Understanding the operational dynamics of Mobility Service Providers: A case of Uber

XINWU QIAN, Lyles School Of Civil Engineering, Purdue University, USA
DHEERAJ KUMAR, Lyles School of Civil Engineering, Purdue University, USA
WENBO ZHANG, Lyles School of Civil Engineering, Purdue University, USA
SATISH V. UKKUSURI, Lyles School of Civil Engineering, Purdue University, USA

The rise of mobility service providers (MSPs) is reforming the traditional taxi service (TTS) market. MSPs differ from TTS with the core idea of using technology to optimally match riders with drivers, features like ride-sharing and surge pricing, and are not entry-regulated. It is of great significance to understand how MSPs operate, and how we can integrate them with TTS for efficient urban mobility. Unfortunately, little is known about MSPs due to limited data revealed by them. In this study, we collect and mine the trajectory data of online drivers who serve Uber (one of the largest MSP) to demystify how Uber drives their drivers. We analyze the trip patterns of different Uber services and reveal their market share, trip metrics, and the spatial distributions of trip origins and destinations. We explore how MSPs improve the driver-rider matching efficiency, and empirically validate the enormous efficiency gap between TTS and MSPs. In the end, we debunk the surge price as an instrument to restore driver-rider balance theory and show that drivers choose to chase or avoid the high surge areas depending on various other factors such as traffic congestion, time and location, and availability of alternate travel options as well. The results of this paper provide insightful knowledge about the supply side of MSPs and contribute to new ideas on improving TTS and regulating MSPs.

Additional Key Words and Phrases: Mobility service providers, Uber data collection, trip pattern, searching efficiency, chasing the surge

ACM Reference Format:
Xinwu Qian, Dheeraj Kumar, Wenbo Zhang, and Satish V. Ukkusuri. 2010. Understanding the operational dynamics of Mobility Service Providers: A case of Uber. *ACM Comput. Entertain.* 9, 4, Article 39 (March 2010), 21 pages. https://doi.org/0000001.0000001

1 INTRODUCTION

The fast expansions of *mobility service providers* (MSP) such as Uber, Lyft, and Didi have greatly challenged the monopolistic taxi market dominated by street-hailing taxicabs, and are reshaping urban mobility. According to the statistics by *New York city* (NYC) *Taxi & Limousine Commission* (TLC), the total daily ridership of the two largest ridesourcing platforms in NYC (i.e. Uber and Lyft) has approached the level of yellow taxicabs this year [10]. A similar tendency of demand shifting from *traditional taxi service* (TTS) to MSP is also observed in other major US cities, such

Authors' addresses: Xinwu Qian, Lyles School Of Civil Engineering, Purdue University, 550 Stadium Mall Drive, West Lafayette, IN, 47906, USA, qian39@purdue.edu; Dheeraj Kumar, Lyles School of Civil Engineering, Purdue University, 550 Stadium Mall Drive, West Lafayette, IN, 47906, USA, kumar299@purdue.edu; Wenbo Zhang, Lyles School of Civil Engineering, Purdue University, 550 Stadium Mall Drive, West Lafayette, IN, 47906, USA, zhan1478@purdue.edu; Satish V. Ukkusuri, Lyles School of Civil Engineering, Purdue University, 550 Stadium Mall Drive, West Lafayette, IN, 47906, USA, sukkusur@purdue.edu.

Manuscript submitted to ACM

as 30% in Austin [20], 22% in Boston [11], and as high as 65% for few taxicabs in San Francisco [5]. Three operational innovations by MSP are believed to contribute to the shift of passengers. First, the smartphone-based application connects passengers with drivers and narrows the information gap. Both sides are more transparent to the other, which saves time and cost spent for bilateral searching in the market. Second, without entry regulation, MSP breaks the financial barrier for TTS, which paired with *surge pricing* (SP) ensures appropriate supply to meet the various levels of demand. Finally, MSP offers various products, from ride-sharing (i.e. UberPool) to economy (i.e. UberX) to luxury (i.e. UberBlack or SUV), which satisfies the needs of different traveler groups.

It is these features that make MSP different from TTS, and understanding how MSP operates could provide insightful views on regulating and managing the emerging taxi market consisting of both street-hailing and ride-sourcing services. Current efforts in the literature to understand MSP consist mostly of theoretical analysis [16; 23], where key assumptions were made on critical behavior of drivers serving MSP due to limited data availability. For instance, it is often assumed that SP will encourage more drivers to join the market, and MSP is more efficient in helping drivers to find passengers than TTS. But neither of them are backed by real-world data and the strengths of such relationships are barely understood. This knowledge gap asserts the needs of an in-depth discussion by mining real life MSP data. This motivated us to collect MSP data to quantitatively investigate their performance.

Though publicly available MSP operational data is limited, there exist plentiful studies on empirical analysis of TTS, which serves same customer base and provides the comparison baseline for our study. On the aspect of TTS trip patterns, Qian et al. [17] explored the unbalanced distribution of taxicab trips, developed the statistical distribution of trip distance, correlated taxi rides with land use, and tagged featured taxi movements with trip purposes. Cai et al. [2] combined occupied and unoccupied taxi trips together as one integrated system and explored spatial and temporal regularities of travel time and travel distance in taxi rides. Regarding ridership of TTS, Zhang et al. [25] identified the influencing factors of temporal characteristics and built environment on traditional taxi ridership. Kamga et al. [9] summarized impacts of the time of the day, the day of the week, and weather condition on traditional taxi ridership. Yang and Gonzales [22] introduced an econometric model to specify taxi ridership, which assumed taxi ridership is normally distributed. Qian and Ukkusuri [15] explored spatial correlations of taxi ridership using geographically weighted regression model. Finally, to understand the rider-driver matching inefficiencies of TTS, Zhan et al. [24] modeled traditional taxicab matching efficiency and suggested that street-hailing taxicabs were far from optimal in finding passengers. Zhang et al. [26] discovered the potential influencing factors on empty trip duration of street-hailing taxicabs.

Through targeted release of trip data from some MSPs and self-efforts in data collection, there are few recent studies that shed some light on driver partners, operational efficiency, surge pricing, and passenger choices of MSP. Hall and Krueger [8] combined survey data with administrative data provided by Uber and explored socioeconomic status, work durations, and earnings of Uber driver partners across major US cities. Cramer and Krueger [4] compared the efficiency of traditional taxicabs with Uber through statistics of revenue hours and miles and concluded the better performance of Uber. Schwieterman and Michel [19] rode 50 trips each of UberPool and Chicago transit, and concluded that “UberPool was an attractive option for far more than extremely time conscious travelers”, but not to general commuters. Chen et al. [3] tracked Uber empty vehicle movement trajectory in middle Manhattan and discussed the impact of dynamic pricing on demand and supply. Guo et al. [6; 7] studied how passenger may respond to SP, measured occupied trip displacement, and investigated the spatial variations of surge multiplier.

In this study, we extend the current literature by analyzing driver behavior of MSP at the city-wide scale, investigating comprehensively the operations of a particular MSP - Uber, and quantitatively evaluate the virtual and the real of Uber from three aspects: trip pattern, matching efficiency, and drivers' response to surge prices. The main contributions of this study can be summarized as follow:

- This is one of the first few empirical studies on city-wide MSP operation using high-resolution trajectory data.
- We introduce data collection method and validated the quality of the data by conducting local field experiments to understand the mechanism on how our data were generated and comparing the inferred trips from our data with for-hire-vehicle trip record released by NYC Taxi and Limousine Commission (TLC).
- We discuss the market share, trip metrics, and the distributions of trip origins and destinations for different Uber services.
- We analyze the searching time of drivers serving MSP, disclose the huge gap in searching efficiency between MSP and TTS, and find that searching efficiency of MSP may be close to theoretical optimal.
- We investigate "chasing the surge" behaviour of drivers and find that such behaviour is not universal and a variety of other factors may affect drivers' reaction to price surge.

The rest of this paper is organized as follows: Section 2 presents the data collection and processing; Section 3 shows the operation patterns of various products and Section 4 compares the driver searching efficiency of MSP with that of TTS. Section 5 discuss the drivers' behaviour to surge price, before summarizing the empirical findings of our study in Section 6.

2 DATA

2.1 Data Collection

We developed a web crawler on Uber mobile platform[1], where we specify the target data collection locations and the crawler will return the trajectories of eight closest available drivers as well as the *surge price* (SP) and *estimated time of arrival* (ETA). We set our study area to be New York City and to cover the entire Uber fleet in our study area, we deployed sufficient data collection stations to emulate ride requests by passengers and exchange pingClient messages[2] with feedback messages from Uber server. The feedback message contains the trajectories of nearby Uber vehicles in the past 10 - 20s, and the recent SP and ETA. Our crawler performs this message exchange every few seconds for all online accounts. In this regard, our data collection was conducted in an ethical manner that neither hacked any driver or passenger privacy information nor sent real ride requests which may disturb Uber operations.

To assess the quality of our data collection, we set two different accounts at same data collection station and checked the similarity of feedback messages by sending a pingClient message at the same time. The results helped to determine the number of accounts needed at each data collection station to cover all nearby vehicles. We randomly deployed some data collection stations spread over the entire NYC area and exchanged pingClient message every 5 minutes for 12 consecutive hours. The test results indicated that over 99.99% of feedback messages between the two accounts were exactly the same. Therefore, we assigned only one Uber account for each data collection station. We conducted another set of experiments to identify appropriate spacing between two collection stations. Since each station may cover up to 8 closest vehicles, this process helps to place fewer number of data collection stations while ensuring sufficient coverage of Uber fleet. We used historical taxi demand distributions to divide the whole study area into three sub-regions based on the demand level as shown in Figure 3, with each region having different levels of spacing, varying from 100m to 1,500m between two data collection stations. We deployed 9 neighbouring stations (see figure 1) in each region and

[1] the web version of Uber for users without smartphone, ref: https://m.uber.com

[2] Each pingClient message contains the tentative ride request location, which is not an actual ride request.

summarize ratio of repeated vehicle observations by any two of all 9 stations for a 12-hour data collection. To avoid missing vehicle observations, we selected station spacing that resulted in high repeating ratio (≥ 40%). The experiment results are presented in Figure 2 and we finalized the station spacing for regions 1 to 3 as 600m, 1,200m, and 1,500m, respectively. After identifying appropriate station spacings, a total of 470 data collection stations were deployed as

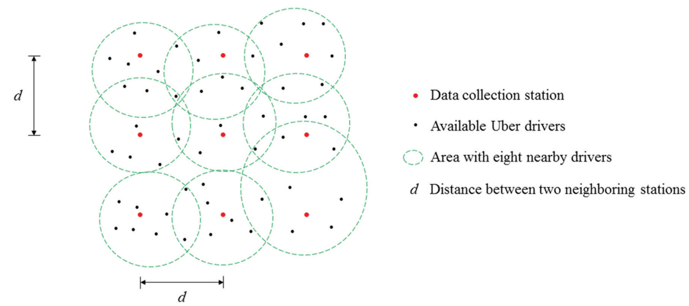
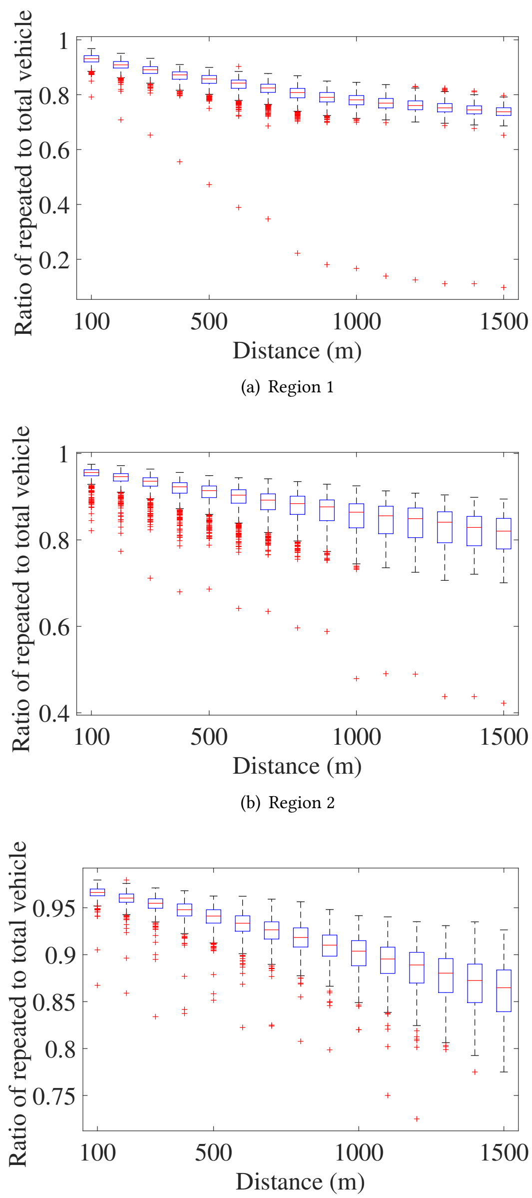


Fig. 1. Illustration of experiments with spacing d between neighboring stations

shown by the red dots in Figure 3, with region 1 having 99 data collection stations, region 2 having 119 data collection stations, and region 3 having 252 data collection stations. Data collection stations at the airport terminals (LaGuardia Airport and John F. Kennedy (JFK) International Airport) were set without considering the spacing level in region 2. Each station was associated with one account, sending a pingClient message every 10 seconds, and fetching feedback messages daily from 5 AM to 1:30 AM (next day) during April 7 to May 1, 2017. To better understand how data collected by our crawler were generated, we conducted additional field experiments with two registered Uber vehicles in West Lafayette, IN. The experiments focused on validating the consistency between the data we collected and the actual status of the vehicle (e.g, if the vehicle is online, serving passengers or offline). On one hand, we confirmed that the same vehicles showed up on both mobile application and the web platform, and verified that both platforms can capture our testing vehicles' offline and online behaviors. On the other hand, we observed that both vehicles were assigned with static vehicle id. The testing vehicles did multiple online and offline activities and their assigned ids remained the same after each online and offline activity. This ensured that we can track the activities of each vehicle following their assigned driver ID.

Two Uber datasets were extracted using our crawler: available vehicle trajectory data and SP & ETA data. Each trajectory record is a tuple of 6 elements: {Uber product id, vehicle id (driver id), epoch time, bearing (i.e. direction), latitude, longitude}. Trajectory records were updated every 2 to 5 seconds for each driver. The second dataset recorded pricing information during the surge period. Each record is a tuple of 7 elements: {data collection station, time, Uber product id, multiplier, minimum estimated waiting time, average estimated waiting time, surge pricing duration}. Besides Uber data, we also experimented on the datasets of yellow taxicab occupied trips during April 2013 [13] as the baseline case. The yellow taxi data recorded the information of trip start time and location, trip end time and location, metered distance, and charged fare. We extracted trip sequences of individual drivers and compare the performance of yellow cabs with Uber vehicles.



(a) Region 1

(b) Region 2

(c) Region 3

Fig. 2. The ratio of the number of repeated vehicles to the total number of observed vehicles for different levels of spacing

2.2 Data processing

The Uber trajectory data include the records of online drivers where online drivers refer to drivers who are available for accepting ride requests. And our crawler will lost track of drivers who go offline or are assigned with riders (either

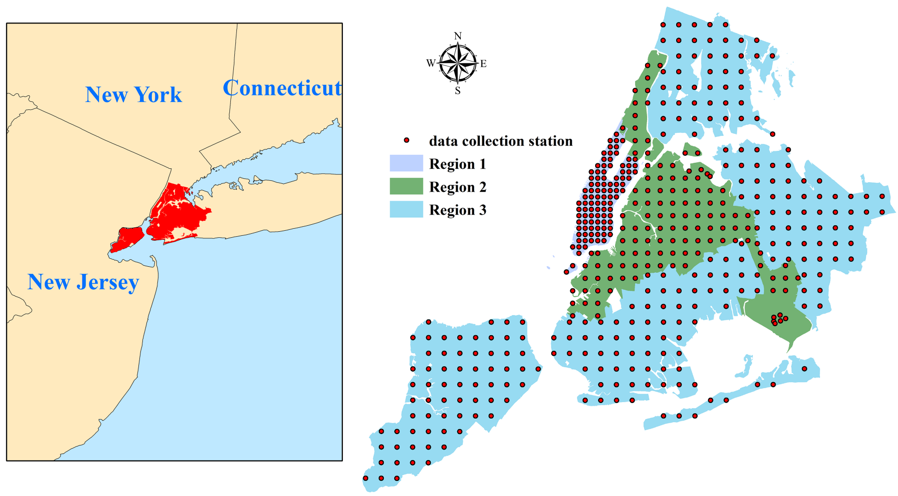


Fig. 3. Study area and the configuration of data collection stations

on their way to pick up riders or during serving passengers). To make best use of the trajectory data of online Uber drivers, we need to mine the data to further differentiate cruise (available to take passengers), trip (serving passengers), and offline (not available for serving passengers) states. For each driver ID, we arrange the data in ascending order with resepct to time epoch. For two consecutive records of each driver, denote Δt as the time lapse (in seconds) between the two time epochs and let Δd be the Euclidean distance (in meters) between two consecutive pairs of coordinates. We define cruise, trip, and offline status according to following rules:

- Cruise : $\Delta t \leq 60s$ OR ($60s < \Delta t \leq 7200s$ AND $\Delta d < 400m$)
- Trip : $60s < \Delta t \leq 7200s$ AND $\Delta d \geq 400m$
- Offline: $\Delta t \geq 7200s$

These criterion is selected based on our understanding of the data generation mechanism, the characteristics of Uber and taxi trips and our observations from reprocessing the trajectory data. Note that the crawler is able to track drivers' locations every 3-5 seconds, and we set Δt to 60 seconds in order to (1) allow buffer for delays in data transmission, (2) account for cases when the vehicle was not captured during certain time intervals and (3) based on the fact that each trip is unlikely to be shorter than 1 minute. As for Δd, the main purpose is to further filter out offline events which are not the result of passenger trips. Typical examples of such events include food/restroom/nap breaks and when drivers need to refill gas. During these events, the driver will be offline and are unlikely to incur large displacements between consecutive observations. On the other hand, the selected distance threshold should avoid false eliminating short trips such as last mile connections. These motivated us to choose 400 meters after processing and analyzing the vehicle trajectories. Finally, we also validate the identified Uber trips by comparing the results with the for-hire-vehicle (FHV) data released by NYCTLC [14]. In particular, the validation is performed for one week period from April 20 to April 26, 2017 and the results are shown in Figure 4. The TLC FHV data contain the information of Uber pickup time and location at taxi zonal level. Due to FHV data being highly spatial aggregated, we only compare citywide inferred number of trips with total number of trips recorded in FHV data for each 30-minute time interval from 6Am to 11:30PM. The total number of trips recorded in TLC FHV data was 1.77 million and we inferred 1.83 million trips, suggesting an

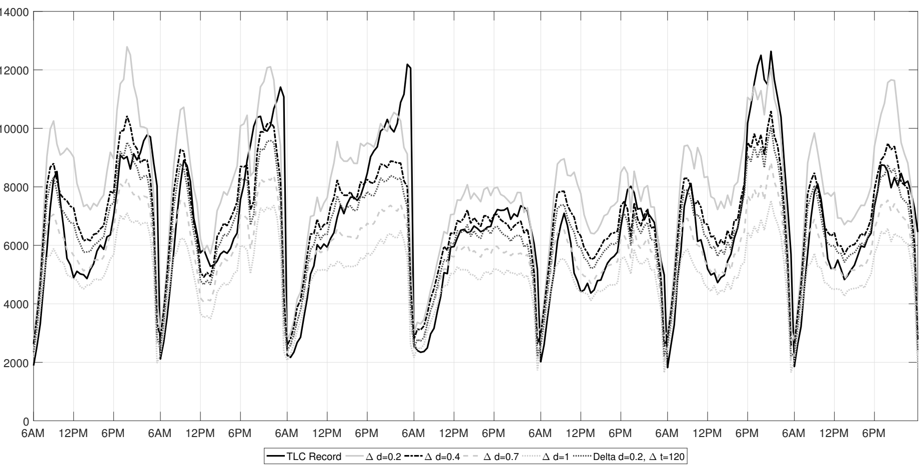


Fig. 4. Data Validation from April 20 (Thursday) to April 26 (Wednesday), 2017

overall overestimation of 3.3%. As for each 30-minute time interval, the inferred trips may slightly overestimate the recorded trips during certain low-demand periods and may underestimate the reported trips during some high-demand time intervals. We suspect the reasons for overestimating being that there are more drivers than passengers during the off-peak time period and hence the 470 data collection stations may fail to track all drivers at high frequency. This can be confirmed by increasing Δt from 60s to 120s which is observed to mitigate the overestimation during off-peak periods. Nevertheless, this also exacerbates the underestimations during peak hours. In general, we observe that 43% of the time intervals have the differences smaller than 10% and over 80% of the time intervals are of differences smaller than 20%, with the average difference per time interval being 8.3%. These results suggest that the selected criterion is able to restore the actual trip patterns with high accuracy and the criterion is therefore used for following analyses in this study.

2.3 Spatial and temporal aggregation

While vehicle trajectories were not directly applicable to understand performance metrics, we first aggregate the trajectories at proper spatial and temporal scale. We divide the study area into 2,557 grid cells of size 600m $\times$ 600m covering entire New York City. We then aggregated the trips within each grid cell in a 15 minute time interval to effectively visualize spatiotemporal trip distribution. However, since the spatial distribution is heavily biased and a significant amount of grid cells may have zero trips, it causes difficulties for further regression analysis. As a consequence, we further cluster the 2,557 grid cells into 100 groups. The method we used for clustering is the weighted K-means, where the weight is set as the number of trips in each grid cell. This helps to balance the number of trips within each cluster, in addition to grouping nearby grid cells.

3 ANALYSIS OF DIFFERENT PRODUCTS

There are a total of 9 Uber services (services) in NYC area. These along with a brief description are listed in Table 1.

Table 1. Summary of Uber services in NYC

Product	Capacity	Fare			Description
		Base ($)	$/miles	$/minute	
Black	4	7	3.75	0.65	Luxury sedan
SUV	6	14	4.5	0.8	Luxury SUV
X	4	2.55	1.75	0.35	Low-cost
Rush	NA	3	4	0	Delivery
XL	6	3.85	2.85	0.5	Low-cost SUV
Family	4	2.55	1.75	0.35	With baby seat
POOL1	2	2.55	1.75	0.35	Share and split
POOL2	2	2.55	1.75	0.35	Share and split (Manhattan only)
WAV	4	2.55	1.75	0.35	Wheelchair friendly

The drivers in the collected data are represented by a driver id. In total there are 4, 641, 896 (4.6 million) unique driver-ids. Many of these driver ids have very few data points. We remove those drivers who have less than 5 GPS records. This gives us a list of 4, 558, 833 unique driver ids. The number of unique driver ids who have serviced different products as a percent of total unique driver ids is shown as a 2D heatmap in Figure 5.

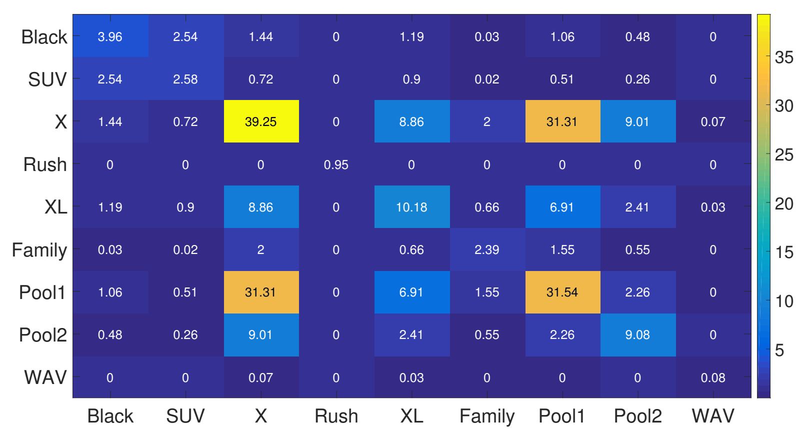
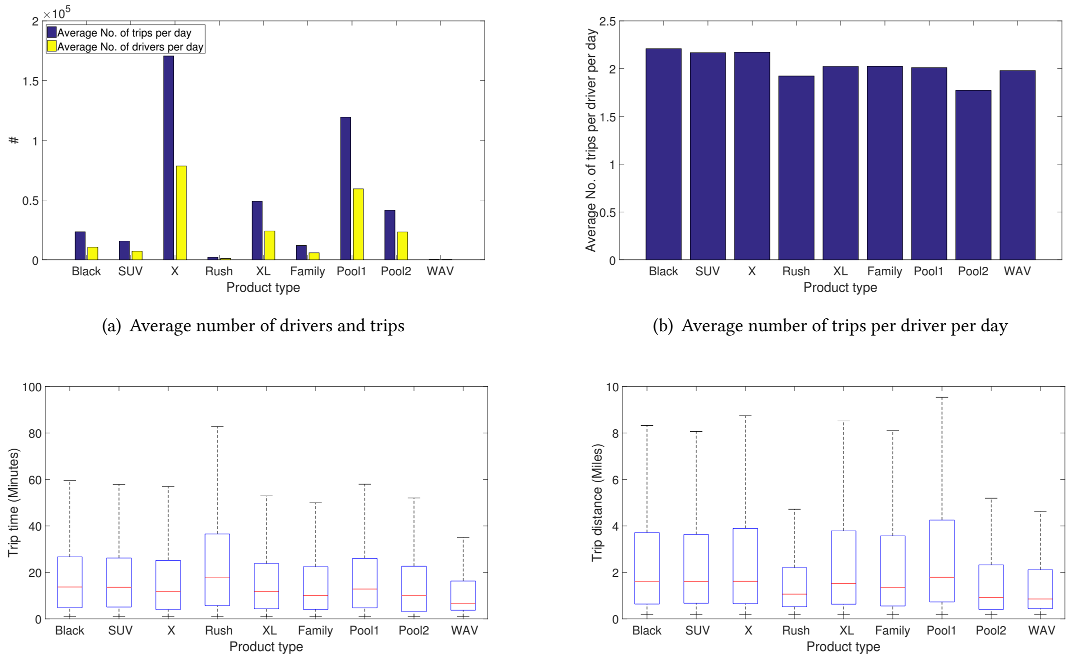


Fig. 5. Percent of all the unique drivers for each product and the common drivers

The sum of all diagonal entries in Figure 5 is more than 100, because many drivers were observed to be available for multiple Uber services. The off-diagonal entries in Figure 5 show the common driver among multiple services as a percent of all drivers. Almost 4 out of 5 drivers who service X or Pool1, service both; this ratio is 3 out of 5 for Black and SUV, indicating that most drivers are simultaneously logged on for multiple products for maximizing their chances of getting a passenger.

The disproportionate distribution of drivers among different services is also evident from the distribution of average number of trips and average number of drivers for all the products as shown in Figure 6(a).



(a) Average number of drivers and trips

(b) Average number of trips per driver per day

(c) Trip time for different products

(d) Trip distance for different products

Fig. 6. Trip metrics of different services of Uber in NYC

The order of popularity among various Uber services from most popular to least popular is X, Pool1, XL, Pool2, Black, SUV, Family, Rush, and WAV. However, the average number of trips per driver per day is almost same for all the products and is close to 2 as shown in Figure 6(b). To see how different Uber services perform in terms of trip time and distance, Figure 6(c)-(d) shows the boxplot of trip time and trip distance for different Uber services. The distribution of both trip time and trip distance has a sharp rise and a very long tail since the minimum value and the first quartile are closer to the median as compared to third quartile and maximum value. Uber Rush, which is a goods delivery service consists of both, bikes and motorized vehicles, hence different from other human mobility services and has the highest median value of trip time and lowest median trip distance indicating slower travel speeds as compared to other products.

To visualize the distribution of pick-up and drop-off locations for all the Uber services and contrast it with Yellocabs, the heatmap of pick-up and drop-off locations is shown in Figure 7.

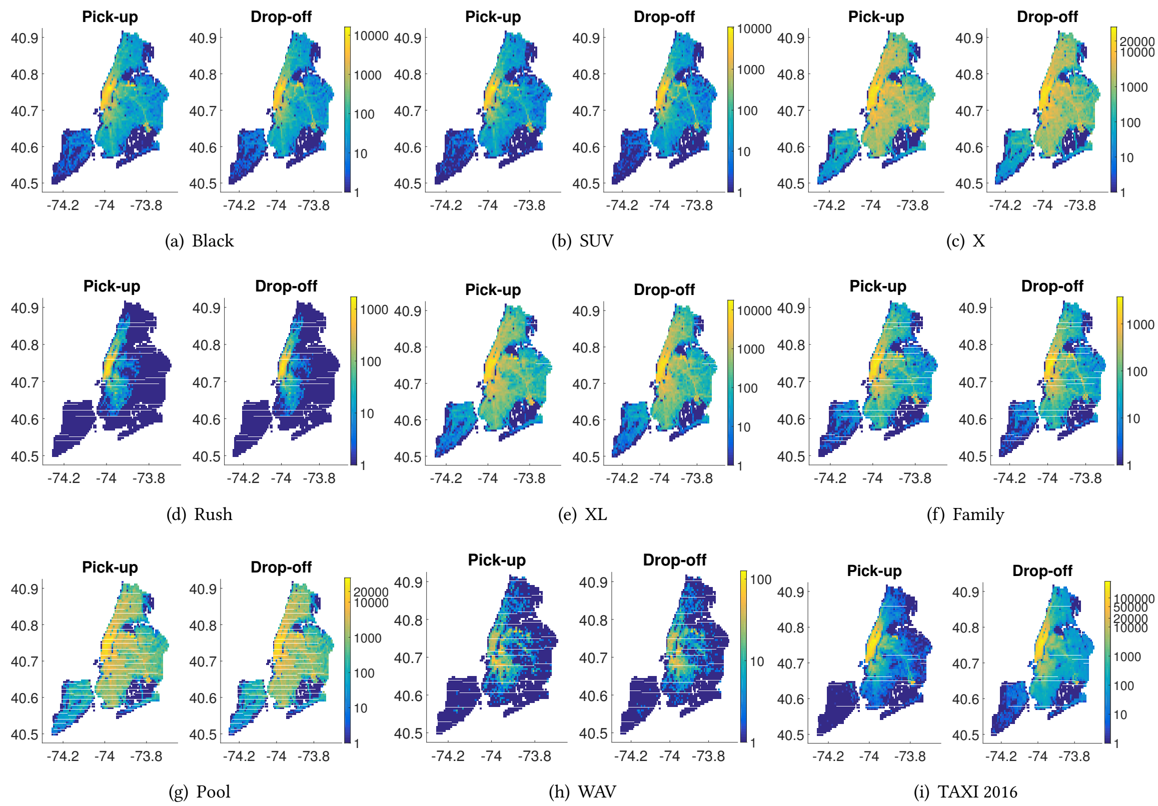


Fig. 7. Location heatmap of pick-up and drop-off locations for different Uber services and NYC Yellowcabs

For almost all the products, the highest intensity of pickups and drop-offs is found in downtown and midtown Manhattan and decreases rapidly as we move further away. The Ranking of products according to the homogeneity of distribution for pick-up and drop-off locations (from more to less) is {X, Pool (Pool1 & Pool2)}, followed by {XL, Family}, {Black, SUV}, and {Rush}. The spatial distribution of pickup and drop off locations for Uber X and Pool is more homogeneous spatially, i.e., evenly distributed in whole of NYC (See Figure 7(c, g)), as compared to the Uber Black, SUV and Rush, with pick-ups and drop-offs highly concentrated in the downtown and midtown Manhattan region, and not many trips originate or end in other parts of NYC (See Figure 7(a, b, d)). Among all passenger products, Uber X and Pool are the most popular and cheapest services and have a higher number of trips in boroughs other than Manhattan. XL and Family being specifically targeting larger groups and groups having young children are quite popular in boroughs other than Manhattan. Uber Black and SUV being the high-end premium services which are more expensive are mostly availed by people on business trips and hence are highly concentrated in Manhattan region and the two airports. WAV has a different pick-up and drop-off profile by location as the highest concentration of pick-up and drop-offs is found in Brooklyn instead of lower-middle Manhattan.

One striking feature of pick-up and drop-off location heatmaps for different Uber services as shown in Figure 7 is the high degree of similarity between pick-up and drop-off intensity for each product. This similarity implies that Uber drivers tend to wait (or move little) for the next ride at the current drop-off location without going back to high-intensity pick-up location (lower-middle Manhattan for most of the products). Contrasting it with the NYC Yellocab pick-up and drop-off intensity by location as shown in Figure 7(i) where the pickup is highly concentrated in lower-middle Manhattan, and the LaGuardia and JFK airports, however, the drop-offs are more homogeneously distributed across other regions as well.

Different products exhibit different serving, cruising, and trip time distribution patterns due to different customer base they cater to and different business model of the driver partners. Table 2 lists the 50, 75, and 90 percentile values for the serving, cruising, and trip times for different products. It can be observed from Table 2 that for a premier product

Product	Serving			Cruising			Trip		
	50	75	90	50	75	90	50	75	90
Black	65	159	180	9	22	50	20	42	74
SUV	70	164	286	9	22	52	20	39	74
X	44	125	221	4	9	17	20	46	81
Rush	57	135	242	17	42	85	31	59	90
XL	46	129	227	4	9	20	20	29	74
Family	55	146	258	7	13	20	17	39	72
Pool1	37	109	192	4	9	17	22	46	81
Pool2	24	76	140	4	9	15	22	46	81
WAV	55	144	258	11	20	31	13	37	72

Table 2. 50, 75 and 90 Percentile values (in minutes) for the serving, cruising and trip times

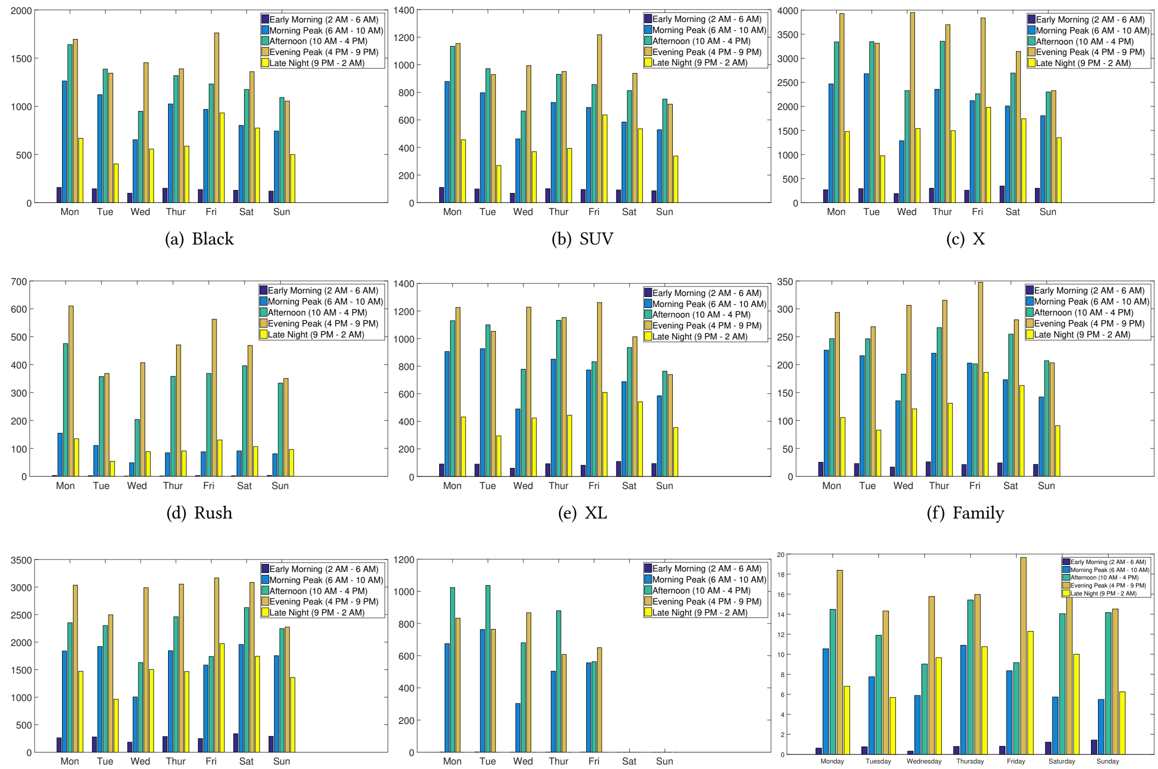
like Black and SUV, a majority of drivers serve for 65-70 minutes at a time, which is significantly higher than low-cost products such as UberX and Pool, suggesting that Uber Black and SUV are generally serviced by professional dedicated drivers having high end vehicles, whereas for UberX and Pool, there are many gig-economy part time drivers who find some time in their schedule to earn extra money. An even more drastic difference is observed in the cruising times of these two categories, where UberX and Pool drivers cruise very little before they either get a ride due to large demand or give up due to impatience suggesting novice experience as an Uber driver, however, professional drivers catering to Uber Black and SUV tend to cruise longer to get a ride. The trip time statistics are more homogeneous due to similar origin-destination pair rides for different categories of products.

3.1 Available vehicles by time and location

As expected for all modes of public transport, the availability of Uber vehicles is also a function of the time of the day and the day of the week based on demand levels. Figure 8 shows the histogram of average number of available vehicles per hour for different days of the week and for different time periods (early morning, morning peak, afternoon, evening peek, and late night). As a general trend, the highest no. of vehicles are available during evening peak (4 pm - 9 pm) for all the service types. Also, the no. of available vehicles during late night (9 pm - 2 am) is high during weekend (Friday, Saturday, and Sunday) as compared to weekdays.

4 SEARCHING EFFICIENCY

TTS is usually criticized for being inefficient in matching passengers with drivers. It is often time-consuming for both passengers and drivers to find each other. One significant advantage of MSP is that they use location-based services and



(a) Black (b) SUV (c) X

(d) Rush (e) XL (f) Family

(g) Pool1 (h) Pool2 (i) WAV

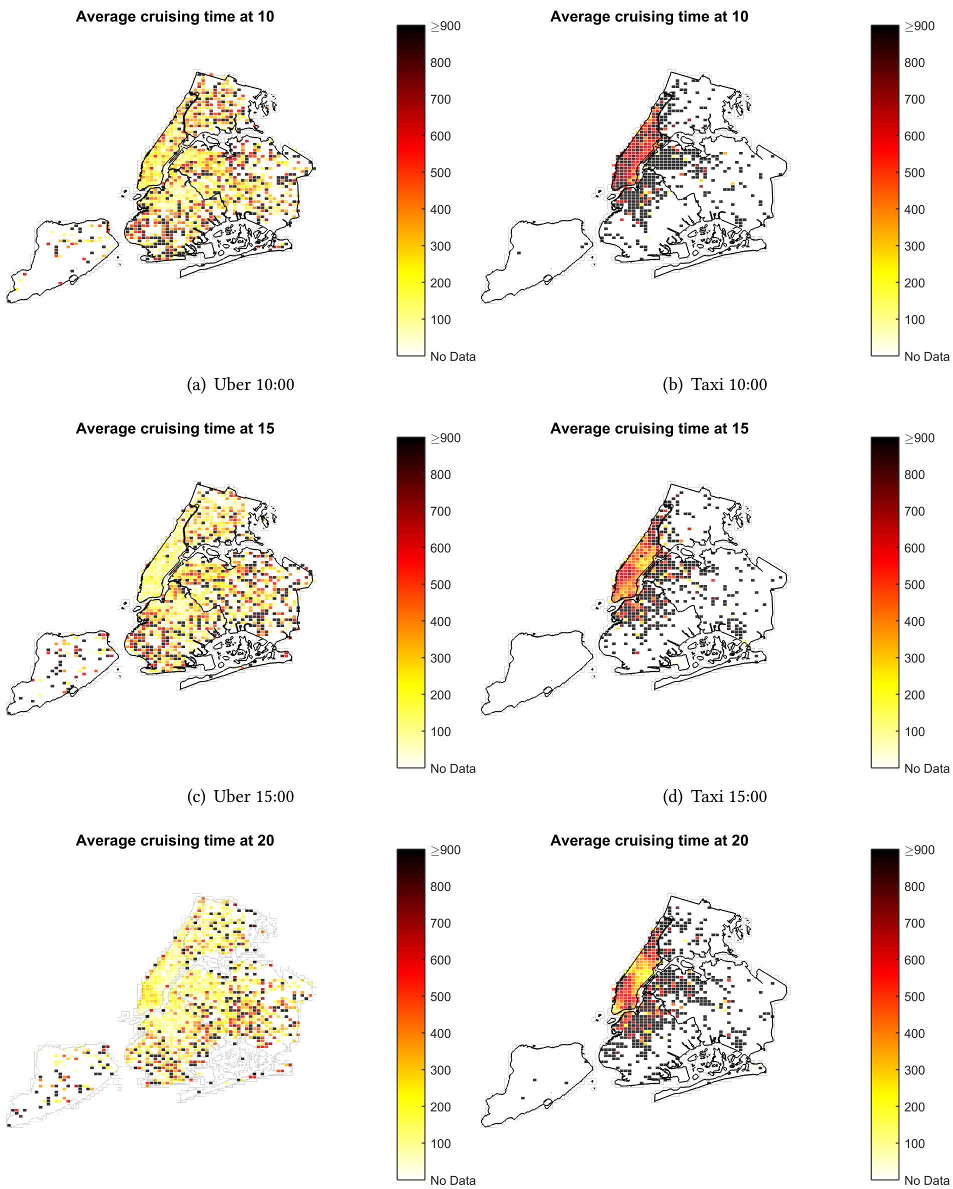
Fig. 8. Average no. of vehicles available per hour

ride-matching algorithms to address the inefficiencies in TTS. Theoretically, searching time for TTS may be reduced by 80 - 90% if drivers and passengers are optimally matched [24].

The analysis in this section focuses on UberX since it has the highest number of trips among all Uber services, however, it can be generalized to all the Uber services. We compare the searching time for UberX drivers with those of taxi drivers. For a fair comparison, we chose to use the April 2013 taxi data, where the market was barely affected by MSP due to Uber being banned during the time [21]. For a particular Uber vehicle, denote (x_t, y_t, t) as the tuple for each data record, the searching time can be measured as the accumulation time lapse between consecutive trip records and can be calculated as:

$$T = \sum_{i=0}^{N} t_{i+1} - t_i, \text{subject to } t_{i+1} - t_i \leq \Delta t \tag{1}$$

where $\Delta t = 60$ following the discussion in section 2, and the searching is considered as starting from the location corresponding to t_0 and terminated at location corresponding to t_N. As for taxi drivers, the searching time is tracked as the time gap between consecutive drop-off and pickup activities for taxi medallion ID and the searching is considered as starting from the drop-off location and ends at the next pickup location. We define searching time at location i during time interval t as the average searching time for all drivers who end their trips and start searching for passengers at location i within the time step t. Figure 9 shows the spatial variation of average searching time for UberX and yellow cabs on a typical Saturday (Uber data corresponds to Saturday, April 22, 2017, and taxi data corresponds to Saturday, April 20, 2013) across NYC. It is observed that the overall searching efficiency of Uber is much higher as compared to



(a) Uber 10:00 (b) Taxi 10:00

(c) Uber 15:00 (d) Taxi 15:00

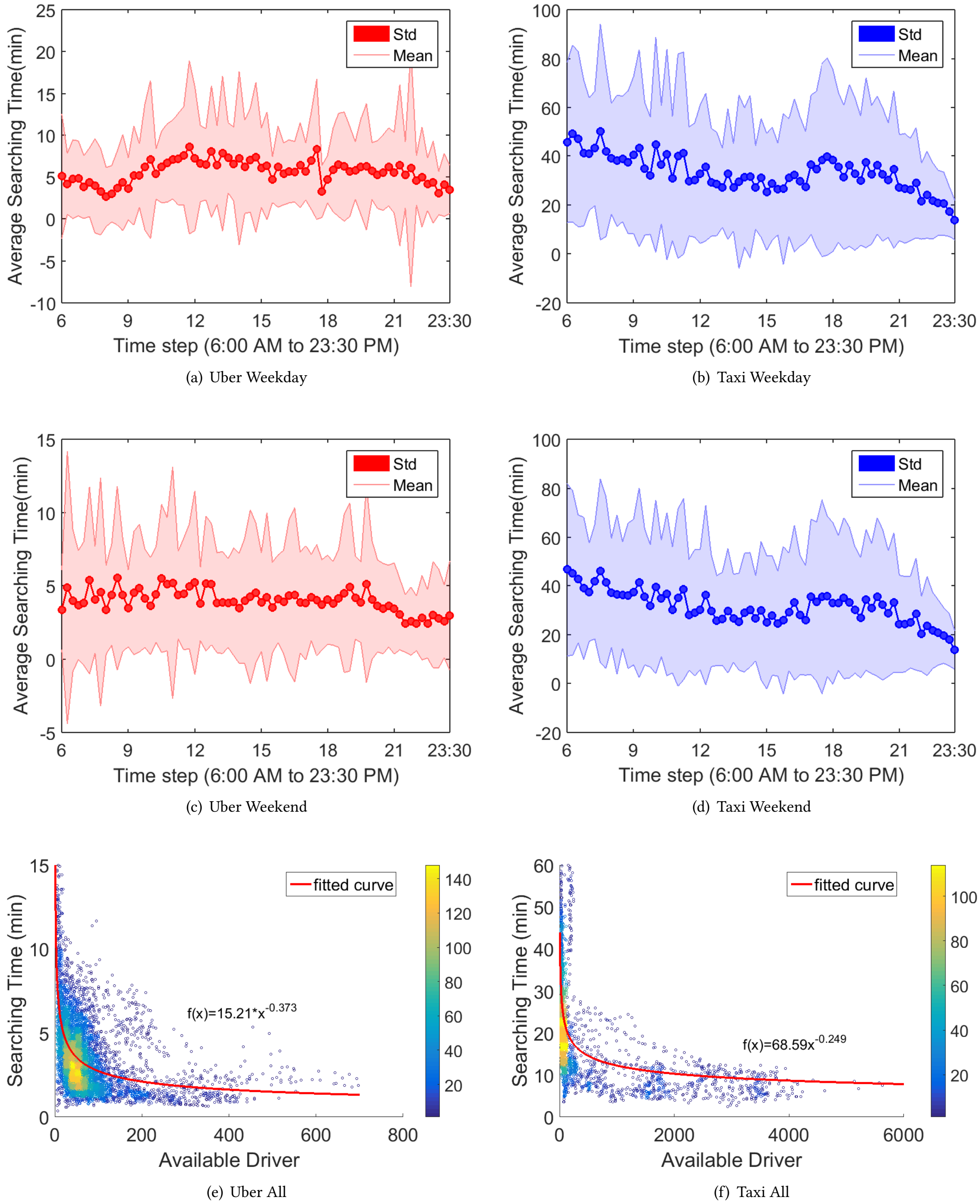
(e) Uber 20:00 (f) Taxi 20:00

Fig. 9. Spatial distribution of average searching time (in seconds) at different time of day

TTS (searching time for Uber being much smaller than that for yellow cabs) across the three time-snapshots: morning, afternoon, and evening. In particular, at 10 AM on a typical Saturday morning, there are fewer passengers than other time periods and taxi drivers find it hard to get riders, with searching time well over 10 minutes in most of the areas. During afternoon (3 PM) and evening (8 PM), only highly populated areas such as parts of lower middle Manhattan are getting better as the number of passengers increases over time. But the searching time for taxi drivers is still around 3-4 minutes at best. On the other hand, UberX drivers, on an average are able to find passengers in less than 5 minutes in the morning (10 AM), which reduces to 1.5 minutes in highly populated areas during afternoon and evening.

Next, we quantitatively assess the differences in searching efficiency between MSP and TTS over the entire study period. For this analysis, we group the 2,557 600m $\times$ 600m grid cells covering the entire NYC area into 100 clusters using weighted K-means clustering as described in Section 2.3. Figure 10 compares the average searching time (enveloped by standard deviation) across all 100 spatial clusters at different time points for MSP and TTS. It can be observed from Figure 10(a)-(d) that the average searching time and standard deviation for Uber are much lower than those of TTS. On an average, the searching efficiency improves by over 80% with the help of ride-sourcing platform, which is close to theoretical analysis results [24]. To eliminate the impact of varying number of drivers on searching time, Figure 10 (e)-(f) shows how searching time varies with the different levels of driver availability. We fit power functions to the data and find that the fitting results (R^2=0.289 for Uber and R^2=0.279 for the taxi data) are much better than other forms including linear, polynomial, and negative exponential models. In particular, the negative value of the power of the fitted function suggests that higher number of drivers results in lower average searching time. The fitting results indicate that the Uber is more efficient than TTS at all levels of driver availability (much lower searching times), which may be over 90% more efficient for low demand areas (which corresponds to low availability of drivers) and is approximately 50% more efficient even when drivers and passengers are densely distributed.

Although we observe that Uber has effectively reduced drivers' searching time, Uber drivers still cruise certain distance for a better chance of getting passengers and do not sit idle waiting for their next ride. Figure 11(a)-(b) shows the average searching distance of UberX drivers on weekdays and weekends respectively. Average cruising distance for UberX drivers on weekdays is in the range of 1-1.5 km, except during morning peak hours (7-10 AM), when the average cruising distance is 600m. During weekend, drivers were observed to cruise for a long distance (1.5 kms on average) during morning hours, however, during evening and late night time (2 PM - midnight) the cruising distance is quite short (close to 600m). In contrast, there is no significant difference for average searching time between weekday and weekend (Figure 10(a),(c)). Figure 11(c)-(f) shows the snapshots of spatial searching distance variation on a typical weekend (Saturday, April 22, 2017). This spatial distribution of searching distance corresponds to the same date and time as in Figure 9(a),(c), and (e). Comparison of the snapshots of searching distance and searching time shows striking similarities. The areas with higher searching time (e.g. >300s) also have longer searching distance (>2km). This implies that only few drivers wait at the drop-off locations, while the majority of them keep cruising until they receive a passenger from the Uber platform. Unfortunately, we are unable to conduct a comparison of the searching distance with TTS, due to lack of trajectory data for cruising taxis.

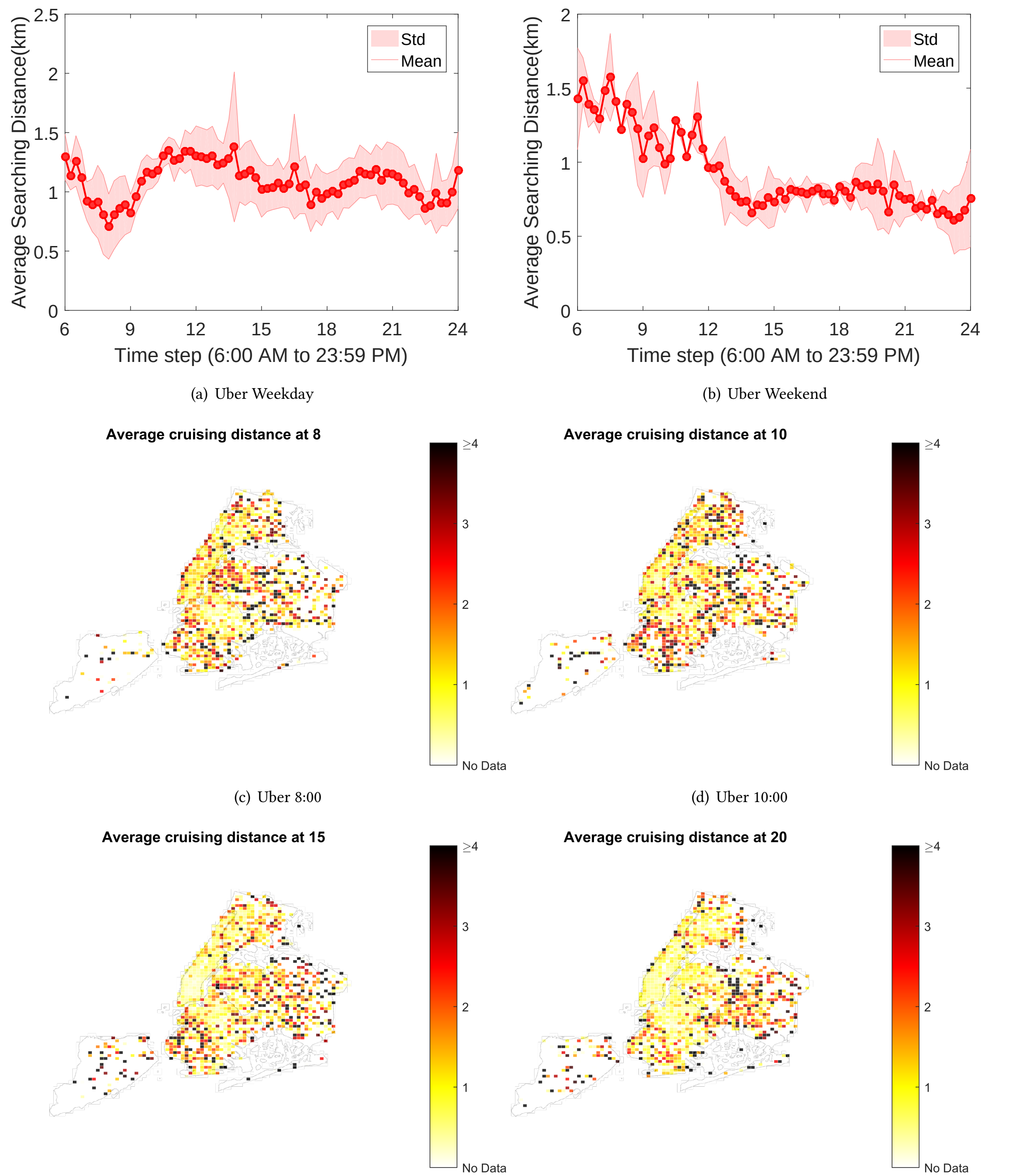


(a) Uber Weekday (b) Taxi Weekday

(c) Uber Weekend (d) Taxi Weekend

(e) Uber All (f) Taxi All

Fig. 10. Average searching time (in minutes) on weekday and weekend



(a) Uber Weekday

(b) Uber Weekend

(c) Uber 8:00

(d) Uber 10:00

(e) Uber 15:00

(f) Uber 20:00

Fig. 11. Average searching distance (in km) on weekday and weekend for UberX (a-b); snapshots of average searching distance (in km) on a typical Saturday (April 22, 2017) for UberX

5 CHASING THE SURGE

Another important feature that differentiates MSPs from TTS is the dynamic pricing being implemented in the form of surge price multiplier (SPM). When demand for rides outstrips the supply of drivers, MSP ride fares may increase to make sure those who need a ride can get one. MSPs claim that surge pricing has two effects: people who can wait for a ride often decide to wait until the price falls and drivers who are nearby can go to that neighborhood to get the higher fares. This helps restore the demand-supply balance and bring the prices back to normal. Efforts have been made to understand how SPM affect passengers' behavior and was found that higher surge price often deters passengers from making ride requests [3]. However, the effect of surge pricing on drivers' behavior is not understood well. While MSPs claim that higher surge lures more drivers by promising them higher fares, one report of SFGATE [18] claimed that most drivers ignore surge since by the time they get to that part of the city the surge is over. Even the drivers who are currently in the high surge area tend to leave the area and drive elsewhere because surge significantly lowers the demand. In this section, we model the underlying mechanisms of surge pricing on drivers' decision-making by analyzing the real-world Uber trip data.

A straightforward approach to investigate the issue is via regression analysis, e.g., using ordinary least-square (OLS) multiple regression model to understand the importance of independent variables on the dependent variable (the dependent and independent variables for the problem at hand are defined later). However, a crucial drawback of OLS is that all independent variables are assumed to be spatially homogeneous, which is unlikely to be the case for analyzing the effect of surge pricing on drivers' chasing-the-surge behavior. As a consequence, we introduce the geographically weighted regression model (GWR) [1], which was successfully used to understand the spatial variation of urban taxi ridership [15]. The model takes the following form:

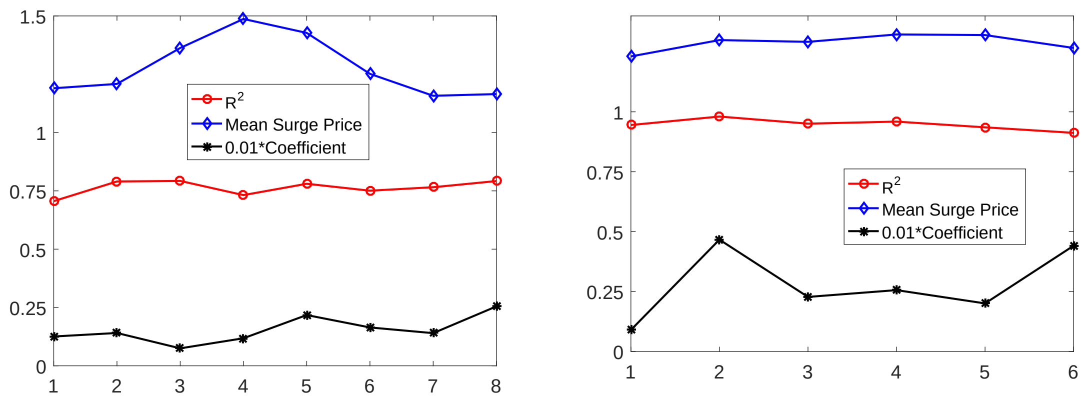
$$W_i y = W_i X \beta_i + \epsilon_i \tag{2}$$

where W_i is the n× n diagonal matrix which measures the distance-based weight between location i and all other locations. y is the n × 1 vector for dependent variables at all locations, X is the matrix of independent variable of dimension $n \times k$, where k is the number of dependent variables at each location. Finally, β_i is the coefficient vector for location i, and ϵ_i is the error term. We use the tri-cube weight function for calculating W_i, which gives the best model fitness as compared to other forms of weight functions:

$$W_i = \begin{cases} (1 - (\frac{d_i}{q_i})^3)^3, & \text{if } d_i \leq q_i \\ 0, & \text{other wise} \end{cases} \tag{3}$$

where q_i is the distance between location i and qth nearest neighbor, and d_i is the vector of distance between location i to all other places.

We set the independent variables to be {*SPM-1, no. of passenger trips*}. Using $SPM - 1$ helps to convert standard rate 1 to 0, which eliminates the bias when no surge is implemented. On the other hand, using the number of passengers as a variable account for the case where additional drivers are induced due to a higher possibility of finding a trip. We define the dependent variable as the *induced supply* of location i, which is the sum of the drivers who are observed to be online during the time at location i and the drivers who get attracted to the location at the time. The attracted drivers are those who find a trip in that particular location, but with a cruising distance higher than a certain threshold (2 kilometers in this study). As a consequence, these two values do not share common drivers and the sum of them does not include the nearby drivers who find a trip at location i. While SP is usually applied during peak hours, we focus on weekday morning peak (7:00-9:00) and weekend evening peak (19:00-21:00) to analyze chasing-the-surge behavior, and the results are presented in Figure 12.



(a) Weekday Morning (4-26-2017,7:00-9:00)

(b) Weekend Night (4-29-2017,22:30-24:00)

Fig. 12. Average surge price multiplier and the fitness of GWR model

It can be seen from the results that GWR fits both weekday and weekend data very well, with an average fitted R^2 being 0.857 and consistently being able to explain over 80% of the variability in the data. The two black lines with star mark show the average of $0.01 * coefficients$ across areas with $tstat \geq 1$. In general, the "chasing the surge" behavior is more obvious during a weekend night, and the reason is likely due to the better traffic condition and better driver availability during the night.

The most unexpected result of the GWR regression analysis is that the chasing-the-surge is not the universal strategy adopted by drivers (contrary to what MSPs expect others believe). Instead, such behavior is observed only at particular places during certain periods of time. For other locations and time intervals, drivers follow the opposite strategy: staying away from the surge. Figure 13 presents the contrasting behavior of drivers towards chasing-the-surge phenomenon at different times and locations. During weekday morning peak hours, the Manhattan region has high SPM (Figure 13(e)), however, the t-statistic value is smaller than -1 (Figure 13(c)) indicating drivers are trying to avoid this region. A possible reason for such a behavior could be high traffic congestion in Manhattan during morning peak hours. However, a totally different picture emerges during weekend night hours. Most places in Manhattan, Brooklyn, Queens, and Staten island have high SPMs (Figure 13(f)), moreover, the t-statistic for the chasing-the-surge hypothesis is observed to be statistically significant with over 99% confidence (≥ 1.96 as shown in Figure 13(d)). And drivers are observed to avoid remote areas as well as $theBronx$ during this time, where the violent crime rate is the highest among all NYC boroughs [12] and safety concern may therefore become one of the contributing factors. The GWR coefficient β_i is observed to be higher than 50 in some areas (Figure 13(b)), indicating that in average 10% increase in surge may attract more than 5 drivers. And the average value of β_i is 1.3 for weekday morning peak and 2.6 for weekday night, suggesting same increase in SPM during weekend night peak hours attracts twice as many drivers as compared to weekday morning peak hours. Two possible reasons for such a drastic difference could be higher availability of drivers during weekends and the lack of alternate travel options We may conclude from the results that chasing-the-surge is more frequently observed during weekend night than weekday morning. Drivers are observed to prefer residential areas with a higher surge and try to avoid congested and unsafe places even though the SPM is high.

6 CONCLUSION

In this study, we crawled the Uber mobile platform to collect the data of online driver trajectories. We explored the operational patterns of different Uber services and estimated their market share, trip statistics, and spatiotemporal trip distributions. We compare operational efficiencies of Uber with Yellow cabs and revealed the existence of a huge efficiency gap. Our analysis shows that the searching efficiency of MSP may be close to theoretical optimal, while for TTS lags far behind. These insightful findings improve our knowledge about the operation of MSP and contribute to framing new ideas on regulating and managing the market with both MSP and TTS.

REFERENCES

[1] Brunsdon, C., Fotheringham, S., and Charlton, M. Geographically weighted regression. *Journal of the Royal Statistical Society: Series D (The Statistician) 47*, 3 (1998), 431–443.

[2] Cai, H., Zhan, X., Zhu, J., Jia, X., Chiu, A. S., and Xu, M. Understanding taxi travel patterns. *Physica A: Statistical Mechanics and its Applications 457* (2016), 590–597.

[3] Chen, L., Mislove, A., and Wilson, C. Peeking beneath the hood of uber. In *Proceedings of the 2015 ACM Conference on Internet Measurement Conference* (2015), ACM, pp. 495–508.

[4] Cramer, J., and Krueger, A. B. Disruptive change in the taxi business: The case of uber. *The American Economic Review 106*, 5 (2016), 177–182.

[5] GARBER, M. After uber, san francisco has seen a 65% decline in cab use, accessed June, 2017. Available online at https://www.theatlantic.com/technology/archive/2014/09/what-uber-is-doing-to-cabs-in-san-francisco-in-1-crazy-chart/380378/.

[6] Guo, S., Liu, Y., Xu, K., and Chiu, D. M. Understanding passenger reaction to dynamic prices in ride-on-demand service. In *Pervasive Computing and Communications Workshops (PerCom Workshops), 2017 IEEE International Conference on* (2017), IEEE, pp. 42–45.

[7] Guo, S., Liu, Y., Xu, K., and Chiu, D. M. Understanding ride-on-demand service: Demand and dynamic pricing. In *Pervasive Computing and Communications Workshops (PerCom Workshops), 2017 IEEE International Conference on* (2017), IEEE, pp. 509–514.

[8] Hall, J. V., and Krueger, A. B. An analysis of the labor market for uberâĂŹs driver-partners in the united states. Tech. rep., National Bureau of Economic Research, 2016.

[9] Kamga, C., Yazici, M. A., and Singhal, A. Hailing in the rain: temporal and weather-related variations in taxi ridership and taxi demand-supply equilibrium. In *Transportation Research Board 92nd Annual Meeting* (2013), no. 13-3131.

[10] New York City Taxi and Limousine Commission (NYCTLC). Fhv base aggregate weekly report, accessed June, 2017. Available online at http://www.nyc.gov/html/tlc/html/technology/aggregated_data.shtml.

[11] Newsham, J., and Adams, D. Amid fight with uber, lyft, boston taxi ridership plummets, accessed June, 2017. Available online at https://www.bostonglobe.com/business/2015/08/19/boston-taxi-ridership-down-percent-this-year/S9dZMELMye6puzTTYoDIrL/story.html.

[12] NYC Mayor's Office of Crime Justice. Violent crime rate by borough, accessed May, 2019. Available online at https://criminaljustice.cityofnewyork.us/individual_charts/violent-crime-rate-by-borough/.

[13] NYCTLC. TLC Trip Record Data, accessed June, 2017. Available online at www.nyc.gov/html/tlc/html/about/trip_record_data.shtml.

[14] NYCTLC. 2017 TLC FHV Record Data, accessed June, 2019. Available online at https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page.

[15] Qian, X., and Ukkusuri, S. V. Spatial variation of the urban taxi ridership using gps data. *Applied geography 59* (2015), 31–42.

[16] Qian, X., and Ukkusuri, S. V. Taxi market equilibrium with third-party hailing service. *Transportation Research Part B: Methodological 100* (2017), 43–63.

[17] Qian, X., Zhan, X., and Ukkusuri, S. V. Characterizing urban dynamics using large scale taxicab data. In *Engineering and Applied Sciences Optimization*. Springer, 2015, pp. 17–32.

[18] Said, C. Report says uber surge pricing has a twist: some drivers flee, accessed June, 2017. Available online at www.sfgate.com/business/article/Report-says-Uber-surge-pricing-has-a-twist-some-6597012.php.

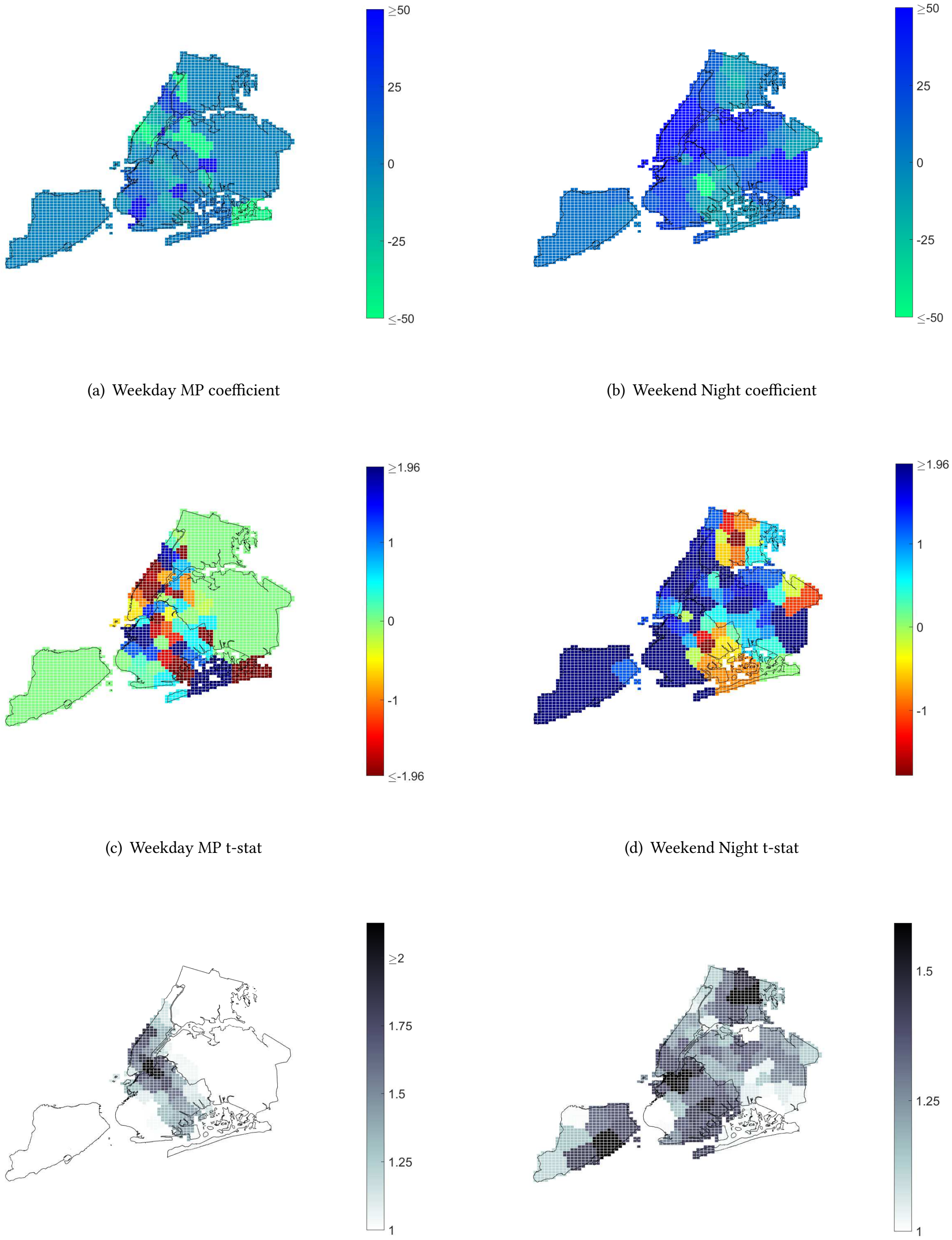
[19] SCHWIETERMAN, J. P., and MICHEL, M. Have app will travel: Comparing the price and speed of fifty cta and uberpool trips in chicago, 2016.

[20] Thesis, M. The uber effect: Austin taxi rides drop dramatically in past year, accessed June, 2017. Available online at https://www.bizjournals.com/austin/news/2016/01/19/the-uber-effect-austin-taxi-rides-drop.html.

[21] Uber Blog. Ubertaxi returns to nyc, accessed June, 2017. Available online at https://www.uber.com/blog/new-york-city/ubertaxi-returns-to-nyc/.

[22] Yang, C., and Gonzales, E. Modeling taxi trip demand by time of day in new york city. *Transportation Research Record: Journal of the Transportation Research Board*, 2429 (2014), 110–120.

[23] Zha, L., Yin, Y., and Yang, H. Economic analysis of ride-sourcing markets. *Transportation Research Part C: Emerging Technologies 71* (2016), 249–266.



(a) Weekday MP coefficient

(b) Weekend Night coefficient

(c) Weekday MP t-stat

(d) Weekend Night t-stat

(e) Weekday MP surge distribution

(f) Weekend Night surge distribution

Fig. 13. Spatial distribution of average searching time at different time of day

[24] Zhan, X., Qian, X., and Ukkusuri, S. V. A graph-based approach to measuring the efficiency of an urban taxi service system. *IEEE Transactions on Intelligent Transportation Systems 17*, 9 (2016), 2479–2489.

[25] Zhang, W., Qian, X., and Ukkusuri, S. V. Identifying the temporal characteristics of intra-city movement using taxi geo-location data. *Enriching Urban Spaces with Ambient Computing, the Internet of Things, and Smart City Design* (2016), 68.

[26] Zhang, W., Ukkusuri, S. V., and Lu, J. J. Impacts of urban built environment on empty taxi trips using limited geolocation data. *Transportation* (2016), 1–29.