{v}_e^{t+1}|\mathbf{v}_e^t, \mathbf{M}^t, \mathbf{m}^t) + \mu(\mathbf{v}_v^{t+1}|\mathbf{v}_v^t, \sigma^t)$.

Let the following objects be defined:

\mathbf{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). \mathbf{m}^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). \mathbf{M}_e^t be the $(L+K) \times (L+K)$ vector of one-period transition probabilities of customers from all locations $\{1, \dots, L\}$ and all in-between positions $\{1, \dots, K\}$. 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). $\mathbf{m}^{t-\tau_{ji}}$ describes how many drop-offs will occur in period t , which is the number of matches made in each pick-up location in τ_{ji} prior periods, and transition matrix $\mathbf{M}_e^{t-\tau_{ji}}$ 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:

$$\mathbf{v}_e^{t+1} = ((\mathbf{v}_e^t + \mathbf{m}^t) \times \mathbf{M}_e^t) - (\mathbf{m}^{t-\tau_{ji}} \times \mathbf{M}^{t-\tau_{ji}}). \quad (13)$$

Next, I define the state transitions of vacant taxis. Let \mathbf{v}_v^t be the $(L + K) \times 1$ vector of vacant taxis in all search locations and in-between locations $\{1, \dots, (L+K)\}$. 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 \mathbf{v}_v^t be the $(L + K) \times (L + K)$ vector of one-period transition probabilities of vacant taxis from all locations $\{1, \dots, L\}$ and all in-between positions $\{1, \dots, K\}$. 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:

$$\mathbf{v}_v^{t+1} = (\mathbf{v}_v^t - \mathbf{m}^t) \times \boldsymbol{\sigma}^t. \quad (14)$$

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

A.5 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 $i \in \{1, \dots, L\}$ to any location $j \in \{1, \dots, L\}$ 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 $t \in \{1, \dots, T\}$, where backwards moves update continuation values and forwards moves update transition paths. The algorithm terminates when all transition paths and continuation values are self-filling and consistent with equilibrium. Below I provide an outline of the taxi equilibrium algorithm.

The Taxi Equilibrium Algorithm begins with an initial guess of the industry state \mathcal{S}_0 (i.e., the number of all vacant taxis across locations and time-of-day).⁶² With \mathcal{S}_0 as well as observations of the empirical distributions of taxi-passenger matches, $\tilde{\mathbf{m}}$, I compute value functions $V_i^t(\mathcal{S}_0; \tilde{\mathbf{m}})$ 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 7. Next, I use the computed policy

⁶²Note that the algorithm will take as given the parameters γ and σ_ε . These are solved for in a second stage, as an “outer loop” to the algorithm.

Algorithm 1 Taxi Equilibrium Algorithm

```

1: Load empirical matches  $\{\tilde{m}_{ij}^t\}$  and  $\{\tilde{m}_i^t\} = \{\sum_j \tilde{m}_{ij}^t\}$ 
2: Fix parameters  $\lambda = \lambda_0$ ,  $\gamma = \gamma_0$ 
3: Set counter  $k = 0$ 
4: Guess  $\mathcal{S}_0^T$  and compute  $V^T(\mathcal{S}_0^T, \lambda^T)$ 
5: for  $t = T - 1$  to 1 do                                 $\triangleright$  Backwards Iteration
6:   Guess  $\mathcal{S}_0^t$  and compute  $V^t(\mathcal{S}_0^t, \lambda^t)$ 
7:   for  $t = 1$  to  $T - 1$  do                       $\triangleright$  Fwd. Iteration to  $T$  for each step back
8:     Derive choice-specific value functions  $W_i^t(j, \mathcal{S}^t)$  for all  $t, i, j$ .
9:     Find policy fcts.  $\sigma_k^t(W_k^{t+1})$  to determine vacant taxi transitions
10:     $\sigma_0^t$  and  $m_{ij}^t$  imply transition to  $\tilde{\mathcal{S}}^{t+1}$ 
11:    Update next period state  $S_{k+1}^{t+1} \leftarrow \tilde{\mathcal{S}}^{t+1}$ 
12:    Update next period continuation values as  $V^{t+1}(\mathcal{S}_{k+1}^{t+1}, \tilde{\mathbf{m}}^t)$ 
13:     $k \leftarrow k + 1$ 
14:  end for
15: end for
16: repeat
17:   Iterate on steps 6 to 15
18: until  $|V_k^t - V_{k-1}^t| \leq \epsilon \quad \forall t$ 

```

functions and, starting at time $t = 1$ at \mathcal{S}_0^1 , I forward simulate the optimal transition paths and update the initial state for $t = 2, \dots, T$, resulting in a new guess of the state, \mathcal{S}_1 . With \mathcal{S}_1 , I again combine the same observations $\tilde{\mathbf{m}}$ to update value functions $V_i^t(\mathcal{S}_1; \tilde{\mathbf{m}})$. This process repeats until value and policy functions converge.

A.5.1 Initial conditions

Recall that \mathcal{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, \mathcal{S}_0^t , is assigned by allocating the exogenous total number of taxis according to the empirical distribution of matches.⁶³ As the algorithm runs, each vector \mathcal{S}_0^t for $t \geq 2$ is updated as $t - 1$ transitions are computed given the $t - 1$ initial state and value functions for $t, t + 1, \dots, T$. Only one term, \mathcal{S}_0^1 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.⁶⁴ Nevertheless, by starting the equilibrium algorithm at 6:00am,

⁶³Recall the total number of taxis equals 11,500, as discussed in Chapter 4.1.

⁶⁴Recall from Figure 4 and earlier discussion that the data do not allow for distinguishing for 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.

Table A2: Alternative Initial Conditions

Initial Condition	Δv_i^t (mean)	% Δv_i^t (mean)
Baseline	0	0
Uniform	2.10	0.0175
Edge	4.52	0.0336

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.⁶⁵

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 A2 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.⁶⁵ 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.6 Proofs

Proof of Proposition 4.1 The expected number of matches $E[m(v, \lambda)|v]$ is one-to-one in λ .

Proof. First decompose $E[m|v, \lambda]$ 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:

⁶⁵Boundary 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.

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

Where the second step follows from the pmf of the Poisson distribution. It is straightforward to show from this step that $E[m(v, \lambda)|v]$ is strictly increasing in λ . Since a strictly increasing function is one-to-one, $\{m(\cdot, \cdot), v\} \leftrightarrow \lambda$ is one-to-one. \square

Proof of Equilibrium Uniqueness

Proof. To show that there is a unique equilibrium to the dynamic oligopoly game, I need to show that players' best response curves intersect only once at every period in time. First note that since any taxi driver plays against the expected distribution of his competitors, and since all transitions are taken to be deterministic with respect to the taxi drivers' decision problem, this game can be viewed as a two-player game: the agent versus his competition. Further, all agents are atomistic, so that no single individual's action has any influence on competitors' strategies. Thus I only need to show that a drivers' best response function (i.e., which location to search in given the distribution of competitors) intersects his competitors' action (i.e., the aggregated policy function) at one point. To do this I will show that a taxi's best response function is strictly decreasing in the level of competition. Specifically, I want to show that for a taxi in location i , his best response curves (given by policy function $\sigma_i(j_a|\mathcal{S}^t)$) are strictly decreasing in v_i^t , the payoff-relevant component of the state (recall \mathcal{S}^t also includes the in-transit status of all employed and vacant taxis who are not actively searching in period t).

Recall that this policy function is given by:

$$\sigma_i(j_a|\mathcal{S}^t) = \frac{\exp(W_i^t(j_a, \mathcal{S}^t)/\sigma_\varepsilon)}{\sum_k \exp(W_i^t(j_k, \mathcal{S}^t)/\sigma_\varepsilon)}.$$

It is straightforward to see that $\frac{\delta\sigma_i(j_a)}{\delta v_{j_a}^t} < 0 \Leftrightarrow \frac{\delta W_i(j_a)}{\delta v_{j_a}^t} < 0 \Leftrightarrow \frac{\delta V_{j_a}}{\delta v_{j_a}^t} < 0$. I will therefore show that value functions are strictly decreasing functions of the state variable. To do this, begin at the

last period, $t = T$. In this period, continuation values are all equal to zero, so that the only profit may be earned from successful search within period T . The value function for location i in the terminal period may be rewritten as:

$$V_i^T(\mathcal{S}) = \mathbb{E}_{p_i} [p_i(u_i^T, v_i^T) | \lambda_i^T] \left(\sum_j M_{ij}^T \cdot \Pi_{ij} \right).$$

The state of competition is summarized by the number of vacant cabs in location i , v_i^T . In Appendix A.3, I prove that $\mathbb{E}[m_i^t | v_i^t, \lambda_i^t] = v \cdot (1 - e^{-\frac{\lambda}{\alpha v}})$. Thus, $p_i(v | \lambda) = (1 - e^{-\frac{\lambda}{\alpha v}})$, a strictly downward sloping function of v . Since the expected profits associated with a fare is independent of the state, this directly implies that $V_i^T(\mathcal{S})$ is downward sloping in the relevant component of the state, v_i^T .

Now let's turn our attention to a generic period t . The goal is to show that V_i^t is decreasing in v_i^t as long as continuation values are also decreasing. Since this is true in period T , by induction this will prove that V_i^t is decreasing in v_i^t for all t . To proceed, suppose $V_j^t(\mathcal{S})$ is decreasing in v_j^k for all j and for all $k > t$. For clarity I display equation 4 below:

$$\begin{aligned} V_i^t(\mathcal{S}) = \mathbb{E}_{p_i | \lambda_i, \mathcal{S}^t} & \left[p_i(u_i^t, v_i^t) \overbrace{\left(\sum_j M_{ij}^t \cdot (\Pi_{ij} + V_j^{t+\tau_{ij}}) \right)}^{\text{Exp. Value of Fare}} + \right. \\ & \left. (1 - p_i(u_i^t, v_i^t)) \cdot \underbrace{\mathbb{E}_{\varepsilon_{j,a}^{t+1}} \left[\max_{j \in A(i)} \left\{ V_j^{t+\tau_{ij}} + \mathbb{I}_{[j=i]} \gamma - c_{ij} + \varepsilon_{j,a} \right\} \right] \Big| \lambda_i^t \right]. \end{aligned} \quad (15)$$

Since we have assumed that $V_j^{t+\tau_{ij}}(\mathcal{S})$ is decreasing, it is easy to see that if the expected value of a fare is greater than the expected value of a vacancy over the next period, then any decrease in the probability of finding a fare will decrease V_i^t . The assumption that taxis can opt to not search in any location ensures this condition will always hold; taxis may always choose to exercise the option to remain vacant and thereby forgo the opportunity to profit in any period.⁶⁶ Thus an increase in v_i^t decreases the probability of matching, p_i , which thereby decreases V_i^t .

I showed above that value functions are decreasing in the level of competition. This directly implies that policy functions are decreasing functions of the competitor's state. Since competitors'

⁶⁶For clarity, this option value is not written. Under this assumption, value functions are more precisely written with an additional max operator in front, i.e. $V_i^t(\mathcal{S}) = \max([\text{Exp. Value of Vacancy}], \tilde{V}_i^t(\mathcal{S}))$, where \tilde{V} is given by equation 4, allowing drivers to choose vacancy with probability one instead of facing a probability of getting a ride. In simulations, taxis very rarely exercise this option.

Table A3: Data and Simulation Moments

Moment	Data Average	Simulation Average
Total Vacant Wait Time (min.)	155.21	184.95
Total Employed Distance Travelled (mi.)	38.37	35.74
$\Pr(\text{pickup in } i \mid \text{drop-off in } i)$.3714	.3809
Match Probability (Section 1)	.4205	.3809
Match Probability (Section 2)	.4912	.4543
Match Probability (Section 3)	.3842	.2214
Match Probability (Section 4)	.2477	.4735

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

policies are independent of an atomistic taxi's actions, best response curves will intersect only once. Therefore equilibrium is unique.

□

A.7 Estimation and Simulated Moments

I identify σ_ε and γ 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 6. Note that these moments will depend at least in part on these parameters; σ_ε reflects how much of a taxi driver's location choice depends on observable features within the model. A high value of σ_ε 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. γ 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 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.⁶⁷ Table A3 displays each simulation moment compared with its observed value.

⁶⁷For 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.8 Details on Estimating α_r

Given parameter values σ_ε and γ , I use matches date $\{\tilde{\mathbf{m}}\}$ and the TEA procedure to numerically solve for the equilibrium state, $\mathcal{S} = \{v_i^{t*}(\tilde{\mathbf{m}})\}$. Proposition 4.1 shows that, given a level of taxis v_i^t , the matching function can be inverted to solve for $\frac{\lambda_i^t}{\alpha_r}$ in location i (within region r) and time t . Next, I use an analytic expression of the variance of matches, given by equation 10. This function depends on both α_r as well as the ratio $\frac{\lambda_i^t}{\alpha_r}$. From here I set up the following estimator:

$$\alpha_r = \arg \min_{\alpha} \sum_{i \in \mathcal{R}_r, t} \left(\text{Var}_d(\tilde{m}_{i,d}^t) - (v_i^{t*})^2 e^{-2\frac{\lambda_{it}}{\alpha_r} \frac{1}{v_{it}^{*2}}} \left(e^{\frac{\lambda_{it}}{\alpha_r^2} \frac{1}{v_{it}^{*2}}} - 1 \right) \right), \quad (16)$$

where \mathcal{R}_r denotes the set of locations within region r , and $\tilde{m}_{i,d}^t$ 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.

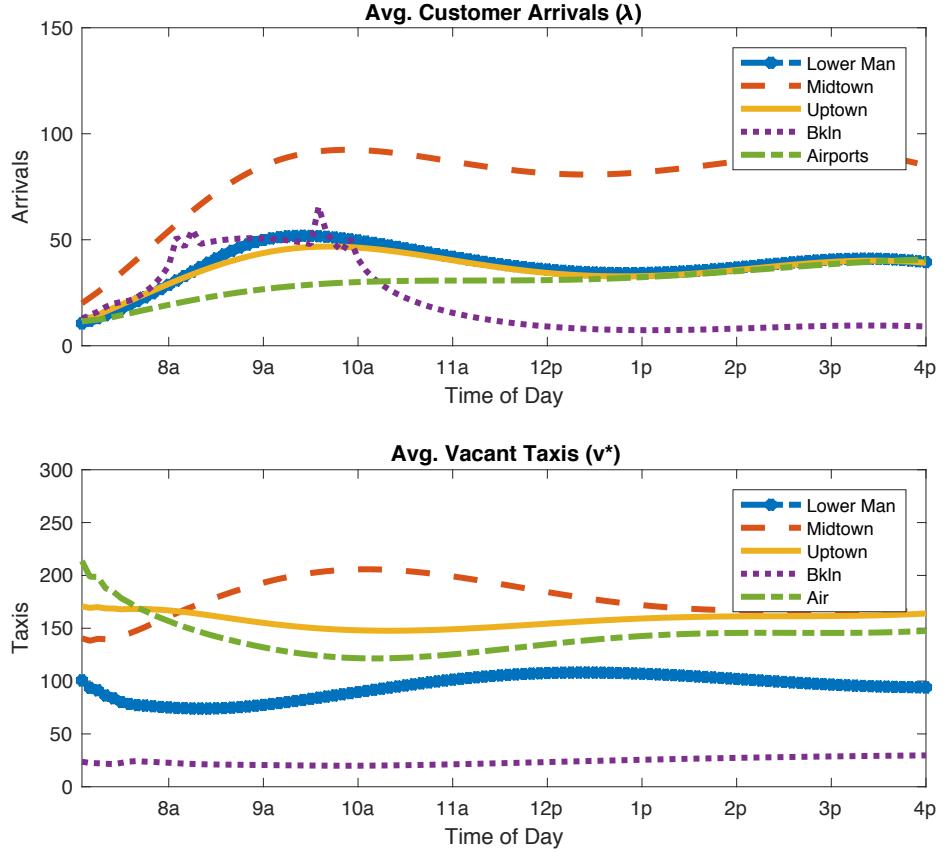
A.9 Detailed Estimation Results

Figure A4 shows aggregate supply and demand results, summing all 39 locations into the five regions corresponding to Figure 6. The results above demonstrate that while taxi supply maintains some coverage across all locations throughout the day, there are intra-day trends in spatial availability and demand. Spatial mismatch is evident, as the relative proportions of supply and demand are not the same across each region.

Figures A5 and A6 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 A7 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.

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



This figure depicts the equilibrium spatial distribution of taxis and arrival rates of customers across the Five Regions shown in Figure 6. 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.

A.10 Welfare Calculation

Consumer welfare is computed by integrating under 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:

Figure A5: Detailed Estimates of v_i^t and λ_i^t (1)

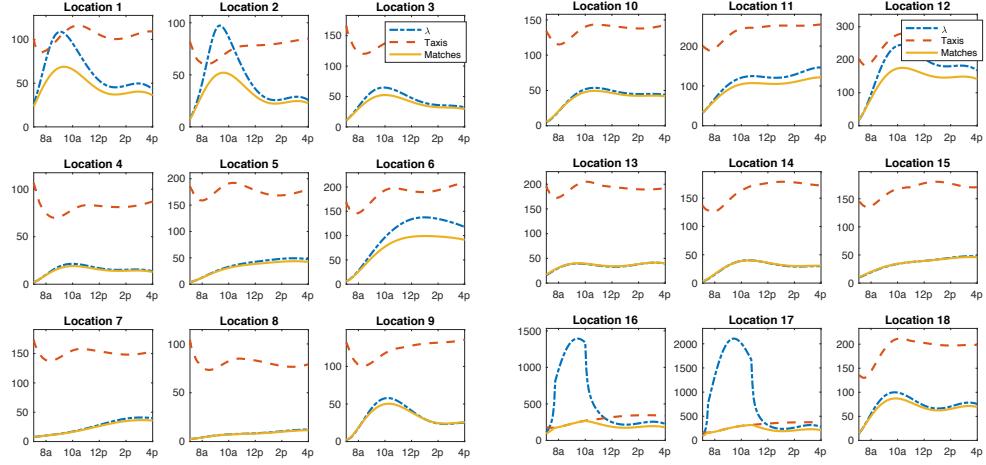
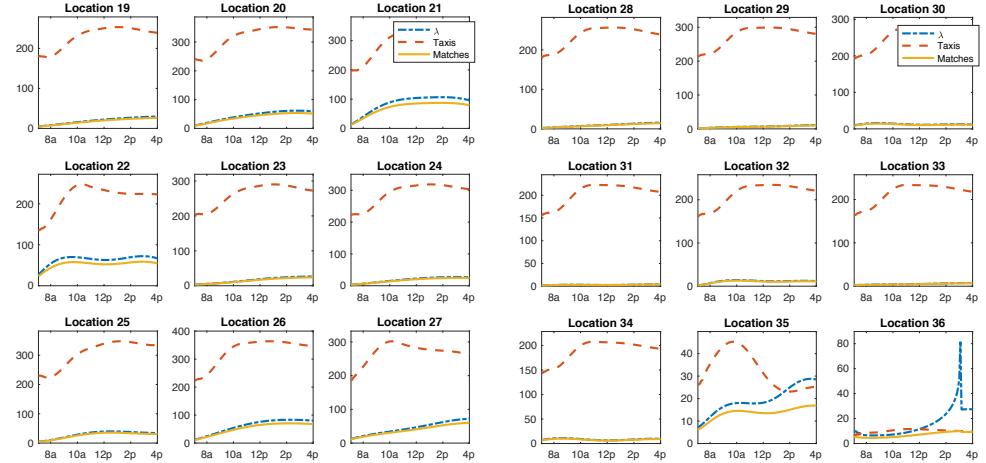


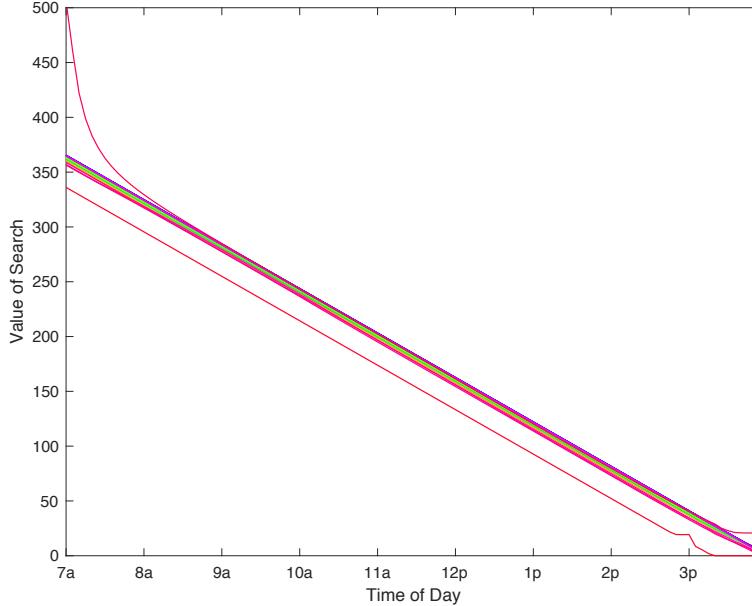
Figure A6: Detailed Estimates of v_i^t and λ_i^t (2)



$$W_{ijt}(m_{ij}^t, \hat{\lambda}_{ij}^t, p_{ij}, \beta) = \underbrace{\frac{m_{ij}^t(\hat{\lambda}_{ij}^t, v_i^t(\lambda))}{\hat{\lambda}_{ij}^t(p_{ij})}}_{\text{frac. successful matches}} \cdot \underbrace{\left(\frac{\alpha_{1,r,a}}{\alpha_{1,r,a} + 1} \cdot e^{-\frac{1}{\alpha_{1,r,a}}(\alpha_{0,r,a} + \delta_{tr} + \xi_{ir})} \cdot \hat{\lambda}_{ijt}(p_{ij})^{\alpha_{1,r,a}^{-1} + 1} - \hat{\lambda}_{ijt}(p_{ij}) \cdot p_{ij} \right)}_{\text{total available surplus at price } p_{ij}}, \quad (17)$$

where $\alpha_{0,r,a}$, $\alpha_{1,r,a}$, δ_{tr} , and ξ_{ir} are the estimated parameters of the demand system, and where $\hat{\lambda}_{ijt}$ is the predicted level of demand (the mean number of customer arrivals given price p_{ij}), and

Figure A7: Equilibrium Value Functions



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.

$v_i^t(\boldsymbol{\lambda})$ 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:

$$W_{ijt}^{\text{taxi}}(m_{ij}^t, \hat{\lambda}_{ij}^t, p_{ij}, c_{ij}) = \underbrace{\frac{m_{ij}^t(\hat{\lambda}_{ij}^t, v_i^t(\boldsymbol{\lambda}))}{\hat{\lambda}_{ij}^t(p_{ij})}}_{\text{frac. successful matches}} \underbrace{\left(\hat{\lambda}_{ijt}(p_{ij}) \cdot (p_{ij} - c_{ij}) \right)}_{\text{trip revenues at price } p_{ij}} \quad (18)$$

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

B Online Appendix

This section presents additional details on the counterfactual exercises.

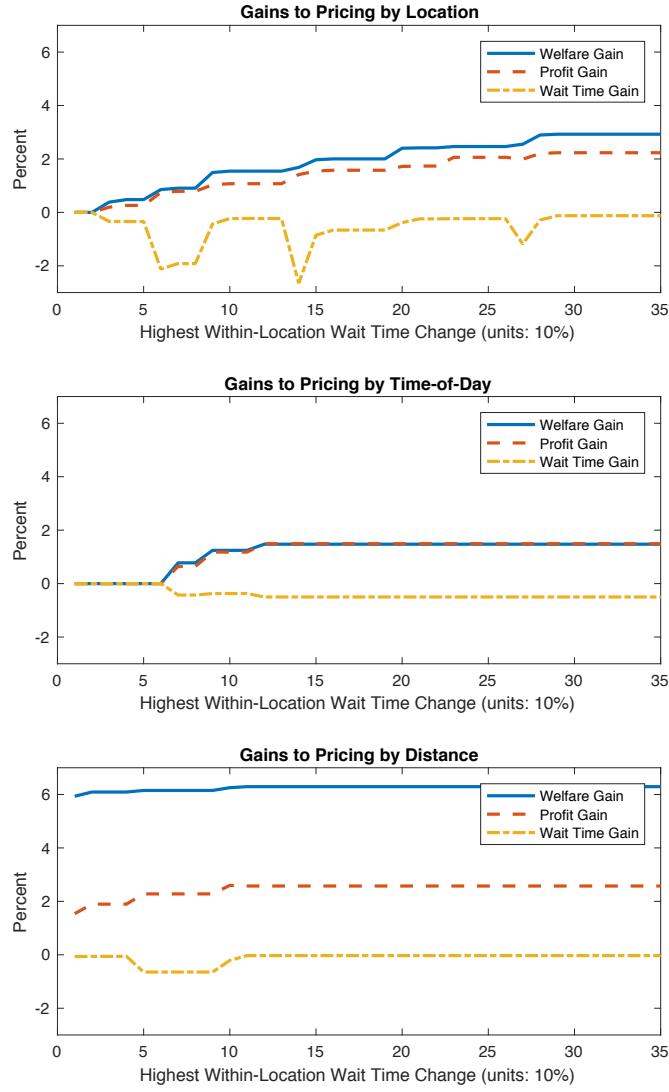
B.1 Consumer Wait Time Heterogeneity

The search times experienced by taxis is internalized into the drivers' dynamic optimization problem, but this is not true of passengers. I compute passenger demand for each trip as the average

number of trips demanded in each period, assuming customers will wait for one period of five minutes. The expected waiting time in each location and time is then computed as $(1 - p^D(\text{match})) \cdot 5$ minutes. When searching cross counterfactual prices, I allow for the estimated demand functions in each i, j, t cell to predict the new arrival rates of demand for every i to j trip at time t . Taxi drivers then re-solve for equilibrium policies, giving rise to a new predicted number of matches and implying a new expected waiting time for customers.

Figure B1 depicts the highest attainable consumer welfare gains, change in total shift profits of taxi drivers, and change in aggregate waiting time to each pricing experiment, measured against an increasing tolerance for impacting particular neighborhoods. This set of figures tell us that the more flexibility we have to reallocate relative supply and demand, to the detriment of some particular places and times-of-day, the more we can enhance efficiency of the entire system. In fact, consumer welfare and taxi profits tend to move in the same direction, as the scarce supply of taxis is being used more efficiently to service more matches. Moreover, distance-based pricing offers the uniformly highest level of efficiency gains, and are less sensitive to increasing the variance of wait times, suggesting this should be a fairly uncontroversial policy to implement when no location suffers from increased frictions.

Figure B1: Aggregate Efficiency Gains vs. Variance of Waiting Times



This figure shows the maximum consumer surplus gains generated by allowing for greater negative impact to location-time-specific wait times, but maintaining the constraint that average wait times are reduced and total profits are increased. Each point represents a different solution to the optimal pricing policy under this increasingly relaxed constraint on location-time waiting times. Thus the affected locations and pricing rules are changing as the constraint is relaxed (i.e., increasing in the x-axis).