Title
On-demand Mobility of the Future: Equity, Behavior and Policy

Author
Xu, Ruoying

Publication Date
2019

On-demand Mobility of the Future：

Equity, Behavior and Policy

By

Ruoying Xu

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy
in
City and Regional Planning
in the
Graduate Division
of the
University of California, Berkeley

Committee in charge:

Professor Daniel G. Chatman, Chair
Professor Daniel Rodriguez
Professor Paul Waddell
Professor Joan Walker

Summer 2019

COPYRIGHT

On-demand Mobility of the Future：

Equity, Behavior and Policy

ABSTRACT

On-demand Mobility of the Future：

Equity, Behavior and Policy

By

Ruoying Xu

Doctor of Philosophy in City and Regional Planning

University of California, Berkeley

Professsor Daniel G. Chatman, Chair

The rapid emergence and growth of transportation network companies (TNCs) such as Uber, Lyft, and Didi Chuxing, operating app-based on-demand ride-sourcing services, has led to a debate over the role of TNCs in the urban transport system. The growth of the ride-sourcing business has brought significant challenges for planners, engineers, and policy makers, due to the magnitude and uncertainty of its impacts. This dissertation focuses on several aspects of on-demand mobility, mostly related to equity and behavior, and answers some of the most debated questions about ride-sourcing to provide important evidence for engineers, planners and policy makers on future ride-sourcing related policy decisions.

The equity analysis investigates the effects of ride-sourcing fare changes to passengers with different socio-economic backgrounds. Using a large GPS-based travel dataset from 2015 in Shanghai, I conducted panel analysis of how ride-sourcing demand was related to average property value, as a proxy for socioeconomic status, measured at small spatial scale at trip origins and destinations. I modeled the ride-sourcing demand (for pick-ups and drop-offs separately) as a product of several spatial and temporal characteristics, using a negative binomial regression with fixed effects whose functional form is appropriate for dispersed count data. The results imply that a decrease in ride-sourcing fares would likely benefit middle to high income travelers more than low-income travelers, by making ride-sourcing an economically competitive mode for those groups. Usage is much higher in neighborhoods with higher property values when fares are lower. At the same time, however, there is still significant though lower use of ride-sourcing in lower-income neighborhoods, and usage in those locations is *less* responsive to the fare. I conclude that ride-sourcing policy which results in fare increases would likely to pose a substantial burden for lower-income travelers, although the number of such lower income travelers may be small compared to the number of middle-to-high income travelers.

The behavioral analysis focuses on three questions: (1) Is parking supply associated with lower ride-sourcing demand? (2) Does better transit access reduce or increase the use of ride-hailing? and (3) Does higher congestion affect ride-sourcing demand? I modeled the ride-sourcing demand (for pick-ups and drop-offs separately) using a generalized additive mixed model (GAMM). The results suggest that first, parking would not necessarily reduce the demand for ride-sourcing unless the parking supply is large enough. Second, whether ride-sourcing compete or complement bus transit depends on the coverage of bus services: they tended to compete when the density of bus stops is high, and complement each other when there are fewer bus stops. Third, ride-sourcing demand was positively correlated with congestion, except that when congestion is severe, there were fewer pick-ups.

In addition to the panel study, I used Google Map API to figure out if each actual ride-sourcing trip has transit alternative. I found that over 90% of the actual ride-sourcing trips have transit alternatives, but transit compete poorly with ride-sourcing because of much longer travel time, need multiple transfer and longer walking. Finally, I discussed policy implication based on these empirical findings.

To those who encourage and enlighten me

Table of contents

Acknowledgement

I would like to take this special moment to express my sincere gratitude to those who have supported me during my PhD endeavor at Berkeley.

First and foremost, I would like to thank Dr. Daniel Chatman for his guidance and mentorship. His encouragement and advises had made me become not just a better PhD scholar with critical thinking, but also a better human being who understand inclusiveness and tolerance. I will always remember his mentorship and friendship in the future. And I sincerely wish him good health.

I greatly appreciate Dr. Daniel Rodriguez, my committee member, for his advises on my research and his contributions in creating better peer relationship within graduate students. His transportation seminar helped me broaden my scope of view. I also deeply acknowledged support and suggestions from Dr. Paul Waddell and Dr. Joan Walker for their help in enhancing my quantitative skills and formulating my research.

Study at Berkeley makes my PhD journey enjoyable, due to the academic environment, respectful professors, and brilliant people around. Over these years, I enjoyed the friendship and mental support from my peers at Berkeley, to name a few: Andrea Broaddus, Abigail Cochran, Yiyan Ge, Yang Ju, Sam Maurer, Lisa Rayle, Rocio Sanchez-Moyano, Manuel Santana, Xize Wang.

I would also like to thank Binliang Guo and Lutuo Gao from MetroData Tech. who provided me the dataset for my study. I found myself fortunate to know these brilliant people from MetroData Tech.

I greatly appreciate that I was born in my family, because I received tremendous care and enlightenment from my parents and grandparents. My family defined my place of happiness and sweetness, as well as my desire to go boldly.

I owe a thanks to my fiancée, Dr. Shanshan Wang. She is my happy “Miss Sunshine”, and I love to spend the rest of our lives happily together. Last but not least, I greatly appreciate the accompany of our lovely cat, Naonao, who is a great stress reliever.

Ruoying Xu

Aug. 12th, 2019

Chapter 1. Introduction

The rapid emergence and growth of transportation network companies (TNCs) such as Uber, Lyft, and Didi Chuxing, operating app-based on-demand ride-sourcing services, has led to a debate over their role in the urban transport system. Ride-sourcing supporters argue that it is a convenient and efficient mode of travel, with the potential to reduce private auto use, decrease the need for parking and solve the first/last mile problem of public transit (Rayle et al. 2016; Shaheen et al. 2016). Other scholars and government officials are less sanguine, concerned that the growing popularity of ride-sourcing might undermine congestion mitigation and replace public transportation, eventually resulting in negative environmental and social consequences (Rayle et al. 2016, SFCTA 2017, Schaller, 2018). The growth of the ride-sourcing business has brought significant challenges for planners, engineers, and policy makers, due to the magnitude and uncertainty of its impacts. For example, the number of full- and part-time TNC drivers in Shanghai exceeded 287,700 in August 2016, a mere two years after ride-sourcing was introduced (Didi Chuxing 2016). In 2017 it was estimated that more than 15% of trips and 20% of vehicle mileage in San Francisco was produced by TNCs (SFCTA 2017).

Despite the exponential growth of ride-sourcing worldwide, information about its relationship with parking, public transit, and congestion is still very limited. These are measures of the relative ease of driving and using transit, and might predict lower or higher use of ride-sourcing as an alternative. Thus one might expect that ride-sourcing demand would be higher in places with restricted parking, at least for auto owners. One might also expect that demand would be higher in places with lower levels of transit access because of the significant time advantages it could offer there, despite ride-sourcing being more expensive than transit. Additionally, it is hypothesized that ride-sourcing may be positively associated with congestion because it may relieve travelers from driving on congested roads, or because it may in fact cause congestion. Also, there are no empirical evidence so far that assess the equity impact on travelers when ride-sourcing fare changes due to policy changes.

This dissertation explores ride-sourcing services from the several perspectives: equity impact of ride-sourcing on passengers; ride-sourcing's relation with public transportation, parking and congestion; and discussion of municipal policy on ride-sourcing. The third chapter investigates whether individual ride-sourcing trips can be replaced by public transportation by descriptive analysis using Google Map Direction API. The fourth chapter explores the impact on ride-sourcing users after price changes. The fifth chapter analyzes the relationship between ride-sourcing and parking supply, public transportation, and congestion. The sixth chapter discusses the implication for policy from the empirical results of data analysis.

1.1 Background of ride-sourcing

There is no widely agreed-upon definition of ride-sourcing partly because of its evolving nature. Generally speaking, ride-sourcing services utilize information and communication technology (ICT) applications to spatially match ride demand and supply, and the vehicles are usually privately owned (Hall and Krueger, 2015). Drivers and passengers within a region can be easily paired up by using the same application. These apps are provided by transportation network companies (TNC), such as Uber, Lyft and Didi Kuaidi. The majority of the fare, which is automatically billed to the passenger's credit card after the ride, goes to the driver, and the remaining amount goes to the TNCs. The TNCs do not directly hire drivers. They are responsible for maintaining the ride-sourcing application and providing travel information. In some areas in the US such as California and New York City, TNC drivers are considered employers and thus require a minimum wage equivalent.

Ride-sourcing services share many similarities with traditional taxis, at least from a passenger's standpoint, as they both offer ride services in exchange for a fare. However, there are distinct differences between ride-sourcing services and taxis. Since specific characteristics of ride-sourcing service differ from place to place and also evolve fairly quickly, in this dissertation I focus on characterizing ride-sourcing in China in the present day. Table 1.1 shows the comparison between the features of ride-sourcing services and taxis. These two services differ in three major aspects: supply, fare, and ease of driver entry. The supply and cost of ride-sourcing services are both responsive to demand while the total price and number of taxis are heavily regulated. From 2003 to 2013, the total number of taxis in Shanghai fluctuated around 50,000, despite the fact that GDP per capita in Shanghai had grown 2.3 times larger (Ministry of Transport of China, 2015). This suggests that there was a huge gap between travel demand and taxi supply.

Table 1.1. Comparing features of ride-sourcing and traditional taxis in China, 2015

	Ride-sourcing	Traditional Taxi
Supply	Regionally dynamic supply depending on demand	Fixed supply of taxis (~50,000)
Fare	Can be demand-responsive	Fixed price, subjected to heavy regulation
Driver entry	Both part-time and full-time drivers: pay fees to TNCs proportionately to ride services provided	Full-time driver: pay fixed fee to taxi company

There are also debates over whether ride-sourcing falls into the category of sharing economy. Sharing economy refers to peer-to-peer sharing of access to goods or services through online application (Hamari et al., 2015). A commonly-used example of sharing economy in transportation is ride-sharing. Ride-sharing groups travelers who have similar destinations in a private vehicle with the goal of reducing cost, congestion and emission (Chan and Shaheen, 2012). It allows the sharing of idle resources, in this case vehicle occupancy, and potentially reduces on-road vehicles. Ride-sourcing differs from ride-sharing because the driver usually does not share the same destination of the passenger. Instead of the sharing of vehicle occupancy, ride-sourcing allows the sharing of idle mobility of the vehicle, which potentially reduces the need of owning additional vehicles. From this perspective, ride-sourcing still falls into the category of the sharing economy. Also, TNCs have introduced shared rides (carpools) as an option for ride-sourcing services.

In China, ride-sourcing services first became available in 2014 (TalkingData 2014). By the end of 2014, two TNCs, Didi and Kuaidi, had claimed the largest market share in China. In order to compete with Uber locally and worldwide, the two TNCs merged, becoming Didi Chuxing in

February 2015. By May 2015, about 15 percent of the population in China had used TNC apps, and 80 percent of those had used Didi Chuxing (TalkingData 2015).

The introduction of TNCs in Shanghai altered the transport landscape. As in other parts of the world, traditional taxis and TNCs in Shanghai differed in three main ways: supply of vehicles, fare, and ease of driver entry. The supply and cost of ride-sourcing services were at this time quite responsive to demand, in contrast to the number of taxis and taxi fares, both of which were heavily regulated. From 2003 to 2013, the number of taxis in Shanghai fluctuated around 50,000, while its population increased from 16 to 24 million, and GDP per capita grew 2.3 times (Ministry of Transport of People's Republic of China 2015). The stagnation of taxi supply over this decade-long period can be explained by high entry costs. Taxi vehicle licenses in Shanghai may cost up to ￥500,000 per vehicle (around $77,000 US), and in addition drivers paid a fixed commission fee of ￥8,200 each month to the taxi company, which was approximately 30 to 40 percent of the driver's monthly income (China National Radio 2015). In contrast, the number of TNC drivers (both full-time and part-time) in Shanghai exceeded 287,700 in August 2016, a mere two years after the ride-sourcing service was introduced in September 2014 (Didi Chuxing 2016). There was no license fee to become a TNC driver in 2015, and the only commission fee was a slight fraction of the fare. The growth of ride-sourcing services in China during this time occurred under minimal government intervention, despite boycott attempts mounted by taxi companies and some policy advocates. The rapid growth was also enabled by ubiquitous smartphone ownership, as more than 94 percent of the adult population possessed smart phones in 2013 (Smart Device Business Center 2013), and there were 137 smartphones in Shanghai per 100 people in 2015 (FORWARD Business Information Co. Ltd. 2019). In comparison, smartphone ownership in the US was about 77% in 2018 (Pew Research Center 2018).

Similar to other cities worldwide, the growth of TNCs and ride-sourcing in Shanghai has triggered great controversy among scholars and policy makers for its potential impacts on travel patterns, particularly road congestion and pollution, parking availability and its potential effects on public transit. The municipal government of Shanghai has long considered promoting public transit and auction for private vehicle license plate as an important part of its strategy for traffic mitigation (Chen and Zhao 2013). Shanghai now has the world's largest public transit system (673 kilometers of metro and 1461 bus routes with more than 16,000 buses), and it has also low auto ownership (3.6 million cars with more than 24 million people) compared to other large Chinese cities (Shanghai Urban-Rural Construction and Traffic Development Academy 2015). Public transit, private vehicles, and walking/biking respectively accounted for about 22, 17 and 57 percent of all person trips in 2015 (Shanghai Urban-Rural Construction and Traffic Development Academy 2015). Although public transit took up a larger share of person trips than private cars, it is likely that private cars occupy more space on the road, making a greater contribution to congestion, due to lower passenger occupancy per vehicle.

What makes Shanghai a particularly interesting place to study ride-sourcing use is the fact that due to its license auction policy the metropolitan area has relatively low auto ownership given its size and relative affluence (Chen and Zhao 2013). In Shanghai, licenses typically cost about ￥80,000 per vehicle in 2015 (about $12,300 in US dollars). As a result of the policy, in Shanghai automobiles are ordinarily owned and used only by residents with medium to high income (Chen and Zhao 2013). In 2014, there were less than 15 private cars for every 100 residents in Shanghai, and private cars accounted for about 17% of trips (Shanghai Urban-Rural

Construction and Traffic Development Academy 2015). Compared to auto ownership, ride-sourcing is a comparatively accessible travel option, as 94.4% of the adult population in Shanghai owned at least one smart phone in 2013 (Smart Device Business Center 2013) and at least during some periods of operation, the average ride-sourcing fare was quite low at ￥1 per kilometer. For comparison, the average hourly wage in Shanghai was around ￥35 in 2015 (Shanghai Bureau of Statistics 2015).

Chapter 2. Literature review

To date, empirical research investigating ride-sourcing from a user perspective remains limited, primarily due to data availability. This literature review covers the existing literature from the following perspectives: ride-sourcing's competition with public transit, equity, and ride-sourcing's relationship with parking and congestion.

2.1 Relationship with public transit

Ever since the introduction of ride-sourcing, its relationship with transit has a focal point of discussion; some have worried that ride-sourcing services would compete with transit, drawing away transit riders; on the other hand, ride-sourcing could help provide an answer to the "last mile problem" of transit, providing connections from transit stations to final origins and destinations (Shaheen and Chan 2016).

A few studies have directly surveyed how TNC users travel along with their attitudes towards travel. One of the first such studies, using survey data from San Francisco in 2014, concluded that ride-sourcing was more a foe than a friend of transit. Analyzing 380 survey responses, the authors found that just 5% of ride-sourcing trips in San Francisco provided access to transit stations, while 33% of survey respondents said that they used ride-sourcing to replace transit, and an additional 6% of the ride-sourcing trips were conducted because public transit was not available (Rayle et al. 2016). The authors also found some additional evidence suggesting that TNCs might replace transit: 81% of ride-sourcing trips ended within 200 meters of a bus stop, and 28% of trips ended within 400 meters of a rail station. Another survey of more than 4,000 respondents in seven major US cities found that whether ride-sourcing substitutes with or complements public transport is dependent on the type of transit: ride-sourcing users reported that they reduced bus trips and light rail trips by 3% to 6%, while increasing heavy rail use by 3%, after they started to ride-hail (Clewlow and Mishra 2017). The same study found that "not enough transit stations" and "transit service is not available" were the second and third most important reasons why ride-sourcing users stated they preferred to use TNCs. Nevertheless, some transit agencies in the US believe that ride-sourcing may benefit public transit given that between 3 and 16% of TNC trips in the US are transit access trips (Feigon and Murphy 2018).

Three studies used panel data and econometric models to explore whether the entrance of ride-sourcing service into a city leads to a reduction in transit use. The results vary. One study applied a regression discontinuity design on 30 urbanized areas in the US and found that the entrance of the first TNC in the metropolitan area was associated with an increase in transit ridership, suggesting an initially complementary relationship, but that the entrance of the second TNC subsequently *reduced* transit ridership (Sadowsky and Nelson 2017). The authors speculated that this effect could have been due to a drop in TNC fares with the entrance of the second company, due to competition. The same study concluded that ride-sourcing is more likely to substitute for rail transit, and that there was no evidence of competition against bus transit. Another study using agency-level ridership data, and employing a difference-in-differences approach with matching pre-treatment trends, showed that ride-sourcing competed with buses but complemented rail transit: the introduction of ride-sourcing seemed to decrease bus ridership by around 1%, but increased rail ridership by between 2% and 7% depending on the type of rail (Babar and Burtch 2017). The third study used a similar difference-in-differences approach to study how Uber entrance affected metropolitan-level transit ridership from 2004-2015 on 196 metropolitan areas (Hall et al. 2018). The researchers found that Uber increased transit ridership for smaller transit agencies and in larger cities, and that the complementarity effect tended to be larger for rail than for bus ridership, possibly due to the fact that rail users are wealthier.

Two studies used aggregated ride-sourcing trip data in New York to explore the association with transit. One cross-sectional analysis using 2015 Lyft and Uber data in New York City showed that the number of ride-sourcing pick-ups was positively associated with transit frequency, and the authors concluded that ride-sourcing was associated with more transit ridership (Mahmoudi and Zhang 2018). Another study using a fixed-effects regression model to explore the relationship between Uber pick-ups and the turnstile traffic at rail transit entrances for 156 taxis zones and for a period of about one year in New York found that each turnstile entrance was associated with an additional 0.005 Uber pick-ups in every 4-hour time interval in each zone, again indicating complementarity of transit and TNCs (Hoffmann et al. 2016).

Since ride-sourcing is, to a certain extent, similar to traditional taxi, studies on the relationship between taxi and transit could shed lights on how ride-sourcing affects public transit. Austin and Zegras (2012) investigated the effects of transit proximity on taxi trip generation, and found that taxi both complements and substitutes urban rail and bus, and the substitution effect is stronger when the service level of nearby transit is lower. King et al. (2012) explored the patterns of taxi trips in New York and found that the origins and destinations of taxi trips have different, asymmetrical distributions, and many taxi trips are accompanied by transit trips, suggesting a complementary relationship between them.

2.2 Equity

Rayle et al. (2016) carried one of the first studies collecting socioeconomic data from ride-sourcing users. From a pool of 380 users in San Francisco in 2014, they found that the majority of users were medium-high income and well educated, with very few lower-income travelers (Rayle et al. 2016). These findings were re-affirmed by a recent study, which found that affluent US travelers with college degrees made twice as many ride-sourcing trips compared to less educated and poorer peers from a survey of more than 4,000 respondents in seven major US cities (Clewlow and Mishra 2017). Another study of 394 online survey respondents across 15 US metros found that ride-sourcing users tended to be younger and well-educated, but that household income was not so significantly associated with use (Dawes 2016). A recent study based on an online survey in Austin, Texas of 1,840 former TNC users after the suspension of Uber/Lyft services there in 2016, found the majority of users were from households with income exceeding $100,000 (Hampshire et al. 2017).

Thus most research has showed ride-sourcing services being used primarily by wealthier residents. However it is possible that in some contexts the demographic composition of ride-sourcing users may be similar to those of taxis, which tends to be somewhat bifurcated, either above average income or very poor (Pucher and Renne 2003; Renne and Bennett 2014). Across metropolitan areas ride-sourcing is often cheaper than taxis, and thus has the potential to improve the mobility of lower income households, especially those without cars. A recent study did in fact observe high ride-sourcing usage among lower income neighborhoods as well as neighborhoods with high minority population in some US cities (Feigon and Murphy 2018). Another recent study using data from the U.S. National Household Travel Survey 2016-2017 showed that ride-sourcing is used disproportionately by travelers from the poorest income groups in smaller urban areas, possibly due to lower vehicle ownership and the lack of public transit in these areas (Schaller 2018).

Two other studies, in Seattle and Atlanta respectively, analyzed the association between the average wait time for Uber service and Census block level socio-economic and built environment characteristics (Hughes and MacKenzie 2016; Wang and Mu 2018). They both concluded that dense areas have better access to Uber services in terms of shorter wait time probably due to higher supply, and it does not appear that block groups with predominantly lower-income households would have significant higher wait time for Uber services. These studies provide valuable information about spatial variation in the potential accessibility to ride-sourcing services, but do not reveal spatial variation in trip-making that would help understand usage by different income levels. Another recent dissertation by Brown investigated Lyft data in Los Angeles found that higher ride-sourcing trips are correlated with lower neighborhood auto-ownership, suggesting that ride-sourcing helped improve people's mobility in zero-car households (Brown 2018). Brown also concluded that Lyft served higher proportion of low income users as well as high income users (Brown 2018). But the question of how changes in ride-sourcing fare may influence use by households of different income levels remains. Such data and analysis could be used to inform policy and regulations on ride-sourcing.

2.3 Parking and congestion

Researchers have speculated that the advent of ride-sourcing, as well as shared autonomous vehicles, would reduce the need for parking, thus generating redevelopment opportunities in cities (Shaheen and Cohen 2018). But empirical evidence of how parking affects ride-sourcing is very limited and the only empirical research available consists of surveys of user's attitudes. Rayle et el. (2016) found that 18% of the 380 respondents in San Francisco reported choosing ride-sourcing to avoid having to park a car. Other studies have also concluded from their surveys that parking is one of the primary reasons why TNC users chose ride-sourcing over driving (Clewlow and Mishra 2017, Schaller 2018).

Another focal point of the ride-sourcing discussion is its relationship with congestion (Shaheen and Cohen 2018). On one hand, ride-sourcing has been criticized by planners and city governments for causing congestion (SF Chronicle, 2018). On the other hand, some researchers believed that ride-sourcing may reduce auto use by allowing vehicles to be shared among multiple passengers and allowing more efficient use of vehicles (Shaheen and Cohen 2018). But to date, there is very little empirical evidence whether ride-sourcing increases or decreases congestion. A survey by Clewlow and Mishra (2017) found that the more frequent a user ride hailed, the more likely she would report a likelihood of reducing auto ownership and driving less in a personal vehicle. A natural experiment of user behavior after the suspension of Lyft/Uber service in Austin, Texas showed that about 40% of ride-sourcing users switched to another TNC while another 40% of users chose to drive themselves, and only 9% of users chose public transit (Hampshire et al. 2017). The study also showed that those who chose personal vehicles made 23% more trips than their counterparts who chose another TNC after the suspension, indicating that ride-sourcing may reduce total vehicle trips. We also identified one study that used a difference-in-differences approach to investigate whether the entrance of Uber in 87 US cities affected the city's overall level of congestion, measured by travel time index, total delay and delay per commuter. The study concluded that Uber entrance *reduced* the congestion level in cities according to each measure (Li et al. 2016).

All the above-mentioned studies provided valuable insights. Studies found that ride-sourcing users substituted trips mainly with autos but also with transit, and evidence suggests that increases in ride-sourcing use are associated with lower overall congestion in cities. Conflicting results from these studies on whether ride-sourcing is associated with a reduction in transit ridership, and if so by how much, indicates that more research is needed, especially taking into account the quality of transit access.

However there are significant gaps in the literature. First, travel behavior within cities is not spatially homogenous, but most previous studies investigate ride-sourcing at the level of cities or metropolitan areas. Those results do not take account of the fact that ride-sourcing may cause decreasing transit ridership or congestion in some neighborhoods but increase the level of traffic in others. It is important to understand transit access and congestion at a local level because they directly affect local accessibility and environmental impacts. Second, the relationship between ride-sourcing and other travel modes are not necessarily linear. For example, the increase of an additional bus stop would improve transit accessibility differently when there is already 5 bus stops nearby compared to 0 bus stops. But so far the most reliable studies using panel data have adopted simple econometric models, testing only for linear relationships.

Chapter 3. Descriptive analysis: what proportion of ride-sourcing trips has transit alternatives

3.1Introduction

This chapter used ride-sourcing GPS travel data in Shanghai to investigate what proportion of actual ride-sourcing trips can be replaced by public transit, namely metro system and regular buses. The data of ride-sourcing trips includes a large sample (a total of 140,854 random samples used in this study, explained in the Data section) of ride-sourcing trips for a period of 10 months in the year of 2015. It is provided by Didi Kuaidi, the largest TNC by market share in China (with 90% of market share). Google Map Direction API was used to find the most convenient transit alternative of each ride-sourcing trip. Since having a transit alternative does not mean the alternative is reasonable, I set criteria for whether a specific ride-sourcing trip can be reasonably substituted by public transit by defining access/egress time to transit stations, wait time, number of transfers, and the total transit travel time compared to ride-sourcing. I then used a descriptive framework to explore whether ride-sourcing trips can be replaced by transit, and also the circumstances under which ride-sourcing is more likely to be a competitor to public transit.

3.2 Data and methodology

The ride-sourcing GPS dataset used in this paper was provided by Didi Chuxing. It consists of 250,000 trips that were randomly sampled by Didi Chuxing from their database from the first seven days of each month from January to October 2015 from the Didi Chuxing database. The data include the date, time of day, addresses and geolocations of origins and destinations of each trip. In order to more accurately represent the changing monthly trip volume, I randomly re-sampled the data for each month based on data provided by the TNC on the total number of trips conducted in Shanghai within each month. The growth and decline of ride-sourcing trips from January to October, 2015 are shown in Figure 3.1. The resampling process resulted in a analysis subsample of 140,854 trips and did not substantially change the characteristics of the data (Table 3.1, below).

Although the dataset does not include the price paid for each trip, Didi Chuxing did provide us a detailed timeline of pricing and subsidy changes during 2015. Table 3.2 outlines the major events of ride-sourcing travel cost changes, including specific dates of programs such as driver subsidy, discount for passengers, and introduction of surge pricing. They are also noted in Figure 3.1. It should be noted that during this period there is no official regulation on ride-sourcing and TNCs.

Generally speaking, from February to June, the travel cost for ride-sourcing was decreasing, and from ate June to October, the cost was increasing. Total usage of the system in these data correspond to the fare changes, with increasing use during the fare reduction period followed by a decrease in use when fares increased and driver incentives were removed (Figure 3.1).

Table 3.1. Comparison of average travel distance and travel time before and after resampling

	N (before resample)	Mean	N (after resample)	Mean
Distance (km)	250,000	5.56	140,854	5.42
Travel time (min)	250,000	17.80	140,854	18.00

Table 3.2. Major ride-sourcing pricing events by TNC in 2015

Date (2015)	Ride-sourcing pricing events
Starting from Feb. 17th	￥10-15 (~$2.5) case back each trip for passengers, ￥50 ($8.5) awards for new drivers
Starting from Mar. 21st	The TNC allocated ￥1 billion for subsidizing passengers; ￥35 ($6) cash back each trip for passengers
Starting from May 13th	New cheap ride-sourcing services started: no initial fee for booking, and ~ ￥1 (17 cents) per km each trips
Starting from Jun. 16th	Introduction of surge pricing; end of passenger promotion; per km cost increased by 50%
Starting from Aug. 16th	End of driver subsidy and incentives
Compare with Price of Taxi, Bus and Metro in Shanghai	
Taxi price	￥14 (initial fee) + ￥2.5 per km
Bus Price	City Bus (operate within the outer rim of Shanghai): ￥2 flat rate Commute Bus (operate between outer rim of Shanghai and Suburbs): ￥2 + ￥0.2 per km, ￥8 cap
Metro Price	￥3 (the first 3 kms) + ￥1 (per additional 10 kms)

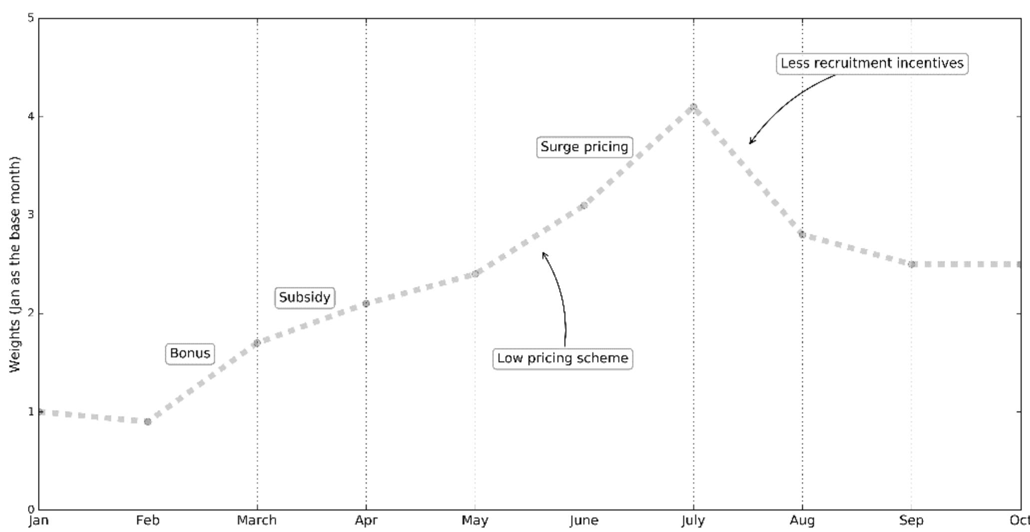

Figure 3.1. Growth and decline of the amounts of ride-sourcing trips each month (in relative number: # of trips in Jan is 1) [1]

[1] Note that October 1st to 7th are national holidays, during when trip patterns might be different.

Figure 3.2 shows the average travel time (min) by each month. As I can see the monthly average travel time varies between 14 minutes to 22 minutes.

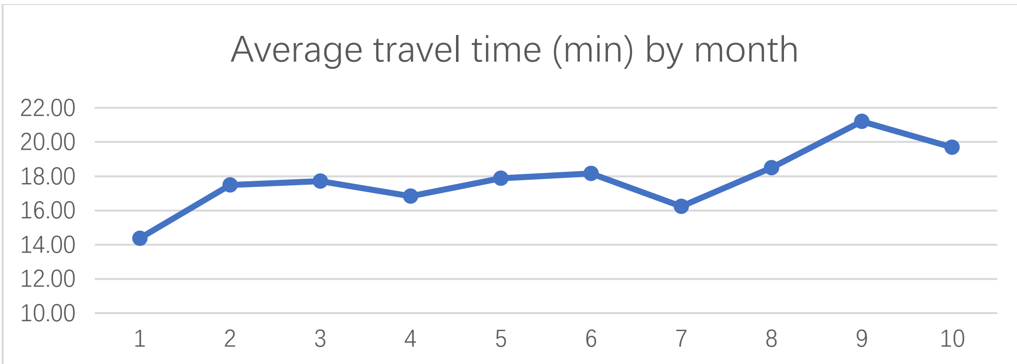

Figure 3.2. Average ride-sourcing travel time (min) each month

I used Google Map Direction API to collect data on transit alternatives for each individual ride-sourcing trip. I used the origin/destination geolocations, travel start time and date as well as travel mode (transit) as inputs and API returned the information on the fastest transit option. Results were stored in a JSON file in our local system. I extracted basic information from the JSON file, including walking time, waiting time, in vehicle time, number of transfer and all types of transit used in this travel (metro and bus). 91% of all 140,854 trips have transit alternatives and 9% of the total trips only requires walking as the travel mode. There are 188 trips that are extremely long in distance, and no transit option was found for these 188 trips.

Since Google map only contains travel information for the most recent two weeks in Shanghai, exact travel dates and time cannot be used as inputs to the API. Instead, I found the most recent day of week that matches with the day of week when the trip took place. For example, a trip was made 19:58 at January 5th, 2015 which is a Monday. Instead of inputting 19:58, January 5th, 2015 as the start time, I inputted 19:58, April 4th, 2016 which is the closest Monday to the date that I run the analysis (April 8th to 11th, 2016). I believe this compromise is acceptable since the data was collected 1 year ago and there is unlikely any significant change in transit route alternative characteristics. In addition, there might be doubts that Google Map data for Shanghai might not accurately represent the travel information there. I used a local Map provider, Gaode Map, and randomly select 20 pairs of trips and find the transit options to compare with the results from Google. I found that the outputs were generally the same across the two Map platforms.

Table 3.3 shows the summary of Google Map Direction API results. The average travel time for transit is 52 minutes, and the average walking time is 15 minutes, which is acceptable given that the walking time includes access time and egress time. The average waiting time is 13 minutes, and the average in vehicle time is 23 minutes. Also, almost 90% of the transit alternatives only require one transfer at most. In addition, the time ratio between transit alternative travel time and ride-sourcing travel time strongly indicates the advantage of ride-sourcing: if choosing transit, on average travelers would need to spend as 3.74 times long as ride-sourcing would take to arrive at their destinations. Among all trips, only 21% would take less than twice as long by transit compared with their ride-sourcing travel time. The results meet our expectation and personal experience of transit services in Shanghai, and it is also similar to studies in San Francisco (Rayle et al., 2016).

Table 3.3. Summary statistics of Google Map Direction API results

Transit alternatives	N	%
Have transit alternatives	127,636	90.6%
Only require walking	13,030	9.3%
No result	188	0.1%
Total	140,854	
Transit travel time (min)	Mean	Std. Dev.
Total transit travel time	52.03	36.32
Walking time	15.58	8.33
In vehicle time	23.31	26.84
Waiting time	13.13	11.17
Number of transfers	N	%
No transfer required	86,179	61.2%
# of transfer = 1	34,225	24.3%
# of transfer = 2	5,907	4.2%
Time ratio (transit travel time/ride-sourcing travel time)		
Average time ratio		3.74
Std. Dev. of time ratio		4.51
% of trips that are 50% longer by transit		90.8%
% of trips that are twice as long by transit		79%

Please note that having a transit alternative does not mean transit is a “viable” option for that specific trip. The google API returns total transit time, walking time, waiting time, and number of transfers for the fastest transit alternative, but that does not mean that the transit alternative is reasonable. In other words, there are transit alternatives that are unlikely to be taken by any traveler. The transit travel time might be much higher than the time spent in a car, or the walking distance to nearest transit station might be too long. Thus I need to identify which transit alternative is considered ‘viable’. The criteria for defining the “viable” competitive transit alternative to ride-sourcing are as follows.

- Walking time <= 30 min
- Waiting time <= 20 min
- Number of transfer <= 1
- Transit travel time / Ride-sourcing travel time <= 2

It is important that I set such criteria for “viable” transit alternative, because the transit network in Shanghai is so dense that the majority of trips can be matched with a transit alternative otherwise (as shown in Table 3.3). In 2015, There are a total 289 metro stations and 14,575 bus stops in Shanghai. The reason why walking time threshold is 30 minutes is that this allows a 15-minute access time at both end of a trip, that is approximately 800-1,000 meters of walking distance at both ends. Setting wait time at 20-minute makes sure that transit alternatives with routes that have large headways are included. I set that transit travel time / ride-sourcing travel time should be less than 2 because the utility of time on board transit is lower than that on a car, but the differences should not be too large so that transit cannot compete with ride-sourcing.

Figure 3.3 shows the percentage of ride-sourcing trips that have reasonable transit alternative each month. As I can see from the figure, the percentage ranges from 15% to about 35%. This percentage is quite comparable with findings from the US (Rayle et al. 2016).

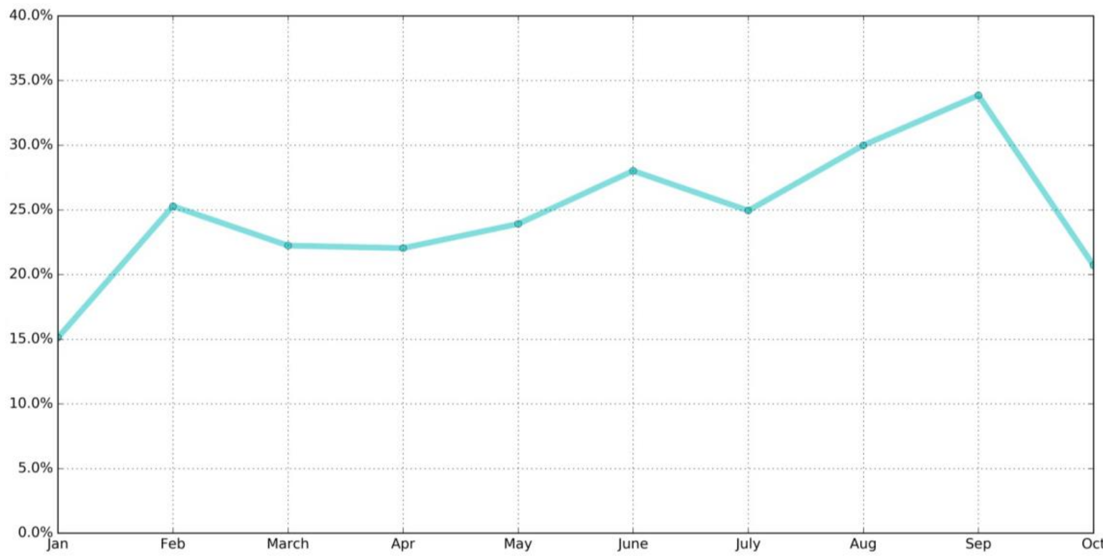

Figure 3.3. Percentage of ride-sourcing trips that have reasonable competing transit alternatives each month

3.3 Descriptive results

The essential question I want to address is *under what conditions is ride-sourcing service more likely to be competing with transit*. To answer this question, I made several assumptions. First, I assume that, when ride-sourcing price decreases, more people take ride-sourcing. This is evidently shown previously (Figure 3.1). Secondly, I assume that these increasing amounts of ride-sourcing trips were either trips that were previously made via other modes, including transit, private cars, taxi, etc., or were trips that would not have been made without access to ride-sourcing service. Transit took up a significant percentage in motorized travel in Shanghai. In 2014, 48% of the trips in Shanghai urban area are taken by transit, and the household level auto-ownership are 28% due to strict restriction on vehicle license (Shanghai Municipal Transportation Commission, 2015). Based on these statistics, I assume that if a ride-sourcing trip has a reasonable transit alternative, it is very likely that the traveler switches from transit to ride-sourcing.

There are many ways to define the kind of condition where I examine the key question. In this paper, I specifically focus on three aspects: the length of travel, time of travel, and the mode of transit. Here this chapter specifically conduct descriptive analysis from the following perspectives: distance of travel, pear hour vs. non-peak hour, and type of public transit.

Short trip vs. Long trip

Our first question is whether transit options are more competitive with ride-sourcing in shorter trips or longer trips. The key parameter explored in this study is the percentage of trips that can be reasonably replaced by transit and the fluctuation of it may its competitiveness with transit. However, there are many possibilities that explain the fluctuation. For example, it can be caused by fluctuation in trips that cannot be replaced by transit, or trips that can be, or both. So I need to look at the total amount of the trips (total number of ride-sourcing trips in January is 1) and the amount of the trips that can be reasonably replaced by transit. Figure 3.4 shows the comparable number of short ride-sourcing trips (ride-sourcing travel time <=20 mins, shown in red), and long ride-sourcing trips (travel time >20 mins, shown in orange). Shorter trips account for the majority of trips, and the total number of shorter trips fluctuates over time as changes in prices. On the contrary, the changes in longer ride-sourcing trips are much less dramatic. I can conclude that short ride-sourcing trips accounted for the majority of growth and decline in the changes of total number of ride-sourcing trips.

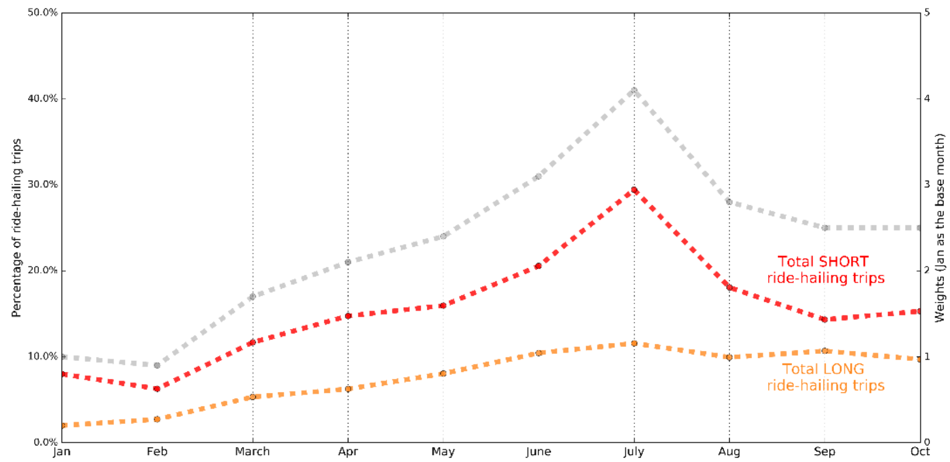

Figure 3.4. The relative amount of short ride-sourcing trips (ride-sourcing travel time <=20 mins), and long ride-sourcing trips (travel time >20 mins) each month

Figure 3.5 shows the percentage of short ride-sourcing trips with viable transit alternatives over the 10-month period. The percentage of short ride-sourcing trips that can be reasonably replaced by public transit are fairly low and stable, ranging from 5% to slightly less than 10%, shown by green line. Quite contrarily, from February to July, I see a sharp increase of short ride-sourcing trips that are without viable transit alternatives, and a dramatic decrease of it after July (Figure 3.5 red). Such dramatic fluctuation indicates that people who made short ride-sourcing trips that had no reasonable transit alternative are very sensitive to price changes. Compared to the large variation of the number of short ride-sourcing trips without viable transit alternatives, the variation of the number (Figure 3.5 orange) of short ride-sourcing trips with viable transit alternatives is fairly small. So does the percentage (Figure 3.5 green) of this type of trip. That is to say, people do not necessarily make more short-distance ride-sourcing trips that can be made via transit, when ride-sourcing is cheap. Neither do people make less of such trips when the cost of doing so is high. This might be indicating that ride-sourcing is not a competitor to public transit when the trip is short in terms of distance. The increase of the number of short ride-sourcing trips without transit alternative in periods when prices were low are either switching from travel modes other than transit (private cars, taxis, motorcycles, etc.), or induced travel demand by the introduction of affordable ride-sourcing.

The pattern in the percentage of long ride-sourcing trips with viable transit alternatives (Figure 3.6) is very different from that of short ride-sourcing trips. The percentage of long ride-sourcing trips (Figure 3.6 blue) that have viable transit alternatives are fairly high compared to short ride-sourcing trips, ranging between 30% to 40% (20% in October, but as I discussed before October 1st to 7th is a national holiday). In other words, ride-sourcing is more likely to be a competitor to transit at longer distance. Growths in both long ride-sourcing trips with and without viable transit alternatives contribute to the increases of the total long ride-sourcing trips from January to July (during which period the prices of taking ride-sourcing is low), and only after July when prices of ride-sourcing were becoming increasingly high, the number of long ride-sourcing trips with transit alternatives started to decrease. Noticeably, long ride-sourcing trips without reasonable transit alternative increased even after travel cost increased. This may suggest that ride-sourcing is complementary to public transit for providing services to smaller and scattered activity centers.

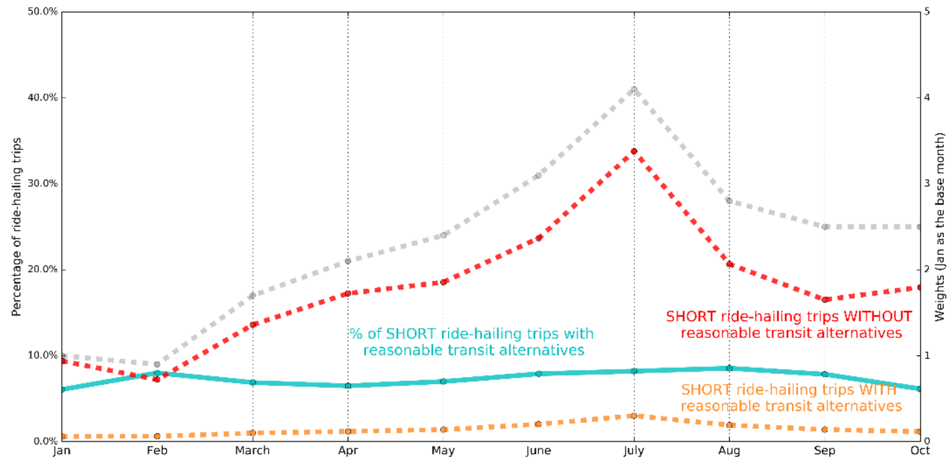

Figure 3.5. The percentage and relative amount of short ride-sourcing trips with viable transit alternatives

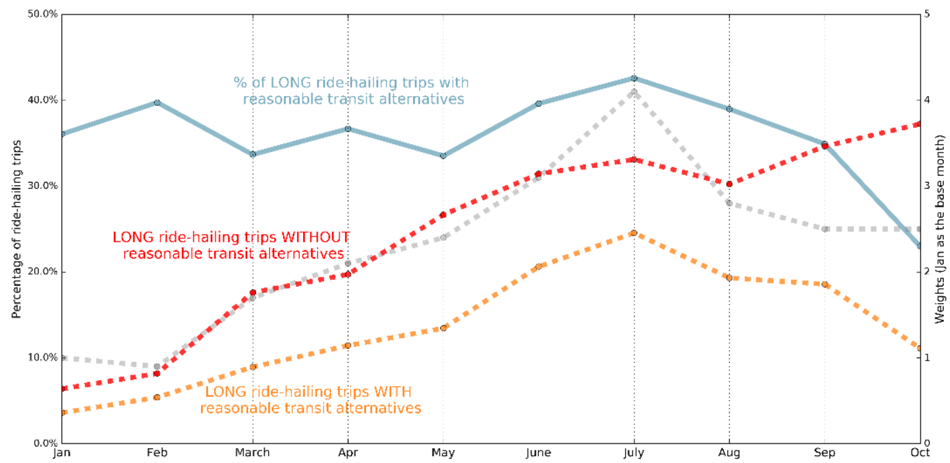

Figure 3.6. The percentage and relative amount of long ride-sourcing trips with viable transit alternatives

To summarize, ride-sourcing seems to be a competitor to public transit over longer distance of travel. This may be due to the fact that a longer ride-sourcing trip would almost always means even longer travel time by transit. But another important implication is that the rise of cheap ride-sourcing option (due to subsidies by TNC, Didi Kuaidi) might induce travel demand for shorter yet transit-inaccessible trips. The mismatch between the increase in total short ride-sourcing trips and the stagnant short ride-sourcing trips without viable transit alternatives suggest that these short trips must have come from other sources. Given the fact that Shanghai has low auto-ownership and expensive, market failed taxi services, induced travel might be a significant portion of these increased short ride-sourcing trips.

Peak vs. Off-peak

Our second question is that whether there is a distinct pattern between ride-sourcing trips that were made in peak hours and off-peak hours. Peak hours are between 6am to 9am, or 4pm to 7pm Monday through Friday. Off-peak hours are other time period Monday through Friday. I chose to focus on weekday trips to simplify the complexity of trip patterns given that it is very likely people behave differently over weekends. Figure 3.7 shows that peak hour trips are roughly half of the volume of off-peak trips, and they have similar growth patterns over time. As what has been mentioned before off-peak trips only include off-peak trips during the weekdays.

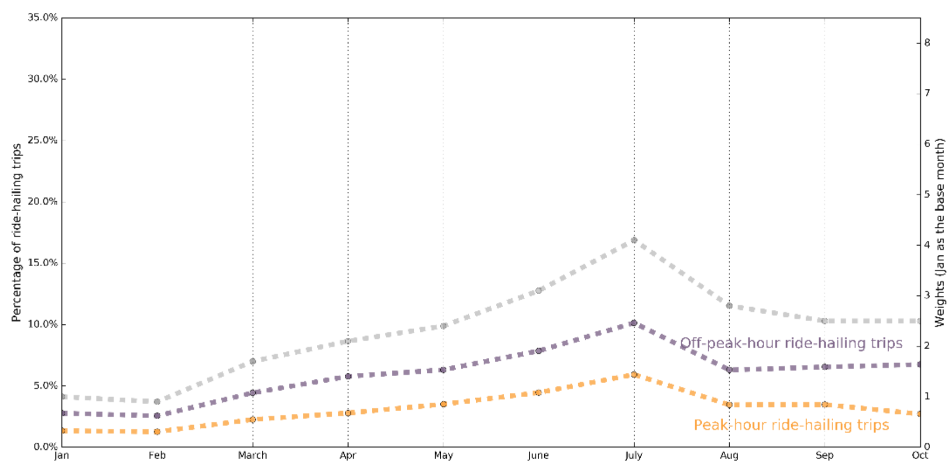

Figure 3.7. The relative amount of peak-hour ride-sourcing trips and non-peak hour ride-sourcing trips each month

During off-peak hour, the percentage of off-peak trips with viable transit alternatives (Figure 3.8) ranges from slightly below 10% to around 20%, and is generally increasing until August. The increase in the number of off-peak trips without viable transit options at low-pricing period (Feb. to July) is clearly shown in Figure 3.8. The number of off-peak trips with viable transit alternatives also increases during this period. However, the percentage of off-peak ride-sourcing trips with viable transit alternatives continue to increase from July to August, due to the rapid decrease of the number of trips without transit alternatives. This suggests that without viable transit alternatives, off-peak ride-sourcing demand is more elastic given price changes. This might suggest that after the increase of rail-hailing price, many of these trips were switched to travel modes other than transit, or the travelers simply stop making the trips. The competition between ride-sourcing and transit is more apparent during peak hour (Figure 3.9), although the differences is small compared to pattern during off-peak hour. The percentage of peak hour trips that have viable transit alternatives are around 20% to 30% (set aside October which is holidays), and the variations of the percentage are more dramatic compared to those of off-peak trips. Some might argue that it is possible that during peak hour, transit services are at a higher frequency, therefore a higher percentage of ride-sourcing trips can find a “viable” transit alternative by the same criteria. But this doesn’t change the other factors in our criteria, such as access time and number of transfer. Also, our viable transit criteria sets wait time at 20-minute, which includes transit alternatives with less frequent routes.

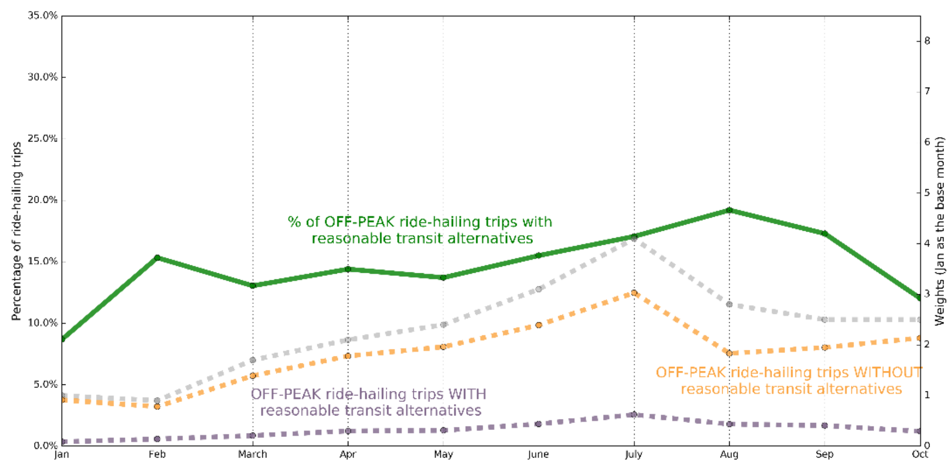

Figure 3.8. The percentage and relative amount of off-peak ride-sourcing trips with viable transit alternatives

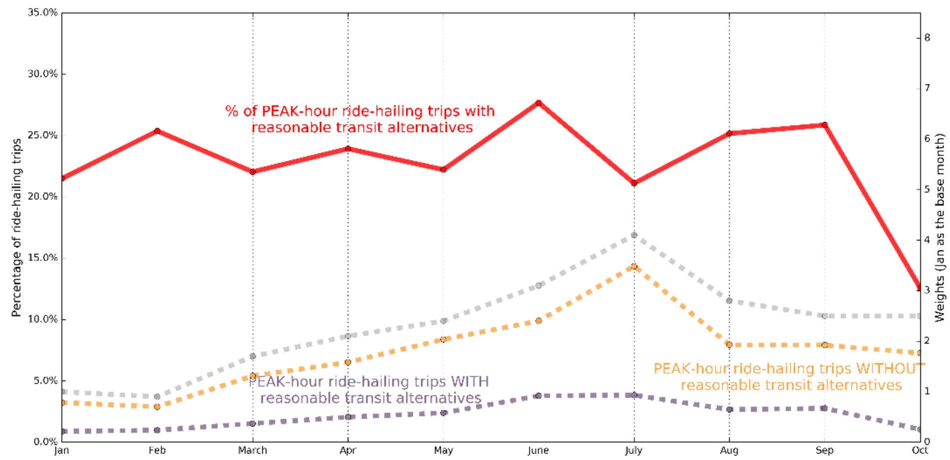

Figure 3.9. The percentage and relative amount of peak ride-sourcing trips with viable transit alternatives

Bus alternative vs. Metro alternative

The next question I explored is whether ride-sourcing is a competitor to bus or metro. I select those trips that have viable transit alternatives that are composed of only bus transit, and compare them with trips that have viable transit alternatives that are composed of only metro. Figure 3.10 shows the difference between the total numbers of the two types of trips. There are generally more ride-sourcing trips that have bus-only alternatives than those with metro only alternative, and the difference is generally larger when ride-sourcing price is low. Also, ride-sourcing trips with bus-only alternative increased faster from February to July, indicating that bus riders are more likely to switch to ride-sourcing than metro riders.

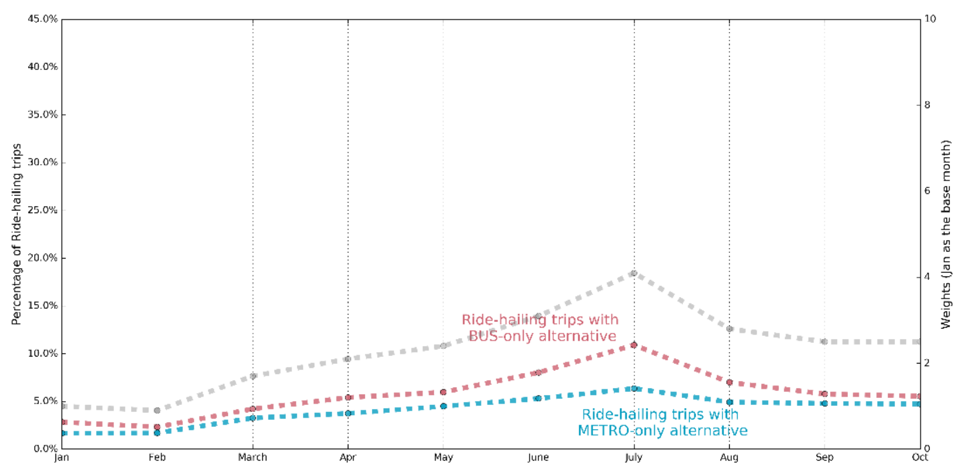

Figure 3.10. The relative amounts of ride-sourcing trips with bus-only alternative and with metro-only alternative

If I look at this question from a percentage perspective, I found that only 5% to 10% of all ride-sourcing trips have a reasonable bus-only alternative (Figure 3.11), indicating that ride-sourcing might not be a direct competitor to bus. However, trips that have metro only alternative are much higher in percentage, ranging from 25% to 40%, shown in Figure 3.12. The variation of the percentage of ride-sourcing trips that have metro-only alternative fluctuates over months, giving a different look compared with the percentage that have bus-only alternative. This indicates that ride-sourcing is more likely to be a competitor to metro. This conclusion seems to be totally different from what I found from Figure 3.10. One possible reason why I observed more bus riders switch to ride-sourcing while at the same time ride-sourcing has a much higher percentage of substitution by metro-only alternative, is that the service and reliability of buses are worse than those of metro, therefore our analysis eliminated most of the bus-only alternatives by setting up a high "viable transit alternative" criteria. In this case bus transit is an inferior good and stands minimum chances when competing with affordable ride-sourcing.

However, the high percentage of substitution between ride-sourcing and its viable metro alternative did reveal that there were competitions between them. One possible reason why ride-sourcing seems to compete with metro is that metro is less accessible compared to buses. I compared the average walking distance for two types of alternatives, and found that the walking time for accessing metro is significantly higher than that for bus (13 minutes compared to 8 minutes). Additionally, Metro stations in Shanghai have security checks at most metro station entries, which may drive people away from taking metro.

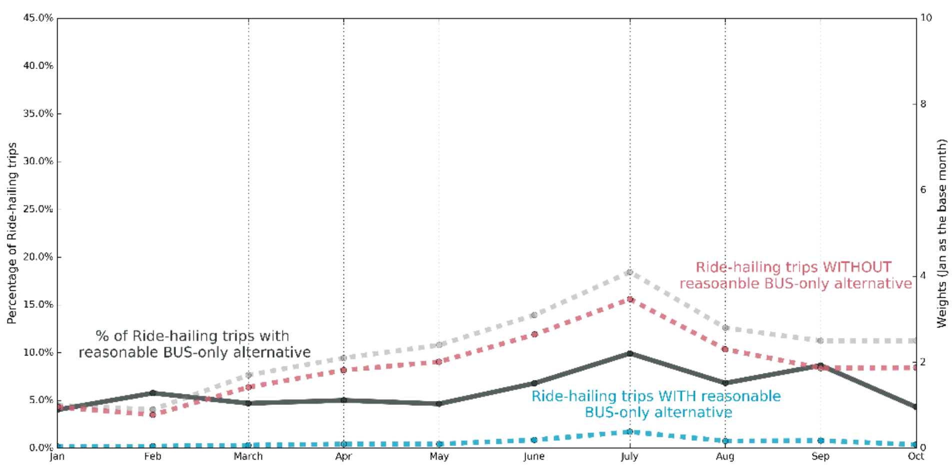

Figure 3.11. The percentage and relative amounts of ride-sourcing trips that have a viable bus-only alternative

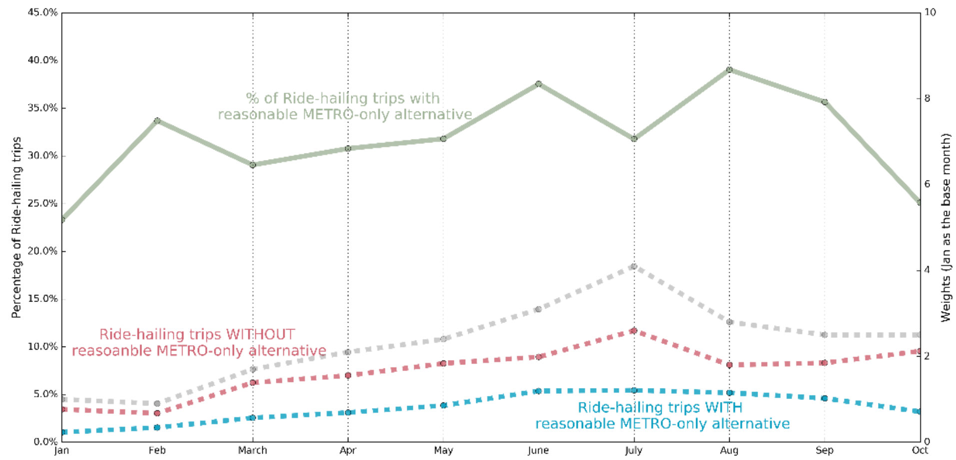

Figure 3.12. The percentage and relative amount of ride-sourcing trips that have a viable metro-only alternative

To summarize the findings above, ride-sourcing is more likely to compete with public transit for longer trips because time saving benefit are more significant in long trips compared to short trips. It seems to be a competitor to public transit, though by a small margin, during peak hour where ride-sourcing could allow travelers to avoid the crowd. Ride-sourcing is more likely to attract passengers from buses, and possibly due to longer walking distance and security checks, it also competes with metro.

3.4 Summary of descriptive results

Although the descriptive analysis did not track the individual behavior of switching between modes after the emergence of ride-sourcing, it did reveal the multi-facet of ride-sourcing travel behaviors. By looking at what proportion of the ride-sourcing trips can be replaced by transit, I discovered that ride-sourcing is more likely to compete with transit in longer travel and peak hour travel. It is likely that ride-sourcing is competing with both bus and metro, but bus riders are more likely to switch to ride-sourcing than metro riders due to low time performance of buses in Shanghai. Metro, on the other hand, stands a higher chance against ride-sourcing, but its longer access time may still push some travelers to use ride-sourcing. However, it is important to note that based on our criteria, the majority of the ride-sourcing trips are complementing the current public transit system. Based on these findings, I understand better why ride-sourcing is so popular: it’s affordable, time-saving, and accessible.

We must keep in mind that these ride-sourcing trips are actual trips that took place via Didi app in 2015. Thus, finding transit alternatives of these trips imply that travelers chose ride-sourcing over its public transit competitor. Despite the seemingly subjectivity of the “viable” transit alternative criteria in this chapter, setting such criteria is important because the transit network in Shanghai is dense that the vast majority of trips, as I shown in the methodology section, can be matched with a transit alternative. Some of the results would certainly change if I alter the criteria, but it is always questionable whether transit that takes more than twice as long as ride-sourcing can still be competitive against ride-sourcing. It is likely that ride-sourcing travelers have different characteristics from transit travelers in Shanghai, or ride-sourcing might be used more often on time-sensitive purposes. Previous research did suggest that the travel purpose of ride-sourcing is different from that of public transit in the US (Shared-use Mobility Center, 2016), but it is used mostly on discretionary purposes. However, this is beyond the capability of this dataset except that I can use some machine learning techniques to presumptuously deduce the trip purpose.

This analysis also suggests that during the “low cost” period of 2015, ride-sourcing enables travelers for longer trips and perhaps “induced” short trips. It would be very interesting to look at the spatial pattern of ride-sourcing trips in the future, perhaps comparing to the spatial distribution of other type of travel. If low-cost, large-supply ride-sourcing did induce significant demand during the first half of 2015 in Shanghai, on-demand ride services might change the landscape of our city.

Chapter 4. Equity implications of on-demand ride-sourcing fare variation: A case study of Didi Chuxing, Shanghai, 2015

4.1 Introduction

The rapid emergence and growth of transportation network companies (TNCs) such as Uber, Lyft, and Didi Chuxing, operating app-based on-demand ride-sourcing services, has led to a debate over the role of TNCs in the urban transport system. Ride-sourcing supporters argue that it is a convenient and efficient mode of travel, particularly compared to the traditional taxi industry which is typically protected via regulatory limitations on market entry and is thus monopolistic, inefficient and not sufficiently customer-focused (Rayle et al. 2016; Shaheen et al. 2016). Other scholars and government officials are less sanguine, concerned about disruptions to the livelihoods of taxi drivers, insufficient insurance protection for drivers and passengers, increased traffic, and competition with public transportation. City governments and regulatory agencies worldwide have begun to regulate ride-sourcing in different ways, including setting licensing criteria for drivers and vehicles, requiring minimum insurance for drivers and passengers, and using part of the revenue from ride-sourcing to compensate taxi drivers. By either increasing the cost of operations or decreasing the supply of vehicles and drivers, some of these regulatory interventions can be expected to lead to fare increases which would undoubtedly have unequal effects on those with different socio-economic backgrounds. This study focuses on these potential equity impacts, asking how the travel demand of users from different socio-economic groups changes when ride-sourcing fares change.

What makes Shanghai a particularly interesting place to study ride-sourcing use is the fact that due to its license auction policy the metropolitan area has relatively low auto ownership given its size and relative affluence (Chen and Zhao 2013). In Shanghai, licenses typically cost about ￥80,000 per vehicle in 2015 (about $12,300 in US dollars). As a result of the policy, in Shanghai automobiles are ordinarily owned and used only by residents with medium to high income (Chen and Zhao 2013). In 2014, there were less than 15 private cars for every 100 residents in Shanghai, and private cars accounted for about 17% of trips (Shanghai Urban-Rural Construction and Traffic Development Academy 2015). Compared to auto ownership, ride-sourcing is a comparatively accessible travel option, as 94.4% of the adult population in Shanghai owned at least one smart phone in 2013 (Smart Device Business Center 2013) and at least during some periods of operation, the average ride-sourcing fare was quite low at ￥1 per kilometer. For comparison, the average hourly wage in Shanghai was around ￥35 in 2015 (Shanghai Bureau of Statistics 2015).

Which segments of the Shanghai population are most likely to benefit from ride-sourcing availability, and particularly lower-priced ride-sourcing services? Wealthy car-owning residents may use ride-sourcing to avoid driving in congested conditions or finding a place to park. Lower-income travelers may use it to avoid crowded public transport vehicles or to access locations that are difficult to reach without a private vehicle. Fare changes thus have the potential to have differential impacts on different socioeconomic classes that depend on how they use and value the service. Using trip data from 2015 provided by Didi Chuxing, the largest TNC by market share in China, this paper explores how ride-sourcing demand is related temporally and spatially to socioeconomic status as proxied by property value.

4.2 Data and methodology

The ride-sourcing GPS dataset used in this paper was provided by Didi Chuxing, the same data used in Chapter 3. Besides the travel data from Didi, I use the average residential property value as a proxy measure for the average household income level of the grid cell. Property value has been used as a socio-economic indicator for place-based analysis, as it is highly correlated with socio-economic status (Coffee et al. 2013). I obtained all the listed residential properties for sales in 2015 and their offer prices (in ¥ per square meter, which is the standard practice in Chinese real estate market) through Metrodata Tech.[2], which includes a total number of 97,811 properties. Table 4.1 shows the descriptive statistics for offer prices of these properties.

Table 4.1. Descriptive statistics for offer prices of listed residential properties for sale in 2015

Offer price per m^2	Mean	Min.	Max.	25 PCTL	Median	75 PCTL	SD	N
RMB ¥ (2015)	35,098	95	991,277	23,331	30,288	40,427	22,101	97,811
USD $ (2015)	5,571	15	157,346	3,703	4,808	6,417	3,508	

In order to conduct the analysis, I aggregated these data into 1-km grid cells as our basic unit of analysis. I deleted grid cells with no records of listed properties, leaving 1,292 grid cells, with 128,037 ride-sourcing pick-ups and 125,970 ride-sourcing drop-offs. Figure 4.1 shows the number of ride-sourcing trips in each grid cell. Figure 4.2 shows the spatial pattern of average property value per grid cell. I can see that both pick-ups and drop-offs have similar spatial pattern: most of the trips were concentrated in the city center.

[2] More information about Metrodata Tech. can be found here: www.metrodata.cn

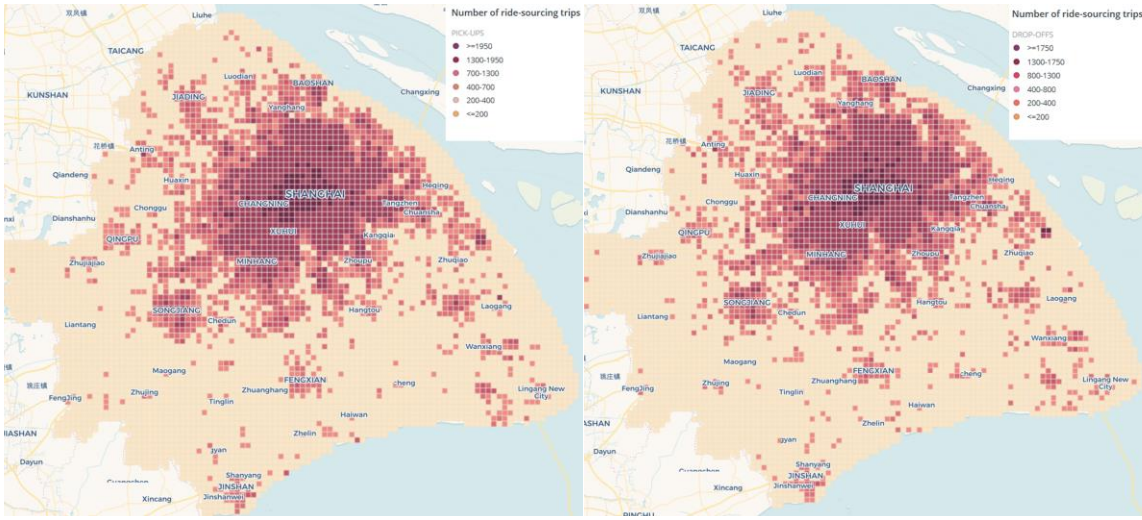

Figure 4.1. Map of ride-sourcing pick-ups (left) and drop-offs (right) in each grid cell

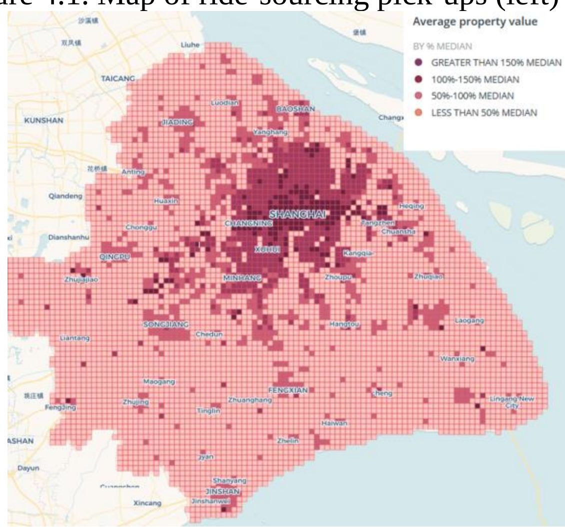

Figure 4.2. Average property value in each grid cell

In a more complete effort to control for other factors I turned to regression analysis. I modeled the ride-sourcing demand (for pick-ups and drop-offs separately) as a product of several spatial and temporal characteristics, using a negative binomial regression whose function form is appropriate for dispersed count data.

$$demand_{it} = exp(\beta_0 + \beta_1 Property\_Value_i + \mathbf{X}'_{\mathbf{it}} \cdot \boldsymbol{\beta} + \boldsymbol{\mu}_{\mathbf{t}} + \beta_2 Property\_Value_i \cdot \boldsymbol{\mu}_{\mathbf{t}} + \varepsilon_{it}). \quad (1)$$

In this formula, $demand_{it}$ is the dependent variable representing ride-sourcing demand in grid cell *i* in month *t*, measured as the number of pick-ups or drop-offs. $\beta_1 Property\_Value_i$ represents the independent correlation of neighborhood property value with that dependent variable. Vector $\mathbf{X}_{\mathbf{it}}$ is a matrix of independent control variables measured at the grid cell (neighborhood) level, including population and employment density as well as activity locations in the neighborhood (see Table 4.2). The model treats $\mathbf{X}_{\mathbf{it}}$ as time-invariant given that these built environment variables are unlikely to have changed very much over the course of the nine-month observation period, and if so, not systematically in such a way as to bias results. Vector $\boldsymbol{\mu}_{\mathbf{t}}$ consists of time-period (monthly) fixed effects. Note that the dataset does not contain a price variable; the monthly fixed effect $\boldsymbol{\mu}_{\mathbf{t}}$ is a proxy for price change, as the ride-sourcing fares changed dramatically on a monthly basis due to passenger promotions (as shown previously in Table 3.2). Finally, I included a set of variables interacting the average property value with each of the fixed effect month dummies ($\beta_2 Property\_Value_i \cdot \boldsymbol{\mu}_{\mathbf{t}}$). As constructed, the model allows a test of whether and how neighborhood property value influences how much variation in demand is associated with variation in the ride-sourcing fare, when controlling for other factors. Understanding this association helps understand if ride-sourcing demand is more correlated with property value when fare of ride-sourcing is changing.

Note that for most fixed effect panel models, an individual fixed effect $\boldsymbol{\upsilon}_{\mathbf{i}}$ is added for each observation to represent all time-invariant characteristics. In this case, however, the independent variables are time constant, and therefore collinear with individual fixed effects. Removing the individual fixed effect $\boldsymbol{\upsilon}_{\mathbf{i}}$ could result in omitted variable bias in the model, but fortunately I was able to include a large set of land use variables that reduce this possibility (data explained later). I clustered standard errors for each month to correct standard errors to reflect the hierarchical structure of the data. Finally, I also calculated Moran's I for both pick-ups and drop-offs to determine whether spatial auto-correlation could bias analysis results. I found little evidence of spatial auto-correlation. The Moran's I result for the number of pick-ups in each grid cell was 0.0915 (with a p-value of 0), while the result for the number of drop-offs in each grid cell was 0.0869 (with a p-value of 0).

I estimated two sets of negative binomial models. The first set used continuous property value, and the second set treated property value as categorical dummy variables: below 50% median, 50%-100% median, 100%-150% median, and above 150% median, while property value below 50% median was used as base category. Both sets of models interacted property value with the monthly fixed effects. Within each set of models I tested two dependent variables: the number of pick-ups and the number of drop-offs.

In additions to the fixed effects which rule out the unobserved spatial and temporal variations, other control variables are included in this study to represent other possible factors that influence the demand for ride-sourcing. Theory of travel demand and empirical evidence suggested that activity centers are likely to generate trips (Hanson & Giuliano 2004). Thus I uses all points of interest (POI) in Shanghai to represent all the activity centers. This POI dataset has over 200,000 activity locations, including all offices, shops, restaurants, convenience stores, recreational facilities (including theatres and cinemas), schools, hospitals, and parks. I used the number of each of these activity location types as control variables in the model. I also used the average residential property value as a proxy measure for the value of land in the grid cell, because previous literature showed that higher land value is associated with more travel demand (Hanson & Giuliano 2004). Population density and employment density at level similar to census tract level the US is included in the study as well since research showed that density is related to the demand of activity (Chatman 2008; Ewing & Cervero 2010). Average road length within each grid cell is also added as a control since previous research showed that road supply is directly associated with motorized travel (Chatman 2008). The transit dataset, property dataset and the points of interest dataset are all public available data in Shanghai, and the authors obtained these datasets through Metrodata Tech. I aggregated the data into 1-km grid cells as well for our analysis. Table 4.2 shows the descriptive statistics of the variables involved in the analysis. Table 4.2 shows the descriptive statistics of the variables involved in the analysis.

I first compared the average number of ride-sourcing trips per grid cell when segmented by the average property value in the grid cell or "neighborhood." Ride-sourcing trips were much more likely to take place in neighborhoods with higher property values (Figure 4.3). Given changes in the average fare during the nine-month observation window, I are also able to observe how trips per neighborhood varied with ride-sourcing fares. Neighborhoods with higher average property values had greater variability with fare changes. Among neighborhoods with property values exceeding 150 percent of median, as fares fell, usage rose more dramatically, and as fares rose, usage dropped off more, when compared to neighborhoods with lower property values (Figure 4.3).

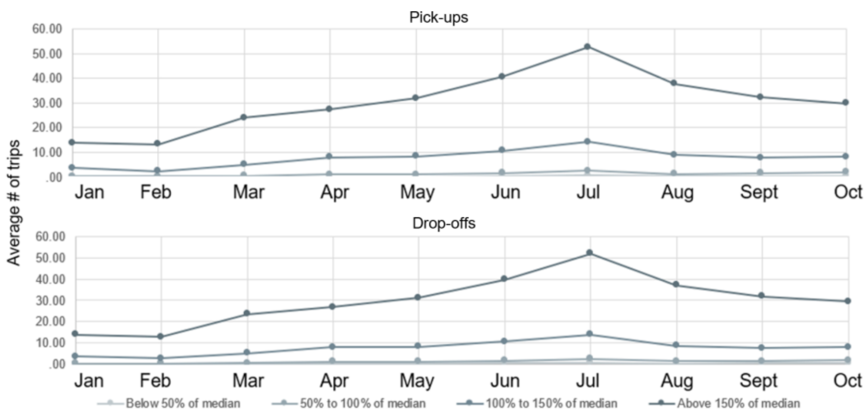

Figure 3.3. Average number of ride-sourcing trips in each grid cell segmented by average residential property value

Higher property value is also strongly associated with higher population density, higher density commercial development, more amenities, and better accessibility. Thus it is helpful to isolate the association between property value and number of trips over time controlling for other relevant factors. I begin to do this by controlling for population density within grid cells. On a per capita basis, ride-sourcing is still predominantly used by higher income travelers; however, in areas with lower property value, per capita demand *increased* after July despite the increase in fare (Figure 4.4).

Table 4.2. Summary statistics

Variable	Mean	Min	Max	SD
# of pick-ups (per month)	9.91	0	457	26.4
# of drop-offs (per month)	9.75	0	463	26.3
# of listed properties	75.63	1	580	82.7
Average property value (￥1000)	29.38	.10	216.75	17.6
# of offices	25.14	0	477	42.5
# of shops	57.75	0	1185	91.9
# of restaurants	27.02	0	427	40.2
# of recreation facilities	7.34	0	88	9.8
# of convenience stores	3.63	0	50	6.0
# of hotels	4.06	0	47	5.6
# of hospitals	.21	0	5	0.6
# of schools	.54	0	6	0.9
# of colleges	.25	0	14	0.9
# of parks	.11	0	2	0.3
# of bus stops	6.55	0	35	4.9
# of metro stations	.19	0	3	0.4
Population density (1000 per km^2)	4.91	0.01	47.84	7.7
Employment density (1000 per km^2)	1.68	.003	32.59	3.5

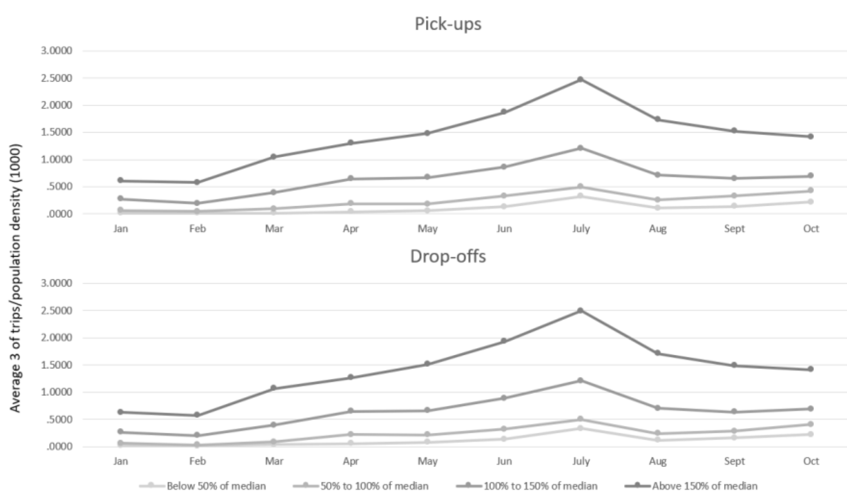

Figure 4.4. Average number of ride-sourcing trips per 1000 population in each grid cell segmented by average residential property value

4.3 Results

The estimation results show that controlling for other factors, wealthier ride-sourcing travelers were always more sensitive to ride-sourcing fare changes. Table 4.3 and Table 4.4 show the outputs of models with property value, its interaction with time fixed effects, and other land use characteristics. Since the coefficients in the negative binomial models with fixed effects are hard to interpret, I also calculated and plotted incident rate ratio (IRR) of property value on ride-sourcing demand which is the change in the value of dependent variable (demand) after the change of an independent variable (property value) by 1 unit. The results of IRR are shown in Figure 4.5 and Figure 4.6. I conducted variance inflation test (VIF) and found that all except three control variables have VIF smaller than 5. I decided that leave all of the control variables in the model because none of them are collinear with our variable of interest: average property value.

If I take a look at the estimated parameters of all interaction terms of property value and time effects for Model 1 in Table 4.3, I can find that the estimated parameters are positive from February up to May, ranging from 0.001 to 0.009, while they are negative from June to October except for August. The results suggest that the association between property value and ride-sourcing pick-ups is higher from January to May, and such association becomes smaller after July. This time frame coincides with the changes of travel subsidies and promotions by Didi as a result of the price battle, described above. I can conclude that ride-sourcing trips were more likely to originate from places with premium residential areas when fare is lower. The same pattern can also be found between ride-sourcing drop-offs and property value, but to a somewhat lesser extent (Model 2 in Table 4.3). Number of drop-offs still has stronger correlation with property value before June, but the magnitude of these associations seems to be weaker compared to those of pick-ups. This indicates that the destinations of ride-sourcing trips are also more likely to occur at premium residential areas when fare is lower.

Figure 4.5 plots the marginal effect of property value on pick-ups/drop-offs each month using the estimated values shown in Table 4.3. The marginal effect here is defined as the change of ride-sourcing demand when property value changed by ￥1000 per square meter (about $13-14 per square foot). As I can see from the figure, the IRR of property value and trip demand (both pick-ups and drop-offs) were generally higher before June, ranging from 1.055 to 1.065, meaning that a neighborhood with a property value greater by the amount of ￥1,000 would be expected to have between 5.5 and 6.5 percent more trips (for the average neighborhood with mean values on the other control variables). After July the IRR became smaller, ranging from 1.038 to 1.053, except in August. At first glance the trend of decreasing effect of property value on ride-sourcing trips with increasing fare is less clear in Figure 4.5, but I must keep in mind that our ride-sourcing dataset is a random sample of the total ride-sourcing trips, thus an increase of 0.02 trips in the sample could represent an increase of thousands of trips. For a more intuitive comparison, I can see from Table 4.1 and Table 4.2 that the average trip demand for each grid cell is around 9 trips and the standard deviation of property value is ￥22,101. Thus if property value increased by one standard deviation, trip demand would rise by one third of the average demand with a marginal effect of 1.05.

The results from the first set of models concludes that the association between property value and ride-sourcing demand (both pick-ups and drop-offs) is higher when ride-sourcing fare was low, while the association is lower after June when ride-sourcing fare increased due to the removal of promotions. This indicates that residents who lived in wealthier communities were more responsive to lower fares and sales promotions. I estimated another set of negative binomial models, replacing continuous property value with dummy variables to capture the non-linearity in the estimated coefficients. The results are shown in Table 4.4.

Table 4.3. Estimation results for negative binomial models with fixed effect

	Model 1: pick-ups		Model 2: drop-offs	
	Coef.	Std. Err.	Coef.	Std. Err.
Intercept	-2.479***	0.034	-2.494***	0.015
# of listed properties	0.004***	0.00	0.004***	0.00
# of office	0.012***	0.001	0.011***	0.001
# of recreation facilities	0.015***	0.001	0.014***	0.001
# of restaurants	-0.004***	0.00	-0.003***	0.000
# of schools	0.125***	0.01	0.126***	0.009
# of colleges	0.036***	0.005	0.039***	0.005
# of shops	-0.002***	0.00	-0.002***	0.002
# of parks	0.161***	0.014	0.142***	0.039
# of convenience stores	0.061***	0.004	0.063***	0.005
# of hospitals	0.026**	0.013	0.023**	0.01
# of hotels	0.018***	0.002	0.019***	0.002
# of metro stations	0.138***	0.015	0.125***	0.011
# of bus stops	0.022***	0.003	0.023***	0.002
Population density (1000 per km^2)	0.010***	0.001	0.009***	0.003
Employment density (1000 per km^2)	0.020***	0.002	0.018***	0.005
Average property value (￥1000)	0.054***	0.001	0.054***	0.004
Time fixed-effect				
Feb	-0.293***	0.003	-0.317***	0.002
Mar	0.045***	0.004	0.160***	0.003
Apr	0.592***	0.008	0.746***	0.005
May	0.599***	0.008	0.735***	0.005
Jun	1.044***	0.013	1.064***	0.008
Jul	1.552***	0.018	1.582***	0.011
Aug	0.888***	0.014	0.855***	0.008
Sept	1.176***	0.017	1.156***	0.011
Oct	1.501***	0.018	1.567***	0.012
Interact property value with fixed-effect				
Average property value (￥1000): Feb	0.000***	0.00	0.002***	0.00
Average property value (￥1000): Mar	0.009***	0.00	0.007***	0.00
Average property value (￥1000): Apr	0.006***	0.00	0.002***	0.00
Average property value (￥1000): May	0.008***	0.00	0.005***	0.00
Average property value (￥1000): Jun	0.003***	0.00	0.004***	0.00
Average property value (￥1000): Jul	-0.002***	0.00	-0.003***	0.00
Average property value (￥1000): Aug	0.003***	0.00	0.004***	0.00
Average property value (￥1000): Sept	-0.007***	0.00	-0.007***	0.00
Average property value (￥1000): Oct	-0.013***	0.00	-0.016***	0.00
Pseudo R^2	0.152		0.154	
# of grid cells	1291		1291	

Note: ***: 99% significant; **: 95% significant

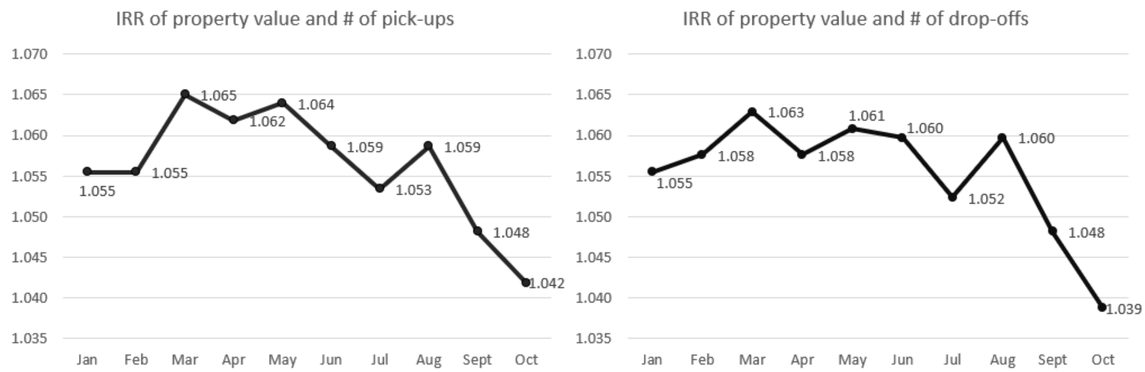

Figure 4.5. IRR of property value and number of ride-sourcing trips per grid cell (pick-ups and drop-offs)

Table 4.4. Estimation results for negative binomial models with fixed effect (property value dummy variables)

	Model 3: pick-ups		Model 4: drop-offs	
	Coef.	Std. Err.	Coef.	Std. Err.
Intercept	-3.557***	0.023	-3.245***	0.016
# of listed properties	0.003***	0.00	0.002***	0.00
# of office	0.015***	0.00	0.014***	0.001
# of recreation facilities	0.010***	0.001	0.008***	0.001
# of restaurants	-0.003***	0.00	-0.002***	0.00
# of schools	0.063***	0.006	0.074***	0.006
# of colleges	0.027***	0.005	0.036***	0.004
# of shops	-0.002***	0.00	-0.002***	0.002
# of parks	0.244***	0.014	0.214***	0.014
# of convenience stores	0.034***	0.002	0.035***	0.003
# of hospitals	0.008	0.01	0.004	0.008
# of hotels	0.013***	0.002	0.014***	0.001
# of metro stations	0.025*	0.014	0.010	0..01
# of bus stops	0.031***	0.003	0.032***	0.002
Population density (1000 per km^2)	0.010***	0.001	0.010***	0.001
Employment density (1000 per km^2)	0.009***	0.002	0.007**	0.003

Note: ***: 99% significant; **: 95% significant; *: 90% significant

Table 4.4 Cont'd

	Model 3: pick-ups		Model 4: drop-offs	
	Coef.	Std. Err.	Coef.	Std. Err.
Time fixed-effect				
Feb	0.004***	0.00	-0.757***	0.001
Mar	0.201***	0.002	0.269***	0.001
Apr	1.063***	0.003	0.944***	0.002
May	1.378***	0.002	1.252***	0.002
Jun	2.022***	0.001	1.714***	0.001
Jul	2.623***	0.005	2.330***	0.005
Aug	1.851***	0.004	1.546***	0.001
Sept	2.082***	0.004	1.848***	0.004
Oct	2.443***	0.008	2.189***	0.006
Property value dummies (<50% as base)				
50% -100% median	1.783***	0.003	1.424***	0.005
100%-150% median	3.532***	0.013	3.139***	0.011
>=150% median	3.904***	0.029	3.629***	0.021
Interact property value with fixed-effect				
Feb: 50% -100% median	-0.396***	0.001	0.326***	0.001
Feb: 100%-150% median	-0.430***	0.002	0.503***	0.002
Feb: >=150% median	-0.151***	0.003	0.587***	0.003
Mar: 50% -100% median	0.156***	0.002	0.141***	0.002
Mar: 100%-150% median	0.100***	0.003	0.121***	0.002
Mar: >=150% median	0.267***	0.001	0.180***	0.001
Apr: 50% -100% median	-0.071***	0.003	0.222***	0.002
Apr:100%-150% median	-0.301***	0.003	-0.082***	0.002
Apr:>=150% median	-0.339***	0.003	-0.255***	0.002
May: 50% -100% median	-0.331***	0.002	-0.073***	0.002
May: 100%-150% median	-0.543***	0.003	-0.363***	0.003
May: >=150% median	-0.515***	0.003	-0.404***	0.002
Jun: 50% -100% median	-0.493***	0.001	-0.152***	0.002
Jun: 100%-150% median	-0.962***	0.002	-0.553***	0.002
Jun: >=150% median	-0.933***	0.002	-0.619***	0.002
Jul: 50% -100% median	-0.684***	0.004	-0.357***	0.005
Jul: 100%-150% median	-1.262***	0.004	-0.906***	0.005
Jul: >=150% median	-1.265***	0.005	-0.992***	0.005
Aug: 50% -100% median	-0.559***	0.003	-0.256***	0.001
Aug: 100%-150% median	-0.955***	0.003	-0.593***	0.002
Aug: >=150% median	-0.864***	0.004	-0.585***	0.002
Sept: 50% -100% median	-0.583***	0.001	-0.413***	0.004
Sept: 100%-150% median	-1.278***	0.002	-1.035***	0.004
Sept: >=150% median	-1.225***	0.003	-1.032***	0.005
Oct: 50% -100% median	-0.722***	0.005	-0.488***	0.005
Oct: 100%-150% median	-1.563***	0.005	-1.269***	0.005
Oct: >=150% median	-1.605***	0.004	-1.442***	0.005
Pseudo R^2	0.185		0.189	
# of grid cells	1291		1291	

Note: ***: 99% significant; **: 95% significant; *: 90% significant

The results from the second set of models shows that in general the association of the ride-sourcing fare with travel demand was much larger for richer neighborhoods, but the differences between neighborhoods with high and low property values decreased over time. In order to better interpret the estimation results, I calculated and plotted the incident rate ratio for property value dummies, shown in Figure 4.6. I can see from the figure that sensitivity to ride-sourcing fares is most strongly associated with property value in February and March when passenger promotions were first launched. During this period ride-sourcing trips were about 60 times more likely to take place in the richest residential neighborhoods (above 150% of the median property value) than the poorest (below 50% of median). The decision to stop subsidizing individual trips in June and July does not seem to differentiate richer and poorer neighborhoods. In September and October, when ride-sourcing fares increased, ride-sourcing trips were still about 10 times more likely to take place in the richest than in the poorest neighborhoods. Interestingly, lower-income travelers were less sensitive to price during the entire study period (as can be seen in Figure 4.6; the lines for the 50% to 100% of median categories are flat).

The estimation results did show that when passenger promotions and benefits came out in February, travel demand at wealthier communities increased much faster than poorer communities. But it is worth noticing that the differences in ride-sourcing trip generation between rich and poor communities reached record high in February and March, but such differences quickly become smaller after March for both trip pick-ups and drop-offs when ride-sourcing fare was still cheap. I can reasonably speculate that such spikes of difference in February and March are due to novelty factor of ride-sourcing as a new and cheap mode, especially for wealthier travelers. And the spike of actual demand happened in July, not February or March, therefore the growth of ride-sourcing demand from March to July can be contributed more on factors such as land use and locations of activities, other than property value.

Figure 4.6. IRR of property value categories (in terms of % median) and number of ride-sourcing trips per grid cell (pick-ups and drop-offs)

4.4 Conclusion

In this chapter I estimated the association between trip demand and average property value, in order to investigate the relationship between ride-sourcing fare and its users. Although in terms of sheer number of trips, our descriptive analysis showed that ride-sourcing is still largely a premium mode serving places with high property value, our estimation results from our controlled analysis showed that travelers from/to areas with lower property value behave differently to travelers from/to areas with higher property value. Ride-sourcing trips are more likely to take place in areas with higher property value when fare is low, but when fare of ride-sourcing increased, places with higher property value become less likely to be origins/destinations of ride-sourcing trips. We can therefore reasonable speculate that wealthier people tend to respond to low fares/promotions more compared to lower income people. That being said, the decrease in ride-sourcing fares would likely to benefit middle-high income travelers more by make ride-sourcing a competitive mode for to replace their driving trips. But the increase of ride-sourcing fare would make lower income travelers baring the consequences of increased travel cost due to their lower elasticity to ride-sourcing fare.

Although this study is one of the first to have access to spatial data on ride-sourcing travel, it has its limitations. First, although property value is a reasonable proxy for the affluence of origin and destination neighborhoods, future study would benefit by analyzing data explicitly describing the income level of residents, although such dataset was not publicly available. Secondly, in this study I investigated trip pick-ups and drop-offs, rather than trips. If I were to investigate the entire trip as a whole, looking at both pickups and drop-offs, I may find similar travel pattern for trips that are both originated and ended at places with low average property value as previous research found that taxis are used by low income travelers to fulfill short distance travel (Renne and Bennett 2014). Thirdly, I only have data on ride-sourcing travel demand, thus limiting the ability to detect substitution or complementarity with other modes such as transit. It would be very helpful to investigate whether the lower demand of ride-sourcing by lower-income travelers is explicitly associated with lack of access to ride-sourcing supply in terms of wait time, with better availability of public transportation, and/or with lower overall travel demand.

Chapter 5. Exploring app-based, on-demand ride-sourcing's relationship with parking, transit access and congestion: empirical evidence from Didi Chuxing in Shanghai

5.1 Introduction

The rapid emergence and growth of transportation network companies (TNCs) such as Uber, Lyft, and Didi Chuxing, operating app-based on-demand ride-sourcing services, has led to a debate over their role in the urban transport system. Ride-sourcing supporters argue that it is a convenient and efficient mode of travel, with the potential to reduce private auto use, decrease the need for parking and solve the first/last mile problem of public transit (Rayle et al. 2016; Shaheen et al. 2016). Other scholars and government officials are less sanguine, concerned that the growing popularity of ride-sourcing might undermine congestion mitigation and replace public transportation, eventually resulting in negative environmental and social consequences (Rayle et al. 2016, SFCTA 2017, Schaller, 2018).

The growth of the ride-sourcing business has brought significant challenges for planners, engineers, and policy makers, due to the magnitude and uncertainty of its impacts. For example, the number of full- and part-time TNC drivers in Shanghai exceeded 287,700 in August 2016, a mere two years after ride-sourcing was introduced (Didi Chuxing 2016). In 2017 it was estimated that more than 15% of trips and 20% of vehicle mileage in San Francisco was produced by TNCs (SFCTA 2017).

Despite the exponential growth of ride-sourcing worldwide, information about its relationship with parking, public transit, and congestion is still very limited. These are measures of the relative ease of driving and using transit, and might predict lower or higher use of ride-sourcing as an alternative. Thus one might expect that ride-sourcing demand would be higher in places with restricted parking, at least for auto owners. One might also expect that demand would be higher in places with lower levels of transit access because of the significant time advantages it could offer there, despite ride-sourcing being more expensive than transit. Finally, I hypothesized that ride-sourcing may be positively associated with congestion because it may relieve travelers from driving on congested roads, or because it may in fact cause congestion.

Using a panel dataset from Didi Chuxing, the largest TNC by market share in China, this study investigates the following questions: (1) To what extent is parking supply associated with lower ride-sourcing demand? (2) Does better transit access reduce or increase the use of ride-sourcing? and (3) Does higher congestion affect ride-sourcing demand?

This chapter addresses these problems by using a statistical modeling method to capture non-linearity using spatially precise data from Shanghai. And unlike some previous studies that relies on retrospective survey data, this chapter relies on disaggregate trip data which are consistent over time. Finally, as far as I am aware, this is the first study to empirically investigate parking availability with objective parking data.

5.2 Data and methodology

The ride-sourcing GPS dataset used in this paper was provided by Didi Chuxing. Didi Chuxing accounted more than 80% of the ride-sourcing market share in China, 2015 (TalkingData, 2015). The dataset consists of 100,000 trips within the outer ring road of Shanghai that were randomly sampled by Didi Chuxing from their database from the first seven days of each month from January to April 2015 from the Didi Chuxing database. The data include the date, time of day, addresses and geolocations of origins and destinations of each trip. I accurately represent the changing monthly trip volume by assigning a weight to samples from different months that represent the total trip volume in that specific month. The trips contained in this dataset were taken on the “Kuai Che” service provided by Didi Chuxing, which is similar to standard Lyft and UberX service in which the driver is considered to be a TNC “subscriber” and is responsible for the vehicle. Table 5.1. compares the ride-sourcing travel cost with the fare of other travel modes in Shanghai, 2015. It should be noted that during this period there were no official regulations on ride-sourcing and TNCs, which means that there were little limitations on the supply or the demand of ride-sourcing.

Our study area is defined as land falling within the Outer Ring Road of Shanghai. This zone is usually considered to be the urbanized portion of Shanghai. To simplify the analysis, I divided the dataset temporally into 6-hour time segments: from midnight to 6 am, 6 am to noon, noon to 6 pm, and 6 pm to mid-night. Since our dataset is collected during the first seven days of each month from January to April, the number of time segments is 4 X 7 X 4 which is 112. I then aggregated pick-up and drop-off locations into 1-km grid cells as our basic spatial unit. There are 709 grid cells within our study area. Figure 5.1 shows the average ride-sourcing demand in each grid cell time of day. I can see the demand are much higher at afternoon and evening. Figure 5.2 illustrates the spatial distribution of ride-sourcing demand in each grid cell.

Table 5.1. Fare of ride-sourcing of Didi Chuxing and other travel modes in Shanghai, 2015

Mode	Fare
Ride-sourcing before July, 2015	~￥1 ($0.17) per km, with various coupons and discounts
Taxi price	￥14 (initial fee) + ￥2.5 per km
Bus Price	City Bus (operate within the outer rim of Shanghai): ￥2 flat rate Commute Bus (operate between outer rim of Shanghai and Suburbs): ￥2 + ￥0.2 per km, ￥8 cap
Metro Price	￥3 (the first 6 kms) + ￥1 (per additional 10 kms)

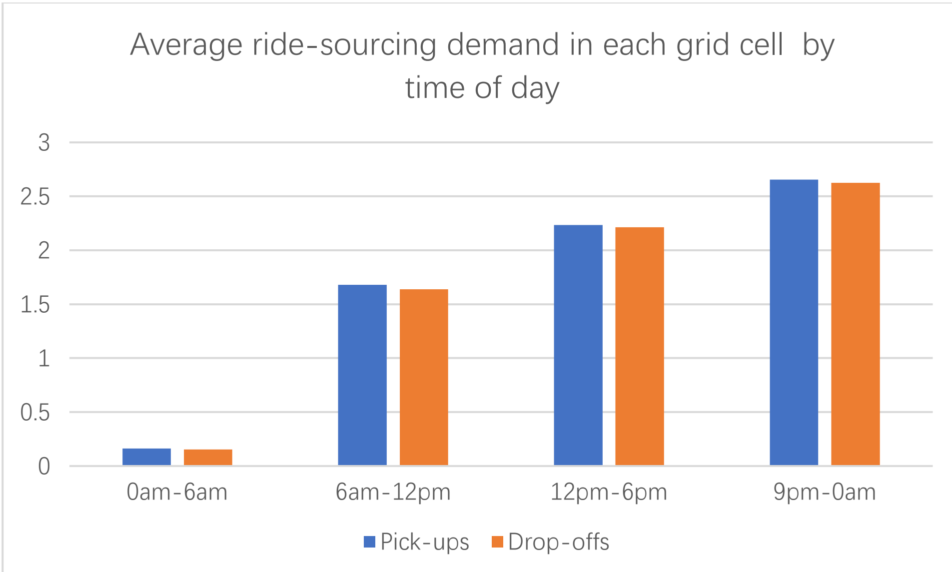

Figure 5.4. Average ride-sourcing demand by time of day

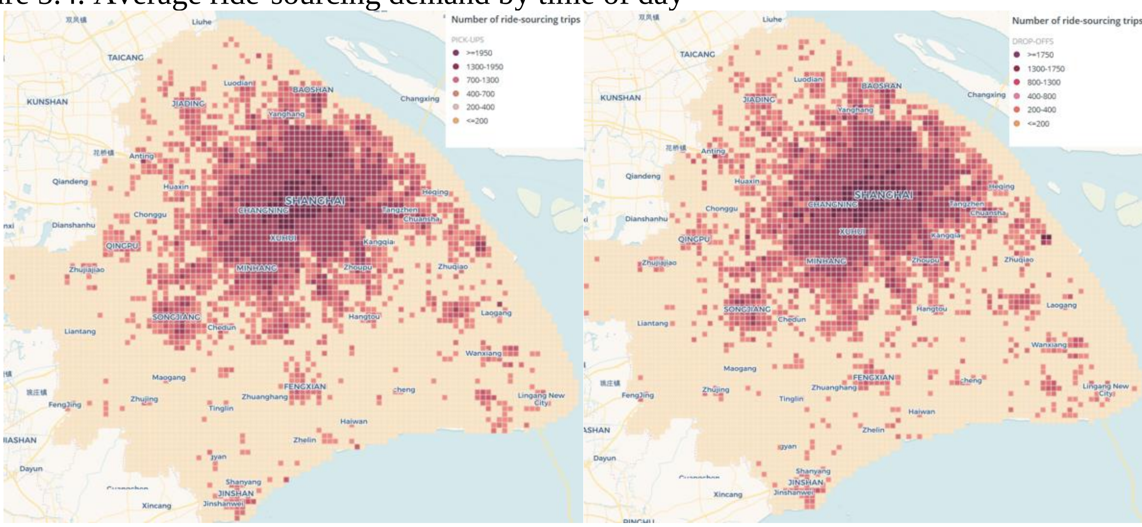

Figure 5.2. Map of ride-sourcing pick-ups (left) and drop-offs (right) in each grid cell

Apart from ride-sourcing data, I also acquired data about road congestion and parking supply. Both of the two datasets I relied upon are publicly available from the Data service website of the Shanghai Municipal Government via http://www.datashanghai.gov.cn/. The data on congestion provide congestion index for 68 travel analysis zones (TAZs) within the outer ring highway of Shanghai (which is usually treated as the urbanized area). This index ranges from 0 to 60, and is calculated every 10 minutes over a period of 9 months by comparing actual travel speed with free flow speed (Guan, 2004). The higher the congestion index, the more congested an area is. For our analysis, I aggregate the congestion index into 112 time-segments and 709 grid cells (consistent with the ride-sourcing data aggregation). The parking data include the number of parking spaces and the addresses of all off-street parking facilities in Shanghai, as well as the number of on-street parking for each road segments within the outer ring highway as of 2013. I then geocoded the address into a map-based shapefile using Google Map API for this analysis. I include the location of all bus stops and metro stations to represent the transit accessibility in different locations.

In a more complete effort to control for other factors I turned to regression analysis. I modeled the ride-sourcing demand (for pick-ups and drop-offs separately) using a generalized additive mixed model (GAMM). GAMM is a variation of the generalized linear model (GLM) in which the linear dependent variable depends on a smooth spline function of independent variables. The fitted spline function would be a non-linear curve which represents the relationship between the dependent variable and the independent variable. GAMM also allows the relaxation of assumptions for linear models, and the model can also include fixed effects and random effects to control for unobserved factors that are associated with spatial and temporal variations. GAMM has recently been used in transportation research to explore the non-linear relationship between the use of electric vehicles and the built environment (Hu et al. 2018). Since our dependent variable would be the number of ride-sourcing pick-ups and drop-offs in each grid cell, I assume that they follow a Poisson distribution.

$$demand_{it} = exp(\beta_0 + \mathbf{X}'_{\mathbf{it}} \cdot \boldsymbol{\beta} + \sum \boldsymbol{f}(\boldsymbol{x_i}) + \boldsymbol{Zb} + \boldsymbol{\mu}_{\mathbf{t}} + \varepsilon_{it}). \quad (1)$$

In this formula, $demand_{it}$ is the dependent variable representing ride-sourcing demand in grid cell *i* in month *t*, measured as the number of pick-ups or drop-offs. Vector $\mathbf{X}_{\mathbf{it}}$ is a matrix of independent control variables measured at the grid cell (neighborhood) level, including population and employment density as well as activity locations in the neighborhood (see Table 5.2). The model treats $\mathbf{X}_{\mathbf{it}}$ as time-invariant given that these built environment variables are unlikely to have changed very much over the course of the nine-month observation period, and if so, not systematically in such a way as to bias results. The term $\sum \boldsymbol{f}(\boldsymbol{x_i})$ represents the spline function of selected variables: parking, transit access and congestion. The spline function in the model allows us to capture the non-linear relationship of these variables of interest. Vector $\boldsymbol{\mu}_{\mathbf{t}}$ consists of time-period fixed effects. I included $\boldsymbol{Zb}$ as random effects for each grid cells. As constructed, the model allows a partial control for unobserved variables associated with time and location by $\boldsymbol{\mu}_{\mathbf{t}}$ and $\boldsymbol{Zb}$, including price of ride-sourcing and demand for other travel modes. I included auto-regressive models for the error terms to control for correlation in time, as our analysis is on panel data. Same as Hu et al. (2016), I include a smooth term for latitude and longitude in our GAMM to control for spatial-autocorrelation.

In additions to the fixed effects which rule out the unobserved spatial and temporal variations, other control variables are included in this study to represent other possible factors that influence the demand for ride-sourcing. Theory of travel demand and empirical evidence suggested that activity centers are likely to generate trips (Hanson & Giuliano 2004). Thus I uses all points of interest (POI) in Shanghai to represent all the activity centers. This POI dataset has over 200,000 activity locations, including all offices, shops, restaurants, convenience stores, recreational facilities (including theatres and cinemas), schools, hospitals, and parks. I used the number of each of these activity location types as control variables in the model. I also used the average residential property value as a proxy measure for the value of land in the grid cell, because previous literature showed that higher land value is associated with more travel demand (Hanson & Giuliano 2004). Population density and employment density at level similar to census tract level the US is included in the study as well since research showed that density is related to the demand of activity (Chatman 2008; Ewing & Cervero 2010). Average road length within each grid cell is also added as a control since previous research showed that road supply is directly associated with motorized travel (Chatman 2008). The transit dataset, property dataset and the points of interest dataset are all public available data in Shanghai, and the authors obtained these datasets through Metrodata Tech[3]. I aggregated the data into 1-km grid cells as well for our analysis. Table 5.2 shows the descriptive statistics of the variables involved in the analysis.

Table 5.2. Summary statistics of variables

Variable	Min	Max	Mean	Std. Dev.
# of pick-ups (every 6 hours)	0	160	1.68	4.69
# of drop-offs (every 6 hours)	0	107	1.66	4.39
# of bus stops	0	33	6.66	4.94
# of metro stations	0	3	0.30	0.53
Total parking	37.00	7607	1002.42	1068.13
Congestion index	7.33	55.86	24.28	8.18
# of offices	0	477	40.29	53.42
# of shops	0	1185	72.47	106.14
# of restaurants	0	427	35.77	48.70
# of recreation facilities	0	88	9.43	11.33
# of convenience stores	0	50	5.58	7.20
# of hotels	0	47	4.84	6.41
# of hospitals	0	5	.28	0.65
# of schools	0	6	.74	1.02
# of colleges	0	14	.39	1.15
# of parks	0	2	.15	0.39
Average road length (km)	0.03	4.28	0.99	0.57
Average property value per m^2 (￥1000)	0	217	29.06	23.28
Population density (1000 per km^2)	0	48	5.71	8.72
Employment density (1000 per km^2)	0	33	2.04	4.29

5.3 Results

The estimation results show that controlling for other factors, the amount of total parking supply is positively associated with ride-sourcing demand when the number of parking spots are lower than 4,000, but is negatively associated with ride-sourcing demand when there are more than 4,000 parking spots. Similarly, the number of bus stops is positively associated with ride-sourcing demand when there are fewer bus stops nearby, and it is negatively correlated with ride-sourcing when the number of bus stops are greater than 10. Congestion level also has a non-linear relationship with ride-sourcing pick-ups, but it has a monotonic increasing relationship with ride-sourcing drop-offs. I estimated multiple models using different model types including

[3] Metrodata Tech. is a private technology company based in Shanghai. More information about Metrodata Tech. can be found here: www.metrodata.cn

generalized linear models, GAMM with/without spline interactions and GAMM with/without series correlation. By comparing the modeling results from these series of models, I believed that our estimation results were robust. I chose to present GAMM with spline interaction and with series correlation by conducting likelihood ratio tests. Additionally, I conducted variance inflation tests in the linear model and found that adding the control variables did not bias the estimation for our variables of interest (parking, transit and congestion). Table 5.3 show the estimation results of our best models.

As I can see from Table 5.3, many control variables (all control variables are in linear terms) are insignificant because the model added time fixed effects and individual random effects. Among all control variables, average road length seems to have the largest impact on ride-sourcing demand. The number of office and number of hotels are positively correlated with ride-sourcing demand, suggesting that offices and hotels tend to attract and generate more ride-sourcing trips. The number of shops, on the other hand, would decrease the likelihood of ride-sourcing trips. Ride-sourcing demand are also likely to generate from areas with higher population density and higher property value.

As for our variables of interest, I estimated a spline function for each one of them and reported an estimated degree of freedom. All estimated spline functions are statistically significant. The higher estimated degree of freedom indicates higher power function used for the smooth terms. For example, the estimated degree of freedom for the spline term of the number of bus stops in Model 1 is 3.85, indicating that the spline function is to the power of 3.85. In order to make the estimation results of the spline term more interpretable, I plotted the partial effect of the spline terms, shown in Figure 5.3, 5.4 and 5.5.

Parking

The GAMM model estimated degree of freedom for the best fit for its spline function, which is estimated by non-parametric methods. The estimated degree of freedom of the spline function of total parking is 7.25 for the pick-ups model (Model 1 in Table 5.3), and 7.07 for the drop-offs model (Model 2 in Table 5.3). This indicates that the relationship between ride-sourcing demand and parking is a non-linear function that has a power of above 7. Figure 5.3 shows the partial effect of parking on ride-sourcing demand. I can see from Figure 5.3 that ride-sourcing demand increases dramatically as parking supply increased from 0 to around 500. Then ride-sourcing demand continue to increase, but the increase is much slower, alongside with parking supply. After parking supply reaches 4,000 spaces, ride-sourcing demand is negatively associated with parking and the more parking available, the less demand for ride-sourcing.

Table 5.3. Estimation results of GAMM

	Model 1: pick-ups			Model 2: drop-offs		
Linear terms	Coef.		Std. Err.	Coef.		Std. Err.
(Intercept)	-3.793	***	0.13	-4.245	***	0.13
# of offices	0.005	***	0.001	0.004	**	0.001
# of shops	-0.002	**	0.001	-0.002	**	0.001
# of restaurants	-0.002		0.001	0.000		0.001
# of recreation facilities	0.005		0.01	0.004		0.01
# of convenient stores	0.012		0.01	0.012		0.01
# of hotels	0.022	*	0.01	0.019	*	0.01
# of hospitals	-0.036		0.07	-0.029		0.07
# of schools	0.029		0.05	0.031		0.05
# of colleges	-0.017		0.04	-0.007		0.04
# of parks	0.033		0.11	0.033		0.11
Average road length (km)	6.575	***	0.87	6.139	***	0.86
# of metro stations	0.029		0.09	-0.005		0.09
Population density (1000 per km2)	0.018	***	0.001	0.021	***	0.001
Employment density (1000 per km2)	-0.007		0.01	-0.008		0.01
Average property value (￥1000)	0.146	***	0.02	0.149	***	0.02
Fixed effects	Yes			Yes		
Spline terms	Estimated degree of freedom			Estimated degree of freedom		
Individual random effect	617.48	***		624.40	***	
s(LAT, LONG)	$7.6*10^{-8}$	***		$6.6*10^{-8}$	***	
s(Total parking)	7.25	***		7.07	***	
s(# of bus stations)	3.81	***		3.69	***	
s(Congestion index)	8.16	***		8.70	***	
Adjusted R square	0.735			0.707		

Note: ***: 99% significant; **: 95% significant; *: 90% significant

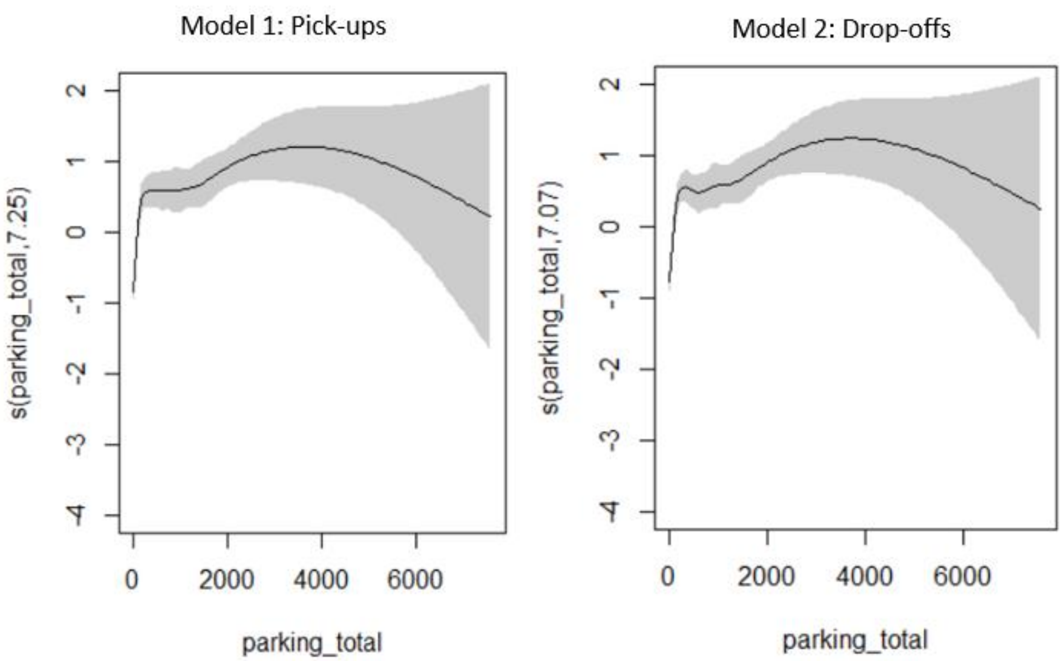

Figure 5.3. Partial effect of total parking on ride-sourcing demand (95% intervals are shown in the grey areas)

Transit

The partial effect of the number of bus stops on ride-sourcing demand (shown in Figure 5.4) is somewhat similar to the effect of parking on ride-sourcing. The shape the spline terms for both pick-ups and drop-offs models are similar to a quadratic function. Ride-sourcing demand is positively correlated with number of bus stops when the number of bus stops is below 10 per grid cell, indicating a complementary relationship, but the relationship reversed when the number of bus stops is greater than 10 in each grid cell, suggesting that ride-sourcing demand are smaller in areas with good bus access.

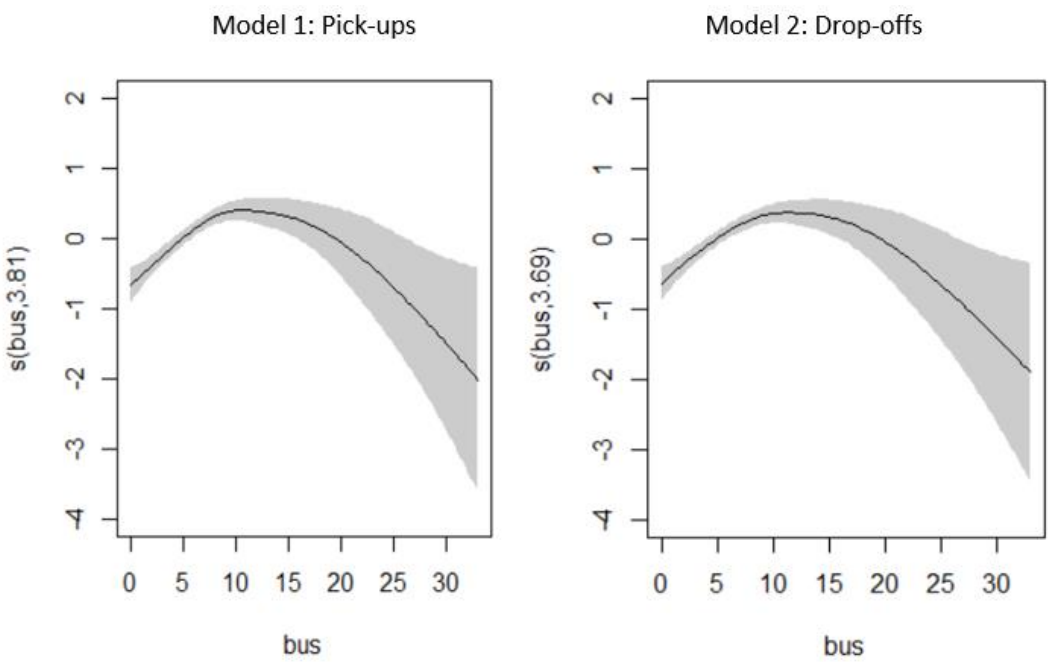

Figure 5.4. Partial effect of number of bus stops on ride-sourcing demand (95% intervals are shown in the grey areas)

I also estimated the relationship between metro access and ride-sourcing, however the coefficients for the number of metro stations turned out to be insignificant in all our models, regardless of whether I treated the number of metro stations as a linear variable or a spline function. The reason why the spline function of metro stations is insignificant could partly contribute the fact that the number of metro stations in each grid cell is very small (ranging from 0 to 3, shown in Table 5.2).

Congestion

Unlike parking and transit, the partial effect of congestion on ride-sourcing demand differed between pick-ups and drop-offs. Ride-sourcing pick-ups is at first positively associated with congestion level until congestion level reached around 35, after that ride-sourcing pick-ups started to decrease with congestion level increased (Figure 5.5 left). Ride-sourcing drop-offs, however, is generally positively associated with congestion level, indicating that the higher the congestion level, the more ride-sourcing drop-offs (Figure 5.5 right). Note that the confidence intervals are wider for models of pick-ups, indicating that the may still may be similar to drop-offs.

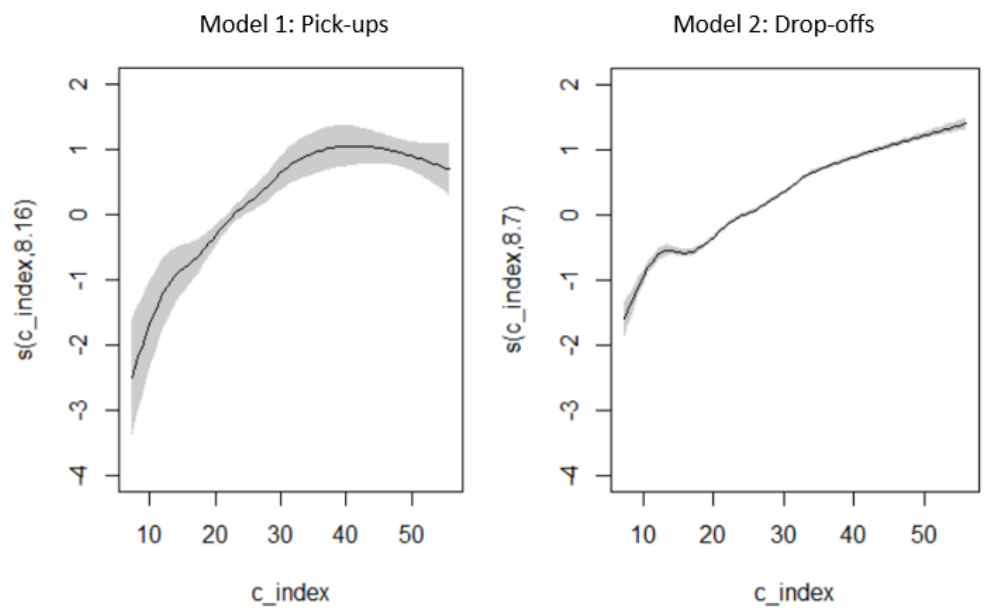

Figure 5.5. Partial effect of congestion index on ride-sourcing demand (95% intervals are shown in the grey areas)

In addition to the spline functions for each individual variable of interest (parking, bus and congestion), I also explored the interaction between each pair of these three variables, and their estimated association with ride-sourcing demand. The GAMM model allows us to include a tensor product between two variables, $\boldsymbol{f(x_1)}\otimes\boldsymbol{f(x_2)}$**,** in lieu of the regular spline function term of the individual variable $\sum \boldsymbol{f(x_i)}$**.** Including the tensor product enables us to estimate the non-linear interaction between the two variables in addition to the non-linear spline functions of each variable by itself. I estimated GAMM models with interactions between parking and congestion, bus and congestion, as well as parking and bus. According to likelihood ratio tests, only the interaction between parking and congestion contributes to more statistical power compared to the models without the interaction (shown in Table 5.3). Because the GAMM model results for this interaction was difficult to interpret, I used an alternative way to test for the interaction by dividing the parking and congestion variables into low, medium and high levels based on the 25th and 75th percentile, and then interacting each one of them to create nine dummy interaction variables. I added eight of them into the models in Table 5.3 (using the other one as the base category), and estimated the parameters. The results were consistent with the GAMM model results for interaction term, but were less complicated and more interpretable. I found that for ride-sourcing pick-ups, if I hold one of the two variables (parking and congestion) constant, the estimated results for the other variable would first increased and then decreased, similar to the patterns in Figure 5.3 and 5.5. But for ride-sourcing drop-offs, congestion is positively correlated with ride-sourcing drop-offs when parking supply is high. This result provides more nuanced interpretation for Figure 5.5, suggesting that when parking supply is in the upper quartile, ride-sourcing drop-offs are positively correlated with congestion. This suggests that ride-sourcing may replace driving when there is limited parking supply and higher congestion. Note that adding the interaction terms does not change the estimated individual spline functions very much.

I also estimated regular mixed-effect models with the standard linear terms to represent parking, transit, and congestion. Although I decided not to show the estimation results of the regular mixed-effect models in this paper due to the space limitation and because I would like to focus on the GAMM model results, I briefly discuss here the estimation results from the regular model for comparison. All estimated coefficients for the variables of interest were statistically significant in the regular mixed-effect models with standard linear terms. The estimated parameters imply that ride-sourcing has a positive relationship with total parking (the estimated coefficients are 0.0035 for pick-ups and 0.0037 for drop-offs), number of bus stops (the estimated coefficients are 0.015 for pick-ups and 0.0145 for drop-offs), and congestion (the estimated coefficients are 0.077 for pick-ups and 0.072 for drop-offs). Thus, without having run a GAMM model I would have erroneously concluded that there are linear and positive relationships across the range of values. In this case the GAMM model has a clear advantage over the regular linear mixed-effect model in providing a more accurate understanding of the data.

5.4 Conclusion

In this study I investigated the relationship between ride-sourcing demand and parking, transit access and congestion using a GAMM on panel data over a four-month period in the city of Shanghai. The main advantage of this paper is that I observed ride-sourcing's non-linear relationship with ride-sourcing and parking, bus accessibility, and congestion, with high spatial resolution, controlling for possible confounding factors using mixed effects and a large number of control variables.

The results suggest that, first, parking supply does not reduce the demand for ride-sourcing in the Shanghai context unless it is sufficiently large. Second, whether ride-sourcing substitutes with or complements public transit depends on service coverage: ride-sourcing tends to increase as transit grows from below-average service levels while it decreases as transit service becomes denser. Third, ride-sourcing demand increases with congestion, but when congestion is at its most severe there are fewer pick-ups. Our methods enable us to find these varying results because we use a nonlinear fitted GAMM model, and we show that a conventional model yields erroneous results.

Chapter 6. Discussion of ride-sourcing behavioral and implication for policy

6.1 Socio-economic accessibility and spatial accessibility of ride-sourcing

<u>Comparison with taxi in terms of accessibility</u>

Previous research indicated that taxi drivers usually concentrated in areas with "perceived higher usage areas", e.g. CBD, train terminals, and airports, where drivers believed they can match with higher demand (Austin and Zegras 2012). This concentration of taxi supply in high demand location was a result of lacking information between drivers and passengers, and may results in underserving low-income neighborhoods (Austin and Zegras 2012). Ride-sourcing on the other hand, match demand and supply by using an online platform. So far evidence in the US did not show uneven distribution against poorer neighborhoods in terms of wait time (Hughes and MacKenzie 2016).

Taxi driver may refuse to serve certain people based on certain characteristics. In China, taxi drivers in some cities were unwilling to serve travelers that travel shorter distance because they believe that serving such customer would decrease earning. In some other cities this situation has improved since taxi companies now required drivers to serve any customer based on first come first serve principle. Compare to taxies, ride-sourcing have theoretical improvements because of better communication and feedbacks from both sides. Ride-sourcing drivers could not discriminate passengers based on destination, although they may still discriminate by refusing pick-up requests from passengers starting from low-income communities or neighborhoods of color. Current ride-sourcing companies would penalize drivers who successively refuse pick-ups request, which can prohibit the discriminatory driver behavior to a certain degree. Further study is needed to fully understand this issue.

The results from chapter 4 imply that a decrease in ride-sourcing fares would likely benefit middle to high income travelers more than low-income travelers, by making ride-sourcing an economically competitive mode for those groups. Usage is much higher in neighborhoods with higher property values when fares are lower. At the same time, however, there is still significant though lower use of ride-sourcing in lower-income neighborhoods, and usage in those locations is *less* responsive to the fare. This suggests that ride-sourcing services are being used more often for non-discretionary trips by those in lower income neighborhoods, perhaps because of the lack of auto ownership and poor public transportation options for those trips.

Thus ride-sourcing policy which results in fare increases would likely to pose a substantial burden for lower-income travelers, although the number of such lower income travelers may be small compared to the number of middle-to-high income travelers. The consumer surplus for these lower income users is likely higher than the consumer surplus for higher income users, as lower income users are willing to pay more (as a fraction of income) to access ride-sourcing. I could reasonably speculate that if a restrictive policy on ride-sourcing supply, such as the one imposed in Shanghai after 2016, would give the fraction of ride-sourcing users who live in lower-income neighborhoods little alternative but to accept the higher fare.

That said, the analysis showed that ride-sourcing is still largely a premium mode serving places with high property value, and there were relatively small number of users traveling to/from lower property price areas. When ride-sourcing fares decreased quickly due to market battles in the first half of 2015, at least for a short period of time, ride-sourcing fare was very cheap and even competitive in comparison to public transit. But wealthier people tended to take advantage of the cheap services, while poorer people generally seemed somewhat inelastic to fare changes. Given the fact that Shanghai has very high smartphone penetration rates (as discussed previously) and low auto ownership, there are likely additional unobserved factors that hinder the use of ride-sourcing by low income travelers. In the US, although lower income travelers can be aware of the availability of ride-sourcing, there are often still several barriers for using them other than cost, among them limited payment methods and low digital literacy (e.g., Dillahunt et al. 2017), even though ride-sourcing companies may market to lower-income travelers by promoting services such as shared rides.

6.2 Ride-sourcing's role in transportation system

I found that ride-sourcing demand (both pick-ups and drop-offs) are positively associated with parking supply when parking supply is low, but negatively associated with parking supply when parking is high. The relationship between ride-sourcing and bus transit depended on the coverage of bus services. Ride-sourcing tended to negatively correlate with each transit access when the density of bus stops was high, and positively correlate with each other when there were fewer bus stops. Finally, ride-sourcing demand was positively correlated with congestion, except that when congestion at the pick-up location was severe, demand declined.

<u>Parking</u>

The estimation results suggest that parking does not necessarily reduce the demand for ride-sourcing unless the parking supply is large enough, and that a lack of parking supply is not necessarily associated with an increase in ride-sourcing demand. Part of this fact may be due to local context. In Shanghai, auto-ownership is relatively low due to its license auction policy (Chen and Zhao 2013). In 2014, there were less than 15 private cars for every 100 residents in Shanghai, and private cars accounted for about 17% of trips (Shanghai Urban-Rural Construction and Traffic Development Academy 2015). In the Shanghai context where auto ownership is relatively low, ride-sourcing may not strongly compete with driving except when congestion is particularly high and parking is particularly scarce at the destination (see below). It is also possible that auto-users and ride hailers are not the same group of travelers, consistent with some existing research that has showed ride-hailers in some US cities are much less likely to be car owners (eg. Rayle et al. 2016). The socio-economic status of ride-hailers in Shanghai is still unclear, given that the nature of our data; this remains an important topic for future research.

Public transit

The descriptive analysis using Google Map API in Chapter 3 showed that the majority of the ride-sourcing trips has transit alternatives. However, transit alternatives compete poorly with ride-sourcing in terms of travel time and convenience: on average, it would take approximately more than twice of the ride-sourcing travel time for a passenger to arrive at the destination if taking transit. It is likely that ride-sourcing is competing with both bus and metro, but bus riders are more likely to switch to ride-sourcing than metro riders due to low time performance of buses in Shanghai. Also, the descriptive analysis discovered that ride-sourcing is more likely to compete with transit in longer travel and peak hour travel by looking at what proportion of the ride-sourcing trips can be replaced by transit.

Moreover, analysis in Chapter 5 suggested that ride-sourcing would be expected to compete with existing travel modes including public transit if it is inexpensive and convenient enough. The analysis suggests that ride-sourcing increases with density when bus access is relatively low, and that it drops off as bus stop density increases, suggesting it could be shunting transit riders away in places where transit accessibility is relatively low, while not being as competitive with transit when transit services are denser. The findings implied that public transit's key competitiveness lies with its system performance, particularly with density and accessibility.

Congestion

Unlike the relationship of ride-sourcing with parking and transit, the analysis in Chapter 5 showed a much more complex relationship with road congestion. On one hand, ride-sourcing pick-ups were positively associated with congestion when congestion level is low, but somewhat negatively correlated with congestion when congestion is severe. On the other hand, ride-sourcing drop-offs are positively correlated with congestion regardless of the congestion level, and particularly so when parking is scarce and congestion is high. One possible hypothesis is that ride-sourcing causes congestion in cities, as many cities have claimed (eg. SF Chronicle 2018). Another hypothesis draws upon research on autonomous vehicles that argues congestion will encourage people to use AV services, because taking a ride as a passenger eases the burden of driving in congestion (Howard and Dai 2014; Fagnant and Kockelman 2015; Wadud et al. 2016). Our analysis cannot perfectly control this potential reverse causal relationship between ride-sourcing and congestion just as any fixed effect models. Nevertheless, our analysis still has interesting implications for ride-sourcing and congestion. Recall that ride-sourcing pick-ups were positively correlated with congestion level at low congestion levels but negatively correlated once congestion is high enough. On the other hand, ride-sourcing drop-offs were positively correlated with congestion regardless of the current congestion level. Based on these findings, I may speculate that the first hypothesis is more likely to be true and ride-sourcing causes congestion at the drop-off end. The second hypothesis that congestion would encourage people to ride hail (due to less burden and lower value of time, etc) is likely to be false, because if it is true then ride-sourcing pick-ups should also be monotonically positively correlated with congestion in Figure 5.5. Thus it is possible that ride-sourcing users refrained from traveling due to the congestion level at the origin end (due to longer wait times caused by congestion, for example), but ride-sourcing still contribute to the congestion level at destination end.

But the second hypothesis could still be valid if the reason for decreasing ride-sourcing pick-ups at higher congested areas is something other than congestion. For instance, if ride-sourcing users decide not to ride hail during congested hours or areas only because the wait time is higher, then people may still willing to take ride-sourcing regardless of the congestion level if the wait time is within an acceptable range. Previous research in San Francisco found that the waiting time for the vast majority of the actual ride-sourcing trips are below 10 minutes (Rayle et al. 2016). Future research could potentially investigate how wait time affects people's decision to ride hail.

6.3 Limitation of the study

There are a few limitations of this study. First, I investigated trip pick-ups and drop-offs separately. Previous research found that taxis are used by low income travelers to fulfill short distance travel (Renne & Bennett, 2014). If I investigate the entire trip as whole, there may be similar travel pattern for trips that are both originated and ended at places with low average property value. Second, I only have data on ride-sourcing travel demand, thus limiting the ability to detect substitution or complementarity with other modes such as transit. In future study I may investigate whether the low demand of ride-sourcing by lower income travelers is associated with lack of access to ride-sourcing supply in terms of wait time, or with the lower overall travel demand. Third, the impact of surge-pricing on ride-sourcing demand during peak hour is under-studied, and due to the sparsity of data point in each day, this analysis did not assess the effect of surge-pricing. Additionally, there may be different ride-sourcing travel pattern by different time of day. Future research would focus on these aspects.

References

Austin, D., and P. Zegras. (2012). Taxicabs as Public Transportation in Boston, Massachusetts. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2277, 2012, pp. 65–74.

Babar, Y., & Burtch, G. (2017). *Examining the Impact of Ridehailing Services on Public Transit Use* (SSRN Scholarly Paper No. ID 3042805). Retrieved from Social Science Research Network website: https://papers.ssrn.com/abstract=3042805

Brown, A. E. (2018). *Ridehail Revolution: Ridehail Travel and Equity in Los Angeles*. UCLA. Retrieved from https://escholarship.org/uc/item/4r22m57k

Chan, N. D., and S. Shaheen. (2012). Ridesharing in North America: Past, Present, and Future. *Transport Reviews*, Vol. 32, No. 1, 2012, pp. 93–112.

Chatman, D. G. (2008). Deconstructing development density: Quality, quantity and price effects on household non-work travel. *Transportation Research Part A: Policy and Practice*, *42*(7), 1008–1030. https://doi.org/10.1016/j.tra.2008.02.003

Chen, X., Zhao, J. (2013). Bidding to drive: Car license auction policy in Shanghai and its public acceptance. *Transport Policy*, *27*, 39–52. https://doi.org/10.1016/j.tranpol.2012.11.016

China National Radio. (2015). Taxi survey from 20 cities in China. Retrieved July 29, 2018, from http://news.cnr.cn/dj/20150803/t20150803_519402657.shtml

Clewlow, R. R., Mishra, G. S. (2017). *Disruptive transportation: the adoption, utilization, and impacts of ride-sourcing in the United States*. University of California, Davis, Research Report UCD-ITS-RR-17-0.

Coffee, N. T., Lockwood, T., Hugo, G., Paquet, C., Howard, N. J., Daniel, M. (2013). Relative residential property value as a socio-economic status indicator for health research. *International Journal of Health Geographics*, *12*, 22. https://doi.org/10.1186/1476-072X-12-22

Dawes, M. (2016). Perspectives on the Ridesourcing Revolution : surveying individual attitudes toward Uber and Lyft to inform urban transportation policymaking (Thesis). Massachusetts Institute of Technology. Retrieved from http://dspace.mit.edu/handle/1721.1/104994

Didi Chuxing. (2016). Employment report in Didi Chuxing's app-based mobile services.

Dillahunt, T. R., Kameswaran, V., Li, L., Rosenblat, T. (2017). Uncovering the Values and Constraints of Real-time Ridesharing for Low-resource Populations. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 2757–2769). New York, NY, USA: ACM. https://doi.org/10.1145/3025453.3025470

Ewing, R., & Cervero, R. (2010). Travel and the Built Environment. *Journal of the American Planning Association*, *76*(3), 265–294. https://doi.org/10.1080/01944361003766766

Fagnant, D. J., & Kockelman, K. (2015). Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations. *Transportation Research Part A: Policy and Practice*, *77*, 167–181. https://doi.org/10.1016/j.tra.2015.04.003

Feigon, S., Murphy, C. (2018). *Broadening Understanding of the Interplay Between Public Transit, Shared Mobility, and Personal Automobiles*. Pre-publication draft of TCRP Research Report 195. Transportation Research Board, Washington, D.C. Retrieved from http://www.trb.org/Publications/Blurbs/177112.aspx

Financial Times. (2017). Didi guts Shanghai fleet following anti-migrant rules. Retrieved April 12, 2017, from https://www.ft.com/content/dee63006-ded3-11e6-9d7c-be108f1c1dce

FORWARD Business Information Co. Ltd. (2019). *Report of market demand forecast and investment strategy planning on China smartphone industry: 2019-2024*. Retrieved from https://bg.qianzhan.com/report/detail/ee7847658f524b0f.html

Guan. (2004). Urban Traffic Index， Travel Index and Mathematic Mode. *Journal of ’Transportation Systems Engineering and Information Technology*, *4*(1), 49–53.

Hall, J. D., Palsson, C., & Price, J. (2018). Is Uber a substitute or complement for public transit? *Journal of Urban Economics*, *108*, 36–50. https://doi.org/10.1016/j.jue.2018.09.003

Hall, J. V., and Krueger, A. B. (2015). An Analysis of the Labor Market for Uber’s Driver-Partners in the United States. Working Paper.

Hamari, J., Sjöklint, M., & Ukkonen, A. (2015). The sharing economy: Why people participate in collaborative consumption. *Journal of the Association for Information Science and Technology*.

Hampshire, R. C., Simek, C., Fabusuyi, T., Di, X., Chen, X. (2017). Measuring the Impact of an Unanticipated Suspension of Ride-Sourcing in Austin, Texas.

Hanson, S., & Giuliano, G. (2004). *The Geography of Urban Transportation*. Guilford Press.

Hughes, R., MacKenzie, D. (2016). Transportation network company wait times in Greater Seattle, and relationship to socioeconomic indicators. *Journal of Transport Geography*, *56*(Supplement C), 36–44. https://doi.org/10.1016/j.jtrangeo.2016.08.014

Hoffmann, K., Ipeirotis, P., & Sundararajan, A. (2016). Ridesharing and the Use of Public Transportation. *ICIS 2016 Proceedings*. Retrieved from http://aisel.aisnet.org/icis2016/DataScience/Presentations/14

Howard, D., & Dai, D. (2014). Public perceptions of self-driving cars: The case of Berkeley, California. *Transportation Research Board 93rd Annual Meeting*, *14*(4502).

Hu, S., Chen, P., Lin, H., Xie, C., & Chen, X. (2018). Promoting carsharing attractiveness and efficiency: An exploratory analysis. *Transportation Research Part D: Transport and Environment*, *65*, 229–243.

King, D. A., J. R. Peters, and M. W. Daus. (2012). Taxicabs for Improved Urban Mobility: Are We Missing an Opportunity? Presented at the Transportation Research Board 91st Annual Meeting, 2012.

Legislative Affairs Office of the State Council of China. (2015). Temporary regulation for app-based, on-demand ride-sourcing. Retrieved April 13, 2017, from http://www.chinalaw.gov.cn/article/cazjgg/201510/20151000479202.shtml

Li, Z., Hong, Y., & Zhang, Z. (2016). *An Empirical Analysis of On-Demand Ride Sharing and Traffic Congestion* (SSRN Scholarly Paper No. ID 2843301). Retrieved from Social Science Research Network website: https://papers.ssrn.com/abstract=2843301

Ministry of Transport of People's Republic of China. (2015). Reform of taxi industry. Retrieved April 13, 2017, from http://zizhan.mot.gov.cn/zhuantizhuanlan/gonglujiaotong/chuzuqchygg/gefangguandian/201507/t20150723_1853535.html

Pew Research Center. (2018). Mobile Fact Sheet. Retrieved March 6, 2019, from http://www.pewinternet.org/fact-sheet/mobile/

Pucher, J., Renne, J. L. (2003). Socioeconomics Of Urban Travel: Evidence From The 2001 NHTS. *Transportation Quarterly*, *57*(3). Retrieved from https://trid.trb.org/view.aspx?id=662423

Rayle, L., Dai, D., Chan, N., Cervero, R., Shaheen, S. (2016). Just a better taxi? A survey-based comparison of taxis, transit, and ridesourcing services in San Francisco. *Transport Policy*, *45*, 168–178. https://doi.org/10.1016/j.tranpol.2015.10.004

Renne, J. L., Bennett, P. (2014). Socioeconomics of Urban Travel: Evidence from the 2009 National Household Travel Survey with Implications for Sustainability. *World Transport Policy & Practice*, *20*(4). Retrieved from https://trid.trb.org/view/1326943

Sadowsky, N., & Nelson, E. (2017). The Impact of Ride-sourcing Services on Public Transportation Use: A Discontinuity Regression Analysis. *Economics Department Working Paper Series*. Retrieved from https://digitalcommons.bowdoin.edu/econpapers/13

San Francisco County Transportation Authority (SFCTA). (2017). *TNCs Today: A Profile of San Francisco Transportation Network Company Activity*. Retrieved from http://www.sfcta.org/sites/default/files/content/Planning/TNCs/TNCs_Today_112917.pdf

Schaller, B. (2018). The New Automobility: Lyft, Uber and the Future of American Cities. Retrieved July 27, 2018, from http://www.schallerconsult.com/rideservices/automobility.htm

SF Chronicle. (2018, October 16). Uber, Lyft cars clog SF streets, study says. *SF Chronicle*. Retrieved from https://www.sfchronicle.com/business/article/Uber-Lyft-cars-clog-SF-streets-study-says-13309593.php

Shaheen, S., & Chan, N. (2016). Mobility and the Sharing Economy: Potential to Facilitate the First- and Last-Mile Public Transit Connections. *Built Environment*, *42*(4), 573–588. https://doi.org/10.2148/benv.42.4.573

Shaheen, S., Stocker, A., Bhattacharyya, A. (2016). Multimobility and Sharing Economy: Shaping the Future Market Through Policy and Research. *Transportation Research Circular*, (210). Retrieved from https://trid.trb.org/view/1416014

Shanghai Bureau of Statistics. (2015). *Shanghai Statistical Yearbook: 2015*.

Shanghai Municipal Transportation Commission. Shanghai app-based, on demand ride service regulation (2016).

Shanghai Urban-Rural Construction and Traffic Development Academe. (2015). *Main findings from the fifth comprehensive travel sursey in Shanghai.*

Shared-use Mobility Center. (2016). *Shared Mobility and the Transformation of Public Transit*. Prepared for American Public Transportation Association. TCRP J-11/TASK 21. March, 2016.

Smart Device Business Center. (2013). *The 5th survey on smarphone users in China*. DY Holding.

TalkingData. (2014). Report of app-based ride users in 2014.

TalkingData. (2015). Report of app-based ride users in 2015.

Wadud, Z., MacKenzie, D., & Leiby, P. (2016). Help or hindrance? The travel, energy and carbon impacts of highly automated vehicles. *Transportation Research Part A: Policy and Practice*, *86*, 1–18. https://doi.org/10.1016/j.tra.2015.12.001

Wang, M., Mu, L. (2018). Spatial disparities of Uber accessibility: An exploratory analysis in Atlanta, USA. *Computers, Environment and Urban Systems*, *67*(Supplement C), 169–175. https://doi.org/10.1016/j.compenvurbsys.2017.09.003

Xu, R. & Ju, Y. (2016). How app-based, on-demand ride-sourcing service area changed overtime: An exploratory, spatial-temporal analysis on how cost and subsidy affects the trips' origins and destinations in Shanghai. Working paper.