Freewheeling: A Spatial Structural Analysis of the Bike-Sharing Industry*

Fangning (Richard) Yan†

November 2, 2021

Click here to download the most recent version

Abstract

This paper studies the spatial mismatch between consumers and bikes in the dockless bike-sharing industry and an externality exacerbating the problem: when a consumer uses a bike for a low and inflexible price, she both displaces another consumer's usage for a potential higher-value trip, and may ride the bike to unpopular destinations. With a trip-level dataset of a bike-sharing company in Beijing, China, I develop a spatial structural model to estimate the demand for bikes with search frictions and local matchings. Compared to the scenario that consumers always get bikes immediately, I find that local spatial mismatch between consumers and bikes reduces the total usage by 29.95%, or a net loss of 332,979 trips. Counterfactual analyses show that (1) doubling the number of bikes increases the trip volume by 28.46% while halving the number of bikes decreases the trip volume by 46.40%; (2) price-discriminating against short trips by 2% increases the total trip time by 0.22%; and (3) changing the frequency of bike reshuffling does not have a significant impact on the total usage of bikes.

Key Words: spatial structural models, search frictions, regulation, ride-hailing industry, dockless bikes

JEL classification: C33; C81; D22; D83; L90; R12; R41

*I am grateful to my committee member Marc Rysman, Jihye Jeon, Jordi Jaumandreu, and Andrey Fradkin for their invaluable advice and support. Thanks to Kevin Lang, Albert Ma, Michael Manove, Robert A. Margo, Yuhei Miyauchi, Daniele Paserman, Johannes Schmieder, Fred Smith, Shyam Gouri Suresh, and participants at the BU Empirical Microeconomics workshop for their valuable comments. I would also like to thank ofo company and Guangyu Cao, Wei Dai, and Jingye Wang, for providing me with the proprietary data.

†Department of Economics, Boston University. Email: riyan@bu.edu.

1 Introduction

Ride-sharing platforms make matches between demand and supply that are spatially distributed across a city and have become a popular topic for research in empirical industrial organization. In this paper, I study the dockless bike-sharing industry in Beijing, China, in which the spatial mismatch between consumers and bikes cannot be adjusted because of an inflexible price and thus causes inefficiency in bike usage. This inefficiency leads to imperfect matches between consumers and bikes, meaning that the excess demand and supply will be simultaneously present in the market. The spatial mismatch problem is aggravated by an externality: when a consumer uses a bike, she does not consider another consumer's usage for a trip of potential higher value or to a popular destination. This externality hinders the efficiency of the bike-sharing market as it is possible for bikes to be allocated to lower-quality rides or unpopular destinations. I ask what the determinants of the demand for bikes are and how much loss in efficiency is associated with the spatial mismatch between consumers and bikes and the externality across consumers. Furthermore, I study the extent to which we can improve the total usage of bikes with measures such as supplying more bikes and price discriminating against low-value trips.

The spatial mismatch between consumers and bikes exists both inside a location and across locations. Inside a location, the parking spots of unmatched bikes did not perfectly accommodate the distribution of consumers, so consumers did not get a bike immediately when they wanted one. Across locations, different numbers of trips were demanded in different locations and at different times, so some areas were filled with unmatched bikes while other areas were almost empty. Furthermore, there was also an externality associated with the inflexible price. There was no arbitrage for consumers with different willingness-to-pay, so when a consumer used a bike, she took the bike away from another consumer who would demand a higher-value trip. Consumers who failed to match with bikes would leave the market, leading to efficiency loss.

In this paper, I develop a structural model to estimate the spatial mismatch and externality of the demand for bikes and provide efficiency-improving measures to enhance the total usage of bikes. With the trip-level data and other data of city characteristics, I discretize the six metropolitan areas of Beijing by census tracts, utilize a Poisson process to simulate the static arrivals of consumers, and formulate a local searching and matching mechanism to incorporate the spatial mismatch and externality. To simulate the geographic transition processes of bikes, I assume that consumers exogenously arrive with destinations in their minds, simultaneously search for bikes, and randomly match with bikes. The matched consumers ride bikes and leave the market, and the unmatched consumers leave the market directly. The unmatched bikes stay put at their origin locations.

For the empirics, I estimate the model with the Simulated Method of Moments (SMM). One empirical challenge is the limited data. Because the trip-level data only reveal matches already made between bikes and consumers as a consequence of searching, I cannot observe all the potential demand for bikes. To reliably estimate the matching and the bike transition processes, I form the moment conditions based on the simulated number of matches between each origin-destination pair at each period, aggregate them over periods and destinations to obtain the average matches at each origin, and then predict them with the average number of matches observed in the data. Another empirical challenge is the computational burden of big data. In each period, the demand is simulated at hundreds of locations, and thus the trips are simulated with tens of thousands of origin-destination pairs, making the computation a prolonged effort.

Model estimation shows that 953 thousand trips accumulate 7.41 million minutes of bike riding per day in Beijing. The spatial mismatch and externality introduce both excess demand and supply, reducing the number of trips by 29.95%, equivalent to a net loss of 330 thousand trips from the perfect condition. Splitting the sample by workdays and off-days and estimating the model respectively, I find that bikes are used more and slightly longer on the workdays. Using the model estimation as a benchmark, I discuss the effects of measures to improve the total usage of bikes with counterfactual simulations. I first check the effectiveness of supplying more bikes to the market. I find that doubling the number of bikes increases the trip volume by 28.46% while halving the number of bikes decreases the trip volume by 46.4%. Then I price discriminate against short trips by concentrating the demand for long trips and discouraging the demand for short trips by the same amount. I find that price-discriminating against short trips by 2% only increases the total trip volume by 0.21% and total trip time by 0.22%. I also price discriminate against unpopular destinations by discouraging trips to them and concentrating trips to popular ones, increasing the total trip time by 3.71% and total trip volume by 3.73%. Overall, the spatial mismatch is costly, but it is challenging to address with simple price changes. I also change the frequency of bike reshuffling and find it does not significantly impact the total usage of bikes.

Related Literature My research adds to the search and matching literature with empirical studies. My matching mechanism is based on the model proposed by Buchholz (2017), who modifies earlier models in Lagos (2000) and Lagos (2003) with non-stationary and stochastic demand and a heterogeneous matching process to study the matching and search frictions of the New York taxi industry. I draw elements from the Buchholz (2017) model, but make several adaptations to reflect the searching and matching process of bike-sharing. One of the most prominent features differentiating the bike-sharing industry from the ride-sharing industry is that there are no drivers to move each bike around. Instead of a two-sided equilibrium in which both consumers and suppliers search for matches, the bike-sharing market is one-sided with consumers searching for bikes in a small area. I add spatial restrictions to the searching process and randomness to the matching process. I aggregate the stochastic markets via a static and localized market-clearing process to an aggregate matching function, using the urn-ball specification as in Hall (1979). Mortensen (1984), Mortensen and Pissarides (1999), and Rogerson, Shimer, and Wright (2004) discuss implementation of aggregate matching functions in labor-search literature. Brancaccio, Kalouptsidi, and Papageorgiou (2020a) and Brancaccio, Kalouptsidi, and Papageorgiou (2020b) study the matching functions in spatial settings with implementations in the bulk shipping industry. Gavazza (2011) and Allen, Clark, and Houde (2019) study search frictions and the resulting distortions in different industries. I model the friction as an efficiency loss of spatial mismatch in the localized matching process, and show that the distortions can be reduced with less spatial mismatch.

The bike-sharing industry and, in fact, the ride-hailing industry as a whole, is a topic of increasing popularity in the industrial organization literature. Buchholz (2017) and Fréchette, Lizzeri, and Salz (2019) study the pricing of the taxi industry and welfare of the spatial allocation. Castillo (2020) and Castillo, Knoepfle, and Weyl (2021) study the benefit of dynamic pricing for ride-hailing services. Camerer et al. (1997), Farber (2008), and Crawford and Meng (2011) study the taxi driver's labor supply choices by determining the labor-leisure trade-off factors. Buchholz, Shum, and Xu (2021) estimate a dynamic labor supply model of taxi drivers to correct for a behavioral bias in a static setting. There are theoretical and empirical studies offering mixed evidence concerning the regulations of the ride-hailing industry. On the one hand, Häckner and Nyberg (1992) suggest that the cost structure of the ride-hailing industry is sufficient to support competition. On the other hand, studies show that regulation is necessary for purposes such as reducing transaction costs (Gallick and Sisk (1987)), supplying sufficient vacant cabs (Flath (2006)), and preventing monopolies (Cairns and Liston-Heyes (1996)). Buchholz (2017) also provides evidence for regulation in a dynamic setting. The discussion of taxi industry regulation provides a reference for the policy implications of the bike-sharing industry. As the bike-sharing industry keeps growing, welfare adjustment and competition regulation will be necessary. I contribute to the literature by estimating the utility changes under different specifications and proposing a measure of negative externality for efficiency loss.

The study of the bike-sharing industry also fits in the literature of two-sided markets and platforms. According to Rysman (2009), a platform serves as an intermediary that makes matches between two sides of the market. Rochet and Tirole (2003) and Weyl (2010) study the price and cost structures, and many theories, such as Cao et al. (2018), Armstrong (2006), and Bryan and Gans (2019), study the competition between platforms. Dinerstein et al. (2018) utilize eBay data to study the trade-off between consumer search and price competition of platform design, and Fradkin (2017) studies the searching behavior and transaction costs of both hosts and guests on Airbnb. The differences that set the dockless bike-sharing platform apart from others that have been studied in the literature are two-fold. On the one hand, dockless bike-sharing platforms differ from docked ones as the bikes are left wherever the previous user chooses. Consumers need to search for bikes themselves and encounter higher search friction in some areas, while more than enough bikes are in other areas. On the other hand, dockless bike-sharing platforms provide bikes directly to consumers and are considered one-sided platforms. My study on the bike-sharing industry will provide a useful reference when autonomous driving technology motivates ride-sharing platforms to move towards a similar one-sided network. In this sense, what I observe in bike-sharing today may have important implications for the future transportation market.

Finally, the bike-sharing industry draws attention from urban planning and transportation literature. On the one hand, technologies have made bike-sharing a welcoming means of zero-emission transportation. Glaeser, Kominers, et al. (2015) discuss how to use big data sources to enable more detailed modeling and improve the study and function of cities. As explained in Demaio (2003), Demaio and Gifford (2004), Demaio (2009), and L. Zhang et al. (2015), the bike-sharing industry has gone through a few generations, driven mainly by technological advancements such as electronic locking stations, GPS-tracking devices, and mobile phone access. On the other hand, figuring out the efficient ways of using bike-sharing as an infrastructure poses a new challenge to urban planners. Glaeser, Ponzetto, and Zou (2016) study how the local amenities, along with other factors, determine the scale of infrastructure and urban networks. Ashraf, Glaeser, and Ponzetto (2016) present a model that provides institutional incentives to connect sanitation infrastructure with its final users. Specifically, for the bike-sharing industry, He et al. (2021) study the consumer preference for docked bikes in the London bike-sharing market. They demonstrate that the locations of existing stations are far from ideal. Shen, X. Zhang, and Zhao (2018) study the usage of dockless bike-sharing in Singapore with an autoregressive model. They find that a large bike fleet promotes high usage with diminishing marginal effect. Kabra, Belavina, and Girotra (2018) find that both bike availability and dock station accessibility determine consumer's willingness to ride. O'mahony and Shmoys (2015) discuss the AI-based solution to the bike rebalancing problem. Cao et al. (2018) discuss the demand for dockless bike-sharing services in China. They study the "floating" nature of the dockless bikes and find that the market expansion effect from competition increases the demand. Pan et al. (2018) provide algorithms that could reduce the imbalance between demand and supply of the dockless bike industry. My paper shows that such imbalance is inherent to the stochastic distribution of demand and supply and fluctuates with different times of the day.

The rest of the paper is organized as follows. In section 2, I provide an overview of the bike-sharing industry and detail the industry characteristics about distribution of bikes and local matching. I also present data collection and sample construction details. Section 3 outlines a spatial structural model of local searching, matching, and bike transition. In section 4, I discuss my empirical strategy for computing spatial distribution and estimating model parameters. I present my results in section 5 and counterfactual analyses in section 6. Section 7 concludes.

2 Data and Sample Construction

In this section, I provide details of my data in four parts. First, I provide a brief background of the ofo company and its bike-sharing service. I also discuss some notable features that I find with ofo's data. Second, I present details of my data preparation and provide summary statistics. Third, I discretize the time and space of my sample. Fourth, I show the reduced form evidence of spatial and temporal distribution of bikes.

2.1 Market Overview

ofo[1] was founded in 2015 and was the first firm in Beijing, China, to provide dockless bike-sharing services. Dedicated to solving the "last-mile" transportation problem widespread in large cities, ofo provided GPS-tracked dockless bikes at *a very low price*. It first started as a two-sided platform on Peking University campus for students to share privately-owned bikes. This initial model of sharing improved the convenience of travel throughout the large campus and became popular soon after its launch. With the early success and the funds raised from venture capital, ofo decided to supply GPS-tracked bikes and expand its service to the entire metropolitan area of Beijing. The bikes ofo delivered to the streets of Beijing were dockless, so they were not fixed to the specific origin and destination locations. Consumers could unlock the bike via ofo's smartphone app wherever and whenever they found it and drop it off once they finished their trips. The payments were made after the bike was locked. The price for using the ofo bike was extremely low in 2017, often less than 2 RMB (about 0.3 USD) per half hour. Customers also had the choice to buy a monthly pass for 20 RMB (about 3 USD) and get unlimited access to bikes. From ofo's transactional data, I observe two distinct surges of bike usage, one at 7-9 am and the other at 5-7 pm throughout each workday. This pattern indicates that bike-sharing is primarily used for daily commutes. I also observe that most trips are finished within 15 minutes, suggesting that bike-sharing is generally used to solve the "last-mile" problem between a transportation hub and a final destination.

The great availability of ofo bikes at high-traffic areas such as schools and subway stations and the extremely low price were the two main contributors to ofo's exponential growth. With a market share of more than 70%, ofo was by far the largest bike-sharing company in Beijing in May 2017. The competition effect and various other issues such as multi-homing that would have been concerns are trivial in this setting. The extremely low price for using ofo bikes also negates the incentives of price shopping. It is beneficial to my research because I can obtain a relatively precise identification and simulation of the matching and distribution of bikes.

The fast expansion was riddled with problems. The availability of ofo bikes was the most critical factor for business growth and, in order to increase the availability, ofo oversaturated the market by providing more bikes than needed. From ofo's

[1] As mentioned in Cao et al. (2018), "ofo" is the trademark of the firm, symbolizing a person riding a bicycle. Following the company's own usage, I do not capitalize the first letter even if a sentence starts with "ofo."

transactional data, I observe that there were always unmatched bikes, implying that many bikes were unused and a significant waste of resources. Moreover, ofo utilized a "consumption-as-supply" mechanism: a user who rides from location A to location B makes the bike available for the following consumer at location B. While easy to execute, this mechanism had an apparent drawback. Consumers do not account for the effect of their ride on other consumers' access to bicycles, and letting the bikes freewheel aggravates the spatial mismatch between consumers and bikes. The inflexible price and the decentralized dockless system prevented ofo from accommodating bikes to the spatial and temporal demand changes in different locations and at different times.

The exigency of regulation on innovative industries is demonstrated by ofo's meteoric rise and catastrophic failure in less than three years. As mentioned above, ofo started its expansion out of Peking University campus in November 2016. At that time, the entire bike-sharing industry only had a few firms. The growth of the industry was staggering: in the following six months, over 80 bike-sharing firms were operating in more than seventy cities in China in April 2017. The total number of bikes deployed increased from 2 million in November 2016 to over 23 million in February 2018, with the largest monthly growth of 634%. The monthly active users increased from 2.8 million people in November 2016 to 50 million in 2017, with the largest monthly growth of 1140%. According to the information that Yan (2021) collects from ofo's website[2], ofo raised about 1.5 billion USD from investors and was valued at 3 billion in July 2017.[3] The sole focus on the fast expansion of the bike fleet for extreme growth led to poor maintenance and vandalism, which created fast capital depreciation and a cash flow crisis. As a result, many bikes were parked carelessly on the streets and they were often inoperable.[4] Many firms soon went out of business in 2018 and 2019, ofo was one of them, leaving behind a huge and derelict bike fleet. The sharp rise and fall, or the "shakeout", of the bike-sharing industry, wasted many resources. Early Regulations such as quota on bikes and chartered entry will help the bike-sharing industry plan for growth more efficiently.

[2]`www.ofo.com`. This website may be defunct as ofo is now completely out of business.

[3]Table C.1 provides the timeline of ofo's funding and valuation.

[4]According to an ofo executive, the highest amortization rate was once above 40%. Refer to Figure 14 in the Appendix for a demonstration of such problems.

2.2 Data

The dataset is comprised of two sets. The first set contains proprietary data of bike trips in Beijing in May 2017 provided by ofo. The second set contains city attributes data for control variables. This latter set is in turn comprised of (1) weather records of daily temperature and precipitation in Beijing in May 2017, collected from the China Meteorological Data Service Center; (2) population density data from the China Statistical Yearbook 2018, which reports the statistics of Beijing in 2017; (3) and locations of subway stations in Beijing. I have chosen Beijing as my focus of study for the following reasons:

1. Large Dataset: As a mega-city in China, Beijing is home to 22 million people with massive daily transportation needs. I observe millions of bike trips in the data.

2. Short Trips: Beijing has a sophisticated public transportation network comprising 24 subway lines and hundreds of subway stations. People often have less than a mile to their destination when they disembark from public transportation but would still prefer not to walk. As a result, bike-sharing has become a popular solution to the "last-mile" problem and provides an ideal setting for analysis of short-term utility changes.

3. Conducive Terrain and Climate: Most of Beijing is flat, so it is bike-friendly. The weather during the data sampling period is mild and likely to encourage bike usage.

4. Free of Competition and Regulation: ofo started in Beijing in 2015 and was the largest bike-sharing service provider in Beijing in May 2017. With over 70% of the Beijing bike-sharing market in 2017, ofo did not have significant competition during the data sampling period. There was also little regulation in place. Lack of regulation and competition permits me to treat the bike-sharing service as a perfect new industry for my research.

5. Observable Regional Trends: Beijing has distinct regions serving different purposes. Facilities such as residential areas and business buildings are each concentrated in different regions and form their own clusters. This pattern provides a reliable reference when I consider different consumer behaviors at different times of the day.

The following two subsections describe the datasets.

2.2.1 Trip-Level Data

The trip-level dataset contains starting and ending geo-locations and the time of each trip. Because the complete dataset is so large, I randomly select 20% of all trip observations made from May 15, 2017 to May 30, 2017 in Beijing. This period includes 11 weekdays (May 15 - May 19, May 22 - May 27), two weekend days (May 20 and May 21), and three national holidays (May 28 - May 30).[5] The data set covers most travel patterns that a transportation firm may face in business, enabling comparative studies of trips on different types of days.

Within the dataset, each observation consists of a unique trip ID, the associated bike ID, and the corresponding user ID. The starting and ending locations are recorded in GPS coordinates. The starting and ending times are recorded as exact unlocking and locking times. The unmatched bikes in each period are identifiable because they have identical starting and ending GPS coordinates and only one-time stamp associated with them.

Although I can count the number of bikes in the data by their ID, I do not have a reliable observation of bikes for the following two reasons. First, the GPS tracker on ofo bikes was not very precise. According to an ofo executive, an error margin below 200 meters when locating a bike was tolerated. Second, as a measure of efficient data storage, ofo only counted the bikes if they had been recently interacted with. Hence, I do not observe a bike if it was available but had been left unattended on the street for a while.

The raw data set contains 57,881,025 observations and the clean version contains 36,787,460 observations. Table 1 provides summary statistics for this dataset. The cleaning routine is provided in Table A.1.

2.2.2 City Attributes Data

The city attributes data is at the sub-district level. These sub-districts, or *Jiedaoban*, are the basic administrative level of municipal government. They are similar in geographical size and functionality to U.S. census tracts that are more widely discussed in the literature. For simplicity, I use the term "census tract", or simply "tract" henceforth, rather than sub-district. I choose the tract-level for three reasons: (1) most city attributes data are collected on the tract-level; (2) the tracts

[5] This period is atypical in that, to accommodate the holiday period, a Monday and a Tuesday (May 29 and 30) were designated days off, and a Saturday (May 27) was made a workday. The holiday was Dragon Boat Festival, and the workday adjustment schedule can be found at `http://www.gov.cn/zhengce/content/2016-12/01/content_5141603.htm` (in Chinese).

Table 1: Bike Trips Summary Statistics by Day

Date	All Trip Time (seconds)			Matched Bikes Trip Time (seconds)					unmatched Bikes
	Obs.	Mean	Std. Dev	Obs.	Mean	Std. Dev	25%ile	75%ile	Obs.
5/15/2017	2,322,767	457.33	836.94	1,177,745	938.56	992.91	358	1,110	1,145,022
5/16/2017	2,309,231	427.25	814.84	1,119,319	921.31	988.21	348	1,093	1,189,912
5/17/2017	2,268,306	342.00	754.40	1,011,521	930.53	1,000.43	354	1,102	1,256,785
5/18/2017	2,131,399	311.83	733.65	906,563	917.94	1,013.93	347	1,086	1,224,836
5/19/2017	2,732,683	332.56	830.16	1,256,687	907.71	1,165.75	320	1,042	1,475,996
5/20/2017	2,707,035	469.56	1,001.93	1,449,824	952.08	1,255.41	323	1,082	1,257,211
5/21/2017	2,809,890	482.48	1,020.96	1,491,338	972.75	1,274.61	324	1,111	1,318,552
5/22/2017	1,534,956	326.61	832.77	569,309	940.61	1,191.52	345	1,109	965,647
5/23/2017	1,990,969	443.16	978.59	985,574	944.90	1,252.11	336	1,106	1,005,395
5/24/2017	1,514,751	449.04	1,027.65	729,305	951.96	1,326.67	332	1,104	785,446
5/25/2017	1,961,563	406.01	978.84	881,324	936.51	1,308.89	336	1,081	1,080,239
5/26/2017	2,805,858	448.97	1,073.28	1,392,306	944.98	1,398.51	326	1,066	1,413,552
5/27/2017	2,498,786	555.72	1,190.09	1,472,903	953.91	1,432.23	318	1,070	1,025,883
5/28/2017	2,094,773	465.26	1,174.88	1,021,613	967.94	1,544.39	298	1,059	1,073,160
5/29/2017	2,548,606	480.11	1,105.11	1,274,525	980.10	1,415.29	310	1,111	1,274,081
5/30/2017	2,555,887	459.23	1,070.37	1,230,001	979.83	1,390.83	316	1,104	1,325,916
All Dates	36,787,460	428.39	974.15	17,969,857	947.82	1,267.79	329	1,089	18,817,603

are small enough that about 58% trips start and end in different tracts, enabling a detailed representation of bike trips; (3) there are not too many tracts to make computation impossible. There are 134 census tracts in the six metropolitan areas in Beijing. On average, a tract is about 10-15 square kilometers in size and has a population density of about 100,000 people per km^2.

Weather Data The weather dataset contains the daily precipitation and temperature observations for each of the six metropolitan areas in Beijing. There are 96 observations for the 16 days. I use the weather data for a comparative study of demand for trips on rainy days and non-rainy days. Since the weather does not change much both across the city and throughout the day, I do not use an explanatory variable and instead infer a *rain* variable controlling for the weather conditions. It takes the value of 1 when the precipitation suggests a rainy day and 0 otherwise. According to the table A.4, only May 22, 2017, was a rainy day. I take out the rainy day and report estimation results with the non-rainy sample for a comparative study. The data collection process is described in Appendix A.3 with corresponding summary statistics in table A.4.

Population Density Data The population density dataset contains the population density of each census tract in Beijing. There are 134 observations for each of 134 census tracts in the Beijing metropolitan area. I take the log value of population density of each tract to serve as a measure that affects the demand for bikes. I expect more demand associated with higher population. The data collection process is described in Appendix A.4 with corresponding summary statistics in table A.5.

Subway Stations Data The subway stations dataset contains the geo-locations of Beijing subway stations. I use the subway stations data because the bike-sharing service is predominantly used to connect the last miles of a trip. Subway is a major component of the public transportation network in Beijing, with tens of millions of passengers every day. I expect a positive relationship between the number of subway stations and the demand for bikes. I map the geo-location of each subway station to the map of the Beijing metropolitan area and count the number of stations in each tract. I double-count the transfer station as an individual station for each subway line because different lines do not share platforms and are instead connected by tunnels. Passengers need to walk some distance to make a transfer. I also double-count the stations on the borders of two or more tracts, as people in both tracts have the same access to these stations and use the bike-sharing service. That is, if a subway station is located on the border of a tract, it is counted as inside the tract. Hence, the sum of subway stations in each tract in my data does not equal the total subway stations in Beijing. The data collection process is described in Appendix A.5 with corresponding summary statistics in table A.5.

2.3 Discretizing Time and Space

I discretize the time within each day and the space across Beijing as follows. Time is divided into 30-minute intervals. This is a data-driven discretization that finds the balance point. That is to say, on the one hand, I would like the time period to be short enough such that most bikes travel only once during the period. On the other hand, longer periods allow most trips to start and end in the same period. Besides, the discretization should also allow for efficient computation. In Table A.3 I present the percent of bikes that are used once and the maximum number of times a single bike is re-used in different time periods. During a 30-minute interval, on average, 90% of all bikes are used only once, or not at all.

I discretize space according to the census tracts of the metropolitan area of Beijing. There are a total of 134 census tracts spanning the six metropolitan areas. The 134 tracts are numbered in alphabetical order as shown in table A.2. I drop the trips that do not start or end inside the metropolitan areas, which make about 8% of all observations. I do not observe trips between tracts that are far away. The sample space I use is 134×134, with origin tracts as rows and destination tracts as columns, for each time period. For each day divided into forty-eight 30-minute periods, the sample space has the size of $134 \times 134 \times 48$.

The distance is the length of trip time, measured in seconds, expected to ride a bike between tracts. I calculate the geometric centroid of each tract and measure the geographical distance between each pair of the centroids. I then divide the distance by the average bike-riding speed (10 mph) to get the travel time. For bike trips inside a tract, I take the average time of within-tract trips observed in the data.

2.4 Reduced Form Evidence

In this subsection I illustrate the market inefficiency due to spatial mismatch. Within-tract spatial mismatch is illustrated in Figure 1 as there are always bikes left unmatched in Beijing in each period. Cross-tract spatial mismatch occurs when consumers cannot find unmatched bikes in their tract, even if there are many unmatched bikes in other tracts in the same period. Figure 3 shows the distribution of bike trips on May 15, 2017. A period-by-period presentation is shown in Figure 2. There are both substantial intertemporal and cross-sectional variations in the number of trips. The busiest tracts are located in the northeast, southeast, and southwest. Tracts in the northwest have large intertemporal variations. Tracts in the city center are often not very busy as they are predominantly central government buildings. These buildings make up a huge government plaza, and most officials use cars. Tracts in the city outskirts are also not very busy because the subway is less accessible, and presumably, it is uneconomical to ride for long distances.

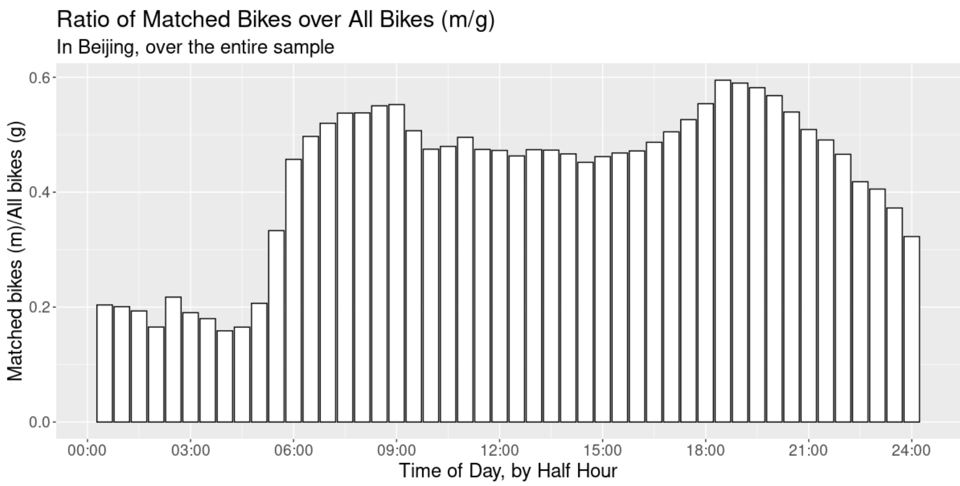
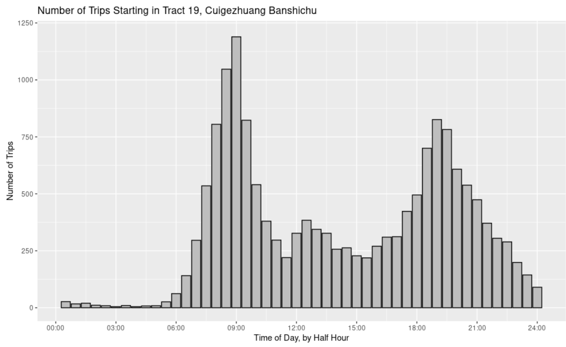
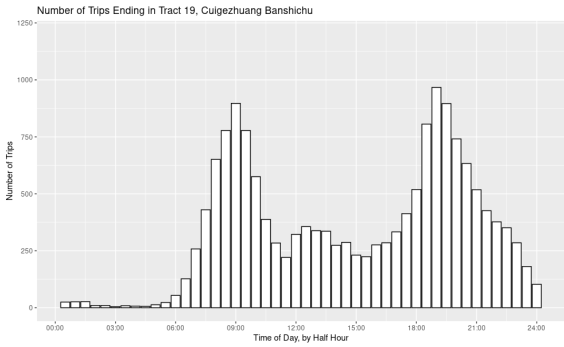


Figure 1: Ratio of the Matched Bikes to the Total Number of Bikes

Figure 2: Daily Variation in Trip Starts and Trip Ends for Three Districts

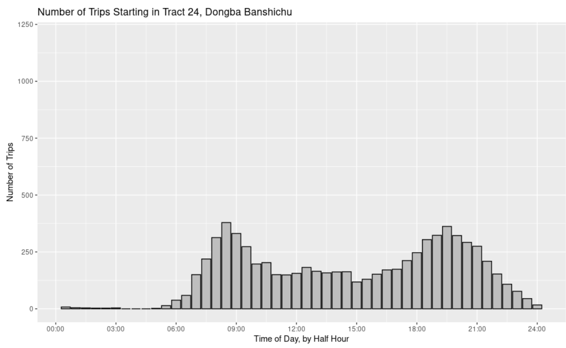


(1) Number of Trips Starting in the Tract

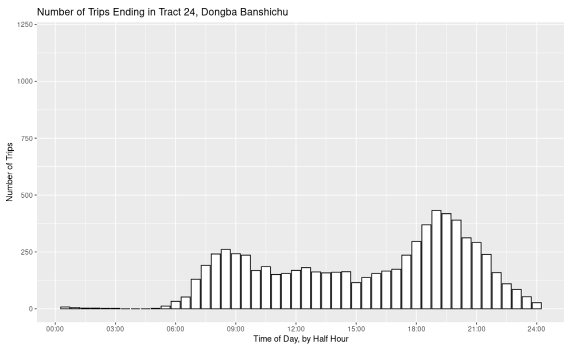


(2) Number of Trips Ending in the Tract

(a) One of the Busiest Tracts: Tract 19 Cuigezhuang Banshichu

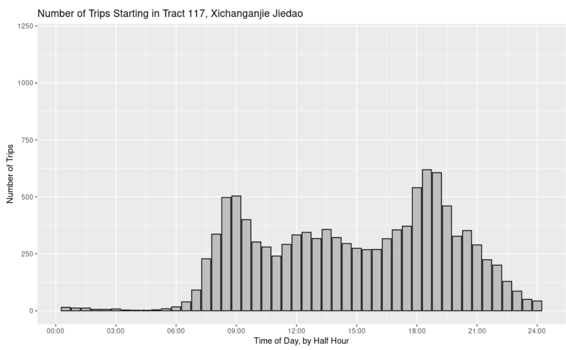


(1) Number of Trips Starting in the Tract

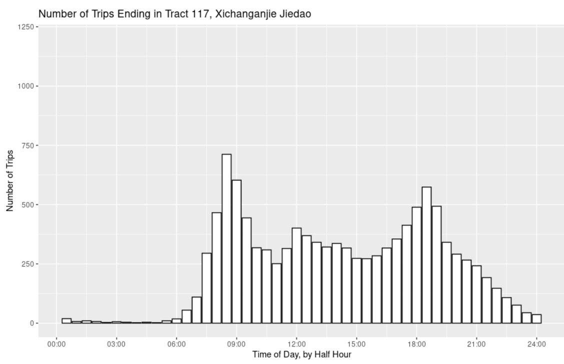


(2) Number of Trips Ending in the Tract

(b) One of the Least Busy Tracts: Tract 24 Dongba Banshichu



(1) Number of Trips Starting in the Tract



(2) Number of Trips Ending in the Tract

(c) An Average Tract: Tract 117 Xichanganjie Jiedao

Notes: While each tract generally shows the same pattern of high activity at morning and evening rush hours, as well as a slight increase in activity at midday, the overall number of trips varies considerably from tract to tract.

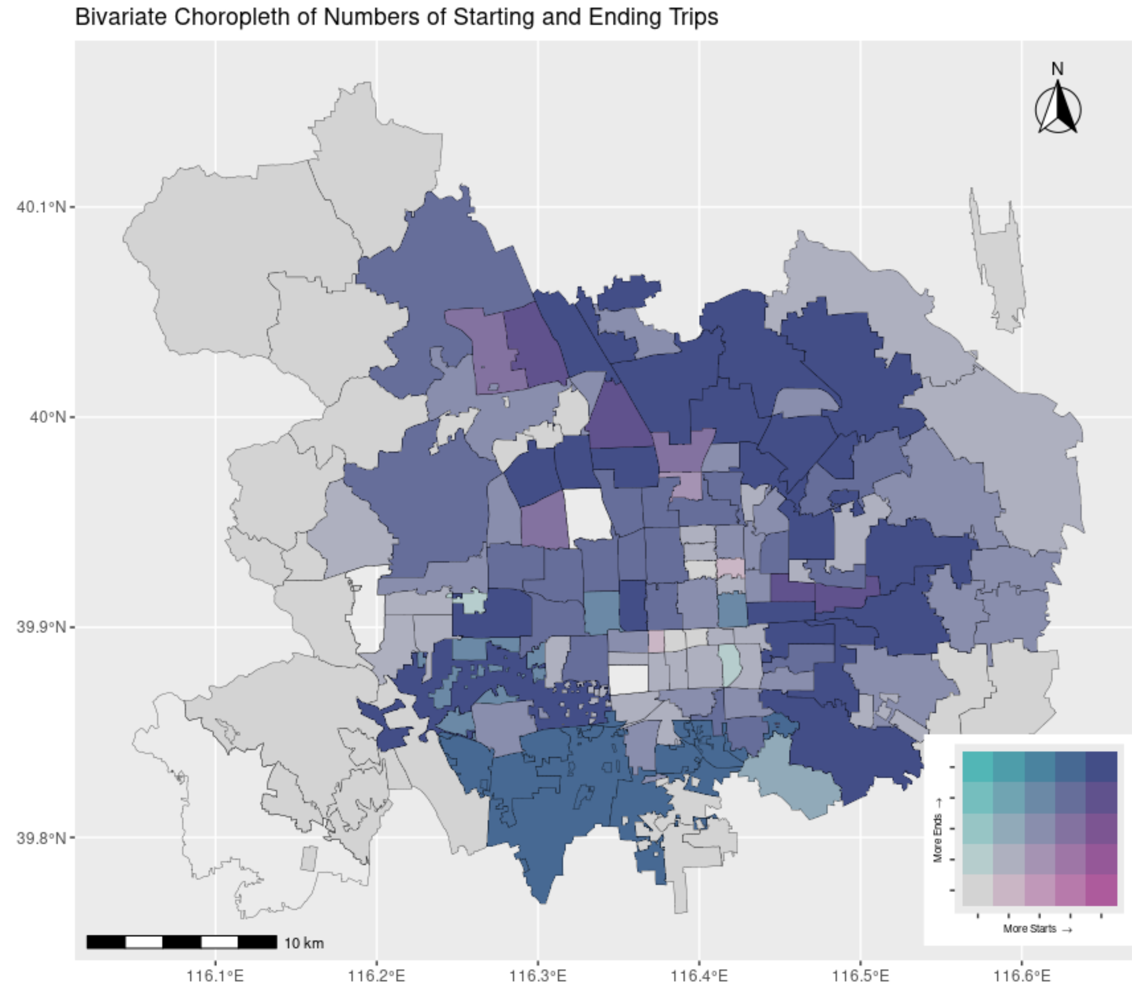


Figure 3: Bivariate Choropleth of Number of Trips in Census Tracts

Notes: Each of the outlined sections of Beijing is one of the 134 locations indexed by i when serving as starting locations, and indexed by j when serving as ending locations. As the legend shows, the color on the horizontal axis darkens with more trip starts; the color on the vertical axis darkens with more trip ends.

The uneven distribution of bike trips is demonstrated by regions with stark different colors. There are three regions with dark colors: the northeast, the southeast, and the southwest. The northeast and the southwest regions have large residential areas and thus have high demand. The southeast region has large business districts. The city center has light colors. Government buildings and historical sites make up the majority of the area; hence the demand for bikes is low. The city outskirts also have light colors. The low number of trips in these regions is mainly attributable to lower population density and less access to public transportation.

There is an animated version of this choropleth to show the period-by-period fluctuations of bikes trips in all tracts in Beijing. It is available on my website at `https://www.yanfn.com/research`.

3 Model

I build a spatial structural model to explain the bike searching and matching mechanism and estimates the demand for bikes. The city of Beijing is defined as a network of 134 tracts, with i denoting the origin location and j denoting the destination location: $i, j \in \{1, \ldots, 134\}$. Time within a day is divided into discrete intervals with finite horizon $t \in \{1, \ldots, 48\}$. At period $t = 1$ the day starts and at $t = 48$ it ends. The measure of bike distribution is $g_i^t \in \mathbb{N}$, denoting the number of bikes in location i and at period t. The total number of bikes at any period is then $g^t = \sum_i g_i^t$. The measure of consumers wishing to use the bike-sharing service is $n_i^t \in \mathbb{N}$, which means that there are n_i^t consumers looking for bikes in location i and at period t. These n_i^t consumers are assumed to arrive according to a Poisson distribution with parameter λ_i^t. The consumers have a probability of getting matched with bikes. They take the bike if they get matched and leave the market if not. The matched bikes transit to the destination locations and unmatched bikes are idle at the origin locations.

I discuss the model in four subsections. First, I provide measuring metrics of time of space. Second, I present the demand and supply of bikes. I use an exogenous demand system in which consumers search myopically for bikes and discuss assumptions that rationalize the Poisson distribution. I utilize a supply system in which the total profit for ofo directly depends on the total time of bike usage. Third, I introduce a matching mechanism between consumers and bikes. The number of matches between consumers and bikes is a random draw from a binomial distribution, depending on demand, supply, and market efficiency. Fourth, I demonstrate the transition of bikes according to the matching system and present a measure of trip time. These four ingredients are presented in more detail below.

3.1 Demand and Supply

One of the challenges of this paper is that I do not directly observe the demand for bikes. I observe the number of trips as a result of the matching between demand and supply of bikes in tracts. I make distribution assumptions of demand and assume that the firm does not adjust the bike supply in the middle of a day. In this subsection, I discuss the details of demand and supply.

Demand For each location i at period t, the measure of consumers searching for bikes n_i^t is drawn from a Poisson distribution with parameter λ_i^t. The parameter

λ_i^t is a sum of Poisson parameters λ_{ij}^t, namely $\lambda_i^t = \sum_j \lambda_{ij}^t$. The parameter λ_{ij}^t is the destination location-j specific Poisson arrival process of consumers in origin location i at time t. That is, each consumer comes to location i with her destination location j in mind. I assume that consumers arrive exogenously, and begin search immediately.

Assumption 3.1. *In each location i, all consumers exogenously and myopically search for bikes.*

This is a practical assumption as the bike-sharing service is mainly used to cover a short distance that can be finished in a short time. Based on this assumption, I make two implications. *First*, the demand is exogenous: every consumer wishing to ride a bike will search for bikes regardless of the probability of matching. *Second*, every consumer searches in only one origin tract and one period. Consumer's utility beyond the current period is zero. This means that the model is static. I choose to make this assumption to focus on the distribution mechanism of bikes. Relaxing the first implication requires a discrete choice model with a complicated fixed-point algorithm solving for utility in every origin tract i and period t. Relaxing the second implication requires relaxation of the first implication and value function iteration for every period and location. They make computation nearly impossible and are tangential to the main focus of this paper.

Supply Let there be a total of g_i^t bikes in location i at period t. I assume that the firm does not supply new bikes or re-shuffle bikes in the middle of a day. The operation cost on the firm side is assumed to be exogenous since I will not focus on firm operations and the operation cost is unobservable. The price of riding a bike is almost zero during the sampling period. The profit is closely related to bike usage, so the total time length of bike riding is the measurement of profit. I also use the total time as the benchmark of total utility for the counterfactual analysis, and discuss the changes of total time with different tract characteristics and policy interventions. More details are discussed in Section 6.

3.2 Matching

At the start of every period t, the n_i^t consumers arriving exogenously and searching myopically over g_i^t bikes in tract i. The g_i^t unmatched bikes are matched randomly among all the consumers, with probability $p_i^t(g_i^t, n_i^t, \alpha)$. The aggregate expected matches $m_i^t(g_i^t, n_i^t, \alpha)$ is drawn from a binomial distribution $B(g_i^t, p_i^t)$. The parameter α is the search efficiency coefficient. The matched consumers ride bikes to their respective destinations. The unmatched consumers leave the market. The unmatched bikes will remain unmatched and stay put. Figure 4 illustrates the within-period search and matching process. The rest parts of this subsection discuss the search efficiency and matching in detail.

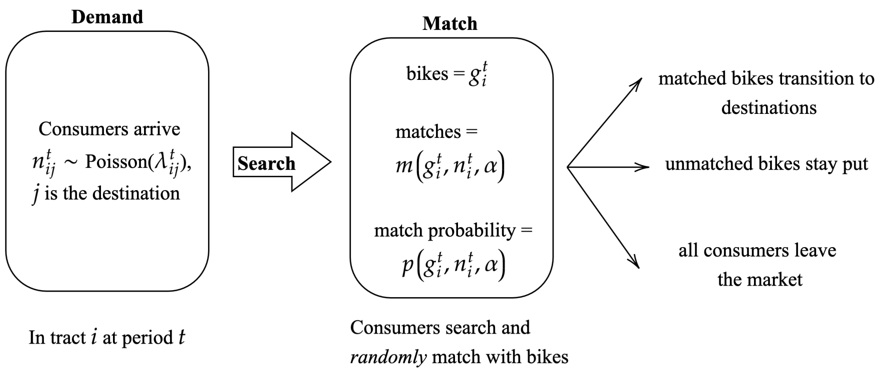


Figure 4: Flow of Demand, Searching, and Matching at Tract i and Period t.

Matching The number of expected matches is given by m_i^t. To derive the matching probability is p_i^t, I follow the solution formulated in Butters (1977) and Hall (1979) for the discrete urn-ball problem. The probability for a bike to be matched to each consumer is

$$p_i^t = p(g_i^t, n_i^t, \alpha) = 1 - e^{-\frac{n_i^t}{\alpha g_i^t}}, \tag{1}$$

where the coefficient α accounts for the market inefficiency due to spatial mismatch between the demand and supply of bikes.

The matching function $m_i^t = m(g_i^t, n_i^t, \alpha)$ is the total number of matchings that happens in location i in period t. The matchings m_i^t is distributed according to the probability density function of a binomial distribution $B(g_i^t, p(g_i^t, n_i^t, \alpha))$:

$$m(g_i^t, n_i^t, \alpha) \sim B(g_i^t, p(g_i^t, n_i^t, \alpha)) = \binom{g_i^t}{m_i^t} (p_i^t)^{m_i^t} (1 - p_i^t)^{g_i^t - m_i^t}. \tag{2}$$

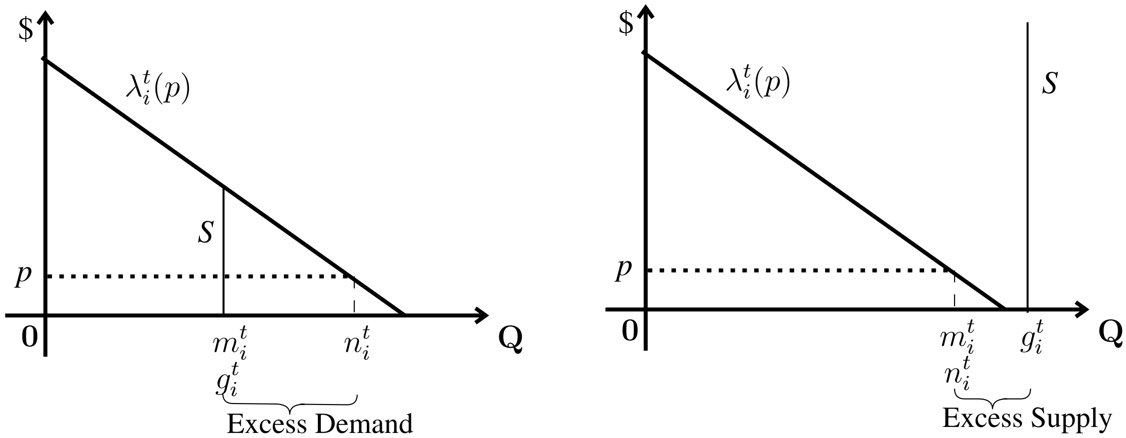
To clarify equation (2), note that $m(g_i^t, n_i^t, \alpha)$ on the left-hand side is the number of matchings in location i in period t. It is a random number of outcomes. The binomial distribution $B(g_i^t, p(g_i^t, n_i^t, \alpha))$ on the right-hand side provides the probability associated with that random outcome. In this way, the matchings m_i^t is always a non-negative integer that fits the discrete setting. Additionally, the number of bikes for one match is $\rho_i^t = \frac{g_i^t}{m_i^t}$.

Spatial Mismatch and Market Inefficiency I denote α in equation (1) as the "market efficiency coefficient" to measure the market inefficiency due to the within-tract spatial mismatch between demand and supply of bikes: within a tract, the parking spots of unmatched bikes did not perfectly accommodate the distribution of consumers. For example, imagine a situation where a consumer gets off the subway and wants to use a bike, but finds that the bikes are on the other side of the street. She may give up using the bike if she finds it costly to cross the road. In this case, a match is not made even though there are both demand and supply. In the following, I generalize this situation to make an economic explanation for α and illustrate how it affects matching mathematically.

In an Arrow-Debreu model with Walrasian demand and perfectly inelastic supply of bikes, along with an extremely low and inflexible price, the matches should always equal the smaller of the consumers and bikes if there is no spatial mismatch. As illustrated in panel (a) in figure 5, when $\alpha = 0$ (probability $p_i^t = 1$), excess demand and supply cannot both be present in the market. However, when $\alpha > 0$ (probability $p_i^t < 1$), there is within-tract spatial mismatch between the demand and supply of bikes and thus both excess demand and supply are present in the market. The market incurs a loss in efficiency with spatial mismatch.

Mathematically, α dictates the ratio $\frac{n(\lambda_i^t)}{\alpha g_i^t}$, which is the power index of the matching probability function in equation (1). The coefficient α decides how many matches m_i^t will be made when either the demand n_i^t or the bikes g_i^t change. A larger α implies a more severe spatial mismatch and a lower matching probability. Figure 6 shows how the probability decreases as α increases.

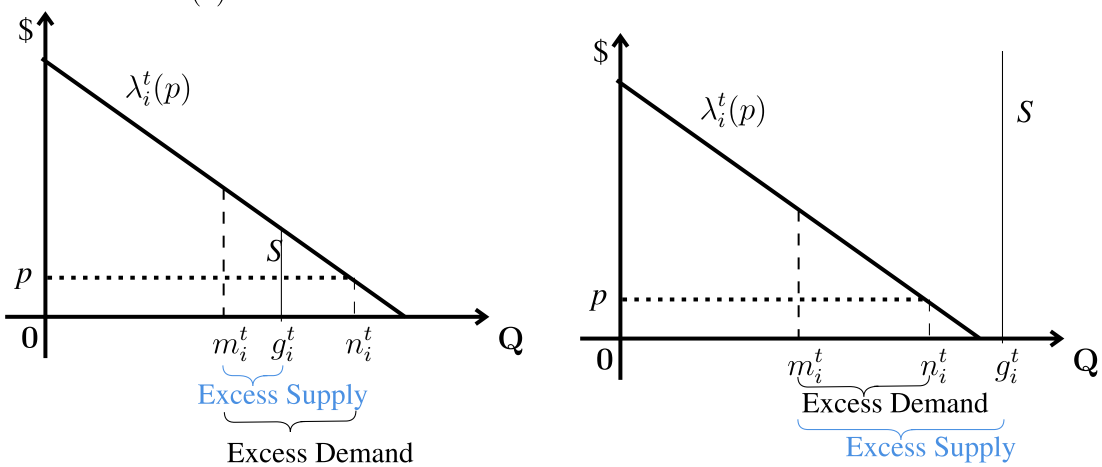
Additionally, in a slight abuse of notation, I connect the economic meaning of α to labor economics literature. The matching efficiency coefficient α combines two commonly seen externalities in the labor search model: the thick market externality (when there are more bikes than consumers) and the congestion externality (when there are more consumers than bikes). Analogous to the Hosios condition that defines the efficient allocation in the labor market, α determines how the matching rate between consumers and bikes changes with the market tightness. One challenge is that I do not directly observe the number of consumers searching for bikes, and thus I cannot measure the market tightness directly. To attenuate the measurement issue of demand, I make assumptions on the parametrization of the demand process.



(1) More Consumers Than Bikes: $n_i^t > g_i^t$

(2) More Bikes Than Consumers: $g_i^t > n_i^t$

(a) Perfect Matches Between Bikes and Consumers: $\alpha = 0$



(1) More Consumers Than Bikes: $n_i^t > g_i^t$

(2) More Bikes Than Consumers: $g_i^t > n_i^t$

(b) Spatial Mismatch of Bikes and Consumers: $\alpha > 0$

Figure 5: Exposition of Excess Demand and Supply

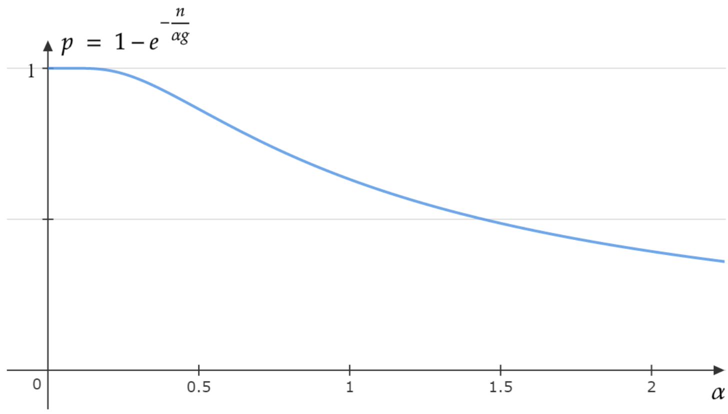


Figure 6: Matching Probability vs. α

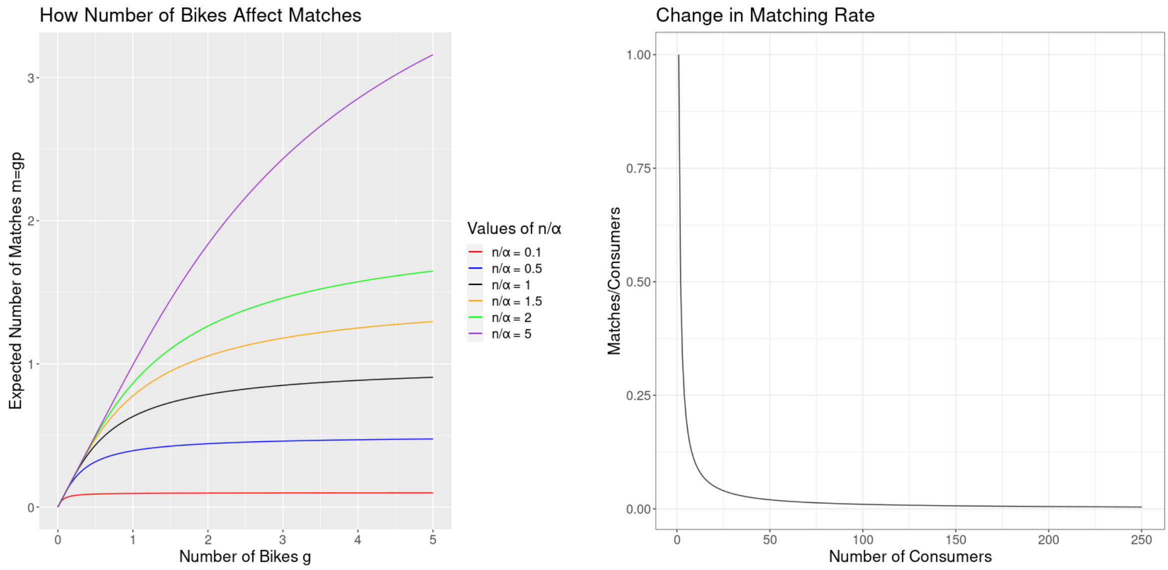
Notes: The figure on the left shows that the matching probability $p(n_i^t, g_i^t, \alpha) \to 1$ as $\alpha \to 0$, and $p(n_i^t, g_i^t, \alpha) \to 0$ as $\alpha \to \infty$. According to the literature, my discussion focuses on the domain $\alpha \in [0, 2]$.

I provide more details in Section 4.2.

Externality The inflexible price prevents the arbitrage for higher-value trips. The consumer exerts a negative externality when riding a bike that could have been used for longer trips. In the following, I demonstrate how my model incorporates the externality by discussing how to match m_i^t changes with n_i^t, g_i^t, and α. According to equations (1) and (2), the match m_i^t increases when demand n_i^t or supply g_i^t increases, and decreases when α increases. First, having more bikes g_i^t increases the number of matches, although the effect may seem unclear upon the first look. In figure 7(a), I show that the number of expected matches increases with more bikes, and simultaneously increasing the demand for bikes boosts the matches further. Second, although increasing the demand for bikes increases the total number of matches, it decreases the matching for each consumer. In figure 7(b), I show that the ratio m_i^t/n_i^t decreases as demand n_i^t grows large. The intuition is that the more consumers are searching for bikes, the less chance for each of them to find a bike. Third, the number of matches m_i^t decreases as α increases, meaning that the market becomes less efficient as the spatial mismatch worsens.

3.3 Bike Distributions

In location i at period t, the matches m_i^t is given by equation (2) as the sum of matches made in location i with different destinations j : $m_i^t = \sum_{j \in L} m_{ij}^t$. As



(a) Matches m_i^t vs. Bikes g_i^t

(b) Match-Demand Ratio m_i^t/n_i^t vs. Demand n_i^t

Figure 7: How Matches m_i^t Changes With Bikes g_i^t And Consumers n_i^t.

Notes: Figure (a) on the left shows the match m_i^t, calculated as the expected number of matches from equation (2), monotonically increases with the number of bikes g_i^t when demand and spatial mismatch are fixed. I take different values of n_i^t/α and plot the curve of the expected match with each of them. All the curves increase with the number of bikes g_i^t, with an increasing marginal rate with a higher n_i^t/α value. It illustrates that more matches are associated with more demand and supply of bikes and less spatial mismatch.

Figure (b) on the right shows the market congestion. For a given number of bikes, the match-demand ratio m_i^t/n_i^t decreases with the demand n_i^t. For this figure, I assume that there are 200 bikes and simulate the matches with the spatial mismatch.

illustrated in Figure 4, a *matched* consumer and his bike will transit to location j, consistent with the Poisson parameter λ_{ij}^t. Only occupied bikes will transit. Unmatched bikes will stay put. The distribution of bikes is given by $\mathbf{g}_e^t$, the matched bikes in period t. The distribution $\mathbf{g}_e^t$ is a 134×1 matrix of the number of matched bikes. The distribution of unmatched bikes is $\mathbf{g}_v^t$, which is also a 134×1 matrix. The matched bikes $\mathbf{g}_e^t$ and the unmatched bikes $\mathbf{g}_v^t$ form the aggregate distribution for bikes at the start of next period:

$$\mathbf{g}^{t+1} = \mathbf{g}_e^t + \mathbf{g}_v^t. \tag{3}$$

Measuring Trip Time The total trip time u_{ij}^t of bikes from location i to j in period t is

$$u_{ij}^t = D_{ij} \times m_{ij}^t.$$

Denote D_{ij} as the distance between locations i and j, measured in the number of seconds the consumer takes to ride from origin to destination. The total trip time of all bikes in one day is then:

$$u = \sum_t \sum_i \sum_j u_{ij}^t. \tag{4}$$

It serves as a benchmark for counterfactual analysis in Section 6.

3.4 Intra-Day Timing

At the start of each day, the market is reset to the initial state.[6] Bikes are exogenously distributed as $\mathbf{g}^1$.The sequence of events is as follows:

1. A number of n_i^1 consumers enter tract i according to Poisson parameter λ_i^1;
2. Consumers are matched according to the probability $p_i^1 = p(n_i^1, g_i^1, \alpha)$ and the matched $m_i^1 = m(g_i^1, n_i^1, \alpha)$ bikes transit to destinations;
3. The unmatched $n_i^1 - m_i^1$ consumers leave the market; the unmatched $g_i^1 - m_i^1$ bikes stay put;
4. Matched bikes arriving in destinations and unemployed bikes stay put, forming new distribution $\mathbf{g}^2$;
5. The process is repeated until $\mathbf{g}^{48}$, the terminal period.

Spatial Structural Model To sum up, the model has the Poisson process of consumers n_i^t, the matching probability p_i^t, and the matching vector of destination choices m_i^t in location i in period t. The fluctuations of arrivals of consumers are captured by the random draw n_i^t. The frictions of consumers searching for a bike in different regions are captured by the matching probability p_i^t. The distribution of matched bikes is captured by m_i^t, which depends on the consumers' trips to different destinations.

4 Empirical Strategy

Three variables of the model are directly identified from the data. The first variable is the distance between different tracts. The second variable contains tract-

[6] I make this assumption to get the daily average of trips between tracts that are used to identify model parameters. As shown in Figure 2, bike trips are very scarce after midnight.

specific characteristics. The third variable is the number of trips between tracts at different time periods. As described in section 2.2.1, I do not have a reliable observation of bikes g_i^t. I assume that there are 200 bikes in each tract at the start of the day. This section shows how I use a set of parameters to simulate the demand and form the moment conditions to identify the model parameters.

4.1 Identifying the Parameters λ_{ij}^t and α

In this subsection, I model λ_{ij}^t as a parameter dependent on the distance, the tract characteristics, and the time effects. With these parameters, and given the number of matches m_{ij}^t, I use the spatial structural model to minimize a set of moment conditions by Simulated Method of Moments (SMM) to recover the Poisson parameter λ_{ij}^t and additionally the market efficiency coefficient α. I present the details as follows.

4.1.1 Estimating λ_{ij}^t

I assume the Poisson coefficient for demand λ_{ij}^t takes the form of

$$\lambda_{ij}^t = \theta^1 \frac{1}{D_{ij}} + \theta_i^2 + \theta_j'^2 + \theta^{3t}. \tag{5}$$

In equation (5), D_{ij} is the distance between geographical centroids of locations i, j, and θ_i^2 and $\theta_j'^2$ are the tract effects at location i, j. Mathematically, $\theta_i^2 = X_i'\theta^2, \theta_j'^2 = X_j'\theta'^2$ where $\theta^2 = [\theta^{21}, \theta^{22}]$ and $\theta'^2 = [\theta'^{21}, \theta'^{22}]$. The variables X_i and X_j include the population density and subway stations measures of tracts i and j, respectively. The third variable θ^{3t} is the time effect that consists of 48 periods of the day. Because setting a time dummy for each of the 48 periods will greatly increase the computation time, I model the time fixed effect as a smoothed curve of a combination of two normal distributions:

$$\theta^{3t} = \theta'^3 \left[\varphi_1 \left(\frac{t - \mu_1}{\sigma_1} \right) + \varphi_2 \left(\frac{t - \mu_2}{\sigma_2} \right) \right]. \tag{6}$$

To identify λ_{ij}^t, I rewrite equations (1) and (2) as

$$m(g_i^t, n_i^t, \alpha) \sim B(g_i^t, n_i^t, \alpha)) = \binom{g_i^t}{m_i^t} (1 - e^{-\frac{n_i^t}{\alpha g_i^t}})^{m_i^t} (1 - e^{-\frac{n_i^t}{\alpha g_i^t}})^{g_i^t - m_i^t}. \tag{7}$$

Given that the demand $n_i^t = \sum_j n_{ij}^t$ and n_{ij}^t is according to the Poisson parameter

λ_{ij}^t, I can match the simulated demand with the trips in each tract i and period t via the simulated method moments (SMM) to estimate λ_{ij}^t and α.

4.1.2 Estimating α

The market efficiency coefficient α is identified by measuring the change in matches m_i^t as a result of the change in bikes g_i^t while holding the demand variable constant. One challenge with the identification of α is that I do not observe the number of consumers searching for bikes. Therefore, I lack exogenous variation that differentiates the market efficiency from other randomnesses such as time- and tract-varying values of bike usage. In the following, I show that although I cannot separate the effects, α is still well-identified after parametrizing the demand. I first present reduced form evidence to provide intuition for identifying α and then discuss the structural estimation strategy.

Reduce Form Identification of α I regress the matches m_i^t on the number of bikes g_i^t and tract characteristics:

$$\ln m_i^t = \gamma_0 + \gamma_1 \ln {g_i}^t + X_i \gamma_2 + \varepsilon_i^t. \tag{8}$$

The intuition is that conditional on demand, I want to see how responsive is matches m_i^t to bikes g_i^t. So I hold demand constant and let bikes g_i^t vary. The variable X_i contains the population density and number of subway stations of tract-i, and ε_i^t is the error term. The coefficient of interest is γ_1 as it captures the responsiveness of m_i^t to g_i^t, and is indicative of the value of market efficiency α. A large γ_1 suggests that a large number of matches m_i^t is related to a large number of bikes g_i^t. This implies a highly efficient market and hence a small α. The same applies to a small γ_1.

	log(match)	(t-stat)
log(bike)	0.427***	(48.66)
log(popden)	0.00737***	(0.47)
stations	0.0176***	(43.64)
_cons	2.631***	(44.83)
N	6432	-

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2: OLS Regression Results

Table 2 above shows the regression results of equation (8). The value of γ_1 is 0.427, meaning that for each 1% increase in the number of bikes, the number of matches increase by 0.427%. This value is relatively small, suggesting the existence of spatial mismatch and a relatively large α.

Structural Identification of α The spatial mismatch coefficient α is identified by fixing α and running the SMM estimation with other parameters. The variation in the objective function identifies α. I provide a plot of the objective function with different values of α in section 5.1 to show that α is well-identified.

4.2 Estimation Details

In the following, I present details of simulation and estimation. First, I present the moment conditions for the SMM estimation. Second, I detail the bike transition simulation algorithm, which recreates the distribution of bikes. Third, I discuss approaches I take during the SMM procedure to reduce the computational burden.

4.2.1 Moment Conditions

With the Poisson demand parameter λ_{ij}^t defined as equations (5) and (6), along with the initial guess for market efficiency coefficient α, the number of matches $\hat{m}_i^t$ in tract i at period t is simulated according to equation (7). The number of observed matches m_{ij}^t is aggregated at the origin tract i to obtain m_i^t. The moment condition is

$$\mathrm{E}[\mathbb{M}(\tilde{\theta})] = 0, \tag{9}$$

where $\tilde{\theta} = \{\lambda, \alpha\}$, and according to equation (5) λ depends on parameters $\{\theta^1, \theta_i^2, \theta_j'^2, \theta'^3, \mu_1, \sigma_1, \mu_2, \sigma_2\}$, so $\tilde{\theta} = \{\theta^1, \theta_i^2, \theta_j'^2, \theta'^3, \mu_1, \sigma_1, \mu_2, \sigma_2, \alpha\}$. Given that there are 134 locations, the sample moment equivalents with each $i \in \{1, \dots, 134\}$ tracts over the average of all periods $t \in \{1, \dots, T\}$ are:

$$\mathbb{M}(\tilde{\theta}) = \begin{bmatrix} \overline{m}_1 - \overline{\hat{m}}(\tilde{\theta}) \\ \overline{m}_2 - \overline{\hat{m}}(\tilde{\theta}) \\ \vdots \\ \overline{m}_{134} - \overline{\hat{m}}(\tilde{\theta}) \end{bmatrix}. \tag{10}$$

The objective function is

$$\hat{\theta} = \arg\min_{\tilde{\theta}\in\Theta} \mathbb{M}(\tilde{\theta})W\mathbb{M}'(\tilde{\theta}). \tag{11}$$

The objective function is a vector of 134 elements that each of them is an average of matches across 48 periods. Appendix B.1 provides details of the Simulated Method of Moments and the construction of the weighting matrix.

4.2.2 Bike Transition Simulation

In the following, I present the simulation algorithm, which takes all model primitives, parameters, and an initial state to recreate the bike transition process. It returns a 134×48 matrix $\hat{\mathbf{M}}$ of matches for each tract $i \in \{1, \ldots, 134\}$ and each time period $t \in \{1, \ldots, 48\}$. In the following, I provide a detailed walk-through of the Algorithm 1 in Appendix B.2. The simulation algorithm begins with initial guesses for parameters of the Poisson demand parameter λ_{ij}^t, and initial distribution of bikes $\mathbf{g}^1$. Starting in period 1, the 134×1 Poisson parameter λ_{ij}^1 is drawn according to $\{\theta^1, \theta_i^2, \theta_j'^2, \theta'^3, \mu_1, \sigma_1, \mu_2, \sigma_2\}$ in each tract i. The exogenous demand n_{ij}^1 is then generated and aggregated at different destination tracts j to form the demand n_i^1. The matching probability is calculated based on the number of unmatched bikes g_i^1, demand n_i^1, and search efficiency α. Then the matches are randomly drawn from the binomial distribution according to equation (7). The process is the same for each tract i to get the vector of matches m_i^1, which is 134×1. The state moves to the next period as the bikes transit to their destinations or stay unmatched at their origins. The entire simulation process concludes at the end of the day, $t = 48$.

4.2.3 SMM Details

The estimation of λ_{ij}^t is challenging as computing $m(g_i^t, n_i^t, \alpha)$ requires simulating bike transitions across periods over a large number of draws in order to obtain a credible estimate for all 134 locations and 48 periods. The estimation also requires the simulation process discussed above for each day in the sampling period. It is equivalent to creating and estimating the bike transition of dimension 134×134 with 48 periods in each of the 16 days of the sampling period. To form the moment conditions, I take the average matches by tracts across both time periods and days. The average number of matches observed is denoted as $\overline{\mathbf{M}}$, and the average number of simulated matches is $\overline{\hat{\mathbf{M}}}$. In accordance with equation (10), the moment condition is then $\overline{\mathbf{M}} - \overline{\hat{\mathbf{M}}} = \mathbf{0}$.

5 Empirical Results

This section provides empirical estimation results for the spatial structural model. First, I discuss the estimation results of the spatial structural model. Second, I discuss the results with the market efficiency coefficient α. Third, I test the model with a few different specifications and discuss the results.

5.1 Discussion of Estimation Results

Table 3 Panel A shows the estimation results for parameters forming the per-period Poisson parameter for customer arrivals $\{\lambda_i^t\}$, as well as market efficiency parameter $\{\alpha\}$. Panel B shows the results calculated from estimation: the total travel time accumulated by bike trips $\{u\}$, the total number of matches $\{m\}$, and the average trip time $\{u/m\}$.

Table 3: Results Summary

Panel A: Parameter Estimates			
Components of Poisson Demand λ	Parameter	Estimate	Std. Err.
Distance, *log(travel time)*	θ^1	0.990***	0.096
Origin tract *log(population density)*	θ^{21}	0.100	0.082
Origin tract *number of subway stations*	θ^{22}	0.148**	0.067
Destination tract *log(population density)*	θ'^{21}	0.036*	0.020
Destination tract *number of subway stations*	θ'^{22}	0.112	0.080
Time function	μ_1	18.317***	1.824
Time function	σ_1	3.398***	0.592
Time function	μ_2	35.405***	1.655
Time function	σ_2	5.818***	0.789
Time function Parameter	θ'^3	0.006	0.010
Market Efficiency Coefficient	α	1.774***	0.582
Panel B: Calculated Estimates		Estimate	Std. Err.
Total Trip Time (mins)	u	7,411,182	-
Total Number of Matches	m	953,332	-
Avg Trip Time (mins)	u/m	7.774	-

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The coefficient for distance parameter θ^1 is positive and significant, suggesting that a longer distance generates more demand for trips. This is not implying that long trips, such as riding a bike across the city, have higher demand. Most of the observations in the data are short trips that either are within-tract or between neighboring tracts. The positive coefficient suggests that, for short trips, the demand for bikes is higher for slightly longer trips.

The coefficients for the time function (6), $\mu_1, \sigma_1, \mu_2, \sigma_2$, suggests that there are two rush hours: one around 9 am ($\mu_1 = 18.317$) and one around 5:30 pm ($\mu_2 = 35.405$). The standard deviations (σ_1, σ_2) also suggest that the morning rush hour is more concentrated than the evening rush hour, reflecting the real-life scenario that working hours have similar starting and various ending times.

The coefficient for the number of subway stations at the origin tract θ^{22} is positive and significant, suggesting that the demand for bikes is positively related to the number of subway stations. This result validates the scenario described in the introduction that the bike-sharing service is mainly used to solve the "last-mile" problem. The coefficient for the population density is not significant because consumers demand bikes wherever they want and are not limited to where they live.

Based on the simulation and estimation results, I calculate that a total number of 953 thousand trips accumulates a total trip time of 7.4 million minutes per day. The average trip time is 7.774 minutes, or 466 seconds.

5.2 Discussion of Market Efficiency Coefficient α

The market efficiency coefficient α is 1.774 and the standard error suggests that it is precisely estimated. To visualize the goodness-of-fit, I fix the values for all other parameters (i.e., $\theta^1, \theta_i^2, \theta_j'^2, \theta'^3, \mu_1, \sigma_1, \mu_2, \sigma_2$) and change α to simulate the objective function. Figure 13 in Appendix B.3 shows that the objective function is convex with local minimum between $\alpha = 1.77$ and $\alpha = 1.78$.

The economic meaning of α measures the loss in market efficiency due to the spatial mismatch between consumers and bikes, so I set α to zero and calculate the total number of trips with the perfect scenario. I find that $\alpha = 1.774$ reduces the total matches by 29.95%, or a net loss of $332,979$ trips, compared to the perfect within-tract matching.

To scrutinize how the total trip time u and the total number of matches m change with α, I simulate the model with $\alpha \in [0, 2]$. I choose a step size of 0.01 and simulate 30 runs to calculate the averages of total matches and total trip time for each α. I find that increasing α by 0.01 reduces matches by 0.197%, suggesting that more severe spatial mismatch leads to worse loss both in matches and total trip time. Figure 8 illustrates how the total matches and total trip time change with α. It shows that (1) an increase in α leads to fewer matches and shorter total trip time, and (2) the marginal decreases in matches and trip time are large at first and flatten out when α passes 1.

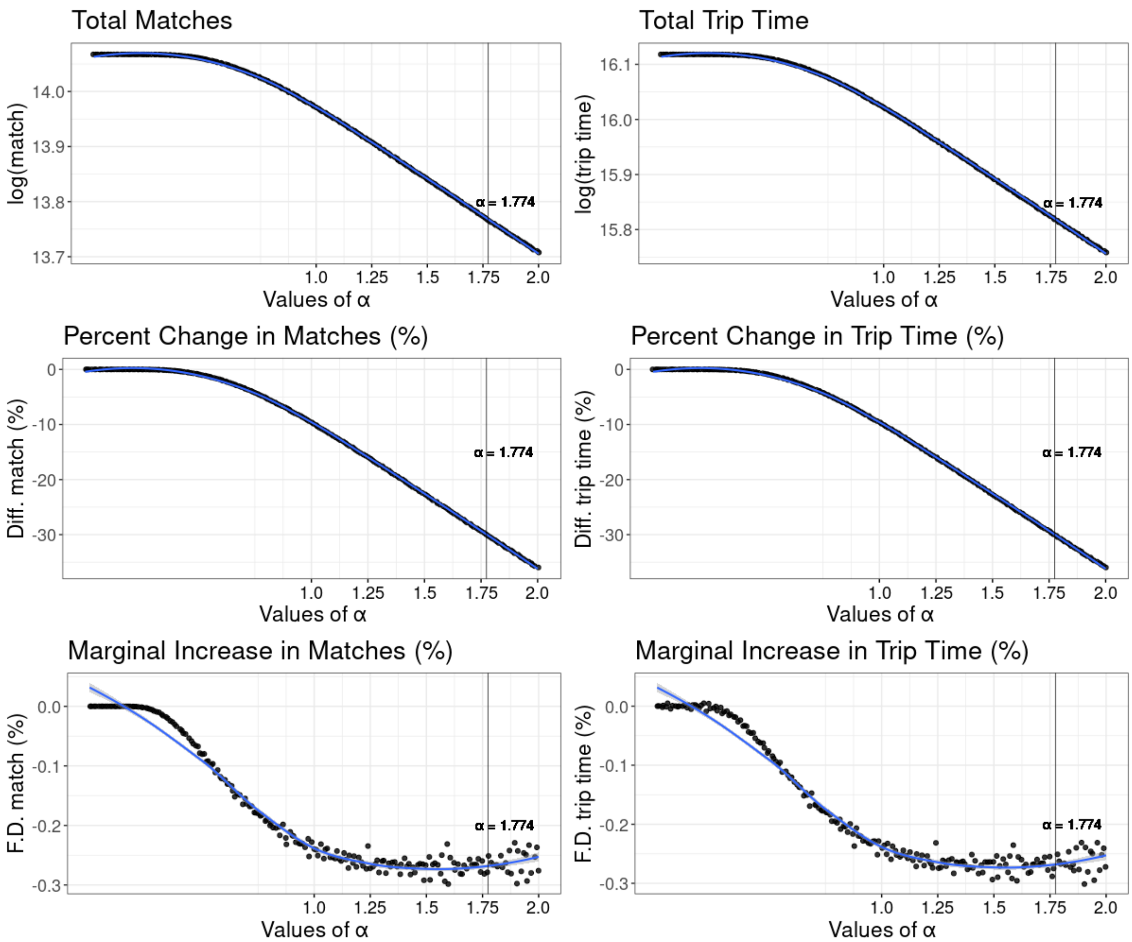


Figure 8: Efficiency Loss due to α

Notes: The left panel pertains to the total matches, and the right panel pertains to the total trip time. The top panel is the log value of total matches and total trip time. The middle panel is the difference and thus reflects the percent change in total matches and trip time. The bottom panel is the first-difference and thus reflects the marginal percent change in total matches and trip time.

5.3 Additional Specifications

In this subsection, I estimate the model with some alternative specifications. I first add a 'rush-hour' dummy variable and a 'busy-tract' dummy variable to control for the spatial and temporal distribution of bikes and find that they do not improve the model fit. Then I simulate the model with three sub-samples, non-rainy days, work days, and off days, and the usage of bikes increases for non-rainy days and work days. In the following I present the estimation results with each of the specifications.

Rush Hours and Busy Tracts The figures 2 and 3 in section 2.4 show that the bike trips have uneven distribution both spatially and temporally. I add a dummy variable τ^t for the rush hours and set it to 1 for periods 15-19 (7:30 am to 9:30 am) and 34-38 (5:00 pm to 7:00 pm) and zero otherwise. I add τ^t to equation (5) and estimate it as an additional parameter for λ_i^t. I report the estimation results in model column (2) in table 4. Compared to the baseline model in column (1), adding a rush-hour dummy does not significantly change the estimation. This result suggests that the time effect function (6) captures the time variation of demand well.

I add dummy variables B_i, B_j to account for the busy origin and destination tracts. The criteria of a busy origin tract are set as follows: I calculate the average number of matches m_i^t of all origin tracts and denote it as $\overline{m}^t$. Then I compare the number of matches of each origin tract m_i^t with $\overline{m}^t$: $B_i = 1$ if $m_i^t \geq \overline{m}^t$ and $B_i = 0$ otherwise. The criteria of a busy destination tract are the same except that I calculate the number of matches ending in the destination tract. I add B_i, B_j to equation (5) and estimate them as additional parameters for λ_i^t. I report the estimation results in model column (3) in table 4. Compared to the baseline model in column (1), although busy-tract dummies have positive coefficients, they do not significantly change the estimation. This result suggests that α captures the cross-tract spatial mismatch well.

Rainy Day, Workday, and Off Day Bike riding is affected by the weather. I infer from my weather data that May 22, 2017, is a rainy day. To see how rainy days affect bike usage, I take the observations of May 22 out and simulate the model with the rest of the data. I report the estimation results in model column (2) in table 5. Compared to the baseline model in column (1), total matches and trip time are both slightly longer for non-rainy days, probably attributable to a slightly lower spatial mismatch and better market efficiency ($\alpha = 1.773 < 1.774$). The change is not significant because I observe only one rainy day in my sample. I expect the change to be more significant with a larger sample that including many rainy days.

To validate the argument that the bike-sharing service is mainly used to solve the "last-mile" problem people often have when they transit to and from work, I split the sample according to work days and off days and simulate the model with them, respectively. According to section 2.2.1, the off days include two weekend days (May 20 and May 21) and three national holidays (May 28 - May 30). The remaining 11 days in my sample are workdays. I simulate the model with each of them separately,

Table 4: Comparative Studies Results: Rush Hours and Busy Tracts

Panel A: Parameter Estimates		(1)	(2)	(3)
Components of Poisson Demand λ	Parameter	Estimate (Std. Err.)	Estimate (Std. Err.)	Estimate (Std. Err.)
Distance, *log(travel time)*	θ^1	0.990^{***} (0.096)	0.988^{***} (0.088)	0.962^{***} (0.084)
Origin tract *log(population density)*	θ^{21}	0.100 (0.082)	0.096 (0.084)	0.091 (0.077)
Origin tract *number of subway stations*	θ^{22}	0.148^{**} (0.067)	0.134^{**} (0.054)	0.128^{**} (0.051)
Destination tract *log(population density)*	θ'^{21}	0.036^{*} (0.020)	0.035 (0.025)	0.038 (0.028)
Destination tract *number of subway stations*	θ'^{22}	0.112 (0.080)	0.108 (0.078)	0.107 (0.088)
Time function	μ_1	18.317^{***} (1.824)	18.317^{***} (1.824)	18.317^{***} (1.824)
Time function	σ_1	3.398^{***} (0.592)	3.398^{***} (0.592)	3.398^{***} (0.592)
Time function	μ_2	35.405^{***} (1.655)	35.405^{***} (1.655)	35.405^{***} (1.655)
Time function	σ_2	5.818^{***} (0.789)	5.818^{***} (0.789)	5.818^{***} (0.789)
Time function Parameter	θ'^3	0.006 (0.010)	0.006 (0.009)	0.006 (0.009)
Market Efficiency Coefficient	α	1.774^{***} (0.582)	1.770^{***} (0.567)	1.772^{***} (0.585)
Rush hour dummy	τ^t	-	0.0688 (0.0442)	-
Origin tract busy dummy	B_i	-	-	0.855 (0.628)
Destination tract busy dummy	B_j	-	-	0.749 (0.608)
Panel B: Calculated Estimates		(1)	(2)	(3)
Total Trip Time (mins)	u	7,411,182	$7,243,191$	$7,453,371$
Total Number of Matches	m	953,332	$931,449$	$958,899$
Avg Trip Time (mins)	u/m	7.774	7.776	7.773

Std. Err. in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5: Comparative Studies Results: Rainy Days, Workdays, and Off Days

Panel A: Parameter Estimates		(1)	(2)	(3)	(4)
Components of Poisson Demand λ	Parameter	Estimate (Std. Err.)	Estimate (Std. Err.)	Estimate (Std. Err.)	Estimate (Std. Err.)
Distance, *log(travel time)*	θ^1	0.990^{***} (0.096)	0.976^{***} (0.091)	0.992^{***} (0.078)	1.049^{***} (0.092)
Origin tract *log(population density)*	θ^{21}	0.100 (0.082)	0.133 (0.094)	0.172 (0.116)	0.162^{*} (0.095)
Origin tract *number of subway stations*	θ^{22}	0.148^{**} (0.067)	0.142^{**} (0.063)	0.167^{***} (0.057)	0.172^{***} (0.067)
Destination tract *log(population density)*	θ'^{21}	0.036^{*} (0.020)	0.036 (0.025)	0.040 (0.026)	0.027 (0.025)
Destination tract *number of subway stations*	θ'^{22}	0.112 (0.080)	0.115 (0.073)	0.118 (0.092)	0.133 (0.109)
Time function	μ_1	18.317^{***} (1.824)	18.516^{***} (1.627)	18.308^{***} (1.776)	18.892^{***} (1.921)
Time function	σ_1	3.398^{***} (0.592)	3.922^{***} (0.587)	3.384^{***} (0.579)	3.582^{***} (0.614)
Time function	μ_2	35.405^{***} (1.655)	35.590^{***} (1.882)	35.771^{***} (1.676)	35.296^{***} (1.634)
Time function	σ_2	5.818^{***} (0.789)	5.725^{***} (0.877)	5.698^{***} (0.732)	5.921^{***} (0.844)
Time function Parameter	θ'^3	0.006 (0.010)	0.006 (0.008)	0.007 (0.007)	0.005 (0.008)
Market Efficiency Coefficient	α	1.774^{***} (0.582)	1.773^{***} (0.542)	1.774^{***} (0.555)	1.662^{***} (0.605)
Panel B: Calculated Estimates		(1)	(2)	(3)	(4)
Total Trip Time (mins)	u	7,411,182	$7,445,941$	$7,644,270$	$7,045,068$
Total Number of Matches	m	953,332	$957,767$	$983,643$	$935,109$
Avg Trip Time (mins)	u/m	7.774	7.774	7.771	7.533

Std. Err. in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

and I report the estimation results for workdays in column (3) and off days for column (4) in table 5. Compared to the baseline model in column (1), total matches and trip time are higher for workdays and lower for off days. Column (3) also reports slightly more concentrated rush hours, suggesting consumers have similar demand for bikes during workdays. Column (4) has a higher value for distance θ^1, implying that consumers use bikes for longer than the "last-mile" trips. The market efficiency coefficient α is also lower for off days, suggesting that the spatial mismatch is less pronounced when consumers have more dispersed usage of bikes.

6 Counterfactual Analyses

In this section, I estimate the externality from consumers' usage of bikes with counterfactual analyses. The model has captured the within-tract spatial mismatch between the demand and supply of bikes. As mentioned in the introduction, the externality arises when a consumer uses a bike and does not consider the bike for a higher-value trip. More consumers in the market introduce congestion as the price is inflexible and the supply of bikes is limited. Consumers who do not get matched with bikes leave the market. The consumer may also ride the bike to an unpopular destination, exacerbating the spatial mismatch problem.

I first simulate the model with the different number of bikes g_i^t to see how it helps with the market congestion and affects the matches and trip time. Then I estimate the model with price discrimination against low-value trips and unpopular destinations. Finally, I test if changing the frequency of bike reshuffling would have any significant impact.

6.1 Counterfactual I: Changing the Number of Bikes

Supplying more bikes increases the market thickness and reduces the market congestion. According to figure 7, given the fixed demand, matches increase with the number of bikes. In order to quantify the effect of supplying more bikes, I fix the demand and simulate the model with different numbers of bikes. The method is as follows. I adjust the number of bikes at the start of the day by a multiplier in $[0.1, 2]$ so that the minimum is one-tenth and the maximum is twice the original number of bikes. I choose a step size of 0.01 and simulate 30 runs to calculate the averages of total matches and total trip time for each multiplier. I find that halving the number of bikes reduces the total number of matches by 46.40%, while doubling the bikes increases by 28.46%. Figure 9 illustrates how the total matches and total trip time change with the multiplier of bikes. It shows that (1) adding more bikes increases the matches and total trip time while reducing the bikes decreases the matches and total trip time. (2) The bottom panel shows decreasing returns to scale for both matches and total trip time on the multiplier of bikes. One policy implication is that by combining the decreasing returns of scale on the multiplier of bikes with an upward sloping marginal cost of supplying more bikes, the social planner can find the optimal number of bikes to supply.

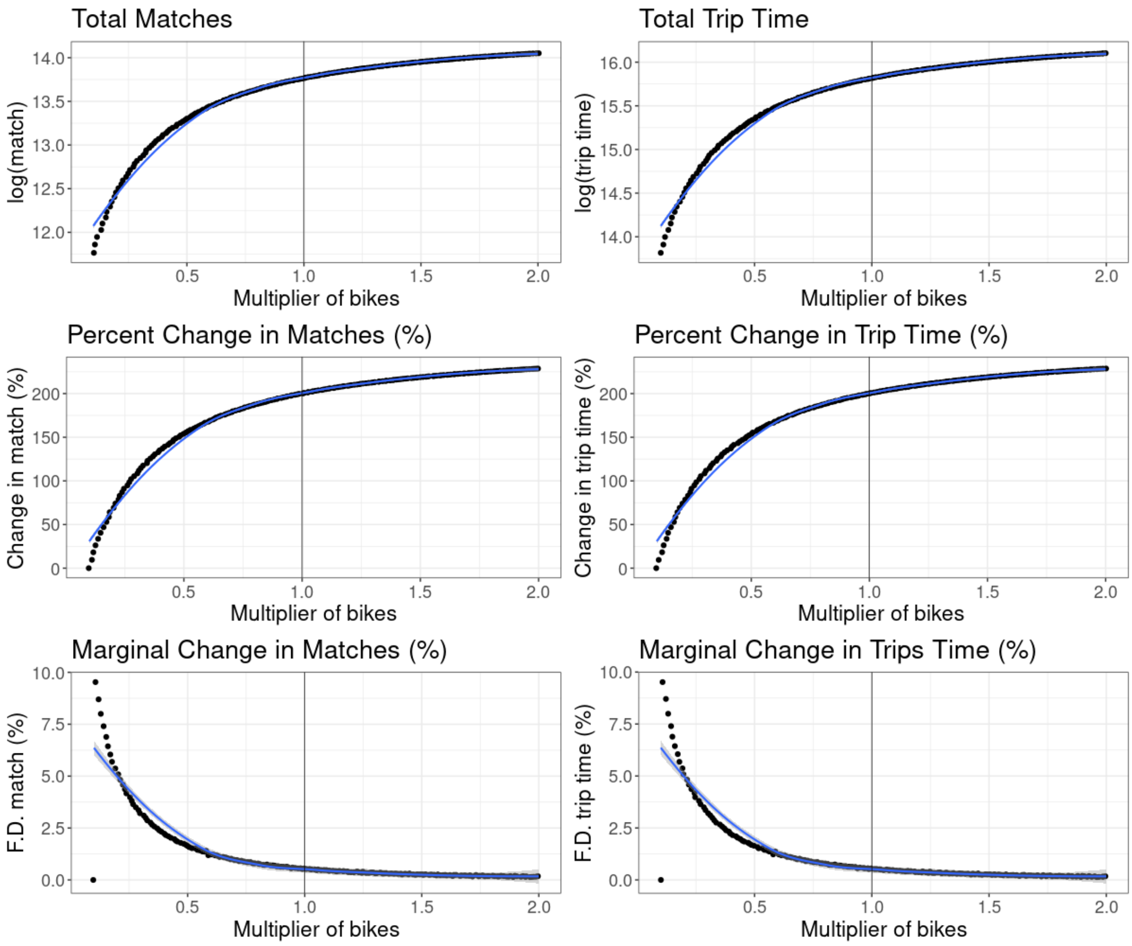


Figure 9: Different Initial Distribution of Bikes

Notes: The left panel pertains to the total matches, and the right panel pertains to the total trip time. The top panel is the log value of total matches and total trip time. The middle panel is the difference and thus reflects the percent change in total matches and trip time. The bottom panel is the first-difference and thus reflects the marginal percent change in total matches and trip time.

6.2 Counterfactual II: Price Discrimination

In this subsection, I test two measures: price discriminating against low-value trips and unpopular destinations. The counterfactual analysis with the former addresses the externality of the consumer's ignorance of other high-value trips, and the latter discusses the externality of the consumer's usage to unpopular destinations. I do not observe demand directly, and the price is inflexible, so I cannot apply and test price discrimination measures in a usual way. I adjust the Poisson parameter λ according to the implications of price discrimination, and in the following, I provide detailed analyses.

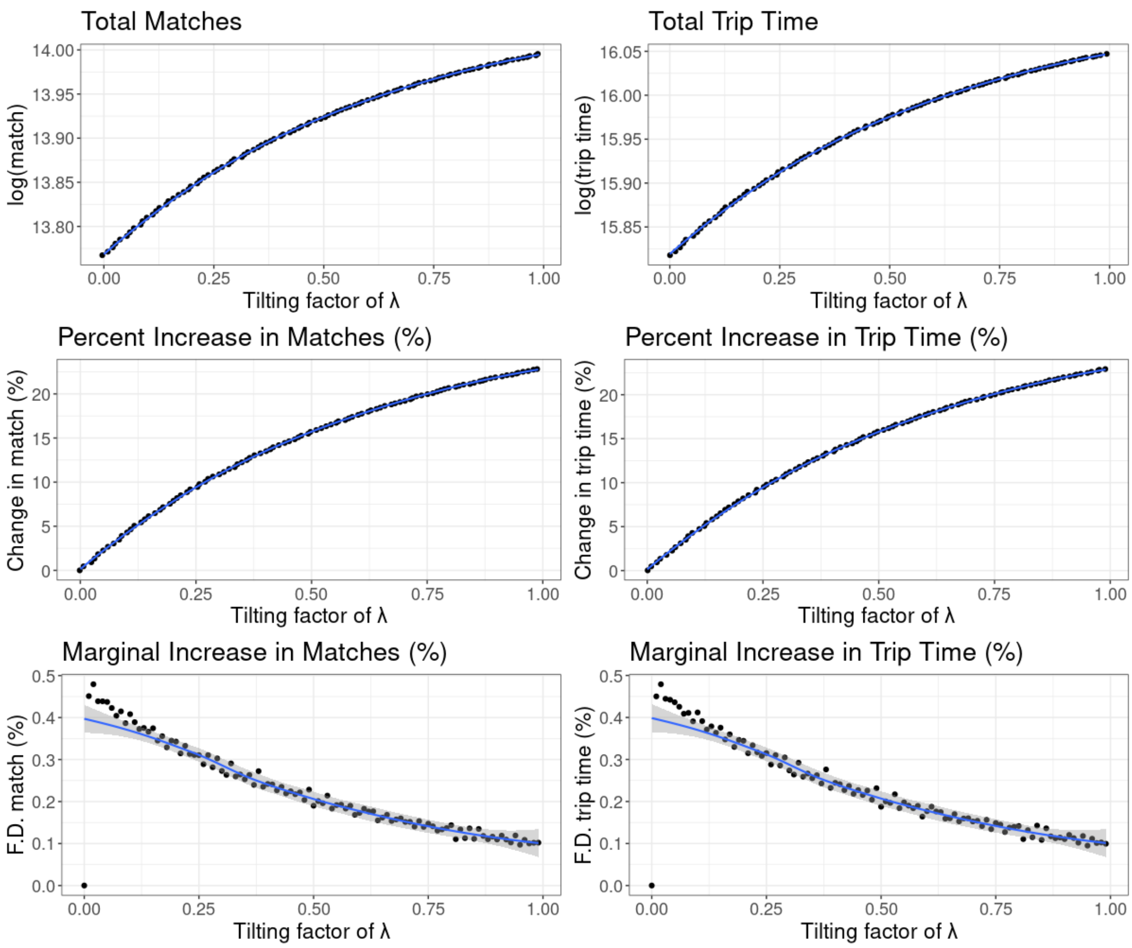


Figure 10: Changes in Matches and Trip Time with Different λ

Notes: The left panel pertains to the total matches, and the right panel pertains to the total trip time. The top panel is the log value of total matches and total trip time. The middle panel is the difference and thus reflects the percent change in total matches and trip time. The bottom panel is the first-difference and thus reflects the marginal percent change in total matches and trip time.

Price Discriminating Against Low-Value Trips According to table 3 and observations from the data, many bikes trips are either within-tract or between neighboring tracts. In order to increase the total usage of bikes, the bike-sharing company can raise the price for extremely short trips and provide discounts for longer trips. To reflect this policy in my model, I adjust the demand parameter λ by a tilting factor $\delta \in [0, 0.99]$. The method is as follows. I multiply the Poisson parameter λ for within-tract and neighboring-tract trip with $(1 - \delta)$ so that the demand for these low-value trips are discouraged by δ. I time the Poisson parameter λ for other trips with $(1 + \delta)$ to incentivize the longer trips by δ. For example, a tilting factor of 0.01 multiplies λ for short trips by 99% and long trips by 101%, thus price discriminating against short trips by 2%. I choose a step size of 0.01 and simulate 30 runs to calculate the averages of total matches and total trip time for each tilting factor δ. I find that price-discriminating against short trips by 2% only increases the total matches by 0.21% and total trip time by 0.22%. Figure 10 illustrates how the total matches and total trip time change with the tilting factor λ. It shows that (1) price discriminating against short trips increases the matches and total trip time, and (2) The bottom panel shows decreasing returns to scale for both matches and total trip time on the tilting factor. One policy implication is that the bike-sharing company can find the optimal price discrimination policy by combining the decreasing returns to scale of the price discrimination with an upward sloping marginal cost of providing discounts and subsidies for long trips.

Price Discriminating Against Unpopular Tracts One externality that aggravates the spatial mismatch is that a consumer rides a bike from a busy tract to an unpopular tract in figure 3. She takes one bike away from the busy tract, intensifying the congestion here, and transits it to an unpopular tract, magnifying the market thickness there. The bike-sharing company can raise the price for trips to the unpopular tracts and provide discounts for the popular ones. To reflect this policy in my model, I adjust the demand parameter λ by a tilting factor ψ, which is based on the ratio of n_i^t/g_i^t : $\psi_i^t = 1 + (0.1 \times \log(n_i^t/g_i^t))$. For busy tracts, there are more consumers n_i^t than bikes g_i^t so $\log(n_i^t/g_i^t) > 0$ and $\psi_i^t > 1$; For unpopular tracts, there are more bikes g_i^t than consumers n_i^t so $\log(n_i^t/g_i^t) < 0$ and $\psi_i^t < 1$. The number 0.1 is an arbitrary tuning parameter. With this setting, the demand parameter λ will be higher for busy tracts and lower for unpopular tracts, concentrating bikes for high-demand tracts. I simulate 30 runs to calculate the averages of total matches and total trip time and find that price-discriminating against unpopular tracts increases the total matches by 3.73% and total trip time by 3.71%.

Based on the two counterfactual price discrimination analyses in this subsection, I find that price discriminating against unpopular tracts is more effective in increasing the total usage of bikes. It implies that cross-tract externality has a larger impact on the market efficiency that the within-tract externality.

6.3 Counterfactual III: Reshuffle Bikes

In my model, I assume that bikes are reshuffled at the end of the day. In the real world, however, ofo could move only a few thousand bikes around each day. I simulate the model with different reshuffling frequencies to see if they make any difference. I elongate the interval of bike reshuffling from 1 day to 30 days with the step size of one day, simulate 30 runs, and calculate the average total number of matches and total trip time for each interval. Figure 11 shows that changing the reshuffling frequency does not have a significant impact on either the total matches or the total trip time.

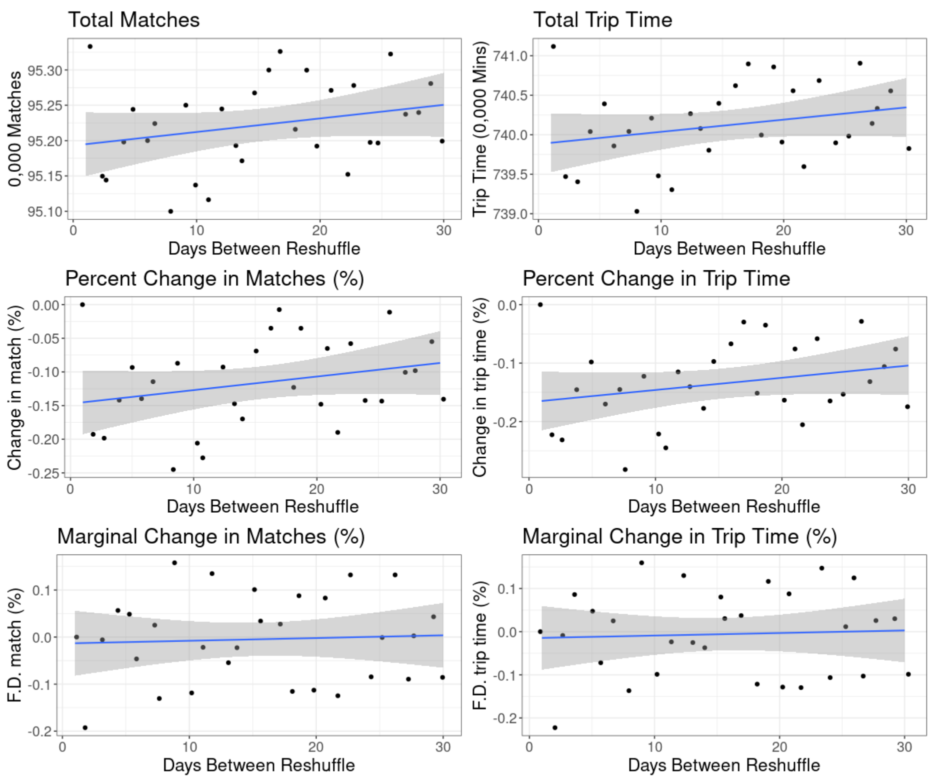


Figure 11: Changes in Matches and Trip Time with Different Bike Reshuffling Period

Notes: The left panel pertains to the total matches, and the right panel pertains to the total trip time. The top panel is the log value of total matches and total trip time. The middle panel is the difference and thus reflects the percent change in total matches and trip time. The bottom panel is the first-difference and thus reflects the marginal percent change in total matches and trip time.

7 Conclusion

The spatial mismatch of consumers and bikes in the bike-sharing industry hinders market efficiency. An externality from the consumer's usage of bikes exacerbates the efficiency loss. This paper develops a spatial structural model with local search and matching process between bikes and consumers, showing how the demand and market efficiency can be recovered from data on tract characteristics and variations of intra-day matches. Using data from a dockless bike-sharing company in Beijing, China, I estimate the model to pin down the parameters determining the distribution of bikes' demand. Using the variation in matches, I estimate the market efficiency loss due to spatial mismatch between consumers and bikes. I estimate that the daily trip volume in Beijing is $953,332$, accumulating a total trip time of $7,411,182$ minutes and an average trip time of 7.8 minutes. Compared to the perfect within-tract matching, within-tract spatial mismatch reduces the total matches by 29.95%, or a net loss of $332,979$ trips.

The estimated model also allows me to perform counterfactual analyses to quantify the efficiency cost of the externality. I show that (1) supplying more bikes attenuates the market congestion, increasing the matches with diminishing returns. (2) Price discriminating against short trips and unpopular tracts helps reduce the externality and increases the total bike usage. (3) Changing the frequency of bike reshuffling does not affect the total usage of bikes.

References

Allen, Jason, Robert Clark, and Jean-François Houde (2019). "Search Frictions and Market Power in Negotiated-Price Markets". In: *Journal of Political Economy* 127.4, pp. 1550–1598. DOI: 10.1086/701684. URL: https://ideas.repec.org/a/ucp/jpolec/doi10.1086-701684.html.

Armstrong, Mark (2006). "Competition in Two-Sided Markets". In: *RAND Journal of Economics* 37, pp. 668–691.

Ashraf, Nava, Edward L. Glaeser, and Giacomo A. M. Ponzetto (May 2016). *Infrastructure, Incentives, and Institutions*. DOI: 10.1257/aer.p20161095. URL: https://www.aeaweb.org/articles?id=10.1257/aer.p20161095.

Beijing Subway (2017). *Beijing Subway Route Map*. data retrieved from https://www.bjsubway.com/station/xltzs/.

Brancaccio, Giulia, Myrto Kalouptsidi, and Theodore Papageorgiou (2020a). "A guide to estimating matching functions in spatial models". In: *International Journal of Industrial Organization* 70.C. DOI: 10.1016/j.ijindorg.2019.1. URL: https://ideas.repec.org/a/eee/indorg/v70y2020ics0167718719300554.html.

— (Mar. 2020b). "Geography, Transportation, and Endogenous Trade Costs". In: *Econometrica* 88.2, pp. 657–691. DOI: 10.3982/ECTA15455. URL: https://ideas.repec.org/a/wly/emetrp/v88y2020i2p657-691.html.

Bryan, Kevin A. and Joshua S. Gans (Jan. 2019). "A theory of multihoming in rideshare competition". In: *Journal of Economics & Management Strategy* 28.1, pp. 89–96. DOI: 10.1111/jems.12306. URL: https://ideas.repec.org/a/bla/jemstr/v28y2019i1p89-96.html.

Buchholz, Nicholas (2017). *Spatial Equilibrium, Search Frictions and Efficient Regulation in the Taxi Industry*. PhD thesis. Princeton University. URL: https://scholar.princeton.edu/sites/default/files/nbuchholz/files/taxi_draft.pdf.

Buchholz, Nicholas, Matthew Shum, and Haiqing Xu (2021). "Semiparametric estimation of dynamic discrete choice models". In: *Journal of Econometrics* 223.2, pp. 312–327. DOI: 10.1016/j.jeconom.2020.01. URL: https://ideas.repec.org/a/eee/econom/v223y2021i2p312-327.html.

Butters, Gerard R. (1977). "Equilibrium Distributions of Sales and Advertising Prices". In: *Review of Economic Studies* 44.3, pp. 465–491. URL: https://ideas.repec.org/a/oup/restud/v44y1977i3p465-491..html.

Cairns, Robert D. and Catherine Liston-Heyes (Jan. 1996). "Competition and regulation in the taxi industry". In: *Journal of Public Economics* 59.1, pp. 1–15. URL: https://ideas.repec.org/a/eee/pubeco/v59y1996i1p1-15.html.

Camerer, Colin et al. (1997). "Labor Supply of New York City Cabdrivers: One Day at a Time". In: *The Quarterly Journal of Economics* 112.2, pp. 407–441. URL: https://ideas.repec.org/a/oup/qjecon/v112y1997i2p407-441..html.

Cao, Guangyu et al. (Aug. 2018). *Market Expanding or Market Stealing? Competition with Network Effects in BikeSharing*. Working Paper 24938. National Bureau of Economic Research. DOI: 10.3386/w24938. URL: http://www.nber.org/papers/w24938.

Castillo, Juan Camilo (2020). *Who Benefits from Surge Pricing?*

Castillo, Juan Camilo, Daniel Knoepfle, and Glen Weyl (2021). *Matching in Ride Hailing: Wild Goose Chases and How to Solve Them.* DOI: 10.2139/ssrn.2890666.

China Meteorological Data Service Centre (2017). *Daily Data From Meteorological Sta-tions in China*. data retrieved from https://data.cma.cn/en/.

Crawford, Vincent P. and Juanjuan Meng (Aug. 2011). "New York City Cab Drivers' Labor Supply Revisited: Reference-Dependent Preferences with Rational-Expectations Targets for Hours and Income". In: *American Economic Review* 101.5, pp. 1912–1932. URL: https://ideas.repec.org/a/aea/aecrev/v101y2011i5p1912-32.html.

Demaio, Paul (2003). "Smart Bikes: Public Transportation for the 21st Century". In: *Transportation Quarterly* 57.1, pp. 9–11.

— (2009). "Bikesharing:History, Impacts, Models of Provision, and Future". In: *Journal of PublicTransportation* 12.4, p. 3.

Demaio, Paul and Jonathan Gifford (2004). "Will Smart Bikes Succeed as Public Transportation in the United States?" In: *Journal of Public Transportation* 7.2, p. 1.

Dinerstein, Michael et al. (July 2018). "Consumer Price Search and Platform Design in Internet Commerce". In: *American Economic Review* 108.7, pp. 1820–1859. URL: https://ideas.repec.org/a/aea/aecrev/v108y2018i7p1820-59.html.

Farber, Henry S. (June 2008). "Reference-Dependent Preferences and Labor Supply: The Case of New York City Taxi Drivers". In: *American Economic Review* 98.3, pp. 1069–1082. URL: https://ideas.repec.org/a/aea/aecrev/v98y2008i3p1069-82.html.

Flath, David (June 2006). "Taxicab regulation in Japan". In: *Journal of the Japanese and International Economies* 20.2, pp. 288–304. URL: https://ideas.repec.org/a/eee/jjieco/v20y2006i2p288-304.html.

Fradkin, Andrey (2017). *Search, matching, and the role of digital marketplace design in enabling trade: Evidence from Airbnb*.

Fréchette, Guillaume R., Alessandro Lizzeri, and Tobias Salz (Aug. 2019). "Frictions in a Competitive, Regulated Market: Evidence from Taxis". In: *American Economic Review* 109.8, pp. 2954–92. DOI: 10.1257/aer.20161720. URL: http://www.aeaweb.org/articles?id=10.1257/aer.20161720.

Gallick, Edward C and David E Sisk (Spring 1987). "A Reconsideration of Taxi Regulation". In: *Journal of Law, Economics, and Organization* 3.1, pp. 117–128. URL: https://ideas.repec.org/a/oup/jleorg/v3y1987i1p117-28.html.

Gavazza, Alessandro (June 2011). "The Role of Trading Frictions in Real Asset Markets". In: *American Economic Review* 101.4, pp. 1106–1143. URL: https://ideas.repec.org/a/aea/aecrev/v101y2011i4p1106-43.html.

Glaeser, Edward L., Scott Duke Kominers, et al. (Dec. 2015). *Big Data and Big Cities: The Promises and Limitations of Improved Measures of Urban Life*. Working Paper. DOI: 10.3386/w21778. URL: http://www.nber.org/papers/w21778.

Glaeser, Edward L., Giacomo A. M. Ponzetto, and Yimei Zou (2016). *Urban networks: Connecting markets, people, and ideas*. DOI: https://doi.org/10.1111/pirs.12216. eprint: https://rsaiconnect.onlinelibrary.wiley.com/doi/pdf/10.1111/pirs.12216. URL: https://rsaiconnect.onlinelibrary.wiley.com/doi/abs/10.1111/pirs.12216.

Häckner, Jonas and Sten Nyberg (Dec. 1992). *Deregulating Taxi Services - A Word of Caution*. Working Paper Series 353. Research Institute of Industrial Economics. URL: https://ideas.repec.org/p/hhs/iuiwop/0353.html.

Hall, Robert E. (Apr. 1979). "A theory of the natural unemployment rate and the duration of employment". In: *Journal of Monetary Economics* 5.2, pp. 153–169. URL: https://ideas.repec.org/a/eee/moneco/v5y1979i2p153-169.html.

He, Pu et al. (Mar. 2021). "Customer Preference and Station Network in the London Bike-Share System". In: *Management Science* 67.3, pp. 1392–1412. DOI: 10.1287/mnsc.2020.3620. URL: https://ideas.repec.org/a/inm/ormnsc/v67y2021i3p1392-1412.html.

Kabra, Ashish, Elena Belavina, and Karan Girotra (2018). *Bikeshare Systems: Accessibility and Availability*. Tech. rep. Chicago Booth Research.

Lagos, Ricardo (Oct. 2000). “An Alternative Approach to Search Frictions”. In: *Journal of Political Economy* 108.5, pp. 851–873. DOI: 10.1086/317674. URL: https://ideas.repec.org/a/ucp/jpolec/v108y2000i5p851-873.html.

— (May 2003). “An Analysis of the Market for Taxicab Rides in New York City”. In: *International Economic Review* 44.2, pp. 423–434. URL: https://ideas.repec.org/a/ier/iecrev/v44y2003i2p423-434.html.

Mortensen, Dale (Apr. 1984). *Job Search and Labor Market Analysis*. Discussion Papers 594. Northwestern University, Center for Mathematical Studies in Economics and Management Science. URL: https://ideas.repec.org/p/nwu/cmsems/594.html.

Mortensen, Dale and Christopher A. Pissarides (1999). “New developments in models of search in the labor market”. In: *Handbook of Labor Economics*. Ed. by O. Ashenfelter and D. Card. Vol. 3. Handbook of Labor Economics. Elsevier. Chap. 39, pp. 2567–2627. URL: https://ideas.repec.org/h/eee/labchp/3-39.html.

National Bureau of Statistics of China (2018). *China Statistical Yearbook 2018*. data retrieved from http://www.stats.gov.cn/tjsj/ndsj/2018/indexeh.htm.

O’mahony, Eoin and David Shmoys (2015). “Data Analysis and Optimization for (Citi) Bikesharing”. In: *Twentyninth AAAI conference on artificial intelligence*.

Pan, Ling et al. (2018). *A Deep Reinforcement Learning Framework for Rebalancing Dockless Bike Sharing Systems*. Tech. rep. eprint: arXiv:1802.04592.

Rochet, J.-C. and Jean Tirole (2003). “Platform competition in two-sided markets”. In: *Journal of the European Economic Association* 1, pp. 990–1029.

Rogerson, Richard, Robert Shimer, and Randall Wright (July 2004). *Search-Theoretic Models of the Labor Market-A Survey*. NBER Working Papers 10655. National Bureau of Economic Research, Inc. URL: https://ideas.repec.org/p/nbr/nberwo/10655.html.

Rysman, Marc (2009). “The Economics of Two-Sided Markets”. In: *Journal of Economic Perspectives* 23, pp. 125–144.

Shen, Yu, Xiaohu Zhang, and Jinhua Zhao (2018). “Understanding the usage of dockless bike sharing in Singapore”. In: *International Journal of Sustainable Transportation* 12.9, pp. 686–700. DOI: 10.1080/15568318.2018.1429696. eprint: https://doi.org/10.1080/15568318.2018.1429696. URL: https://doi.org/10.1080/15568318.2018.1429696.

Weyl, Glen (2010). “A Price Theory of Multi-Sided Platforms”. In: *American Economic Review* 100, pp. 1642–1672.

Yan, Fangning (2021). *Financing and Investment by a Platform Start-up: An Analysis of ofo.*

Zhang, Lihong et al. (2015). *Sustainable Bike-sharing Systems: Characteristics and Commonalities across Cases in Urban China.*

A Data Preparation

A.1 ofo Data Cleaning

The trip-level data from ofo are subject to some errors due to technology limits. I first remove duplicated records (those with the same combination of trip ID, user ID, and bike ID within each day). Next, I drop observations that are missing time stamps and geo-locations. Then, I drop any apparently erroneous observations, such as those well outside the Beijing Area. Finally, I drop trips that both start and end outside the locations of interest, which are the six metropolitan areas of Beijing. The table below shows the data cleaning routine and changes in observations after each step.

Table A.1: Data Cleaning Routine

Date	Category of Days		Procedures					Final Data Set
	Day of Week	Type of Day	Initial Data	Remove Duplication	Missing Time/Locations	Invalid Time/Locations	Outside Locations	
5/15/2017	Monday	Weekday	3,507,338	-668,075	-361,302	-60,890	-94,304	2,322,767
5/16/2017	Tuesday	Weekday	3,746,237	-820,900	-333,024	-178,630	-104,452	2,309,231
5/17/2017	Wednesday	Weekday	3,325,757	-66,836	-320,200	-186,514	-483,901	2,268,306
5/18/2017	Thursday	Weekday	4,490,182	-1,213,730	-303,045	-304,724	-537,284	2,131,399
5/19/2017	Friday	Weekday	6,184,028	-2,192,938	-383,966	-177,116	-697,325	2,732,683
5/20/2017	Saturday	Weekend	3,563,863	-95,302	-453,264	-75,650	-232,612	2,707,035
5/21/2017	Sunday	Weekend	3,771,089	-213,856	-471,875	-78,656	-196,812	2,809,890
5/22/2017	Monday	Weekday	1,991,860	-147,039	-183,296	-21,939	-104,630	1,534,956
5/23/2017	Tuesday	Weekday	3,096,696	-622,001	-338,440	-34,806	-110,480	1,990,969
5/24/2017	Wednesday	Weekday	3,558,383	-1,728,114	-263,807	-20,352	-31,359	1,514,751
5/25/2017	Thursday	Weekday	2,526,816	-152,755	-307,259	-33,943	-71,296	1,961,563
5/26/2017	Friday	Weekday	5,039,737	-1,609,429	-443,230	-56,610	-124,610	2,805,858
5/27/2017	Saturday	Weekday	3,275,065	-201,451	-494,883	-50,472	-29,473	2,498,786
5/28/2017	Sunday	Holiday	2,925,386	-371,408	-374,908	-53,708	-30,589	2,094,773
5/29/2017	Monday	Holiday	3,299,138	-187,508	-464,247	-45,560	-53,217	2,548,606
5/30/2017	Tuesday	Holiday	3,579,450	-451,786	-458,182	-45,118	-68,477	2,555,887
Total	-	-	57,881,025	-10,743,128	-5,954,928	-1,424,688	-2,970,821	36,787,460

A.2 Discretize the Geographical Space

The census tracts are numbered in alphabetical order as shown in table A.2 below. The census tracts outside the six metropolitan areas of Beijing are all assigned zero. As mentioned in appendix A.1, the trips that both starting and ending outside the metropolitan area are dropped. Trips with only starting or ending points outside the metropolitan area are considered.

Table A.2: Census Tracts Encoding

Name	Number	Name	Number	Name	Number
Andingmen Jiedao	1	Hepingjie Jiedao	46	Shuguang Jiedao	91
Anzhen Jiedao	2	Hepingli Jiedao	47	Sijiqing Banshichu	92
Aoyuncun Jiedao	3	Heyi Jiedao	48	Sujiatuo Banshichu	93
Babaoshan Jiedao	4	Huaxiang Banshichu	49	Sunhe Banshichu	94
Baizhifang Jiedao	5	Huayuanlu Jiedao	50	Taipingqiao Jiedao	95
Bajiao Jiedao	6	Hujialou Jiedao	51	Taiyanggong Banshichu	96
Balizhuang Jiedao (Chaoyang)	7	Jiangtai Banshichu	52	Taoranting Jiedao	97
Balizhuang Jiedao (Haidian)	8	Jianguomen Jiedao	53	Tiancunlu Jiedao	98
Beitaipingzhuang Jiedao	9	Jianwai Jiedao	54	Tianqiao Jiedao	99
Beixiaguan Jiedao	10	Jiaodaokou Jiedao	55	Tiantan Jiedao	100
Beixinqiao Jiedao	11	Jindingjie Jiedao	56	Tiyuguanlu Jiedao	101
Changxindian Jiedao	12	Jingshan Jiedao	57	Tuanjiehu Jiedao	102
Changxindian Zhen	13	Jinrongjie Jiedao	58	Wangjing Jiedao	103
Changying Banshichu	14	Jinsong Jiedao	59	Wangjing Dev Jiedao	104
Chaowai Jiedao	15	Jinzhan Banshichu	60	Wangsiying Banshichu	105
Chaoyangmen Jiedao	16	Jiuxianqiao Jiedao	61	Wangzuozhen	106
Chongwenmenwai Jiedao	17	Laiguangying Banshichu	62	Wanliu Banshichu	107
Chunshu Jiedao	18	Laoshan Jiedao	63	Wanpingcheng Banshichu	108
Cuigezhuang Banshichu	19	Liulitun Jiedao	64	Wanshoulu Jiedao	109
Dahongmen Jiedao	20	Longtan Jiedao	65	Wanquan Banshichu	110
Dashilan Jiedao	21	Lugouqiao Banshichu	66	Wulituo Jiedao	111
Datun Jiedao	22	Lugouqiao Jiedao	67	Xiangheyuan Jiedao	112
Desheng Jiedao	23	Lugu Jiedao	68	Xiangshan Jiedao	113
Dongba Banshichu	24	Maizidian Jiedao	69	Xiaoguan Jiedao	114
Dongfeng Banshichu	25	Majiapu Jiedao	70	Xiaohongmen Banshichu	115
Donggaodi Jiedao	26	Malianwa Jiedao	71	Xibeiwang Banshichu	116
Donghuamen Jiedao	27	Nanmofang Banshichu	72	Xichanganjie Jiedao	117
Donghuashi Jiedao	28	Nanyuan Banshichu	73	Xiluoyuan Jiedao	118
Dongsheng Banshichu	29	Nanyuan Jiedao	74	Xincun Jiedao	119
Dongsi Jiedao	30	Niujie Jiedao	75	Xinjiekou Jiedao	120
Dongtiejiangying Jiedao	31	Panjiayuan Jiedao	76	Xisanqi Jiedao	121
Dongzhimen Jiedao	32	Capital Airport PEK	77	Xueyuanlu Jiedao	122
Dougezhuang Banshichu	33	Pingfang Banshichu	78	Yangfangdian Jiedao	123
Fangzhuang Banshichu	34	Pingguoyuan Jiedao	79	Yanyuan Jiedao	124
Fatou Jiedao	35	Qianmen Jiedao	80	Yayuncun Jiedao	125
Fengtai Jiedao	36	Qinghe Jiedao	81	Yongdinglu Jiedao	126
Ganjiakou Jiedao	37	Qinghuayuan Jiedao	82	Yongdingmenwai Jiedao	127
Gaobeidian Banshichu	38	Qinglongqiao Jiedao	83	Youanmen Jiedao	128
Guanganmennei Jiedao	39	Sanjianfang Banshichu	84	Yuetan Jiedao	129
Guanganmenwai Jiedao	40	Sanlitun Jiedao	85	Yungang Jiedao	130
Guangning Jiedao	41	Shangdi Jiedao	86	Zhanlanlu Jiedao	131
Guanzhuang Banshichu	42	Shangzhuang Banshichu	87	Zhongguancun Jiedao	132
Gucheng Jiedao	43	Shibalidian Banshichu	88	Zizhuyuan Jiedao	133
Haidian Jiedao	44	Shichahai Jiedao	89	Zuojiazhuang Jiedao	134
Heizhuanghu Banshichu	45	Shuangjing Jiedao	90		

Table A.3: Bike Usage Over Different Time Periods

Time interval (seconds)	21600	14400	10800	7200	3600	2400	1800	1200	900	600	300
	% distinct bikes (max. no. re-use)	% distinct bikes (max. no. re-use)	% distinct bikes (max. no. re-use)	% distinct bikes (max. no. re-use)	% distinct bikes (max. no. re-use)	% distinct bikes (max. no. re-use)	% distinct bikes (max. no. re-use)	% distinct bikes (max. no. re-use)	% distinct bikes (max. no. re-use)	% distinct bikes (max. no. re-use)	% distinct bikes (max. no. re-use)
5/15/2017	55.10	62.85	66.23	72.91	82.89	88.15	91.29	94.94	96.82	98.53	99.65
	(20)	(19)	(14)	(12)	(8)	(8)	(6)	(5)	(5)	(4)	(4)
5/16/2017	58.34	63.65	67.56	73.61	83.34	88.44	91.49	95.05	96.86	98.51	99.59
	(24)	(22)	(22)	(18)	(14)	(12)	(11)	(11)	(10)	(6)	(5)
5/17/2017	59.45	63.40	68.55	74.46	84.09	89.12	92.10	95.47	97.20	98.74	99.72
	(16)	(14)	(14)	(10)	(10)	(10)	(10)	(6)	(5)	(4)	(4)
5/18/2017	61.24	65.10	70.16	75.83	85.01	89.76	92.60	95.75	97.36	98.82	99.96
	(17)	(17)	(17)	(17)	(9)	(9)	(7)	(5)	(4)	(4)	(3)
5/19/2017	54.95	59.08	64.20	70.53	81.09	86.76	90.24	94.21	96.36	98.26	99.58
	(34)	(31)	(29)	(29)	(26)	(21)	(16)	(12)	(9)	(7)	(5)
5/20/2017	53.25	58.93	63.82	70.62	81.67	87.32	90.72	94.61	96.58	98.33	99.55
	(24)	(21)	(21)	(21)	(16)	(11)	(10)	(9)	(8)	(7)	(5)
5/21/2017	51.27	57.53	62.42	69.52	81.05	86.89	90.46	94.48	96.55	98.34	99.57
	(40)	(39)	(36)	(35)	(21)	(21)	(18)	(13)	(10)	(7)	(7)
5/22/2017	67.16	71.52	74.80	79.80	87.45	91.50	93.81	96.52	97.85	99.09	99.82
	(13)	(13)	(12)	(9)	(8)	(6)	(6)	(4)	(4)	(3)	(3)
5/23/2017	59.22	63.62	68.82	74.89	84.61	89.55	92.52	96.75	97.41	98.91	99.82
	(14)	(14)	(11)	(9)	(7)	(6)	(6)	(5)	(4)	(4)	(3)
5/24/2017	63.23	69.82	72.95	78.93	87.70	91.75	94.08	96.70	98.00	99.16	99.86
	(13)	(11)	(10)	(9)	(6)	(5)	(5)	(4)	(4)	(3)	(2)
5/25/2017	61.75	68.21	71.87	77.79	86.73	91.03	93.64	96.41	97.82	99.09	99.85
	(13)	(11)	(11)	(8)	(7)	(6)	(6)	(6)	(6)	(6)	(5)
5/26/2017	54.60	60.89	65.91	72.48	82.73	87.99	91.21	94.86	96.76	98.47	99.64
	(32)	(20)	(20)	(20)	(13)	(10)	(10)	(9)	(8)	(6)	(5)
5/27/2017	54.76	61.52	65.64	72.16	82.66	87.96	91.21	94.84	96.74	98.48	99.65
	(28)	(21)	(21)	(20)	(13)	(12)	(10)	(7)	(6)	(5)	(4)
5/28/2017	62.24	67.86	72.07	77.46	86.42	90.74	93.27	96.17	97.56	98.84	99.68
	(23)	(23)	(16)	(13)	(11)	(8)	(8)	(8)	(7)	(5)	(3)
5/29/2017	55.54	61.97	66.87	73.56	84.11	89.17	92.23	95.54	97.24	98.67	99.64
	(34)	(34)	(31)	(23)	(16)	(13)	(9)	(8)	(6)	(5)	(5)
5/30/2017	55.75	61.57	66.37	73.02	83.46	88.64	91.73	95.26	97.04	98.60	99.64
	(25)	(21)	(20)	(13)	(12)	(11)	(10)	(7)	(6)	(5)	(4)

A.3 Weather Data Collection

The weather data is collected from the Daily Data From Meteorological Stations in China from the China Meteorological Data Service Centre (2017). It provides the daily observations from the six metropolitan areas in Beijing. Since the surface weather observation stations are not evenly distributed in these six areas, I use the inverse distance weighted interpolation to obtain a complete set of observations. The resulting dataset has the average daily temperature and precipitation observations for each of the six metropolitan areas, and the total number of observations for 16 days is 96.

Table A.4: Weather Fixed Effects Summary Statistics

Date	Precipitation (0.1mm)		Temperature ($^{\circ}C$)	
	Mean	St. Dev	Mean	St. Dev
5/15/2017	5.306	1.740	20.06	0.579
5/16/2017	2.057	0.658	22.32	0.547
5/17/2017	0.181	0.059	24.43	0.544
5/18/2017	0.617	0.206	26.27	0.512
5/19/2017	0.989	0.316	27.28	0.734
5/20/2017	4.427	0.837	26.17	0.679
5/21/2017	6.819	2.778	22.80	0.316
5/22/2017	210.4	27.32	17.56	0.232
5/23/2017	15.09	4.662	19.83	0.711
5/24/2017	3.279	1.045	23.06	0.843
5/25/2017	1.452	0.445	22.58	0.621
5/26/2017	0.114	0.037	21.87	0.176
5/27/2017	0.095	0.031	24.03	0.064
5/28/2017	13.04	7.773	25.96	0.505
5/29/2017	4.577	1.282	22.81	0.454
5/30/2017	19.75	3.857	21.35	0.227

A.4 Population Density Data Collection

The population density data is collected from China Statistical Yearbook 2018 from National Bureau of Statistics of China (2018), which reports the statistics by the city in 2017. I cross-validate the data with China Statistical Yearbook 2017 to consolidate and avoid any structural changes in population density. Both of the statistical yearbooks are compiled by the National Bureau of Statistics of China.

A.5 Subway Locations Data Collection

The subway location data is collected from the Beijing Subway (2017), which operates the subway system in Beijing. The subway system is continuously updated with new lines and stations, and I use the 2017 subway map data.

Figure 12: Beijing Subway Map

Table A.5: Tract Characteristics - Population Density and Number of Subway Stations

Census Tract Number	Population Density (per km^2)	Number of Subway Stations	Census Tract Number	Population Density (per km^2)	Number of Subway Stations	Census Tract Number	Population Density (per km^2)	Number of Subway Stations
1	44,358	2	46	108,406	2	91	102,397	6
2	68,454	12	47	112,058	16	92	170,579	2
3	105,263	0	48	40,477	8	93	46,786	6
4	53,606	2	49	143,041	8	94	23,876	6
5	95,737	0	50	148,829	4	95	63,588	2
6	116,416	6	51	61,588	10	96	70,367	2
7	107,542	2	52	71,341	10	97	43,455	0
8	135,984	0	53	57,170	4	98	108,744	8
9	201,614	8	54	34,692	0	99	46,385	6
10	158,776	10	55	49,196	0	100	50,304	0
11	82,273	0	56	75,736	6	101	40,303	6
12	159,357	2	57	40,308	6	102	37,674	0
13	47,986	6	58	67,888	0	103	168,167	0
14	48,470	0	59	106,983	0	104	19,872	0
15	39,999	8	60	57,618	8	105	84,394	6
16	36,702	2	61	62,127	4	106	53,808	0
17	48,817	0	62	123,665	8	107	28,062	0
18	30,547	8	63	42,606	0	108	49,180	2
19	101,996	8	64	107,943	6	109	172,456	4
20	193,382	0	65	56,257	18	110	50,800	6
21	36,997	10	66	173,690	0	111	30,462	6
22	141,433	5	67	181,666	6	112	51,800	6
23	116,768	0	68	95,455	8	113	28,535	12
24	88,541	0	69	31,741	8	114	68,363	0
25	86,525	4	70	119,595	4	115	57,751	12
26	44,912	10	71	106,585	0	116	142,664	8
27	61,366	4	72	129,844	8	117	51,477	4
28	52,775	4	73	140,155	2	118	83,430	6
29	49,852	2	74	48,076	2	119	159,357	10
30	43,731	8	75	51,877	1	120	95,497	2
31	144,894	4	76	113,115	2	121	144,126	8
32	46,018	2	77	20,512	6	122	243,307	0
33	32,535	6	78	120,605	2	123	127,134	8
34	83,454	6	79	101,775	2	124	37,548	4
35	54,872	2	80	12,924	0	125	72,415	6
36	144,185	10	81	139,752	8	126	49,346	0
37	118,455	10	82	51,785	4	127	84,693	10
38	118,094	4	83	128,887	2	128	83,936	12
39	73,692	0	84	126,183	6	129	116,543	4
40	179,536	4	85	35,394	0	130	32,711	4
41	17,395	6	86	102,105	4	131	130,925	2
42	105,407	13	87	44,814	6	132	159,637	4
43	59,783	4	88	200,884	8	133	138,411	10
44	144,700	4	89	95,433	7	134	80,249	2
45	57,257	4	90	96,898	8			

B Mathematical Appendix

B.1 SMM Details

The estimation procedure is typical with exception for the construction of the weighting matrix. I construct the GMM with fixed centered weighting matrix. The moment condition in equation (??) can be rewritten as $E(\mathbf{h}_i) = \mathbf{0}$, where $\mathbf{h}_i = \mathbf{z_i} \cdot [\mathbf{m} - \hat{\mathbf{m}}]$. Assume that $\mathbf{h}_i$ is a martingale difference sequence with finite second moments. Then $\{\mathbf{h}_i\}$ has $E(\mathbf{h}_i|\mathbf{h}_{i-1}, \ldots, \mathbf{h}_1) = \mathbf{0}$ for $i \geq 2$. There is no serial correlation in $\mathbf{h}_i$ and the matrix of cross moments $E(\mathbf{h}_i\mathbf{h}_i')$ is non-singular. Hence I use the Gram-Schmidt process for the QR decomposition to get the matrix of variances of moment conditions and then the efficient weighting matrix is the inverse of moment conditions given by $\mathbf{W}_n^* = \mathbf{S}^{-1} = E(\mathbf{h}_i\mathbf{h}_i')^{-1}$.

B.2 Bike Transition Simulation Algorithm

Algorithm 1 Bike Transition Simulation Algorithm

Input initial guess values for parameters $\{\theta^1, \theta_i^2, \theta_j'^2, \theta'^3, \mu_1, \sigma_1, \mu_2, \sigma_2, \alpha\}$
Set initial bike distribution $\mathbf{g}^1$
set counter $t = 1, i = 1$
Require: $n_{ij}^t \geq 0, g_i^t \geq 0$
for $t = 1$ to T **do**
    for $i = 1$ to 134 **do**
        Draw λ_{ij}^t according to $\{\theta^1, \theta_i^2, \theta_j'^2, \theta'^3, \mu_1, \sigma_1, \mu_2, \sigma_2\}$
        Generate demand n_{ij}^t by Poisson parameter λ_{ij}^t
        Aggregate n_{ij}^t across j to get n_i^t
        Calculate matching probability p_i^t
        Obtain simulated matches m_i^t ▷ Random Binomial draw
        Compute bike transition $\mathbf{g}^{t+1}$ by equation (3)
    end for
end for

B.3 The Identification of α

Figure 13 below shows local minimum of the objective function with respect to α, with all other parameters fixed. The convex curve and the local minimum shows that α is identified.

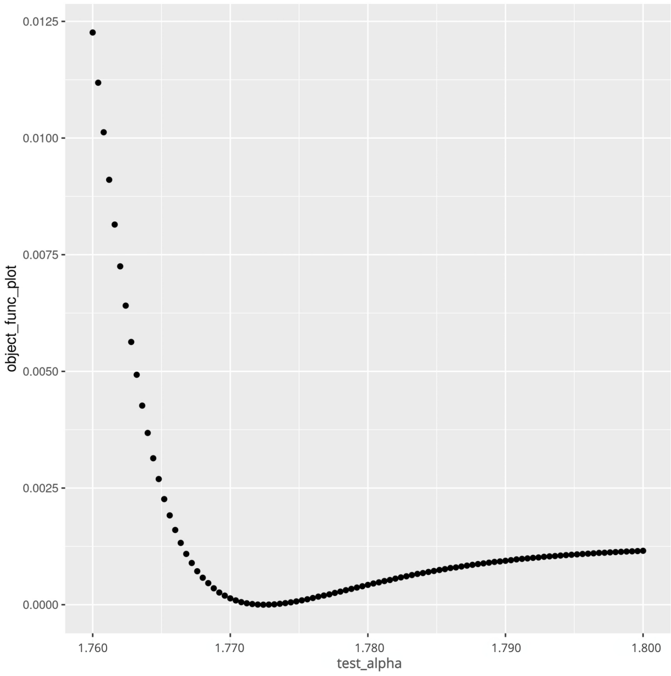


Figure 13: Objective Function vs. α.

Notes: The y-axis is the objective function value and the x-axis is the values of α, holding all other parameters constant. I plot 100 evenly spaced values for α from 1.76 to 1.80.

C Business Reference

Table C.1: ofo Funding and Valuation Timeline

Date	Round	Investors	Funding (million USD)	Valuation
03/17/2015	Angel	Will Hunting Capital	0.2	-
04/28/2016	Pre-A	Will Hunting Capital, Hongdao Capital	1.3	-
9/2/2016	A&B	Will Hunting Capital, Matrix Partners China, etc	at least 10	-
9/26/2016	B+	Matrix Partners China, GSR Capital, etc.	Undisclosed	-
10/10/2016	C	DiDi, Xiaomi, Citic PE, DiDi, etc.	130	1 billion USD
03/01/2017	D	DiDi, Citic PE, Xiaomi, DST, etc.	450	1.16 billion USD
07/01/2017	E	Alibaba, DiDi, Citic PE, etc.	700	3 billion USD
03/13/2018	E2-1	Alibaba, Ant Financial, etc.	866	$<$2 billion USD

Notes: The purpose of this table is to display an approximate timeline. The numbers and information are collected by the author from ofo's media releases.

(a) Congestion: Bikes Occupying Parking Spaces

(b) Vandalism: Bikes Thrown in the River

(c) Bike Graveyard: Top View

(d) Bike Graveyard: Front View

Figure 14: The Bike-Sharing Problem

Notes: The problems are caused by the reckless expansion of the bike-sharing service. Figure (a) shows that bikes were parked on the streets, taking the space from cars. Cars had to park next to the bikes and on the driving lanes, causing traffic problems. Figure (b) shows one form of vandalism. People threw bikes into the river, either for fun or getting rid of too many bikes. Figures (c) and (d) provide views of a "sharing-bike graveyard". Bikes were piled in layers in a large vacant land. The bikes were identifiable by their colors as to which company they belonged: yellow - ofo, orange - Mobike, and blue - Hellobike. These three companies were the largest bike-sharing service providers in China in 2017.