Information Economics and Policy 57 (2021) 100941

Contents lists available at ScienceDirect

Information Economics and Policy

journal homepage: www.elsevier.com/locate/iep

The geography of ridesharing: A case study on New York City☆

Chungsang Tom Lam[a], Meng Liu[b], Xiang Hui[b,*]

[a] Department of Economics,320L Wilbur O. and Ann Powers Hall, Clemson, SC, 29634, USA
[b] Olin Business School 1 Brookings Drive, Knight Hall 406St. Louis, MO 63130, USA

ARTICLE INFO

Article history:
Received 7 January 2021
Revised 29 June 2021
Accepted 25 July 2021
Available online 29 July 2021

Keywords:
Ridesharing
Geography
Transportation

ABSTRACT

Despite the popularity of ridesharing, there is limited empirical evidence on how ridesharing activities differ across regions with different levels of accessibility and the implication for consumers. In this paper, we study the market for rides across New York City neighborhoods. We construct a novel data set that contains massive API queries on route-specific estimates of pricing, wait time, and travel time of Uber, Lyft, and the public transit. After linking this data with actual trip records of taxis, Uber, and Lyft, we document a strong pattern that ridesharing has a larger market share relative to taxis in neighborhoods with lower accessibility, defined either in terms of geographic distance to Midtown Manhattan or "economic distance" to job opportunities. Next, we estimate a discrete-choice model of demand for rides and interpret the geography of ridesharing through the lens of the model. We find that consumer surplus from ridesharing varies drastically across geography: passengers that are 5 to 15 miles (resp. more than 15 miles) from Midtown experience a 60% (resp. 19%) larger consumer surplus relative to passengers that are within 5 miles from Midtown. Over half of these gains comes from reduced wait time.

1. Introduction

Ridesharing platforms have gained popularity around the world in recent years and, alongside, scrutiny from policy makers. Policies and regulations about ridesharing across cities seem to be predominantly a binary decision: either ridesharing is allowed or it is banned. However, when policy makers ignore the potential distributional impacts of ridesharing across geographies, a one-size-fits-all policy will benefit some neighborhoods while being detrimental to others. Studying the geography of ridesharing and its implication for consumers helps us understand *where and when* ridesharing could benefit consumers, therefore creating opportunities for designing nuanced policies tailored to different geographies.

This paper studies how ridesharing affects consumer welfare across neighborhoods in a metropolitan setting. We aim to answer two research questions: First, how do ridesharing activities differ across neighborhoods with different levels of accessibility? Second, how might this geography translate into consumer welfare in different neighborhoods? To answer these questions, we construct a comprehensive data set that describes the market of rides in New York City (NYC hereafter). Specifically, we gather massive API queries of Uber, Lyft, and the NYC public transit, which include their pricing, wait time, and travel time. We link this data with actual trip records of taxis, Uber, and Lyft published by NYC Taxi and Limousine Commission (TLC hereafter). We further enrich the data by conducting a field collection of around 70,000 historical trip records from a sample of Uber and Lyft drivers. The resulting data set provides a comprehensive picture of the market that allows for analysis at the route level in real time.

We first summarize and visualize geographic patterns in ridesharing and taxi rides. We find that for both ridesharing and taxis, the total number of pickups is smaller in less accessible regions, where accessibility is defined either in terms of geographic distance to the city center or "economic distance". The geographic distance is calculated as the distance between the pickup location and Manhattan Midtown, and economic distance is measured by the neighborhood's total number of jobs accessible within one hour's commute by public transit (Kaufman et al., 2014). However, the rate at which the number of pickups decreases with respect to distance, which we will refer to as the trip elasticity of distance henceforth, is smaller than that of taxis. Consequently, the market share of ridesharing relative to that of taxi is higher in low accessibility neighborhoods under both definitions of distance. In addi-

☆ Lam and Liu are co-first authors and Hui is the second author of the paper. The authors would like to thank Erik Brynjolfsson, Babur De los Santos, Dean Eckles, Chiara Farronato, Andrey Fradkin, Avi Goldfarb, Daniel Miller, Catherine Tucker, Joel Waldfogel, Patrick Warren for their invaluable comments and suggestions. We also thank conference and seminar participants at Stanford Workshop at Marketplace Innovation, Marketing Science Conference, Toulouse Conference on the Digital Economy, AEA, MIT, WashU, Clemson, and Erasmus University. This study is not sponsored by Uber, Lyft, or any other organizations.

* Corresponding author.
E-mail addresses: tomlam@clemson.edu (C.T. Lam), mengl@wustl.edu (M. Liu), hui@wustl.edu (X. Hui).

https://doi.org/10.1016/j.infoecopol.2021.100941
0167-6245/

tion, we further highlight the geography of ridesharing by using green boro taxis as a benchmark, which follow the same pricing rule as yellow medallion taxis but are allowed to pick up passengers only outside Manhattan Core[1], i.e., predominantly low accessibility neighborhoods. We find that the market share of ridesharing remains larger than that of green taxis in low accessibility neighborhoods, despite that green taxis are designed to serve low accessibility neighborhoods.[2]

The geographical distribution of market shares is an equilibrium outcome, driven by both consumer preference toward various transportation modes and drivers' decisions on when and where to provide the service. We provide empirical evidence that is consistent with a high degree of matching friction between taxi drivers and passengers, which likely gives rise to the low presence of taxi pickups in low accessibility neighborhoods. First, we show that low accessibility neighborhoods more frequently exhibit excess demand (the incidence of more people needing a ride than the number of available drivers in the area) instead of insufficient demand for taxis. Second, the trip elasticity of distance is larger for yellow taxis than for green taxis in low accessibility regions, even though both taxi services face the same demand due to their identical pricing rules outside of the Manhattan Core. This difference suggests that the matching friction in low accessibility neighborhoods relative to the Manhattan Core is even higher for yellow taxis than for green taxis.

The matching friction present among taxis may be alleviated by ridesharing through technology-enabled matching, incentives created by dynamic pricing, or both. Ridesharing platforms seem to reduce this information gap in low accessibility neighborhoods mainly through the use of advanced matching algorithms: According to our data, surge pricing is rare in low accessibility neighborhoods, although it is more frequently activated in Manhattan.

The second part of the paper aims at translating the observed geography to consumer welfare, leading us one step closer to answering relevant policy questions. We model consumer's decision as a discrete choice problem between various competing transportation modes for a given route that is defined by a particular origin-destination-time combination. We use the logit framework to map the observed rides into consumer preference toward price, wait time, as well as other observed and unobserved characteristics. In the demand estimation, endogeneity arises when the price of a region is correlated with the unobserved demand condition in the same region, conditional on service type, route, and time fixed effects. To obtain a consistent estimate of price elasticity, we use an instrumental variable (IV) that affects the market share of a transportation mode on a given route *only through* its impact on price, but it is otherwise uncorrelated with local demand shocks. We construct such an IV using a design feature of the ridesharing Apps that were present during our sample period: passengers see and commit to a surge multiplier on Uber and Lyft Apps *prior to* entering their destination on their phones.[3] Exploiting this feature, we use the surge prices of trips into the focal zone to instrument for price in this zone. On the one hand, because these ridesharing Apps do not possess information on where the consumer is going, they cannot adjust their surge multiplier according to the consumer's destination, which creates the uncorrelatedness of surge prices at the origin and the underlying demand condition at the destination. On the other hand, surge pricing at the origin will affect the price in the focal zone through its effect on the number of available drivers in the focal zone after their drop-offs.

Using this instrument, we identify heterogeneous consumer preference toward price and wait time. The parameter estimates show that consumers prefer lower price and less wait time. Also, the elasticity estimates across regions and different times of the day are consistent with the intuition that people are less price-sensitive and care more about short wait time during rush hours and in destinations with more offices. The fact that consumers prefer less wait time in certain locations during certain times of the day suggests that ridesharing platforms may increase consumer surplus if the service could effectively reduce wait time in these regions.

We use a highly stylized model to characterize taxi supply, where the equilibrium predictions are consistent with the data. Through the lens of the estimated demand and supply models, we translate the geography of ridesharing to the geography of consumer surplus. Conceptually, we adopt the idea of compensating variation and estimate how much consumers should be compensated if Uber and Lyft were to be removed from the market in order to maintain the same level of utility for consumers in different regions. We find that consumer surplus from ridesharing varies drastically across geography: passengers that are 5 to 15 miles (resp. more than 15 miles) from Midtown experience a 60% (resp. 19%) larger consumer surplus relative to passengers that are within 5 miles from Midtown. Additionally, over half of these gains comes from reduced wait time for getting a ride.

The distributional results on the relative market share and the uneven consumer gains from ridesharing have important implications for public policy makers. The results suggest that the impact of ridesharing can be highly uneven across regions with different levels of accessibility. It is important to note that metropolitan areas in the United States exhibit significant geographical disparity in transit access (Tomer et al., 2011; Owen and Murphy, 2018), even for New York City, which has the nation's best public transit system. Its neighborhoods at the 90th percentile in accessibility have access to 18 times more jobs via public transit within an hour than the neighborhoods at the 10th percentile in accessibility (using data provided by Kaufman et al. (2014)). This geographical disparity is unlikely to change in the near future, because it depends on long-run factors such as city planning and re-designing of public transit networks. Taxis are another important component of metropolitan transit systems, yet we show that they are highly concentrated in the dense areas, even for taxis that are specifically designed to pick up passengers in less accessible regions, due to the lack of technology that provides real-time information of demand.

Based on our findings, urban and transportation planners should recognize the important role of ridesharing in connecting neighborhoods with low accessibility and disproportionately create incentives for ridesharing in these neighborhoods, holding the social costs constant. A more inclusive transit network is important because mobility and proximity to markets could affect individuals' labor market outcome and overall well-being, especially those of disadvantaged demographic groups and individuals without car ownership (e.g., Ihlanfeldt and Sjoquist, 1990, Ihlanfeldt and Sjoquist, 1991, Holzer et al., 1994, Hering and Poncet, 2010, among others).

1.1. Related literature and contribution

Our paper contributes to three literature strands. First, it contributes to the growing literature on estimating the impact of ridesharing technology on consumers. The closest paper to ours is

[1] Manhattan Core is defined as the part of Manhattan below the north edge of Central Park.

[2] This finding is consistent with Mitchell L. Ross' argument: "(green taxis) basically cluster at transit and retail hubs because they are more likely to find passengers there than if they cruise the streets they are authorized to cruise." https://www.nytimes.com/2018/09/03/nyregion/green-cabs-yellow-uber.html

[3] After our sample period, this old version of Uber was updated to upfront pricing, and now riders need to input their destination to get a fixed price. Lyft also went through a similar design change.

(Cohen et al., 2016), which uses internal data from Uber and exploits a series of regression discontinuity designs (RDD) around the surge multiplier cutoffs and estimate a consumer surplus of $1.6 per dollar spent on Uber. Our paper is different in two ways. First, our goal is to understand the distributional gain in consumer surplus across regions with different levels of accessibility. Second, instead of focusing on changes in short-run consumer surplus from Uber, we allow for consumer substitution among transportation alternatives, and study consumers' substitution pattern towards taxi and public transit when ridesharing platforms are permanently removed from the market.

Researchers have also examined other important welfare topics associated with ridesharing and the sharing economy in general, such as flexible work arrangements (Chen et al., 2017; Hall and Krueger, 2016; Hall et al., 2017), routing efficiency and driver moral hazard (Cramer and Krueger, 2016; Liu et al., 2018; Wang et al., 2019), reduced matching friction (Buchholz, 2015; Frechette et al., 2016; Zhang et al., 2019), reduced drunk driving (Greenwood and Wattal, 2017), local consumption patterns (Zhang and Li, 2017), traffic congestion (Li et al., 2016), and the use of public transit (Hall et al., 2018; Babar and Burtch, 2017). We add to this literature by analyzing the role of ridesharing in inclusive mobility.

Second, this paper contributes to the research on geographical impact in transportation in U.S. urban areas. There is a large literature that studies the disparity in transportation in U.S. urban areas, which finds mobility inequality across U.S. metropolitan areas as well as the correlation between neighborhood accessibility and measures such as unemployment and economic opportunities (Tomer et al., 2011; Kaufman et al., 2014; Owen and Murphy, 2018). In addition, these studies agree that only marginal improvements in public transit systems can be made in light of deep budget cuts at all levels of government. Building on the measures and insights provided by these studies, we offer empirical evidence that ridesharing can fit into the economic and social fabric of metropolitan areas and promote inclusive mobility.

Lastly, researchers have studied the geographical impact of technology in other contexts, such as in the area of international trade (e.g., Blum and Goldfarb, 2006; Hortaçsu et al., 2009; Brynjolfsson et al., 2019), crowdfunding (e.g., Agrawal et al., 2015), and in the accommodation market (e.g., Farronato and Fradkin, 2017; Zervas et al., 2017; Calder-Wang, 2019; Schaefer and Tran, 2020). Our results are consistent with the qualitative finding that the welfare impact of technology is larger in geography and groups with a larger mismatch between supply and demand. The non-uniform effect suggests that the optimal regulation of technology should depend on the extent of this mismatch in different geography.

2. Data

Our data set focuses on the five boroughs of NYC, namely the Bronx, Brooklyn, Manhattan, Staten Island, and Queens from June 1, 2016 to August 31, 2016. The city is divided by TLC into 263 taxi zones that vary in size.[4] For example, an average taxi zone in Manhattan is approximately equivalent to an area of 6 by 6 avenue blocks. Throughout this paper, we treat these taxi zones as our basic geographical units of analysis, given the nature of the data. Our data set contains data from four sources:

(1) Real-time Uber and Lyft dynamic pricing and wait time

The first part of our data is a massive set of API queries on dynamic pricing and wait time at the geographical centroids of all 263 taxi zones at the minute level, for all service types on Uber and Lyft in our sample period, namely UberX, UberXL, UberBlack, UberSUV, UberPOOL, Lyft, LyftLine, and LyftPlus.[5] We also queried trip distance, trip duration, and trip cost for a route between a given pickup zone and a randomly chosen drop-off zone, in any given minute. Therefore, for each unique route, i.e., a pickup-drop-off pair, we have information on trip distance, duration, and cost approximately every 4 h.[6] We use this route-level information mainly in the demand estimation.

(2) TLC-published Taxi, Uber, and Lyft trip records

NYC TLC publishes taxi, Uber, and Lyft trip records. Taxi trip records (summarized in Appendix Table A.4) contain detailed trip information, such as pickup and drop-off date and time, the GPS coordinates of the pickup and drop-off locations, number of passengers, trip fares, etc. Uber and Lyft trips are identified from other For Hire Vehicles (FHV hereafter) trip records using the dispatching base numbers.[7] Far from the level of detail of taxi trip records, Uber and Lyft trip data only contain pickup date, time, and locations in the form of taxi zones. Restricted by this, we then map the GPS coordinates of taxi trips into their corresponding taxi zones so that taxi and ridesharing trips have the same geographical unit.

(3) Field-collected Uber and Lyft trip records

Uber and Lyft trip records provided by TLC lack information on drop-off locations, drop-off time, and service types. This data limitation hinders our analysis at the service type-route-time level. To address this problem, we conducted field data collections to acquire 70,277 historical trip records from 443 Uber and Lyft drivers in NYC. The purpose is to get a representative sample in order to estimate the empirical distribution of drop-off time and locations. We then use these probabilities to impute the service-type-route-time trip counts (the detailed procedure in Appendix D). The representativeness of the field data is supported by a comparison between their pickup distribution and that of the TLC-published Uber and Lyft trips.

(4) Route-specific subway and bus travel time

We queried Google Maps API to obtain total travel time via public transit (subway and bus). This data is collected at the level of 263×263 routes, by night/day *and* by weekend, where night is defined as 0:00am to 6:59am. Specifically, the total travel time for a given route via public transit includes (1) walking time from the zone centroid to the subway station or bus stop that Google Maps assigns, (2) subway or bus wait time, (3) subway or bus travel time, and (4) walking time out of the subway station or from the bus stop to the centroid of the destination zone. Note that the total public transit time is the best route suggested by Google Maps, which may include solely subway, solely bus, or a combination of both, depending on specific routes and times.

3. A description of the geography of ridesharing

In this section, we summarize and visualize model-free evidence using our data set. The average share of trips that experience a price surge is 7% – 8% for Uber non-luxury services and is around 19% for Lyft services. In terms of wait time, both Uber and Lyft's flagship services, namely UberX and Lyft, are around 7 min. To focus on our research question, we visualize how price and wait time vary across geography, and leave the reporting of other summary statistics in Table A.3 in the Appendix.

[4] Refer to the following link for a shape file of taxi zones provided by TLC: http://www.nyc.gov/html/tlc/html/about/trip_record_data.shtml.

[5] UberX and Lyft are the regular and mostly used service types on each platform, respectively. UberXL and LyftPlus are service types with cars of larger capacity, which usually seat up to six passengers. UberBlack and UberSUV are the luxury options on Uber. UberPool and LyftLine are the carpool options on Uber and Lyft, respectively.

[6] More accurately, if we query on each route in exactly one minute, it takes on average about 263 min to query the same route again.

[7] Refer to http://www.nyc.gov/html/tlc/downloads/pdf/find_a_ride.pdf. See the data appendix Appendix C for how the base numbers are identified.

Table 1
Heterogeneous Consumer Preference.

	Price(α), Time(β)	Lower	Midtown	UE & UW	Non Core
Morning rush	α	0.199***	0.058*	0.136***	0.242***
		(0.044)	(0.031)	(0.016)	(0.023)
	β	0.554***	0.655***	0.479***	0.501***
		(0.091)	(0.090)	(0.046)	(0.080)
Evening rush	α	0.026**	0.094***	0.127***	0.099***
		(0.012)	(0.016)	(0.013)	(0.020)
	β	0.807***	0.676***	0.487***	0.644***
		(0.053)	(0.038)	(0.042)	(0.083)
Weekday night	α	0.124***	0.168***	0.335***	0.233***
		(0.011)	(0.018)	(0.041)	(0.016)
	β	0.570***	0.512***	−0.173	0.320***
		(0.037)	(0.031)	(0.127)	(0.077)
Weekday late night	α	0.170***	0.149***	0.457***	0.094***
		(0.019)	(0.015)	(0.141)	(0.021)
	β	0.421***	0.447***	−0.642	0.840***
		(0.067)	(0.050)	(0.544)	(0.095)
Weekday day time	α	0.113***	0.062***	0.266***	0.210***
		(0.012)	(0.012)	(0.018)	(0.014)
	β	0.663***	0.830***	0.102	0.561***
		(0.058)	(0.064)	(0.074)	(0.117)
Weekend night	α	0.209***	0.274***	0.294***	0.210***
		(0.019)	(0.025)	(0.022)	(0.009)
	β	0.272***	0.115**	−0.014	0.471***
		(0.046)	(0.054)	(0.057)	(0.046)
Weekend late night	α	0.180***	0.137***	0.383***	0.237***
		(0.011)	(0.014)	(0.043)	(0.010)
	β	0.587***	0.551***	−0.484***	0.519***
		(0.044)	(0.032)	(0.152)	(0.059)
Weekend day time	α	0.173***	0.135***	0.176***	0.222***
		(0.018)	(0.014)	(0.013)	(0.015)
	β	0.442***	0.533***	0.458***	0.512***
		(0.054)	(0.056)	(0.071)	(0.069)
Airport	α	−0.080***			
		(0.013)			
	β	0.668***			
		(0.128)			
Rain	β	0.474***			
		(0.027)			
Luxury (per service minute)		−0.104***			
		(0.009)			
Capacity (per service minute)		−0.110***			
		(0.010)			
N		14,464,715			

Note: This table presents the demand estimation results. Throughout the table, α indicates a row of price sensitivity estimates and β indicates a row of time sensitivity estimates. "Lower" stands for lower Manhattan; "UE & UW" stands for Upper East and Upper West; "Non Core" stands for non Manhattan core. "Airport" is a dummy for to-airport trips. The time blocks are defined as: morning rush (weekdays 7am–9am), evening rush (weekdays 4pm–7pm), weekday daytime (weekdays 10am–3pm), weekday night (weekdays 8pm–11pm), weekday late night (weekdays0am–6am), weekend day time (weekends 5am–5pm), weekend night (Friday 8pm–11pm and weekends6pm–11pm), and weekend late night (weekends 0am–4am). Standard errors are in parentheses. * represents statistical significance at 10% level, ** 5%, and *** 1%.

In Fig. 1, we plot the average surge multipliers and wait time across different regions at different time of the day. Note that conceptually, the same region at different time of the day can be regarded as a different region, because they could be subject to different underlying market conditions. Two features of the data emerge: First, prices and wait time change rapidly across geography and time of the day. Second, surge multipliers and wait times are usually greater in Manhattan than in outer boroughs, and are greater in rush hours than in non-rush hours. These variations suggest that the prevalence of ridesharing and its implication for consumers could vary substantially across space and time.

Next, we plot the number of taxi trips and the share of ridesharing trips out of all trips (taxi plus ridesharing) in Fig. 2. The graph shows that the number of taxi pickups are concentrated in high accessibility region, i.e., the Manhattan Core, and it decreases as the distance to the Manhattan core increases. Meanwhile, the share of ridesharing pickups relative to taxi is larger in regions with low accessibility, namely the outer boroughs.[8]

To see this pattern more clearly, in Fig. 3a, we plot the relative share of ridesharing to yellow taxis against different neighborhoods that differ in the distance to Midtown Manhattan measured in miles. Consistent with Fig. 2, we see that the market share of ridesharing is greater in neighborhoods that are farther away from Midtown Manhattan, suggesting that ridesharing supplements taxis in neighborhoods that are underserved by taxis. This substitution seems to happen not only across space, but also across time: in Fig. B.11 in the appendix, we find evidence that passengers substitute toward ridesharing more during the hours when taxis become less available due to shift change.

In Fig. 3b, we performed the same analyses by comparing ridesharing with pickups by green taxis, which was introduced by TLC in August 2013 to increase taxi coverage in neighborhoods outside Manhattan. Note that green boro taxis use the pricing structure of yellow medallion taxis, but they have limited pickup areas.[9] We see two interesting patterns. First, although green taxis

[8] The pattern agrees with anecdotes https://www.nytimes.com/2017/10/12/nyregion/uber-taxis-new-york-city.html

[9] Specifically, they can only accept street hails anywhere in NYC except south of West 110th Street and East 96th Street in Manhattan, and they can take prearranged trips from JFK and LGA airports. Additionally, green taxis may drop pas-

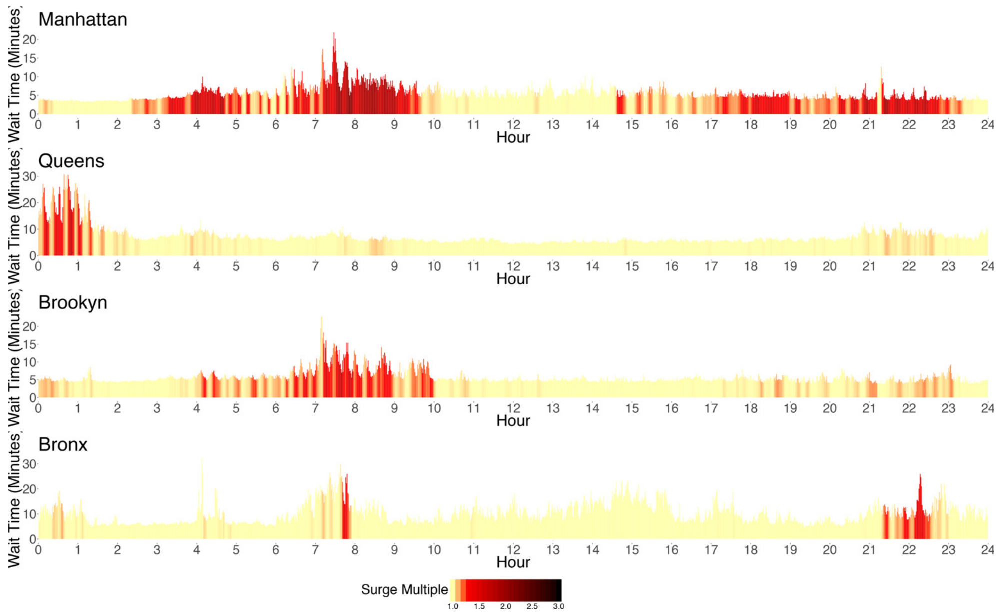


Fig. 1. Surge Multiple and Wait Time Change Rapidly across Space and Time. (UberX, Monday, June 6, 2016). *Note:* This graph plots how UberX minute-level surge multiple and wait time, averaged over all zones of a given NYC borough, vary across time of the day, on Monday, June 6th 2016. Surge multiples are in varying shades as illustrated by the legend, and wait time are measured by the bar length.

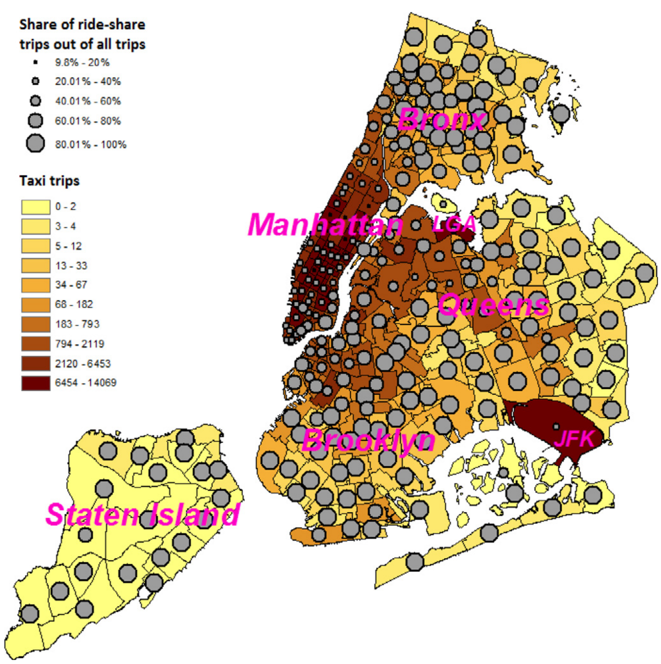


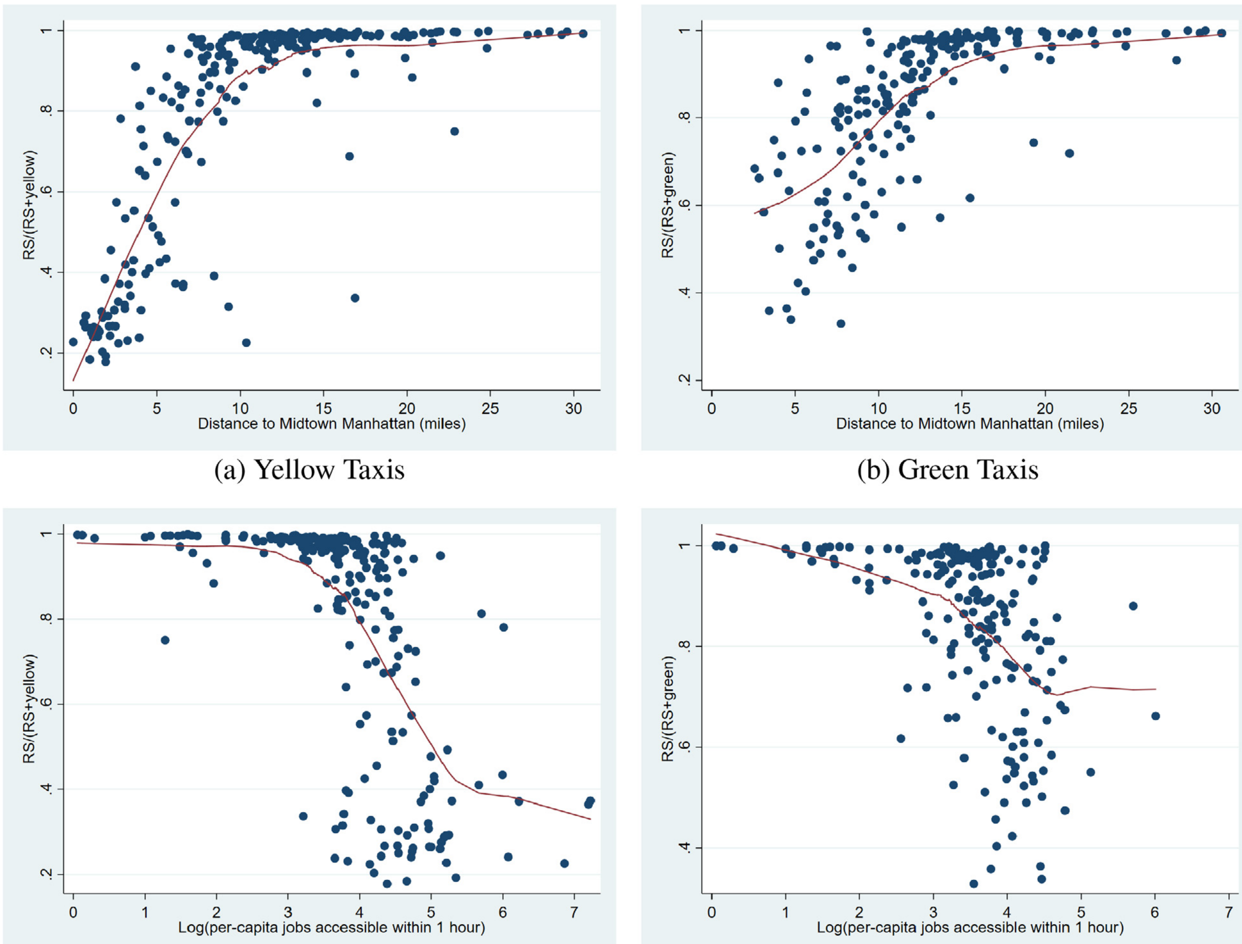
Fig. 2. The Geography of Ridesharing and Taxi Rides.

are specifically designed to serve outer boroughs, ridesharing still has a larger market share than green taxis in neighborhoods that are farther away from Midtown Manhattan. Next, comparing across Fig. 3a and b, we see that the second graph has a less steep slope, which means that green taxis have a lower trip elasticity of distance compared to yellow taxis. This difference reflects different drivers' cruising strategy, because the two types of taxis essentially face the same demand due to their same pricing rules outside of the Manhattan Core. The different cruising strategy is consistent with drivers responding optimally to different matching frictions: Yellow taxi drivers prefer to cruise towards the Manhattan Core likely because their matching friction of finding passengers is higher in outer boroughs, relative to that of green taxi drivers, possibly due to differences in the knowledge of demand in these regions.

Subsequently, we show that the geography of ridesharing activities is robust to the use of "economic distance". Specifically, we follow the literature to measure the accessibility of a neighborhood by the amount of employment and other economic activities it is connected to. We adopt the measure proposed by Kaufman et al. (2014) that counts the number of jobs accessible, within an hour of travel by public transit from a given neighborhood of NYC's 177 neighborhoods. In constructing the measure, the authors drew on Census data and Google Maps API. To fit the measure into our analyses that use taxi zones as the unit of analysis, we built a crosswalk between NYC neighborhoods and NYC taxi zones using a geographical software which returns the portion of a given neighborhood that overlaps with a taxi zone in terms of geographical area. We then computed the weighted sum of jobs accessible across all neighborhoods that (partially) overlap with a given taxi zone, using the above-mentioned portions as weights. In addition, we chose to use per-capita jobs accessible as our preferred measure of accessibility, since it captures the density of jobs better than the absolute number of jobs. The distribution of this accessibility measure exhibits a clear long tail — the neighborhood at the 90th percentile has access to 18 times more jobs than the neighborhood at the 10th percentile. In addition, this accessibility measure negatively correlates with the geographical distance measure (with a correlation coefficient at −0.7454, statistically significant at 1% level).

Using this measure of accessibility, we visualize the correlation between the relative share of ridesharing with respect to taxis and the level of accessibility of a region. Fig. 3c shows a clear negative correlation between accessibility and the prevalence of ridesharing relative to yellow taxis. In Fig. 3d, we observe a similar pattern for the relative share of ridesharing to green taxis. Similar to the evidence using geographical distance, we see that ridesharing has a larger market share than taxis in less accessible neighbor-

sengers off anywhere in NYC. By the end of 2015, 7676 green boro taxi licenses were active in the market.



(a) Yellow Taxis (b) Green Taxis

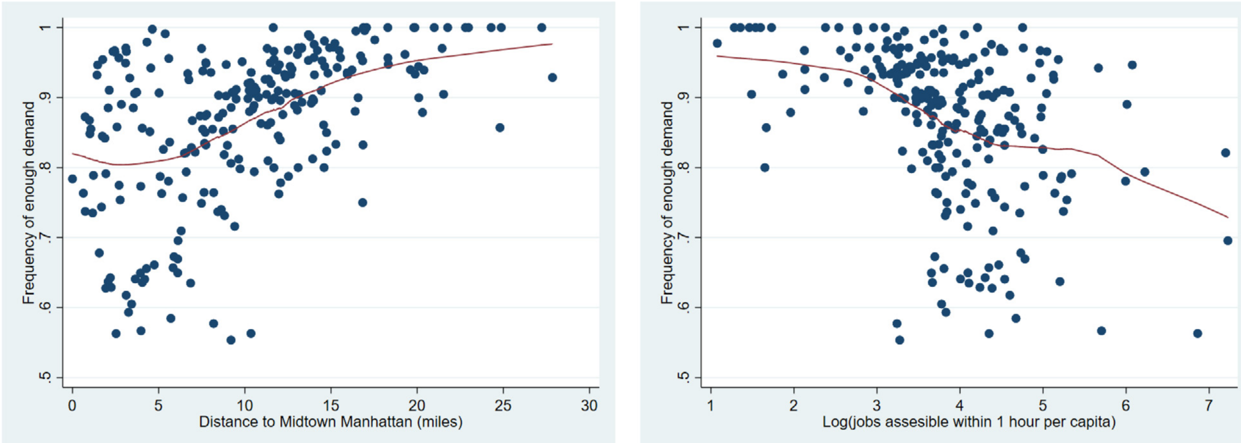
(c) Yellow Taxis (d) Green Taxis

Fig. 3. Ridesharing and Taxi Rides by Regions' Accessibility.

hoods. Additionally, comparing across Fig. 3c and d, we see again that green taxis are less elastic with respect to economic distance when it comes to picking up passenger in low accessible region. This suggests that the difference in the trip elasticity with respect to economic distance is due to differences in drivers' cruising strategy, which is likely driven by different levels of matching friction.

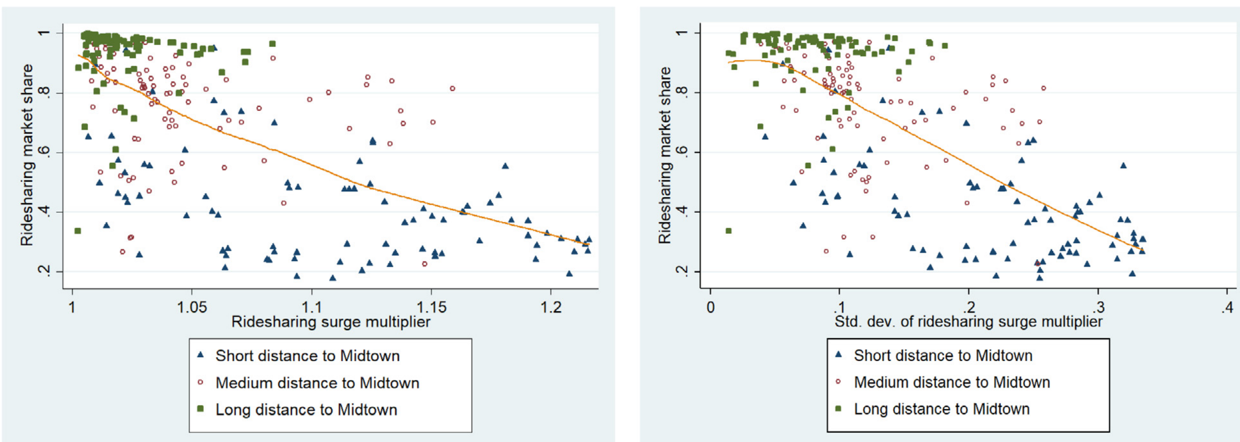
One alternative hypothesis is that the smaller share of taxis in less accessible neighborhoods is driven by the lack of demand for rides. However, this is not consistent with the evidence that there is more unfulfilled demand for taxis in less accessible neighborhoods. We use realized pickups that we observe in the data as a proxy for demand of rides. Specifically, we use July 2016 data and count the total number of ridesharing and taxi pickups in a given 15-minute period in a given zone. We use the number of taxi drop-offs in the previous 15-minute period as a proxy for the taxi supply in the focal period (call this X), and the number of pickups by both taxi and ridesharing in the focal period as a proxy for demand for rides (call this Y). We define a dummy variable "enough demand" that equals 1 if $Y \geq X$ and 0 otherwise. Fig. 4a plots each zone's frequency of "enough demand", averaged across all 15-minute intervals of the month. We see that as we move away from Midtown, the frequency of "enough demand" increases. Similarly, Fig. 4b shows a higher rate of enough demand in less accessible regions. These existence of excess demand, combined with the finding that yellow and green taxis have different trip elasticity of distance in outer boroughs, suggest that low demand for taxi is unlikely to be the sole reason for a relatively small share of taxi pickups in the low accessibility regions. Admittedly, Taxi and ridesharing services are different products, and therefore the observed differences could in principle be driven by product differentiation. In Section 4, we explicitly model and estimate consumer preferences for these attributes and reach a similar conclusion in our counterfactual analyses.

Lastly, we study the role of technology-aided matching and surge pricing in explaining the geography of ridesharing. Using zone-minute level dynamic pricing on Uber and Lyft from the data on ridesharing APIs, we plot the relative market share of ridesharing and surge multipliers of the taxi zones across regions with different levels of accessibility in Fig. 5a. In the graph, triangles, circles, and cubes represent regions that are within a short, medium, and long distance to Midtown, respectively (here "short", "medium", and "long" are defined by the terciles of the distance to Midterm). We find that the average share of ridesharing is positively correlated with the region's distance to Midtown, but it is negatively correlated with surge multiplier both within and across each distance bracket. Additionally, if we look at long-distance trips (squares), the relative share of ridesharing is more than 90% even in regions where the average surge multiplier is close to 1. In Fig. 5b, we repeat the analysis on the standard deviation of surge multipliers, instead the average. We similarly find that a lower variation in surge multiplier predicts higher share of ridesharing within a distance bracket. These patterns suggest that, even though dynamic pricing may play a big role in matching demand and supply in highly accessible regions (Hall et al., 2015 and



(a) Yellow and Green Taxis
(b) Yellow and Green Taxis

Fig. 4. Enough Demand across Geography.



(a)
(b)

Fig. 5. Ridesharing's Market Share Negatively Correlates with Surge Multipliers and Its Standard Deviation.

Castillo et al., 2017), it cannot be the main factor that leads to a larger market share of ridesharing in less accessible regions. In other words, technology-aided matching seems to play an important role in matching drivers and passengers in less accessible regions.

4. Consumer surplus

The descriptive results in Section 3 discover geographic patterns of ridesharing rides relative to taxi rides across regions with different accessibility, measured in terms of the distance to Manhattan and in terms of economic distance to job and other economic opportunities. While these results are novel and interesting, one could argue that they are steps away from informing policy making because it is unclear how to interpret the relative market shares.

In this section, we develop a simple discrete-choice model to estimate consumer preference and demand for taxis and ridesharing. The purpose of this exercise is to translate the observed geographic difference in the market share of ridesharing to consumer surplus through a counterfactual analysis where ridesharing services are hypothetically removed. Additionally, the exercise also sheds light on relative strength of different mechanisms of changes in consumer surplus, such as from changes in price or wait time. We should note that the benefits come at the cost of making more assumptions on the choice model, the data compatibility, and the validity of the instrument for demand estimation, many of which can only be indirectly tested. Therefore, we view this part of our analysis as complementary to the descriptive evidence in Section 3.

4.1. Consumer demand for ridesharing

Market conditions constantly change in the market of rides: factors such as prices, wait time, and conditions of competing alternatives vary at a high frequency across time within a location. Consumers make transit decisions by evaluating these relevant time-varying characteristics. Therefore, we set up the demand model at a granular route-time level, which we believe is a reasonable unit of analysis in this market.

The product in this market is a ride, which is a transportation service that moves a consumer i from a starting point j to a destination k in time t. We restrict attention to taxis, various ridesharing service types, as well as the public transit.[10] For a

[10] We implicitly assume that consumers multi-home across the two ridesharing services in that they compare the prices and wait times of these service (along with those of taxis) before making a decision. Based on the study by Earnest Research (https://medium.com/earnest-research/how-strong-is-ubers-lead-over-lyft-in-the-u-s-5df79dd0f3d0), there were about 25% of ridesharing users who are multi-homers in the third quarter of 2016, which overlaps with our sample period. The share of multi-homers increased steadily over time and reached 34% by the end of 2018, as the study finds. While the numbers

given route jkt, the consumer's utility of transportation mode depends on price, travel time, wait time, other observed and unobserved service-specific characteristics, as well as the consumer's own idiosyncratic taste, which is specified as

$$U_{sijkt} = -\alpha_{jkt} P_{sjkt} + \beta_{jkt} (Total_{ojkt} - WT_{sjt} - TT_{sjkt}) + X'_{sjkt}\Theta + \xi_{sjkt} + \phi_{jkt} + \epsilon_{sijkt} \quad (1)$$

where s denotes one of the "inside options", namely taxis and all service types on Uber and Lyft, while o stands for the outside option — the public transit. We take the public transit as the outside option because demand for rides is a "derived" demand — once a trip need (to go from Point A to Point B at Time T) is identified, the agent has to choose one of the transportation options to make the trip. We let $Total_{ojkt}$ denote the total travel time of the public transit, or the door-to-door time, which is the sum of the walking time to the subway station or bus stop, wait time, time en route, and walking time from the station/stop to the destination. Let WT_{sjt} denote the wait time of s at location j and time t, and TT_{sjkt} is the time en route of s. Therefore, $Total_{ojkt} - WT_{sjt} - TT_{sjkt}$ represents the amount of time s can save compared to the outside option, and β_{jkt} is the marginal utility of time saved. For example, for a route that is 35-minute long in total via the public transit, if the UberX wait time is 5 min and the travel time is 20 min (thus total time is 25 min), then the consumer's utility of UberX is $\beta_{jkt} \times 10$ min higher than that of the baseline.

We let P_{sjkt} denote the price of the trip, and $-\alpha_{jkt}$ is the marginal utility of price. Let X_{sjkt} represent a vector of observed service-specific characteristics that affect utility, Θ being the associated vector of parameters. The other variables are defined as follows: ξ_{sjkt} is the unobserved (to the researcher) route-service-specific utility component; ϕ_{jkt} is the utility difference between all car options and the public transit, and it measures the comfort of not having to walk to the subway station/bus stop and being able to sit for the entire trip, as opposed to not being able to sit down with the public transit sometimes; ϵ_{sijkt} is the consumer idiosyncratic error term. We further assume that the travel time TT_{sjkt} is the same for all inside service types, since they travel on the same route at the same time and thus they are subject to the same traffic and road conditions. As a result, the subscript s in TT_{sjkt} is removed from here on. The purpose of regulating a common travel duration will become clear later. Finally, we normalize the utilities of all inside options against that of the public transit, i.e., $U_{oijkt} = \epsilon_{oijkt}$.[11]

4.2. Empirical framework

Usually, the preference parameters of discrete choice models are obtained by estimating a mapping from the characteristics space to realized market shares of options (e.g. Berry, 1994, Berry et al., 1995, Nevo, 2000; Nevo, 2001, and Petrin, 2002). However, market shares of service types cannot be computed because we do not observe route-level (or, jkt-specific) public transit ridership. With a market jkt defined so finely, we do not have a reasonable benchmark of market size to impute the market shares like authors who study other settings (e.g. Berry et al., 1995 use population size of a geographical area to proxy the market for automobiles). Nonetheless, Chevalier and Goolsbee (2009) demonstrate that odds ratios of the logit framework can be used to establish identification. We follow this approach to avoid the problem due to unavailable market shares, while at the same time flexibly allow for preference heterogeneity across markets. Specifically, we assume a Type 1 extreme value distribution of the error term, which amounts to a standard logit model. Then the market share of s in the market jkt, though not observable, has to satisfy the following:

$$MarketShare_{sjkt} = \frac{exp(\delta_{sjkt})}{1 + \sum_{n=1}^{S} exp(\delta_{sjkt})}, \quad (2)$$

where $\delta_{sjkt} = -\alpha_{jkt} P_{sjkt} + \beta_{jkt}(Total_{ojkt} - WT_{sjt} - TT_{jkt}) + X'_{sjkt}\Theta + \phi_{jkt} + \xi_{sjkt}$ is the mean conditional utility of service s at jkt. To ease illustration, let taxi cabs be denoted as c. Then taking logs of the predicted odds ratios of taxis' share and the share of any ridesharing service type yields

$$log(\frac{D_{cjkt}}{D_{sjkt}}) = \alpha_{jkt}(P_{sjkt} - P_{cjkt}) + \beta_{jkt}(WT_{sjt} - WT_{cjt}) + (X_{cjkt} - X_{sjkt})'\Theta + \xi_{cjkt} - \xi_{sjkt}, \quad (3)$$

where D_{cjkt} and D_{sjkt} are trip counts of taxi and service type s, respectively. Note that the terms $Total_{ojkt}$ and TT_{jkt} are canceled out in the algebra, because they are common across service times within the same route. Eq. (3) indicates that the relative market shares of taxi and service type s at the route-time level should be affected by their differences in price, wait time, as well as other observed and unobserved characteristics.

We allow heterogeneity in consumer preference for price and wait time to vary across jkt's.[12] We assume that consumer preference for price and wait time is the same within a route jkt. We make this assumption because we have defined routes at a very granular level, and consumers with the same pickup and dropoff location within a narrow time window likely have similar preferences. The specification is given as:

$$\alpha_{jkt} = Y'_{jkt}\Theta_A + \epsilon_\alpha \quad (4)$$

$$\beta_{jkt} = Y'_{jkt}\Theta_B + \epsilon_\beta, \quad (5)$$

where Y_{jkt} is a row vector of dummy variables that contain various combinations of pickup areas and time blocks, and Θ_A and Θ_B are the vectors of the corresponding coefficients for price and time, respectively. Specifically, the areas include Lower Manhattan (dummy), Midtown Manhattan (dummy), Upper East and West Manhattan (dummy), and Non-Manhattan Core (dummy). The time blocks include morning rush (weekdays 7 a.m. - 9 a.m.), evening rush (weekdays 4 p.m. - 7 p.m.), weekday day time (weekdays 10 a.m. - 3 p.m.), weekday night (weekdays 8 p.m. - 11 p.m.), weekday late night (weekdays midnight - 6 a.m.), weekend day time (weekends 5 a.m. - 5 p.m.), weekend night (Friday 8 p.m. - 11 p.m. and weekends 6 p.m. - 11 p.m.), and weekend late night (weekends midnight - 4 a.m.).

Note that consumers do not know the precise taxi wait time at a given location-time, unlike in the case of ridesharing. We assume that consumers act according to their expectations of WT_{cjt} for taxi time when choosing transportation modes, and $-\beta_{jkt}WT_{cjt}$ can be absorbed by the many location and time fixed effects in the regression.

We choose to use the simple logit framework primarily due to its tractability in generating analytical solutions that allow for identification in the absence of market shares. The simple logit is well known for the independence of irrelevant alternatives problem, namely that the odds ratio of two options need to stay the

reported by these studies are not overwhelmingly large, it does lend support to our assumption on multi-homers. While we were not able to find such numbers across US cities, it is reasonable to believe that share of multi-homers in NYC is likely to be higher than the national average, given the demographic composition and technological savviness of NYC residents.

[11] Liu et al. (2017) study the welfare impact of DiDi in China. They assume that consumers make choices at the origin level, as opposed to the origin-destination pair level in our model.

[12] This flexibility allows us to capture the effect of route- and time- specific characteristics on the sensitivity parameters, such as the different car ownership across regions as discussed in Leard and Xing (2020).

same regardless of the availability of an alternative. In principle, we could estimate a model that allows for more flexible substitution pattern, such as a nested logit model. However, this would require an instrumental variable that shifts demand in one nest but not in the other, which is difficult to find. We will discuss the validity of our current instrument under the simple logit framework in Section 4.3.

Additionally, we need to make assumptions on the data compatibility. Specifically, Uber and Lyft trip records published by TLC lack drop-off information and trip service types, which prevents us from getting the trip counts of various Uber and Lyft service types at the *jkt* route, or D_{sjkt}. To address this issue, we use the surveyed 70,277 Uber and Lyft trips in the same time period to construct proxies. This sample consists of a random subsample and a convenience subsample. In Appendix D, we show that the convenience subsample resembles the random subsample closely, so we can rely on the randomness of the full sample to infer D_{sjkt}. To infer D_{sjkt}, we first estimate a probit model to predict the probability of a certain trip that takes place in a certain *jkt* cell, using the full sample. We then impute D_{sjkt} by distributing the total Uber/Lyft pickups at a *jt* to various service types and destinations, using these empirical probabilities (details in Appendix E).

The level of analysis is at origin-destination-15 min level. We choose a 15-minute time window because it can approximate the real-time setting, while not dividing the time too finely such that consumer information collection and decision making are split into different time periods. Accordingly, all the explanatory variables are collapsed into the 15-minute averages, while trip counts are sums over 15 min. We set the origin and destination at the taxi zone level and PUMA (Public Use Microdata Areas) level, respectively, where taxi zones (263 in total) are a more granular geographical unit than PUMA (55 in total). The reason for choosing such a level of analysis is that prices and wait times vary across taxi zones so pickup locations need to be at the taxi zone level to fully exploit the richness of the API data. On the other hand, having 263 destination zones means studying 263 × 263 markets, which is a stretch for imputation of D_{sjkt}. Therefore, we study 263 × 55 markets at the 15 min interval.

4.3. Identification and estimation

Our identification, like other demand estimation studies, is subject to price and wait time endogeneity, given that Uber and Lyft pricing algorithms can potentially take into account many factors that affect demand, both observed and unobserved to us. A simple OLS estimation of Eq. (3) would lead to biased estimates of variables of interest. To deal with this potential endogeneity problem, we control for service type, route, and time fixed effects at a granular level. Besides, we implement an IV strategy, where the IV affects the market share of a service type on a given route *only through* its impact on price, but it is otherwise uncorrelated with other demand shocks.

Our IV leverages a design feature of the ridesharing Apps that were present during our sample period: passengers see and commit to a surge multiplier on Uber and Lyft Apps *prior to* entering their destination on their phones. We instrument for $P_{sjkt} - P_{cjkt}$ using the average surge prices of trips into the focal zone in the previous time period. Because the ridesharing Apps do not possess information on where the consumer is going, they cannot adjust their surge multiplier according to the consumer's destination, which creates the uncorrelatedness of surge prices at the origin and the underlying demand condition at the destination.[13] The exclusion assumption can be violated if platforms can somehow infer the distribution of destinations for trip requests at a given origin. We rely on the many location and time fixed effects to help alleviate this concern. On the other hand, surge pricing at the origin will affect the price in the focal zone through its effect on the number of available drivers in the focal zone after they drop off their consumers.

To deal with the endogeneity of WT_{sjt}, we use the total number of drop-offs at all neighboring zones of the focal zone as the instrument. This essentially measures the stock and availability of cars close to the focal zone, which affects the wait time at the focal zone. The exclusion restriction assumption that we make is that the availability of cars in the neighboring zones is not correlated with demand shocks in the focal zone.

We perform first-stage regressions to test the strength of the instruments. Since we have dozens of endogenous variables (because of interactions of time-location dummies with price and wait time), the first-stage tests entail an equal number of regressions, each of which regresses an endogenous variable on all regressors. F statistics and partial R^2's are reported in Appendix Table G.6. Given that all F-statistics are large (Staiger and Stock, 1997; Lee et al., 2020), and that the partial R^2's are relatively high (Shea, 1997), our instruments are not weak instruments.

The stochastic unobserved utility component ξ is assumed to be normally distributed with mean zero and independent across both service types and *jkt*'s. The error term $\xi_{cjkt} - \xi_{sjkt}$ in Eq. (3), however, creates a within-*jkt* correlation among the observations, as these observations share a common part ξ_{cjkt} in the error. Specifically, $COV(\xi_{cjkt} - \xi_{sjkt}, \xi_{cjkt} - \xi_{s'jkt}) \neq 0$ for two on-demand service types s and s' in the same *jkt*. Given that we use instrumental variables that are not correlated with ξ_{cjkt}, a random effects estimator would be appropriate to deal with the correlation in errors. One complication to the problem, however, is that the analysis sample is unbalanced — as illustrated in Section 4.1, there are varying number of observations within a given *jkt* due to the sub-sampling filters applied ($D_{cjkt} > 0$ *and* $D_{sjkt} \geq 0.1$). We choose to follow the general method proposed in Balestra and Varadharajan-Krishnakumar (1987) to estimate Eq. (3) by Feasible Generalized Two-Stage Least Squares. Details of the application to generate our estimator are reported in Appendix F.

In the estimation, we drop airport pickups from the sample because only very few trips end in neighboring zones of airports, and the correlation between these trips and airport wait time is very weak to justify the use of the wait time IV. To-airport trips, however, are kept in the sample. We include the dummy variable "to-airport" in the vector Y_{jkt} to allow consumers on trips to the airport to have additional price and time sensitivities. Similarly, we include the dummy variable "rain" to allow for extra dis-utility of wait time. The vector of observed characteristics X_{sjkt} includes luxury and capacity. Luxury measures the units of luxury service provided by the trip, which is a dummy variable on UberBlack and UberSUV, multiplied by the duration of the *jkt* trip. Capacity is defined similarly, except for service types UberXL, UberSUV, and LyftPlus. Finally, the set of fixed effects includes pickup zone, drop-off PUMA, pickup hour by weekend (dummy), pickup PUMA by time block, pickup PUMA by drop-off borough, and drop-off PUMA by time block, dummies of trips at various trip duration levels (e.g. 10-minute trip, 15-minute trip, etc.).

[13] This may not be the case anymore given that Uber and Lyft now practice "upfront pricing", which requires rider destinations before showing a fixed price for the trip. In this case, origin price can be affected by the destination demand shock, if the ridesharing platforms practice dynamic programming and discriminate riders based on destinations. For example, the demand shock at the destination may lead the platform to willingly offer a low price at the origin, just to have the driver relocated to the destination.

4.4. Consumer preference for price and wait time

The estimation results are shown in Table 1. First of all, our results identify consumer preference toward lower prices and shorter wait time — across almost all location-time-block combinations, the price effects on utility are estimated to be significantly negative ($-\alpha_{jkt}$), and the marginal utility of time (β_{jkt}) is estimated positive and strong. These estimates directly indicate that consumers are price-sensitive and value less wait time. This finding is consistent with recent studies on value of time in the context of taxis and ridesharing platforms (Buchholz et al., 2020; Goldszmidt et al., 2020). In addition, these estimates suggest that consumers are willing to make price-time trade-offs, and these economic fundamentals rationalize the use of dynamic pricing by ridesharing companies.

Second, we find sensible heterogeneity in consumer price and time preference across locations and times. For example, consumers in Midtown Manhattan during morning rush hours tend to be more time-sensitive and less price-sensitive, compared to consumers in most other location-times. This may reflect the preference of relatively high-income workers who rush into their workplaces on weekday mornings. A very similar pattern appears in Lower Manhattan during evening rush hours, which may be driven by Wall Street workers. Also, consumers going to the airport value time more and are additionally less sensitive to price. Similarly, extra disutility of wait time is found on consumers who hail a ride during the rain. These heterogeneities are consistent with intuition and provide confidence on the identification. In addition, the coefficient estimates of luxury and capacity are sizable, indicating that NYC consumers value these features that are made conveniently available on ridesharings compared to the offline options.

4.5. Taxi supply

To study consumer welfare gain by a counterfactual with no ridesharings, it is important to understand how taxi supply would respond in this hypothetical situation. In this section, we propose a model of taxi market equilibrium. The uniqueness of the taxi market, well articulated in Orr (1969), is the difference between operating hours and passenger service units supplied: taxi drivers' costs depend on the hours of driving and searching for passengers, while their revenues depend on the fare multiplied by the number of passenger service units supplied. This is due to the nature of the taxi matching technology. Here, we modify the framework used in Orr (1969) to fit our purpose. Let the taxi demand be specified as,

$$D = f(F^a, q; M, \Theta_D)$$

where D is the number of passenger service units demanded; F^a is the fixed administered fare; q is the total operating hours of taxi drivers; M is the number of potential consumers; and Θ_D is a vector of demand parameters. Also, D is continuous in F^a and $\frac{\partial D}{\partial F^a} < 0$; D is continuous in q, and $\frac{\partial D}{\partial q} > 0$, because as more taxi hours are provided, consumers are more quickly matched to drivers and the average consumer wait time decreases; $D(q=0, F^a; M, \Theta_D) = 0$, $\frac{\partial D}{\partial q}|_{q=0} >> 0$, $\frac{\partial^2 D}{\partial q \partial q} < 0$, and $\frac{\partial^2 D}{\partial q \partial M} > 0$.

The taxi market is a market with a rather elastic supply of labor; only modest skills are required to operate taxi cabs, and the practice of daily lease of medallions to the drivers imposes quite low entry and exit costs.[14] Cab drivers respond positively to wage increases by working longer hours (Chen and Sheldon, 2016; Farber, 2015; Hall et al., 2017). Under competitive conditions, the market equilibrium is characterized as a steady state where marginal cost and average revenue are equalized:

$$\frac{F^a * D(F^a, q; M, \Theta_D)}{q} = MC(q) \tag{6}$$

where $MC(q)$ is the marginal cost of operating a taxi, $MC(q) > 0$, $MC(q) << \infty$, $\frac{\partial MC}{\partial q} \geq 0$, and $\frac{\partial^2 MC}{\partial q \partial q} \geq 0$. The fixed medallions impose a hard constraint on the number of operating hours available in the market, that is, the maximal amount of daily operating hours is the number of medallions multiplied by 24 h. Let this maximum be $\bar{q}$.[15] Then the algebra leads to that the equilibrium operating hours supplied is concave in D up till $\bar{q}$.[16] Let wait time for taxi passengers be defined as,

$$WT = f(q, D; \Theta_{WT})$$

where WT is a continuous function, twice differentiable in q and D, with parameters denoted by the vector Θ_{WT}. Particularly, $\frac{\partial WT}{\partial q} < 0$ and $\frac{\partial^2 WT}{\partial q \partial q} > 0$: holding other things constant, more taxi service supplied leads to less wait time for consumers, yet this decrease in wait time diminishes as service units increase. $\frac{\partial WT}{\partial D} > 0$ and $\frac{\partial^2 WT}{\partial D \partial D} > 0$: holding other things constant, more trips demanded lead to longer consumer wait time, and this increase in wait time is greater as more trips are demanded.

A graphical characterization of the taxi market equilibrium is presented in Fig. 6a, where the equilibrium path is the combination of the part of q^* before $\bar{q}$ and the part of vertical line above q^* (the curve in red). An immediate implication on the equilibrium service units and equilibrium taxi wait time is depicted qualitatively in Fig. 6b — wait time increases as equilibrium service units increase after the capacity constraint. This is because supply cannot further adjust after the capacity constraint. Therefore, $\frac{\partial WT}{\partial D*} = \frac{\partial WT}{\partial q*}\frac{\partial q*}{\partial D*} + \frac{\partial WT}{\partial D} = 0 + \frac{\partial WT}{\partial D} > 0$.[17] The intuition is simple: as demand increases and maximal taxi capacity is reached, more and more consumers compete with each other to get matched to a fixed number of operating taxi cabs, which leads to longer average consumer wait time.[18]

One direct way to test the model prediction is to do a scatter plot of taxi pickups with taxi wait time, and check whether the empirical pattern resembles Fig. 6b. Unfortunately, in this study we cannot observe, estimate, or simulate the actual taxi wait time. However, Frechette et al. (2016) are able to simulate taxi wait time from observed taxi cabs, taxi search time, and exogenous time-varying factors, combined with a simulated matching function (Fig. 6 of their paper). We contrast their wait time estimates with UberTaxi wait time from our API queries in Fig. 7, and find that UberTaxi wait time follows a similar trend as their estimates across hours of day, although UberTaxi wait time is less volatile. We believe that UberTaxi is a reasonable proxy for taxi wait time and use it in the test of the taxi market equilibrium.

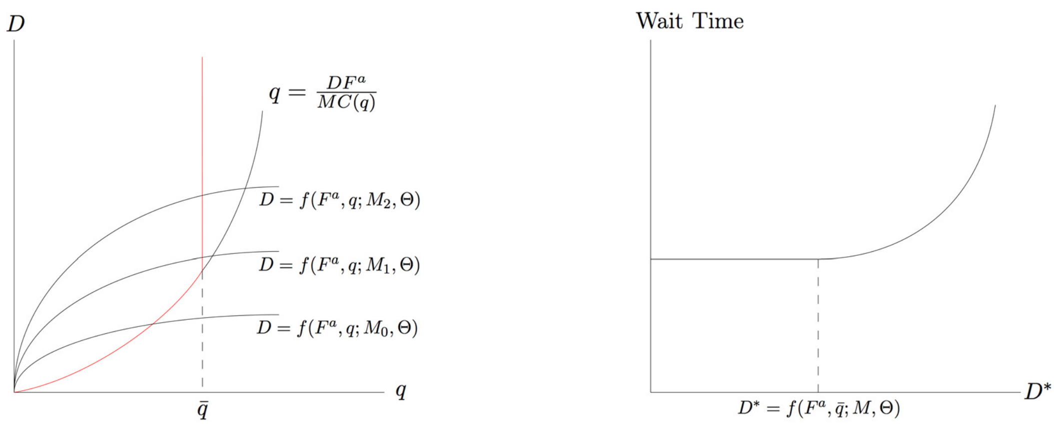
[14] Farber (2015) documents "a fair amount of entry, exit, and reentry among taxi drivers". Hall et al. (2017) demonstrate the horizontal labor supply curve for Uber drivers, which may as well be the case for cab drivers.

[15] This assumption is justified because the number of taxi medallions in NYC is relatively fixed over time: the number is 13,605 in March 2014, and by July 2016, this number dropped slightly to 13,587 medallions and stays the same as of May 2021. Source: https://www1.nyc.gov/site/tlc/businesses/yellow-cab.page.

[16] Differentiating both sides of Eq. (6) with respect to D leads to $\frac{\partial q}{\partial D} = \frac{F^a}{MC(q) + \frac{\partial MC}{\partial q} q} > 0$. Then $\frac{\partial^2 q}{\partial D \partial D} = -\frac{F^a \frac{\partial q}{\partial D}[2MC'(q) + MC''(q)q]}{[MC(q) + \frac{\partial MC}{\partial q} q]^2} < 0$.

[17] Before the capacity constraint $\bar{q}$, the term $\frac{\partial WT}{\partial q*}\frac{\partial q*}{\partial D*}$ is negative, so the sign of $\frac{\partial WT}{\partial D*}$ is undetermined. We use a flat line in Fig. 6b to describe the relationship between q and D before $\bar{q}$, but it should be noted that this relationship can be either positive, zero, or negative.

[18] Note that the market equilibrium proposed here abstracts away from the spatial equilibrium models such as (Lagos, 2000), (Lagos, 2003), and (Buchholz, 2015). It is possible that when there is an exogenous shock of demand, taxi cabs relocate spatially and form a new spatial equilibrium, which results in a different average wait time than implied by our model.



(a) Taxi Market Equilibrium

(b) Model Prediction of Equilibrium Service Units and Wait Time

Fig. 6. Taxi Market Clearing.

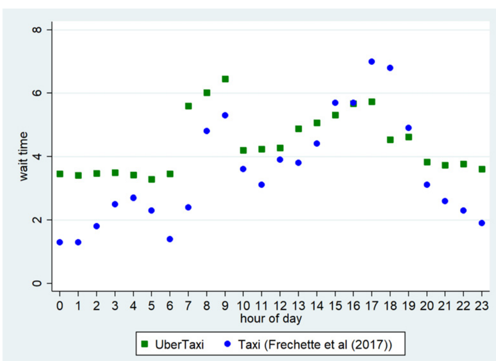


Fig. 7. UberTaxi Wait Time and Taxi Wait Time from Frechette et al. (2016).

The data strongly support the model prediction, as shown in Fig. 8b. Overall, there is a positive correlation between taxi trips and wait time. In particular, the average wait time is relatively low below some trip quantity threshold but becomes much higher after the threshold with a greater variation. In addition, there is a sharp contrast between rush hours and non-rush hours: first, rush hour wait time is on average higher; and second, the correlation between trip volume and wait time after the threshold at 10,000 trips, is greater during rush hours (correlation coefficient is 0.42) than non-rush hours (correlation coefficient is 0.19). This is likely due to the certain spatial distribution of commuting routes during rush hours, which exacerbates the within-location imbalance of demand and supply. We leverage these empirical correlations in the counterfactual analysis.

4.6. Consumer surplus

To evaluate the change in consumer surplus due to ridesharing, we adopt the concept of compensating variation – how much consumers should be compensated if Uber and Lyft were to be removed from the market such that the consumers can maintain the same level of utility? In the counterfactual, we assume that taxis and the public transit are the only viable options. We also assume that the public transit remains the same operation, without capacity constraint when more riders substitute toward it; the taxi system remains the current fixed number of medallions and administered fares.

There are two major reasons why consumers could be worse off in the absence of ridesharing services: first, existing ridesharing users would lose all the amenities from ridesharing services which they value more than other alternatives, due to revealed preference (that is, they would not have used ridesharing services had these services not provided the users with the highest utility); second, as shown in Section 4.5, when more consumers willingly substitute toward taxis, the equilibrium average wait time increases, which makes existing taxi users worse off. We then take the following procedure to compute consumer surplus:

1. First, we estimate ϕ_{jkt}, the utility difference between all car service types and the public transit, which measures the comfort of not having to walk to the subway station/bus stop and being able to sit for the entire trip. We allow for this utility in rush hours to differ from the utility in non-rush hours to account for the extra disutility of riding the public transit during rush hours, which is denoted as ϕ^r. In the search for ϕ and ϕ^r. We rely on the fact that once Uber and Lyft are removed from the market, the number of taxi trips will increase because of sheer substitution. That is, the values of ϕ and ϕ^r must be such that the corresponding counterfactual taxi ridership is greater than or equal to the current ridership. It is important to note that using this boundary equality (counterfactual ridership is equal to current ridership) leads to conservative estimates of ϕ and ϕ^r, given the monotonic relationship between these values and consumers' preference for taxis, which then leads to a conservative estimate of taxi wait time change and a lower bound for the welfare estimate.

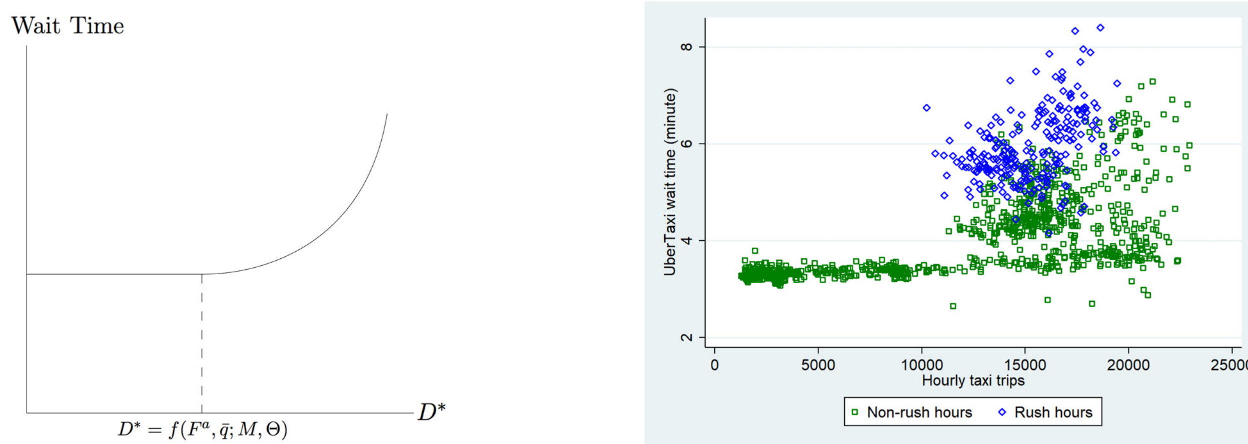
2. We use taxi wait time estimates of Frechette et al. (2016) as the counterfactual taxi wait times, and directly compute the mean conditional utilities of the public transit and taxis in the counterfactual. Comparing these mean conditional utilities yields the counterfactual taxi trips.

3. We then infer current taxi wait times. We use the empirical relation between UberTaxi wait time and taxi trips in Fig. 8b to approximate the true relation between taxi wait times and trip volumes, where we estimate the following OLS regression on day-hours when taxi trips exceed the capacity threshold 10,000:

$$\textit{Taxi wait time} = \pi_0 + \pi_1 \textit{Taxi trips (1000s)} \tag{7}$$

where we estimate the equation separately for rush hours and non-rush hours; π_1 is estimated at 0.155 for rush hours ($N = 233$, $t = 7.09$), and at 0.057 for non-rush hours ($N = 559$, $t = 4.51$). The estimated coefficients, counterfactual wait time (Frechette et al., 2016), counterfactual taxi trips (from Step (2)), and current taxi trips (directly from data) help us compute the current taxi wait times. For example, for rush hours,

$$\begin{aligned} \textit{current wait time} = {} & 0.155 * (\textit{current taxi trips-counterfactual taxi trips}) \\ & + \textit{counterfactual wait time} \end{aligned} \tag{8}$$



(a) Model Prediction of Equilibrium Service Units and Wait Time

(b) Taxi Wait time and Hourly Taxi Trips in Manhattan Core, Weekdays

Fig. 8. Taxi Trips and Taxi Wait Time.

Table 2
Consumer Surplus Due to Entry of Ridesharing Platforms.

	Non-Manhattan Core Taxi Wait Time		
	10 min	15 min	20 min
Consumer Surplus of Ridesharing Users (Unit: Dollar)			
Per dollar spent on RS platforms	0.72	0.86	0.96
Per RS trip	14.05	16.79	18.67
Per RS service minute	0.92	1.10	1.22
Ridesharing Welfare Sources			
Time	56.4%	58.2%	57.9%
Price	8.3%	6.5%	5.4%
Luxury	18.8%	16.0%	14.4%
Capacity	3.5%	3.0%	2.7%
Comfort	13.0%	16.2%	19.5%
		Consumer Surplus of Taxi Users (Unit: Dollar)	
Per dollar spent on taxis	0.16	0.21	0.21

4. With the inferred current taxi wait time, we compute the utility difference of current ridesharing as well as taxi users by comparing consumers' current options with counterfactual best options. It is important to note that taxi wait time estimates in Frechette et al. (2016) are for Manhattan core only. As a result, our inferred key values in Step 2, 3, and 4 so far, namely counterfactual taxi trips and current taxi wait times, are for Manhattan core only. Therefore, we make assumptions on Non-Manhattan core taxi wait times and computer consumer surplus accordingly. We take three alternative values, namely 10 min, 15 min, and 20 min, as various approximation of the true outer boroughs taxi wait times.

We present the counterfactual results in Table 2. The consumer surplus of ridesharing is about 72 cents *per dollar*, or \$14 for an average trip, when the taxi wait time in non-Manhattan core is a conservative 10 min. As the taxi wait time increases, the value of ridesharing also increases as the incumbent options becomes even less appealing. We find the majority of the consumer surplus comes from shortened wait time, compared to taxis and the public transit. Luxury cars are also valued significantly. Similarly, consumers value the utility of sitting in a car (thus not having to walk and possibly stand in the public transit), as 13% of welfare comes from "comfort". A very thin share of the consumer welfare increase comes from price, yet this is not surprising since in NYC, Uber and Lyft prices overall compare to taxi prices at the base-price level. In addition, taxi riders gain 16 cents per dollar spent, because taxi wait time becomes shorter as a result of consumer substitution toward ridesharings.

Given that the goal of this paper is to understand the geography of ridesharing and its implication on consumers, we compare the per-dollar consumer surplus across neighborhoods represented by their distance to Midtown Manhattan (recall that this distance measure is highly correlated with the accessibility measure). In Fig. 9, we first plot the density of taxi and ridesharing trips across distance to Midtown. Consistent with the previous results, we see that although both taxi and ridesharing pickups are concentrated in regions that are closer to Midtown Manhattan, the level of concentration is smaller for ridesharing trips. Next, we plot the geography of consumer surplus from ridesharing. We find that it varies drastically across geography and exhibits a U-shape: passengers that are 5 to 15 miles from Midtown experience 60% larger consumer surplus (average per-dollar consumer surplus is \$1.15) relative to passengers that are within 5 miles from Midtown (average per-dollar consumer surplus is \$0.72). For passengers that are more than 15 miles from Midtown, their surplus (\$0.86) is 19% larger than trip that start within 5 miles from Midtown.

Furthermore, we decompose the geography of consumer surplus of ridesharing into distinct utility components. As shown in Fig. 10, the U-shaped pattern observed in the total consumer surplus in Fig. 9 is mainly driven by the consumer surplus from saved time for consumers in the medium-range outer boroughs. Specifically, the welfare gain is higher in medium-range outer boroughs than in Manhattan because, even though the wait time of ridesharing is similar for both regions, the wait time of taxi is much shorter in Manhattan, making ridesharing a less attractive option. In comparison, the welfare gain medium-range outer boroughs is higher

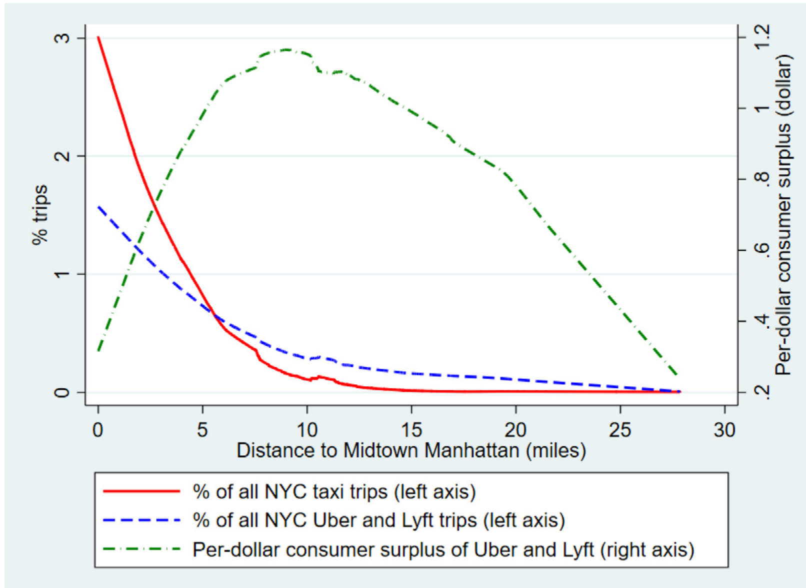


Fig. 9. Per-dollar Consumer Surplus of Ridesharing Varies across Neighborhoods. *Note:* In the graph, the density of taxi and ridesharing trips across distance to midtown is plotted by lowess smoothing against the left axis. The per-dollar consumer surplus across distance is plotted by lowess smoothing against the right axis.

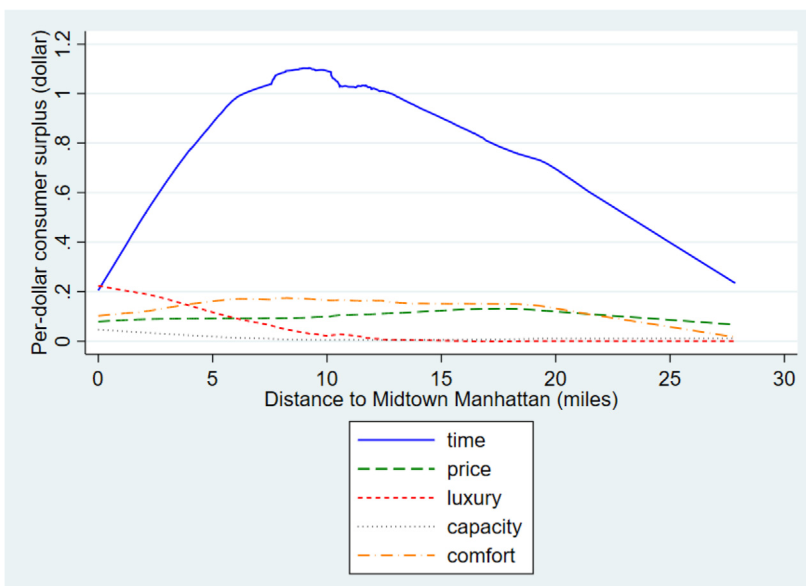


Fig. 10. The Geography of Consumer Surplus by Utility Component. *Note:* This graph plots the per-dollar consumer surplus against distance to midtown, separately for each of the five utility components. All curves are plotted by lowess smoothing.

than that of long-range outer boroughs because the wait time of Uber and Lyft in the latter regions is longer, even though the wait time for taxi is assumed to be the same in our simulation exercise. The geographical distribution of price-induced consumer surplus is relatively uniform. Consumer surplus from luxury ridesharing options appears greater in Manhattan than in outer boroughs, which is consistent with the business-oriented positioning of luxury options. Similar to price, the consumer surplus due to capacity and comfort seems to be evenly distributed across geographies.

Taken together, our consumer surplus calculation suggests a highly uneven distributional benefit of ridesharing across geographies. Neighborhoods that are within a medium distance to Midtown Manhattan (i.e., 5 to 15 miles from Midtown) benefit the most from the presence of ridesharing, and the major welfare source is the reduced time waste.

5. Discussion

Mobility inequality correlates with economic inequality. Yet achieving inclusive mobility is a difficult task, primarily due to the lack of incentives for public transit expansion. Taxis are a good way of point-to-point transportation, but they tend to cluster in dense areas and leave non-dense areas underserved, because this is an equilibrium response to the geographical distribution of passengers and lack of real-time demand information. Therefore, the proposition that aims to make taxi services more evenly distributed across space by expanding the medallion capacity constraint, restricting pickup areas, or both is unlikely to produce satisfactory outcomes, as exemplified by the case of NYC green taxis.

In this paper, we find that ridesharing could promote inclusive mobility through tech-aided matching, which makes the ridehailing service conveniently available for consumers in areas that are underserved by other transportation modes. Specifically, we use data from actual trips to present evidence that ridesharing coverage dominates that of taxis in less accessible areas. We then provide evidence that the advanced driver-passenger matching of ridesharing companies could mitigate the lack of information about demand in these areas. Finally, we interpret the geography of ridesharing through the lens of a choice model, and the estimation results suggest a disproportionate larger gain in consumer surplus of ridesharing services in low accessibility areas.

The literature that explores the economics behind ridesharing platforms has primarily focused on their average effects, i.e., how ridesharing platforms integrate massive real-time market information to enhance capacity utilization, in a generic dense metropolitan market. Our findings shed light on the importance of the underexplored distributional effects, i.e., the ability of ridesharing to connect and serve various metropolitan neighborhoods with inferior access to other mobility resources. Our findings suggest that technology can play a key role in mitigating geographical disparity in transportation, and this calls for impact studies in other domains to explore other ways to foster a more inclusive society. Even though our empirical setting is the transportation market in NYC, we believe our qualitative distributional results can potentially generalize to regions where demand for transportation is concentrated in the city center but dispersed outside the city center, and there is a limited supply of taxis and/or a large real-time information friction between supply and demand.

Appendix A. Summary Statistics

Table A1
Uber and Lyft Dynamic Pricing and Wait Time.

Variable	Mean	Std. Dev.	Min	Max	N
Surge Frequency					
UberX	0.073	0.260	0	1	32,398,537
UberXL	0.072	0.259	0	1	32,398,537
UberBlack	0.012	0.112	0	1	32,359,662
UberSUV	0.013	0.114	0	1	32,421,149
UberPool	0.080	0.272	0	1	32,359,652
Lyft	0.180	0.384	0	1	31,257,800
LyftLine	0.180	0.384	0	1	31,257,800
LyftPlus	0.180	0.384	0	1	31,257,800
Surge Multiple					
UberX	1.037	0.1622	1	4.2	32,398,537
UberXL	1.037	0.1619	1	4.2	32,398,537
UberBlack	1.008	0.0856	1	2.9	32,359,662
UberSUV	1.008	0.0884	1	2.9	32,421,149
UberPool	1.022	0.1028	1	3.4	32,359,652
Lyft	1.100	0.2727	1	5.0	31,257,800
LyftLine	1.100	0.2727	1	5.0	31,257,800
LyftPlus	1.100	0.2727	1	5.0	31,257,800
Wait Time (minutes)					
UberX	6.949	8.912	1	45	32,421,149
UberXL	13.509	14.794	1	45	32,421,149
UberBlack	8.821	10.176	1	45	32,421,149
UberSUV	14.059	15.324	1	45	32,421,149
UberPool	7.496	9.870	1	45	32,421,149
Lyft	7.034	9.882	1	45	33,133,051
LyftLine	6.884	9.869	1	45	33,133,051
LyftPlus	9.510	9.993	1	45	33,133,051

Note: The data for this table come from Uber and Lyft API queries, June-August 2016, where both price and wait time of 263 NYC zones are queried in approximately one-minute intervals. The small variation in variable sizes reflects rare cases of missing values and/or duplicated queries. There is no variation in surge frequency or surge multiples across service types within Lyft platform.

Table A2
NYC Taxi Trips Records.

Variable	Mean	Std. Dev.	Min	Max	N
Trip duration (minutes)†	14.378	11.436	1	180	47,483,432
Trip distance (miles)	2.985	3.515	0.01	199	47,483,432
Base fare	12.858	10.039	0	314	47,483,432
Extra fee	0.337	0.435	0	21.5	47,483,432
MTA tax	0.498	0.023	0	2.34	47,483,432
Tip	1.685	2.267	0	300	47,483,432
Tolls	0.270	1.249	0	111.65	47,483,432
Improvement fee	0.299	0.011	0	0.6	47,483,432
Total fare	15.955	12.293	2.54	360.34	47,483,432
Passenger count	1.631	1.283	0	9	47,483,432
Yellow taxi	0.887	0.315	0	1	47,483,432
Manhattan pickup	0.850	0.356	0	1	47,483,432
Queens pickup	0.084	0.277	0	1	47,483,432
Bronx pickup	0.005	0.074	0	1	47,483,432
Brooklyn pickup	0.060	0.237	0	1	47,483,432
Staten Island pickup	0.000	0.006	0	1	47,483,432
Manhattan dropoff	0.827	0.377	0	1	47,483,432
Queens dropoff	0.076	0.266	0	1	47,483,432
Bronx dropoff	0.012	0.110	0	1	47,483,432
Brooklyn dropoff	0.082	0.275	0	1	47,483,432
Staten Island dropoff	0.000	0.014	0	1	47,483,432

Note: Trip duration is calculated as the difference between the pickup time and the drop-off time. Unreasonable trips from the raw data are dropped using the following filters: trips with any negative cost components (cost components include base fare, extra fee, MTA tax, tip, tolls, improvement fee), trips with negative distance, trips with negative duration, trips greater than 200 miles, trips longer than 180 min. In total, less than 0.5% of the raw sample are dropped by these filters.

Appendix B. Ridesharing and Taxi Rides Across Hour of the Day

Besides the substitutability between ridesharing and taxi across geography observed in Fig. 2, we can also see this across time. Fig. B.11 shows that ridesharing gain larger market shares around 4 a.m. and 4 p.m., likely due to the morning and afternoon taxi shift changes[19] that create an outflux of taxi cabs from Manhattan.[20] This figure suggests that consumers substitute toward ridesharing more during the hours when taxis become less available.

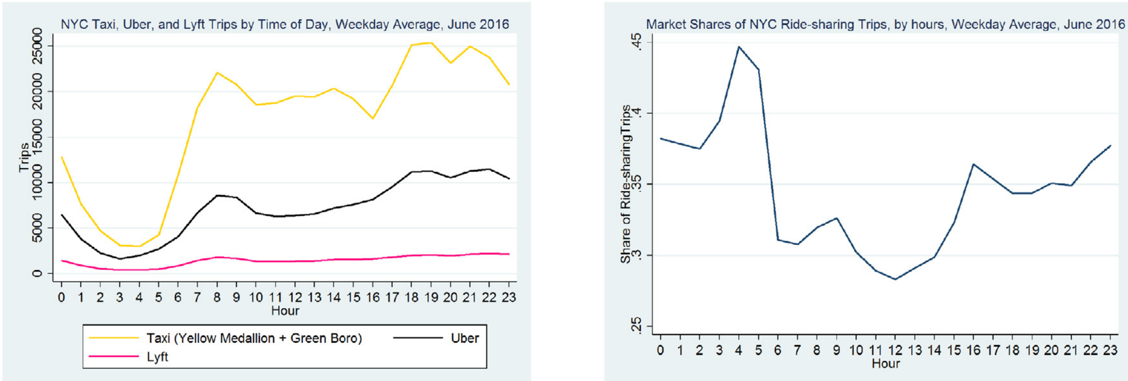
Appendix C. Identifying Uber and Lyft Trips from TLC FHV Trip Records Data

The FHV trip data does not specifically indicate the company name of each trip, instead it shows the trip's dispatching base number. Using the official TLC list of FHV bases[21], we are able to identify Uber and Lyft trips by the correspondence between base numbers and company names. Specifically, the base numbers associated with Uber are B02512, B02395, B02617, B02682, B02764, B02765, B02835, B02836, B02864, B02865, B02866, B02867, B02869, B02870, B02871, B02872, B02875, B02876, B02877, B02878, B02879, B02880, B02882, B02883, B02884, B02887, B02888, and B02889. The base numbers associated with Lyft are B02510 and B02844.

Appendix D. Field Collection of Uber and Lyft Trip Records

We conducted two rounds of field surveys of Uber and Lyft drivers. The first round took place in January 2017 and the second route took place in March 2017. In both rounds, we requested

19 The majority of yellow medallion cabs are operated two shifts per day.
20 Most of taxi leasing garages are located outside of Manhattan.
21 See http://www.nyc.gov/html/tlc/downloads/pdf/find_a_ride.pdf



(a) NYC taxi, Uber, and Lyft trips by time of an average weekday

(b) Shares of Uber and Lyft trips by time of an average weekday

Fig. B1. More ridesharing Trips during Taxi Shift-change Hours.

Table D1
Comparison between TLC Ridesharing Pickups and Field-collected Ridesharing Pickups.

	TLC data	Field
All Time		
LM	0.1884	0.1749
Midtown	0.2667	0.2863
UE & UW	0.0998	0.1448
NMC	0.4451	0.3940
Morning rush		
LM	0.1606	0.1487
Midtown	0.2331	0.2770
UE & UW	0.1482	0.1831
NMC	0.4580	0.3912
Evening rush		
LM	0.1921	0.1681
Midtown	0.3239	0.2875
UE & UW	0.1099	0.1538
NMC	0.3742	0.3906
Weekday night		
LM	0.2171	0.2096
Midtown	0.3051	0.3283
UE & UW	0.0804	0.1402
NMC	0.3973	0.3219
Weekday late night		
LM	0.1839	0.2009
Midtown	0.2303	0.3327
UE & UW	0.0717	0.1257
NMC	0.5142	0.3407
Weekday day time		
LM	0.1692	0.1863
Midtown	0.2736	0.2946
UE & UW	0.1155	0.1677
NMC	0.4417	0.3598
Weekend night		
LM	0.2002	0.1549
Midtown	0.2537	0.2296
UE & UW	0.0862	0.1213
NMC	0.4599	0.4942
Weekend late night		
LM	0.2356	0.1686
Midtown	0.2241	0.2584
UE & UW	0.0485	0.0705
NMC	0.4918	0.5437
Weekend day time		
LM	0.1671	0.1349
Midtown	0.2282	0.2348
UE & UW	0.1052	0.1381
NMC	0.4995	0.4923

for drivers' historical trip records from June to August 2016 so that all of our data sources are from the same time period.

We employed two sampling strategies in the data collection: a random sample and a convenience sample. In the collection of the random sample, the research team split into 4 groups, where two groups started around 9am at 2 locations in Manhattan, and the other two groups started at 2 locations in Brooklyn. Each group requested a trip to a randomly-selected borough out of Manhattan, Bronx, Brooklyn, and Queens. Upon arrival and before the driver answered a subsequent trip, the group made request to the driver for voluntary data disclosure. If the request was declined, they then offered a small sum of money in exchange for the data. The group collected as many trips as possible when the driver willingly accepted the request, either voluntarily or with a small sum of money, and the group chose to walk away when the monetary offer was rejected. The group repeated the same process throughout the day untill 9pm. In total, the research team collected 10,333 trips from 56 drivers out of 76 attempted.

For the convenience sample, we approached Uber and Lyft drivers at places where they either were taking a break and/or were between trips. These places include restaurants, coffee shops, street corners, and parking lots. We followed the same data request procedure as for the random sample. Because we had more time interacting with drivers and recording data this way, the convenience sample is multiple times as large as the random sample.

The validity of this paper's estimation approach crucially relies on how representative the field-collected ridesharing trips are of the population of ridesharing trips. We find support of the representativeness by demonstrating that the distribution of field-collected ridesharing pickups aligns with that of the TLC ridesharing pickups, across times and geographies (Table D.5).

Appendix E. Inference of D_{sjkt}

Due to the data limitation of Uber and Lyft trips published by TLC, the finest level one can get is D_{jt}^{Uber} and D_{jt}^{Lyft}, where D_{jt}^{Uber} measures the total trip counts of all 5 service types on Uber at the pickup location j in time t, and D_{jt}^{Lyft} is the total trip counts of all 3 service types on Lyft at the pickup location j in time t. To infer trip counts at a finer $sjkt$ (type-pickup-dropoff-time) level, we exploit the field-collected sample of 70,277 Uber and Lyft trip records using the following procedure:

Step 1: Construct a vector of zeroes, whose length is $s \times j \times k \times t$, i.e., at the type-month-day-time-hour-15min level. Fill any $sjkt$ cell with 1, if a trip is observed in that particular cell from the field-collected sample.[22] Then the vector contains0′s and1′s.

Step 2: Estimate a probit model of the vector in Step 1 to predict the probability of a trip in $sjkt$ by a number of location-time fixed effects:

$$\begin{aligned} &Pr(1 \text{ trip in } sjkt) \\ &\quad = f(\text{pickup zone, dropoff puma, service type, pickup hour,} \\ &\qquad \text{pickup borough} \times \text{pickup hour, pickup borough} \times \text{dropoff puma,} \\ &\qquad \text{pickup borough} \times \text{service type, dropoff puma} \times \text{service type}) \end{aligned}$$

Step 3: Calculate D_{sjkt} by distributing D_{jt}^{Uber} and D_{jt}^{Lyft} into service types and drop-off locations. This requires constructing weights using the estimated p_{sjkt} in Step 2, and applying the following formulas (Note that $s = 1, 2, 3, 4, 5$, for 5 service types on Uber, and s = 6, 7, 8, for 3 service types on Lyft):

For Uber:

$$D_{sjkt} = \frac{p_{sjkt}}{\sum_{k=1}^{263} \sum_{s=1}^{5} p_{sjkt}} D_{jt}^{Uber}$$

For Lyft:

$$D_{sjkt} = \frac{p_{sjkt}}{\sum_{k=1}^{263} \sum_{s=6}^{8} p_{sjkt}} D_{jt}^{Lyft}$$

These weights ensure that the inferred D_{sjkt}'s return the value of D_{jt}^{Uber} and D_{jt}^{Lyft}, when summed over service types and drop-off locations.

Appendix F. Feasible Generalized Two-Stage Least Square Estimator

Denote $e_{sjkt} = \xi_{cjkt} - \xi_{sjkt}$, where the variances of ξ_{cjkt} and ξ_{sjkt} are σ_c^2 and σ_s^2, respectively. Let S_{jkt} be the set of ridesharing service types available at a particular jkt; T_{jk} is the set of time periods for a particular route jk; K_j is the set of destinations for a particular pickup location j; J is the set of all available pickup locations. Further denote C_{sjkt} as the number of unique $sjkt$'s and C_{jkt} as the number of unique jkt's, which are calculated by summing the cardinality the relevant sets

$$C_{jkt} = \sum_{J} \sum_{K_j} |T_{jk}| \tag{9}$$

$$C_{sjkt} = \sum_{J} \sum_{K_j} \sum_{T_{jk}} |S_{jkt}| \tag{10}$$

Then the estimator can be constructed by the following procedure:

Step 1: Estimate Eq. (3) by a simple two-stage least squares regression without accounting for correlated errors, which leads to the composite residual $\hat{e}_{sjkt} = \widehat{\xi_{cjkt} - \xi_{sjkt}}$.

Step 2: Decompose the composite residual

$$\hat{\xi}_{cjkt} = \frac{1}{|S_{jkt}|} \sum_{S_{jkt}} \hat{e}_{sjkt} \tag{11}$$

and

$$\hat{\xi}_{sjkt} = \hat{e}_{sjkt} - \hat{\xi}_{cjkt} \tag{12}$$

[22] In several rare cases we observe two trips within the same $sjkt$ cell. In these cases, we randomly drop one of the two trips.

Step 3: Compute the variance estimates $\hat{\sigma}_c^2$ and $\hat{\sigma}_s^2$, using the decomposed residuals $\hat{\xi}_{cjkt}$ and $\hat{\xi}_{sjkt}$

$$\hat{\sigma}_c^2 = \frac{1}{C_{jkt}} \sum_{J} \sum_{K_j} \sum_{T_{jk}} (\hat{\xi}_{cjkt} - \frac{1}{C_{jkt}} \sum_{J} \sum_{K_j} \sum_{T_{jk}} \hat{\xi}_{cjkt})^2 \tag{13}$$

and

$$\hat{\sigma}_s^2 = \frac{1}{C_{sjkt}} \sum_{J} \sum_{K_j} \sum_{T_{jk}} \sum_{S_{jkt}} (\hat{\xi}_{sjkt} - \frac{1}{C_{sjkt}} \sum_{J} \sum_{K_j} \sum_{T_{jk}} \sum_{S_{jkt}} \hat{\xi}_{sjkt})^2 \tag{14}$$

Step 4: Construct a $C_{sjkt} \times C_{sjkt}$ block diagonal matrix with C_{jkt} number of blocks

$$\hat{\Omega}^{-\frac{1}{2}} = \begin{bmatrix} \ddots & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \left[\hat{\mathbb{Q}}_{jkt}\right]_{|S_{jkt}| \times |S_{jkt}|} & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \ddots \end{bmatrix}_{C_{sjkt} \times C_{sjkt}} \tag{15}$$

where each block element $(\left[\hat{\mathbb{Q}}_{jkt}\right]_{|S_{jkt}| \times |S_{jkt}|})$ is a square matrix with diagonal elements equal to $\frac{1}{|S_{jkt}|}(\frac{1}{(|S_{jkt}|\hat{\sigma}_c^2 + \hat{\sigma}_s^2)^{\frac{1}{2}}} - \frac{1-|S_{jkt}|}{\hat{\sigma}_s})$ and off-diagonal elements equal to $\frac{1}{|S_{jkt}|}(\frac{1}{(|S_{jkt}|\hat{\sigma}_c^2 + \hat{\sigma}_s^2)^{\frac{1}{2}}} - \frac{1}{\hat{\sigma}_s})$. $\mathbb{O}$ are matrices of zeros.

Step 5: The random effects estimator can be constructed explicitly as

$$(\hat{\alpha} \ \hat{\beta} \ \hat{\Theta})' = (\mathbb{X}^{*'} \mathbb{Z}^* (\mathbb{Z}^{*'} \mathbb{Z}^*)^{-1} \mathbb{Z}^{*'} \mathbb{X}^*)^{-1} \mathbb{X}^{*'} \mathbb{Z}^* (\mathbb{Z}^{*'} \mathbb{Z}^*)^{-1} \mathbb{Z}^{*'} \mathbb{D}^* \tag{16}$$

where $\mathbb{Z}$ denotes the matrix that contains all the instruments, and

$$\mathbb{X} = \begin{bmatrix} \vdots & \vdots & \vdots \\ P_{sjkt} - P_{cjkt} & WT_{sjt} - WT_{cjt} & (X_{cjkt} - X_{sjkt})' \\ \vdots & \vdots & \vdots \end{bmatrix}$$

$$\mathbb{D} = \begin{bmatrix} \vdots \\ \log(\frac{D_{cjkt}}{D_{sjkt}}) \\ \vdots \end{bmatrix}$$

$$\mathbb{X}^* = \hat{\Omega}^{-\frac{1}{2}} \mathbb{X}$$
$$\mathbb{D}^* = \hat{\Omega}^{-\frac{1}{2}} \mathbb{D}$$
$$\mathbb{Z}^* = \hat{\Omega}^{-\frac{1}{2}} \mathbb{Z}$$

Appendix G. First Stage Tests for Instrumental Variables

Table G.6 reports first-stage regressions to test the strengths of the instruments. Each number represents either the F-statistic or the partial R^2 associated with the regression of an endogenous variables on all regressors. For example, the fist number in the table, 431.85, represents the F statistic of all instruments with respect to prices of Lower Manhattan in morning rush hours. The partial R^2, 0.097, measures how much of the variation in prices of Lower Manhattan in morning rush hours is explained by all instruments. Given that all F-statistics are bigger than 10 (Staiger and Stock, 1997), and the partial R^2's are relatively high (Shea, 1997), our instruments are not weak instruments.

Table G1
First Stage F-test and Partial R^2.

	Price		Wait Time	
	F-stat	Partial R^2	F-stat	Partial R^2
Morning rush				
LM	431.85	0.097	5039.903	0.035
Midtown	389.582	0.155	3394.364	0.083
UE & UW	3952.628	0.118	2682.597	0.062
NMC	33.927	0.080	12.481	0.044
Evening rush				
LM	1113.876	0.146	1538.263	0.096
Midtown	586.41	0.168	1691.842	0.114
UE & UW	3062.924	0.123	64973.491	0.139
NMC	28.592	0.092	16.06	0.078
Weekday night				
LM	709.163	0.132	7716.904	0.115
Midtown	1601.706	0.140	2400.526	0.126
UE & UW	4963.373	0.143	130478.994	0.147
NMC	36.802	0.095	42.819	0.102
Weekday late night				
LM	2758.778	0.146	3040.974	0.074
Midtown	13701.841	0.178	4484.22	0.091
UE & UW	2965.271	0.163	6476.689	0.103
NMC	33.965	0.110	26.895	0.06
Weekday day time				
LM	1342.167	0.200	1055.851	0.077
Midtown	1942.77	0.244	1425.041	0.111
UE & UW	3207.152	0.185	4107.147	0.074
NMC	39.43	0.076	10.827	0.038
Weekend night				
LM	3511.998	0.147	5324.336	0.136
Midtown	2488.762	0.158	1489.467	0.143
UE & UW	20343.801	0.144	170436.431	0.145
NMC	47.637	0.123	46.362	0.094
Weekend late night				
LM	734.557	0.124	3788.657	0.147
Midtown	7870.996	0.174	6663.631	0.147
UE & UW	44411.786	0.163	12545.38	0.188
NMC	58.275	0.132	30.523	0.169
Weekend day time				
LM	1815.033	0.147	4033.533	0.094
Midtown	7080.772	0.181	3072.913	0.106
UE & UW	4326.934	0.142	14718.355	0.100
NMC	55.469	0.096	27.511	0.078
To Airport	165.552	0.225	100.863	0.077
Rain			1128.828	0.045

CRediT authorship contribution statement

Chungsang Tom Lam: Software, Formal analysis, Resources, Data curation. **Meng Liu:** Conceptualization, Methodology, Software, Formal analysis, Writing – original draft. **Xiang Hui:** Writing – original draft, Writing – review & editing.

References

Agrawal, A., Catalini, C., Goldfarb, A., 2015. Crowdfunding: geography, social networks, and the timing of investment decisions. J. Econ. Manage. Strategy 24 (2), 253–274.

Babar, Y., Burtch, G., 2017. Examining the impact of ridehailing services on public transit use.

Balestra, P., Varadharajan-Krishnakumar, J., 1987. Full information estimations of a system of simultaneous equations with error component structure. Econom. Theory 3 (2), 223–246.

Berry, S., Levinsohn, J., Pakes, A., 1995. Automobile prices in market equilibrium. Econometrica 841–890.

Berry, S.T., 1994. Estimating discrete-choice models of product differentiation. Rand. J. Econ. 242–262.

Blum, B.S., Goldfarb, A., 2006. Does the internet defy the law of gravity? J. Int. Econ. 70 (2), 384–405.

Brynjolfsson, E., Hui, X., Liu, M., 2019. Does machine translation affect international trade? Evidence from a large digital platform. Manage. Sci. 65 (12), 5449–5460.

Buchholz, N., 2015. Spatial Equilibrium, Search Frictions and Efficient Regulation in the Taxi Industry. Technical Report. Working paper.

Buchholz, N., Doval, L., Kastl, J., Matějka, F., Salz, T., 2020. The Value of Time: Evidence from Auctioned Cab Rides. Technical Report. National Bureau of Economic Research.

Calder-Wang, S., 2019. The distributional impact of the sharing economy on the housing market.

Castillo, J.C., Knoepfle, D., Weyl, G., 2017. Surge pricing solves the wild goose chase. In: Proceedings of the 2017 ACM Conference on Economics and Computation. ACM, pp. 241–242.

Chen, M.K., Chevalier, J.A., Rossi, P.E., Oehlsen, E., 2017. The Value of Flexible Work: Evidence from Uber Drivers. Technical Report. National Bureau of Economic Research.

Chen, M.K., Sheldon, M., 2016. Dynamic pricing in a labor market: Surge pricing and flexible work on the Uber platform.. In: EC, p. 455.

Chevalier, J., Goolsbee, A., 2009. Are durable goods consumers forward-looking? Evidence from college textbooks. Q. J. Econ. 124 (4), 1853–1884.

Cohen, P., Hahn, R., Hall, J., Levitt, S., Metcalfe, R., 2016. Using Big Data to Estimate Consumer Surplus: The Case of Uber. Technical Report. National Bureau of Economic Research.

Cramer, J., Krueger, A.B., 2016. Disruptive change in the taxi business: the case of Uber. Am. Econ. Rev. 106 (5), 177–182.

Farber, H.S., 2015. Why you can't find a taxi in the rain and other labor supply lessons from cab drivers. Q. J. Econ. 130 (4), 1975–2026.

Farronato, C., Fradkin, A., 2017. The Welfare Effects of Peer Entry in the Accommodations Market: The Case of Airbnb. Technical Report. Working paper.

Frechette, G. R., Lizzeri, A., Salz, T., 2016. Frictions in a competitive, regulated market evidence from taxis.

Goldszmidt, A., List, J.A., Metcalfe, R.D., Muir, I., Smith, V.K., Wang, J., 2020. The Value of Time in the United States: Estimates from Nationwide Natural Field Experiments. Technical Report. National Bureau of Economic Research.

Greenwood, B.N., Wattal, S., 2017. Show me the way to go home: an empirical investigation of ride-sharing and alcohol related motor vehicle fatalitie. MIS Q. 41 (1).

Hall, J., Horton, J., Knoepfle, D., 2017. Labor market equilibration: evidence from Uber.

Hall, J., Kendrick, C., Nosko, C., 2015. The Effects of Uber's Surge Pricing: A Case Study. The University of Chicago Booth School of Business.

Hall, J.D., Palsson, C., Price, J., 2018. Is Uber a substitute or complement for public transit? J. Urban Econ. 108, 36–50.

Hall, J.V., Krueger, A.B., 2016. An Analysis of the Labor Market for Uber's Driver–Partners in the United States. Technical Report. National Bureau of Economic Research.

Hering, L., Poncet, S., 2010. Market access and individual wages: evidence from China. Rev. Econ. Stat. 92 (1), 145–159.

Holzer, H.J., Ihlanfeldt, K.R., Sjoquist, D.L., 1994. Work, search, and travel among white and black youth. J. Urban Econ. 35 (3), 320–345.

Hortaçsu, A., Martínez-Jerez, F.A., Douglas, J., 2009. The geography of trade in online transactions: evidence from eBay and MercadoLibre. Am. Econ. J. 53–74.

Ihlanfeldt, K.R., Sjoquist, D.L., 1990. Job accessibility and racial differences in youth employment rates. Am. Econ. Rev. 80 (1), 267–276.

Ihlanfeldt, K.R., Sjoquist, D.L., 1991. The effect of job access on black and white youth employment: a cross-sectional analysis. Urban Stud. 28 (2), 255–265.

Kaufman, S., Moss, M. L., Tyndall, J., Hernandez, J., 2014. Mobility, economic opportunity and new york city neighborhoods.

Lagos, R., 2000. An alternative approach to search frictions. J. Polit. Economy 108 (5), 851–873.

Lagos, R., 2003. An analysis of the market for taxicab rides in new york city. Int. Econ. Rev. 44 (2), 423–434.

Leard, B., Xing, J., 2020. What Does Ridesharing Replace? Resources for the Future working paper 20–03.

Lee, D. L., McCrary, J., Moreira, M. J., Porter, J., 2020. Valid t-ratio inference for iv. arXiv preprint arXiv:2010.05058.

Li, Z., Hong, Y., Zhang, Z., 2016. An empirical analysis of on-demand ride sharing and traffic congestion.

Liu, M., Brynjolfsson, E., Dowlatabadi, J., 2018. Do Digital Platforms Reduce Moral Hazard? The Case of Uber and Taxis. Technical Report. National Bureau of Economic Research.

Liu, M., Tunca, T., Xu, Y., Zhu, W., 2017. An empirical analysis of price formation, utilization, and value generation in ride sharing services.

Nevo, A., 2000. Mergers with differentiated products: the case of the ready-to-eat cereal industry. Rand. J. Econ. 395–421.

Nevo, A., 2001. Measuring market power in the ready-to-eat cereal industry. Econometrica 69 (2), 307–342.

Orr, D., 1969. The " axicab problem": a proposed solution. J. Polit. Economy 77 (1), 141–147.

Owen, A., Murphy, B., 2018. Access across America: Transit 2017.

Petrin, A., 2002. Quantifying the benefits of new products: the case of the minivan. J. Polit. Economy 110 (4), 705–729.

Schaefer, M., Tran, K., 2020. Airbnb, hotels, and localized competition.

Shea, J., 1997. Instrument relevance in multivariate linear models: a simple measure. Rev. Econ. Stat. 79 (2), 348–352.

Staiger, D., Stock, J.H., 1997. Instrumental variables regression with weak instruments. Econometrica 65 (3), 557. doi:10.2307/2171753. http://www.jstor.org/stable/2171753?origin=crossref

Tomer, A., Kneebone, E., Puentes, R., Berube, A., 2011. Missed opportunity: transit and jobs in metropolitan america.
Wang, Y., Wu, C., Zhu, T., 2019. Mobile hailing technology and taxi driving behaviors. Market. Sci. 38 (5), 734–755.
Zervas, G., Proserpio, D., Byers, J.W., 2017. The rise of the sharing economy: estimating the impact of airbnb on the hotel industry. J. Market. Res. 54 (5), 687–705.
Zhang, K., Chen, H., Yao, S., Xu, L., Ge, J., Liu, X., Nie, M., 2019. An Efficiency Paradox of Uberization. Available at SSRN 3462912.
Zhang, Z., Li, B., 2017. A quasi-experimental estimate of the impact of P2P transportation platforms on urban consumer patterns. In: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, pp. 1683–1692.