Title
Mind the Gap: a Case Study of Demand Prediction and Factors Affecting Waiting Time at Uber NYC and Washington D.C.

Author
Ye, Tianyu

Publication Date
2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Mind the Gap: a Case Study of Demand Prediction and Factors Affecting Waiting Time at Uber NYC and Washington D.C.

A thesis submitted in partial satisfaction

of the requirements for the degree Master of Science

in Statistics

by

Tianyu Ye

2019

ABSTRACT OF THE THESIS

Mind the Gap: a Case Study of Demand Prediction and Factors Affecting Waiting Time

at Uber NYC and Washington D.C.

by

Tianyu Ye

Master of Science in Statistics

University of California, Los Angeles, 2019

Professor Ying Nian Wu, Chair

As the ride-hailing economy has been rapidly growing, case studies of Uber's forecasting and pricing frameworks have received much attention among researchers and investigative journalists. This paper 1) attempts to forecast Uber's fulfilled pickups via predictive modelling, and 2) conducts a retrospective study to examine the role of socio-economic factors and surge pricing on riders' waiting time along with their implications. By seeking to answer questions regarding these two topics, this paper hopes to shed light on the demand-supply dynamics of

Uber service, and provides empirical analysis and potential suggestions for future improvement. The scope of this paper focuses on New York City and Washington D.C., where public data is available. This paper evaluates the performance under various predictive statistical learning models based on the lowest error rate (RMSE being the major criteria), while the explanatory modelling section adopts more traditional statistical tools, as well as a Bayesian network. The results suggest that LASSO has outperformed other models used in this thesis for the demand forecast. Moreover, there is a lack of evidence that surge pricing has a consistent effect in reducing riders' average waiting time. Furthermore, various socio-economic factors are found to have significant associations with waiting time, and in particular, districts with higher ethnic minority rates are associated with longer average waiting time, which calls for further examination on the dynamics of Uber's service delivery.

The thesis of Tianyu Ye is approved.

Chad J. Hazlett

Juana Sanchez

Ying Nian Wu, Committee Chair

University of California, Los Angeles

2019

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

According to The New York City Taxi and Limousine Commission (NYC TLC), at the end of December 2017, For-Hire Vehicles (FHVs) including Uber and Lyft provided 1.65 times more pickups than taxis in New York City; in NYC's outer boroughs, those FHVs enjoyed 10 times as much higher usage as taxis (Schneider, 2015). Indeed, the on-demand gig economy has become a part of our lives, at least in the peer-to-peer ride-sharing market, and Uber in particular has been the largest player in this market so far. Although its market share has been slightly declining because of Lyft and other new competitors' fierce service and pricing competition, Uber still has more than 65% of market share in NYC in 2018 (Schneider, 2015), and similar patterns could be observed in other major cities and throughout the United States.

The demand-supply matching and pricing algorithm has long been at the crux of Uber’s profitability and success in daily operation. Notably, Machine learning (ML) and Artificial Intelligence (AI) have been “critical to supporting Uber’s mission of developing reliable transportation solutions,” whether from “a Bayesian neural network architecture that more accurately estimates trip growth, [...] a real-time features prediction system,” or Uber’s internal machine learning platform Michelangelo (Turakhia, 2017, para. 1). Essential to the interest of the company, researchers, and users, the actual algorithms for supply-demand forecast and pricing generation, however, have long been a black box to the public. Nonetheless, cooperating with Uber or independently using Uber API and Client Applications to gather information for analysis, researchers and investigative journalists have made considerable progress in unveiling

Uber's algorithm.

Passenger pickup volume forecast is an ideal question for statistical learning, and by forecasting the supply-demand gap in real time, ride-hailing companies can balance the supply-demand gaps for profit and service quality optimization. Notably, economic, geographical, temporal, social and natural factors could all have potential influences on the supply-demand equilibrium (i.e. the *fulfilled* number of pickups).

Another prominent feature of Uber's pricing scheme is the dynamic pricing (also called the surge pricing). At its basic level, the pricing tool should be a function of the supply-demand curve. Uber uses this price multiplier to shift the points on the supply and demand curve to achieve a smaller gap between the two. It is then natural to ask: effective at least in theory, to what extent does surge pricing actually help to solve the demand-supply re-balance problem in order to improve service quality?

Yet another interesting question also arises: does service quality such as wait time have systematic variance across geographic areas with different socio-economic features? In addition to affecting the company's business operation and users' welfare and experience, the answer to these questions concerns researchers, policymakers, and investigative journalists analyzing the business and data transparency of Uber.

Motivated by the above narratives and past literature, this thesis attempts to study 1) the fulfilled demand prediction and 2) the relationship between wait time and surge pricing and socio-economic factors. In particular, building on previous literature, this paper will use more updated data and various statistical learning methods to enhance the forecast and explanatory power of the study. The outline of the paper is as follows: Section 2 will render a literature review on current studies of the ride-sharing economy, in particular on Uber's demand forecast and service quality examination. Section 3 will describe the dataset and Section 4 will provide a summary of methodology. Section 5 will conduct preliminary data exploration and identify notable patterns. Section 6 will provide analysis where predictive models with various state-of-the-art methods will be applied on the demand forecast, and regression, Bayesian networks, and correlation tools will be adopted to examine the explanatory effects of surge price and socio-economic factors on service quality, of which wait time serves as a proxy. Section 7 will conclude the thesis with some discussion, as well as suggestions for potential future work.

2 LITERATURE REVIEW

2.1 Uber Demand Forecast

Statistical and machine learning methods are at Uber's core to address its marketplace forecasting, hardware capacity planning, and marketing. Some popular models in development include ARIMA and exponential smoothing methods, as well as machine learning models such as Recurrent Neural Networks (RNNs) for time-series forecasting (Laptev, Yosinski, , Li, & Smyl, 2017) and hybrids algorithm that combines Holt-Winters method and long short-term memory units (Bell, F., & Smyl, 2018).

In addition, researchers cooperating with Uber have examined the social welfare of Uber's service. Cohen et al. (2015) suggested that UberX alone has contributed $2.9 billion in consumer surplus in New York City, San Francisco, Chicago, and Los Angeles in 2015. Chen and Sheldon (2016) conducted empirical studies with UberX data in Chicago, Washington D.C., Miami, San Diego, and Seattle, concluding that "Uber partners both drive at times with higher demand for rides, and dynamically extend their sessions when surge pricing raises earnings" (p. 13), which "supports the idea that dynamic pricing significantly increases the efficiency of on-demand service markets" (p. 15).

Meanwhile, non-Uber-affiliated agencies and researchers have used public and API data or client applications to probe into Uber's ride-share patterns. New York City Taxi and Limousine Commission (NYCTLC) has taken advantage of data to conduct various analyses and visualizations: the agency is the first to have concrete numbers showing that at the end of 2016 in NYC, High Volume For-Hire Services (HVFHS) trips, which include Uber, Lyft, Juno and Via, began to outnumber taxi services, and the gap has been rapidly expanding (Schmidt, 2018).

In Schnider (2015)'s empirical study, while extreme weather such as hurricanes and blizzards has a significant impact on Taxi and Uber ridership, rain does not appear to have much influence. Brodeur and Nield (2018) also used the NYCTLC data to study the effects of rain and surge price on Uber supply and fulfilled pickup. On the contrary, they have concluded that Uber pickups increase by 22% during the raining season; additionally, traditional taxi rides, passengers and fare income has all declined since Uber's introduction to the market. In addition, using data collected by San Francisco County Transportation Authority (SFCTA), Charlton

(2016) has illustrated temporal and geographical trends through 2D and 3D visualization of TNC activities in downtown SF (e.g. volume, flow, and price dynamics).

Researchers have also been exploring various algorithms to predict the fulfilled pickup demand. Faghih et al. (2017) assessed the performance of a new spatio-temporal prototype composed of least absolute shrinkage and selection operator applied on spatial-temporal autoregressive (LASSO-STAR), along with the inclusion of proximate areas' Uber demand information and weighting matrix. The paper found the approach outperformed basic STAR models and several other benchmark time-series models. Wang et al. (2018) attempted with deep learning Long Short-Term Memory (LSTM) Networks model and found it has outperformed Poisson model and regression tree model in real-time Uber demand prediction. In addition, Chen, Ampountolas, and Thakuriah (2019) developed a multi-level Wavenet-based deep neural network for hourly Uber pickup forecast.

2.2 Effects of Surge Price and Socio-economics on Uber’s Service Quality

Chen and Sheldon (2016) found Uber’s dynamic pricing tool has "large and pervasive positive supply elasticities" (p. 15), using economics frameworks and discontinuity design to support the causality of this relationship. Chen, Mislove, and Wilson (2015) exploited Uber API and the Uber Client App to assess the effect of surge pricing on users (riders and drivers). Having collected data for about one month in downtown San Francisco and midtown Manhattan, the empirical analysis suggested that the surge price has "a strong, negative impact on passenger demand, and a weak, positive impact on car supply" (p. 13), along with many other dynamics of the surge price in its cost, timing, and locations. Using the information on estimated demand, supply, and waiting time, the paper also applied regression models to forecast the surge. However, the study concluded that even with considerable supply and demand information, surge multiplier forecast has rendered poor performance and was thus not reliable.

In cooperation with SFCTA, Castiglione et al. (2016) used the same dataset to estimate the lower bound supply and demand of UberX service in downtown SF. Subsequently, Jiang, Chen, Mislove, and Wilson (2018) further extended the study to provide "the first head-to-head comparison of Uber, Lyft, and taxis" (p. 9) in downtown San Francisco. The study revisited the surge price's relationship with geo-spatial, demographic, and socio-economic indicators, and 1) introduced Spatial Descriptive Statistics such as Bivariate Ripley's K Function to bolster the credibility of the study's estimation, and 2) conducted competition studies between Uber and Lyft to differentiate the pricing algorithms between the two companies.

Diakopoulos (2015) studied Uber's business operation in Washington D.C. and suggested that on the city level, higher surge price seems to reduce average waiting time, which in turn can be regarded as a proxy of service quality. Nonetheless, the study also found several counter-evidences that eroded the effectiveness of surge pricing tools. Some notable reasons are: 1) the drastic change in surge prices discouraged drivers to rely on the pricing tool, and 2) "rather than getting more [new] drivers on the road in the short-term, Uber's surge pricing instead depletes drivers in adjacent areas" (Diakopoulos, 2015, para. 16). The author then used U Street and Navy Yard as an example to support his claim.

In an extended analysis, Jennifer and Diakopoulos (2016) looked into the correlation between wait time and racial as well as economic compositions in a neighborhood. The study found that areas with higher non-white resident rates or poverty rates are correlated with longer waiting time, even after controlling other demographic and socio-economic factors.

In addition, Ge, Knittel, MacKenzie, and Zoepf (2016) looked into UberX and Lyft drivers' response in accepting or canceling trips after perceiving African American riders' profiles, and their longer wait time is "robust across numerous model specifications" (p. 19), and in "extreme cases drivers are more than four times as likely to cancel on an African American male passenger than on a white male passenger […] This discrimination is not the result of any policy by ride hailing providers, but rather the behavior of individual TNC drivers" (p. 19). In response to the findings, the authors then offered several policy suggestions to the ride-hailing companies.

3 DATA

For fulfilled demand prediction, I use Uber data made available by NYCTLC, which regulates taxis and for-hire vehicles. Before 2015, NYCTLC only collected the yellow and green taxi trips, including their pick-up and drop-off times, trip distances, and fares. In 2015, NYCTLC began to release Uber and Lyft data. In this thesis, the training data spans from April 11th, 2015 to August 23rd, 2016, and the testing data is from August 24th, 2016 to September 27th, 2017. Although NYCTLC provides Uber's information since January 2015, after experimenting with the dataset, the pickups before April 11th, 2015 had very low volume and very slow growth, which is not likely representative to the booming Uber business since late 2015. Thus, I excluded the first three months of 2015 data for better trend projection.

The variables, along with their descriptions can be referred to in Table A.1 in the Appendix section. Moreover, I have included the weather and geographic metrics. The information is obtained from Weather Underground, the historical weather data records by National Centers for Environmental Information (NOAA). The weather data was then pre-processed to match the format of NYCTLC's location and time categorization. In the end, the dataset is consolidated to 1000 daily data points.

To answer the second part of my question, I use the dataset obtained by Nicholas Diakopoulos' team, which uses Uber API to collect wait time, surging price and pickup/dropoff locations in 276 locations in Washington D.C. every three minutes (Diakopoulos, 2016). The data spans from November 14th to November 28th, and from December 8th to December 26th, 2016. The geographic traits are then converted into corresponding 2010 census tracts. Also provided as readily available by Nicholas Diakopoulos' team is the socio-economic and population information, obtained from the 2014 American Community Survey conducted by the U.S. Census Bureau. The census data includes information on census tracts, poverty rates, median household income, and ethnicity composition.

Using *BASE* (census tracts) as the join key, the Uber data and the Census data have been merged into the final dataset for analysis: the consolidated dataset has 2506 data points, 1253 of which contain weekday information and the other 1253 entries contain weekend information. Meanwhile, 358 of the data entries contain UberX information (which will be the main study interest as UberX services constitute the vast majority of the pickups), while the rest contains information for other kinds of Uber services (i.e. uberXL, UberBLACK, UberSUV, uberX + Car Seat, BLACK CAR + Car Seat, SUV + Car Seat) across the census tracts. The variable descriptions can be referred to in Table A.2 in the Appendix Section.

4 METHODOLOGY

4.1 Demand Forecast

The models include regularized regressions (LASSO, ridge, and elastic nets), trees (Extreme Gradient Boosting and Gradient Boosting Decision Tree), neural networks, and ARIMA. The evaluation criteria is RMSE under 10-fold cross-validation.

4.1.1 Regularized Regressions

OLS is one of the most basic regression models, but it could have two issues for prediction. First, the least square methods will incline to overfit if n (sample size) is close to p (number of variables) or $n < p$ when the independent and dependent variables have a strong linear relationship. Secondly, if multicollinearity or very weak relationships between certain independent and dependent variables exist, OLS is less likely to offer much as meaningful feature selection tools and requires necessary corrections.

To address these issues, ridge regression has included an L2 penalty term into the OLS loss function. Here, the adjustment parameter λ in the penalty term is larger than 0. The addition of the penalty term helps to let the estimated parameter approach closer to 0 compared to the OLS case. Often, information criterion such as AIC or Cp and cross-validation is utilized to determine the lambda. As lambda is increasing, the variance of the model is getting smaller with a light increase in bias. Nevertheless, one of the issues with regard to ridge regression is that, since the L2 penalty term cannot shrink the estimated parameters to 0 unless lambda is positive infinity, it does not ultimately conduct automatic feature selection, whereas LASSO is superior in solving this issue.

LASSO (Least Absolute Shrinkage and Selection Operator) uses an L1 penalty term defined as the sum of absolute values of the squared coefficients. As lambda is sufficiently large, some estimated parameters can be shrunk to 0: namely, LASSO conducts variable selection and regularization during GLM fitting. Through adjusting the parameter complexity, it alleviates the overfitting issue compared to OLS, especially when the variables are correlated or when there are too many dependent variables that have been fitted.

Elastic net has combined the L1 and L2 regularization; therefore, LASSO and ridge regression can be considered two specific cases of elastic net. Again, it is an optimization task in which residual sum of square is the target function and the penalty term is the constraint. Regularization helps to trade off some bias to decrease the variance, so as to better minimize the model error terms. Elastic net has two parameters λ and α, which control the model maximization with highly correlated data. In fact, $\alpha=1$ for LASSO, $\alpha=0$ for Ridge, and $0<\alpha<1$ for elastic net. Often, elastic net performs better for $p>>n$ or data with considerable multicollinearity, where it takes into consideration both sparsity and variable correlation.

However empirically speaking, it is hard to say which regularization method will be decisively the best among the three.

4.1.2 Trees

Gradient Boosting Decision Tree (GBDT) and Extreme Gradient Boosting (XGBoost) are both ensemble learning algorithms that aim to improve individual learning machine's generalization ability and robustness. Often, ensemble methods can be categorized in two ways: (1) boosting, where individual learning machines have strong dependence and are generated in serials, and (2) bagging, where individual learning machines do not have strong dependence to one another, and thus can be generated in parallels.

The calculation in GBDT in each series is to reduce the residues from the last time, and in gradient boosting, the key is to use the gradient's direction in the loss function to approximate the local minima of the residues in the model, which is different from the sample weighting in the traditional boosting methods. GBDT will add up the results from all of the trees, and therefore GBDT are regression trees, not classification trees. Compared to Random Forest, it is more flexible and often has higher predictive accuracy. Nonetheless, since GBDT is a boosting algorithm, it is hard to conduct parallel training with the data.

XGBoost is another highly efficient realization of gradient boosting algorithm, where the base learner can be either CART or linear classifier in which it serves as a logistic or linear regression with L1 or L2 regularization. Compared to GBDT, the optimization algorithm of

XGBoost uses Newton's method instead of gradient descent, and it has included regularization terms to penalize the complexity of the trees. Under shrinkage. it alleviates the overfitting issue.

4.1.3 Neural Networks

Artificial neural network (ANN) is a branch of perception learning in machine learning algorithm, and a classical ANN has three main compositions: architecture, activation function, and learning rule. The architecture is composed of input layer, hidden layer, and output layer. Each hidden layer extracts one or more higher-level features. The introduction of activation function enables non-linear transformation in each layer, which offers ANN powerful flexibility in model fitting. The learning rules that adjust the parameters minimize the loss function by adjusting the weights and threshold. Various optimization rules can be used including gradient descent or Newton's method. Neural Network algorithms evaluate the contribution of each parameter to the errors in prediction, and then adjust the weights; after multiple iterations, the loss function reaches its minima which render the final fitting of the parameters.

4.1.4 ARIMA

Auto Regressive Integrated Moving Average (ARIMA) is a time series model that learns from time-related pattern, and the model only needs endogenous not exogenous variables. In ARIMA(p,d,q), p is the auto-regressive term, d is the differencing term, and q is the moving average term. Using ACF and PACF, one obtains the optimal p and q, and d is obtained for stationarity. The stationarity here means that the series' mean and variance (or the series after transformation) do not change with regard to time. Nonetheless, the disadvantage of ARIMA is that it requires the time series to be stationary in itself or after transformation.

4.2 Wait Time, Surge Price, and Urban Socio-economics

4.2.1 Generalized Linear Model with Gaussian Distribution

The second part of the analysis aims to examine the extent to which waiting time for an Uber pickup is related to 1) surge price and 2) neighborhoods with different income levels and composition in race. Gaussian GLM is adopted using Proportion of Mean Expected Wait Time over 360 seconds (70th percentile of the wait time variable) as the response variable. The independent variables are demographic and socio-economic metrics such as percentage of People of Color, population, median household income, and poverty rates. Extending from Stark and Diakopoulos (2016)'s investigative journalism work, the regressors will examine separately for the cases of UberX, UberXL, and Uber BLACK.

4.2.2 Bayesian Networks

This paper will also use Bayesian networks (alternatively called Causal Probabilistic Networks) to study the connections between wait time, surge price, and socio-economic factors. In a Bayesian network, each node represents a finite-state random variable, and the single-directed arrows represent the causal relationships. After evaluating the Conditional Probability Table (CPT) in each node, Bayesian networks are then able to estimate the joint or conditional probability distributions for some or all of the other nodes. Notably, the estimated parameters of "the local conditional distributions associated with the learned structure [...]

determine the causal effect sizes between the parents and their children" (Aragam, Gu, and Zhou, 2018, p.10).

Simulated distributions usually require a considerable amount of variables and dependencies among the variables. Often, modelling by enumeration is costly and time-consuming, which calls for the use of graphs to present the complex structures of the relationships. Consequently, directed acyclic graph (DAG) is used to illustrate a Bayesian Network, where DAG stands for a directed graph without cycles.

In this paper, I will use the R package "sparsebn." Developed by Aragam, Gu, and Zhou (2018) at UCLA, the package is designed for Bayesian Networks with high-dimensional data, and for graphs with either continuous or discrete nodes. In addition, the package takes into consideration experimental interventions and learns with sparse regularization.

In the Analysis section, parameter estimates on UberX data will be provided to examine the causal flow among the various variables, especially the impact of *mean_surge_price*, *percent_POC*, and *percent_poverty* on *mean_expected_waitTime*. The model fitting procedure will include a whitelist, ensuring that edges will be directed to the study of interest *mean_expected_waitTime*. Moreover, the relationship structure will be visualized to offer a straightforward picture of the connectivity among the variables.

4.2.3 Further Examination on Surge Price

Even if the results in Section 4.2.1 suggest a general trend that higher surge price is correlated with less waiting time, on the local level such an effect should still be carefully evaluated. As Diakopoulos (2015) pointed out, surge prices might not succeed in generating new supply of Uber service because 1) the abruptive behavior of the dynamic pricing could discourage drivers' response and 2) the increased supply is through redistribution from adjacent neighborhoods, not from new drivers. Therefore, I will revisit the methods used in Diakopoulos (2015)'s work, but with alternative dates and more specific location/time granulation.

In the Data Exploration section, I will briefly look at the general trend between surge price and wait time in Washington D.C. Again, as surge prices could change very abruptly and are highly localized, studies on surge price effects should be further granularized into specific regions, and the distribution on weekdays and weekends will be inspected.

In the Analysis section, the wait time - surge price relationship in paired and adjacent areas will be further illustrated. Figure 4.1 provides an example where the horizontal axis denotes the difference in surge price between Area B and Area A, and the vertical axis denotes average waiting time in Area A. If the trend is positive, it may suggest that "when prices are higher in one neighborhood it saps driver capacity away from an adjacent neighborhood" (Diakopoulos, 2015, para. 14), whereas a negative trend indicates that a higher surge price in the neighboring Area B does not noticeably affect the supply in Area A, at least in the correlational sense.

I have chosen four areas to study this waiting time - surge price dynamics, and the four areas are: 1) Shopping & University Area (M Street, Georgetown University, and George Washington University), 2) Residence & Private Business Area (Dupont Circle), 3) Federal

Government District (federal government branches and bureaus, landmarks, and the Smithsonian Museums), and 4) Southeast Washington D.C. (Congress Heights, Bellevue, and Washington Highlands). Also, there are two time windows corresponding to the morning and afternoon rush hours (7am to 9am, 5pm to 7pm).

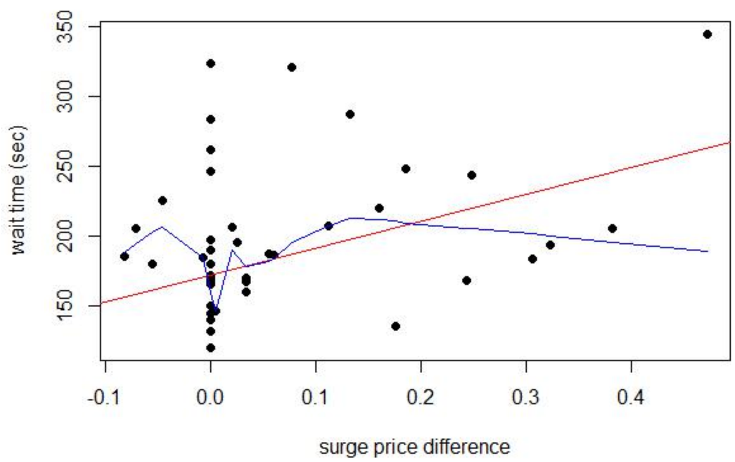


Figure 4.1 Uber Wait Time at Area 1 & Area 4 under Different Surge Price Differences, 7-9 AM

As an example, in Figure 4.1, Area A is the Shopping Area and University, which is adjacent to Area B, the Federal Government District. A point (287.3684, 0.15) on the plot means that in shopping and university area, the average waiting time during the month between 7 and 9 am is about 287 seconds, when the average surge price is 0.15 higher in area B than in area A (for instance 1.35x in area A and 1.2x in area B). The red line is the linear regressor, whereas the blue curve is a regression with locally estimated scatter plot smoothing (LOESS). If the linear regression's slope is positive and statistically significant, it means that the higher Area B's surge price than that in Area A, the longer the riders will wait to get the UberX in area A. Therefore, surge price in Area A seems not to be effective when the adjacent Area B has an even higher surge price, which in turn could serve as evidence that instead of driving new supply, surge prices likely just redistribute the existing supply from one neighborhood to an adjacent one.

5 PRELIMINARY DATA EXPLORATION

5.1 Pickup Volume and Relevant Descriptive Statistics

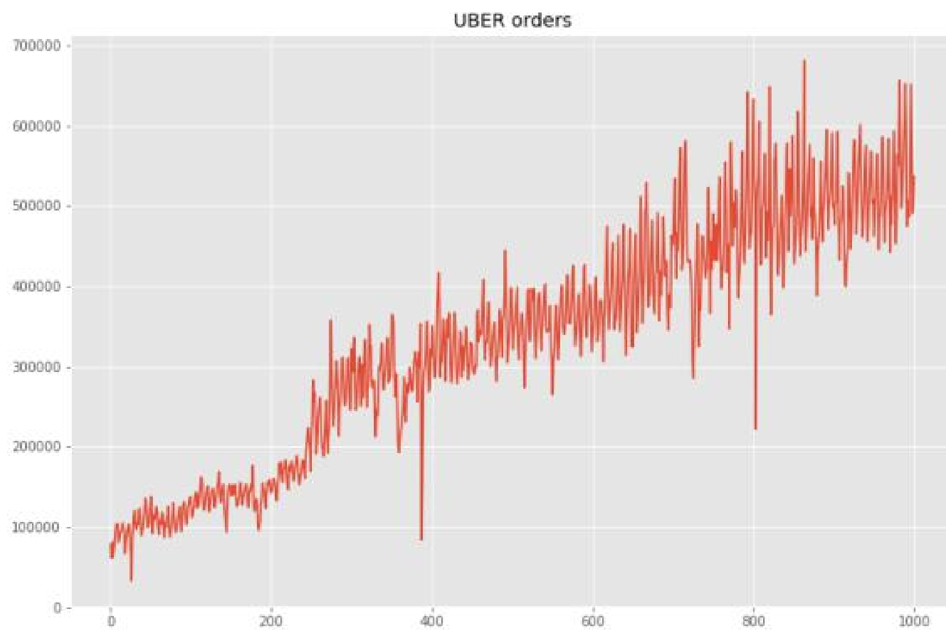


Figure 5.1: Uber pickup series, January 1st, 2015 - September 27th, 2017

The pickup demand series, illustrated in Figure 5.1, spans from January 1st, 2015 to September 27th, 2017. The first 250 series inclines slower than the latter series, mainly because Uber was not a significant market player until late 2016. Therefore, in the Analysis section, I will omit the first 250 data points so that the trend will be more representative. The data has strong semi-weekly and weekly cycles, and based on ACF plots I created moving average terms *MA7* and *MA3* to take into consideration such patterns for further analysis. Also notable is that within a day, the highest pickup volume is often reached around 7-9am and 5-7pm, which correspond to the rush hours.

The descriptive statistics of the continuous variables can be referred in Table 5.1 below:

Table 5.1. Summary Statistics for NYC Pickup Data (n = 1000)

Variable Name	Mean	SD	Min	Max
Orders	331391	145811.8	33171	681253
AWND	2.364	1.0807	0.4	8.2
PRCP	3.052	8.1993	0	76.7
SNOW	2.844	27.1669	0	693
SNWD	15.79	62.8156	0	560
TAVG	15	9.8521	-13.75	31.4
TMAX	18.9	1.028	-9.3	36.1
TMIN	10.6	9.6204	-18.2	27.8
WDF2	194.2	10.13	10	360
WDF5	197.8	99.68	5	360
WSF2	6.272	1.7756	2.7	14.8
WSF5	10.11	2.998	4.5	21.5
MA7	328942	137850.4	77789	579095
MA3	331782	143354.2	62691	611870
STD	35215	20028.86	5149	131662

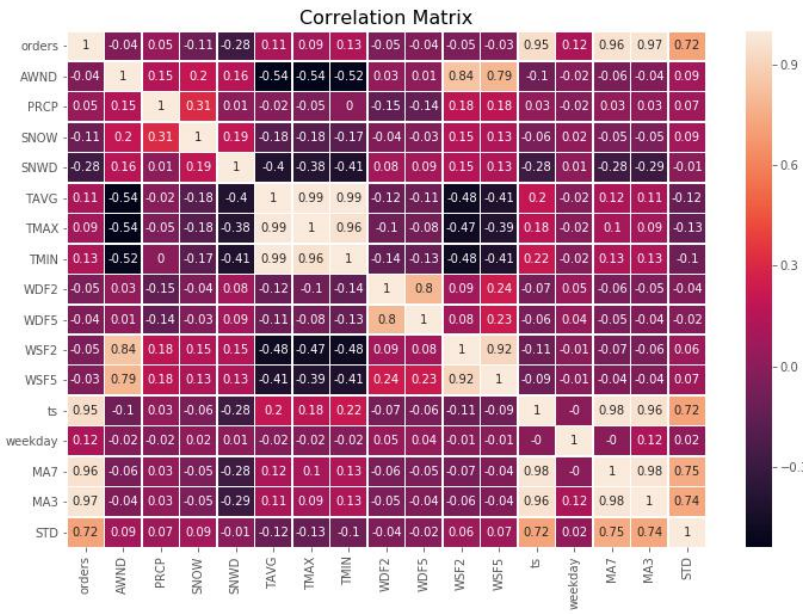


Figure 5.2: Correlation Matrix for Uber Pickup Dataset

Observing the correlation matrix in Figure 5.2, one may find that the daily fulfilled pickup (orders) has the strongest correlation with MA3 and MA7, and therefore time-series models may be of particular fit. The series also is highly correlated with STD, indicating that the fluctuation in data has been increasing as time goes by. This might suggest ARCH or GARCH to be possible modelling choices if the data has certain variance structure or squared residuals pattern.

In addition, weekday and higher temperature are positively correlated with more pickup orders, which meets with common sense. The negative correlation between pickup rate and snowfall/snow depth might be counter-intuitive, as more severe snow should have incentivized riders to call Uber. The correlation between pickup rate and other weather factors such as precipitation and wind is not noticeably high.

5.2 Waiting time, surge price, and socio-economics

5.2.1 Descriptive Statistics and General Trend

Table 5.2 provides the descriptive statistics of the second dataset:

Table 5.2: Summary Statistics for Washington D.C. Wait Time and Surge Price Data (n = 2506)

Variable Name	Mean	SD	Min	Max
mean_expected_waitTime	480.8862	262.3411	136.3594	1435.491
mean_expected_waitTime_surge	577.119	323.9524	150	1516.025
mean_surge	1.560597	0.1682912	1.2	2.5
proportion_waitTime	0.0612907	0.0396731	0.0000618	0.169833
proportion_surge	6.380745	9.038185	0	42.47004
POP100	3361.581	1297.866	33	7436
percent_poverty	19.82386	14.90376	2.372685	95
medianHouseholdIncome	74026.07	40932.98	14813	231042
percent_color	57.22261	32.70954	4.712939	100

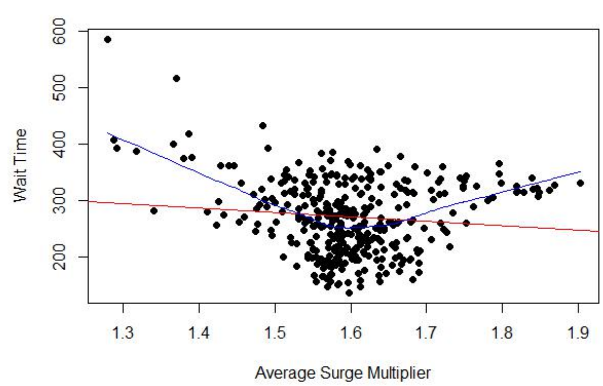


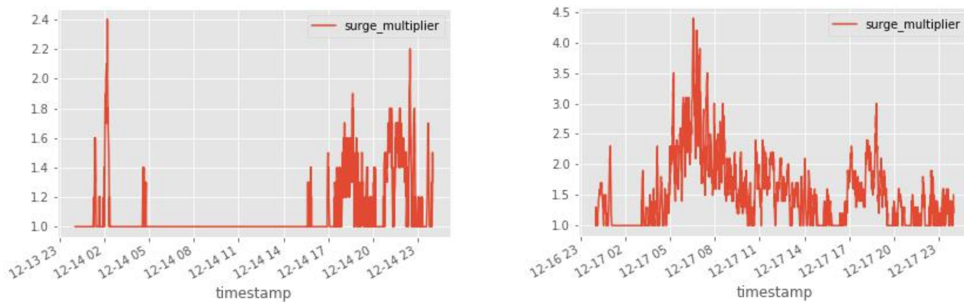
Figure 5.3: Linear and LOESS Trend of the Wait Time-Surge Multiplier Relationship

Figure 5.3 illustrates the general trend between surge price and wait time in New York City.

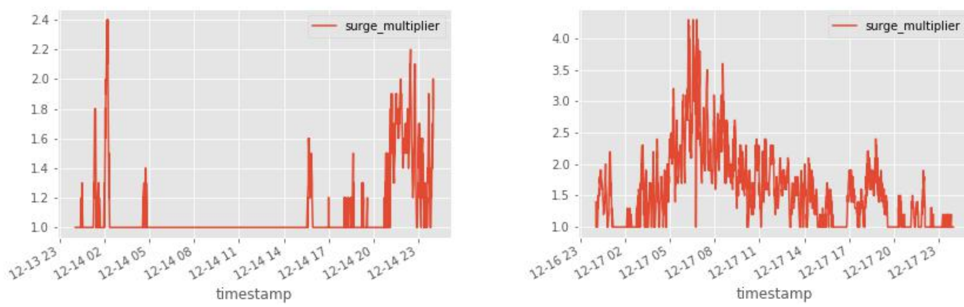
The slope is negative and statistically significant, indicating that on a general scale in Washington D.C., there is evidence that surge pricing has been effective in reducing waiting time. However, as Diakopoulos (2015) pointed out, "price surging can work in any of three ways: by reducing demand for cars [...], by creating new supply [...], or by shifting supply (drivers) to areas of higher demand" (para. 5). Additionally, surge pricing could change very abruptly and is highly localized. Therefore, this general trend should be granulated into specific regions and time windows for further examination, which will be rendered in the Analysis section.

5.2.2 Surge Price Distribution

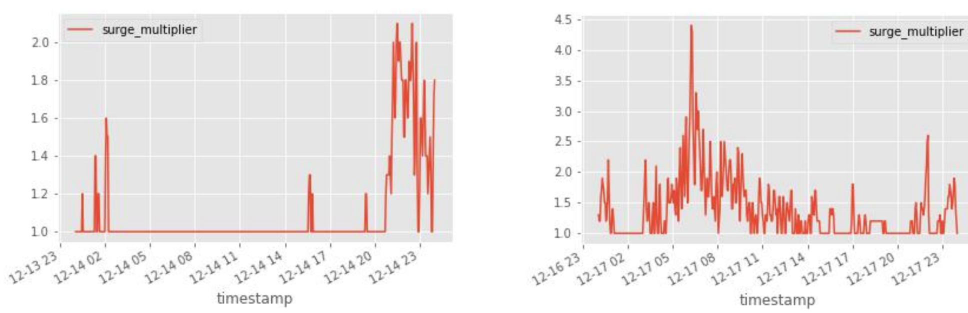
Figure 5.4 below illustrates surge pricing trends on a weekday (Wednesday, December 14th, 2016) and a weekend (Saturday, December 17th, 2016) in the four areas defined in the Methodology section. Recall again that the four areas under examination stand for: 1) Shopping Area & University Area, 2) Residence & Private Business Area, 3) Federal Government District, and 4) Southeast Washington D.C. Notably, Areas 1, 2 and 3 are adjacent neighborhoods in the Northwest Washington D.C., and these three areas are more affluent and have much lower ethnic minority rates and poverty rates than Area 4. On December 14th (a weekday), surge price mostly happened during the afternoon rush hours and late night to early morning (8pm to 5am) in Areas 1, 2 and 3. At other times, there was almost no surge price, likely because they were hours in which most residents were staying for work and there was not as much pickup demand. Area 4, the poorer neighborhoods with higher ethnic minority rates, had two peaks during the morning and afternoon rush hours, and almost no surge price at other times of the day.



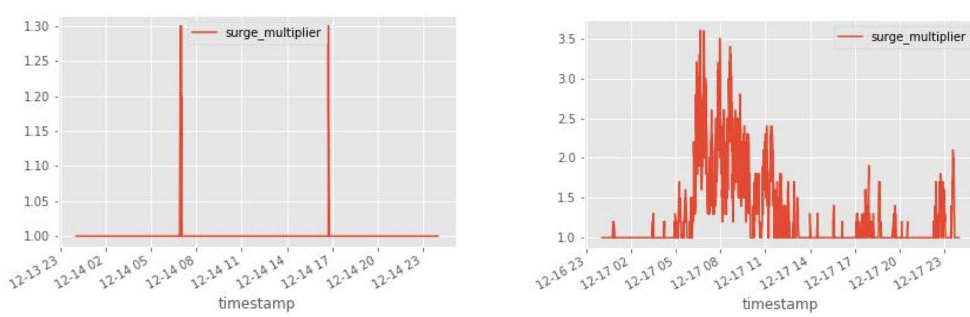
Shopping Area & University Area



Residence & Private Business Area



Federal Government District



Southeast Washington D.C.

Figure 5.4: Distribution of Dynamic Pricing on December 14th and December 17th, 2016

Dynamic pricing is usually high when the neighborhood has high demand, such as during holidays, severe weather, or important events. It is then very interesting yet puzzling to see that high price surge occurs at late night (8pm to 2am) in Area 1, 2, and 3 during the weekdays: this pattern did not only occur on December 14th but also on many other weekdays in this dataset. One possible explanation is that at night, FHV companies usually elevate the surge price to urge more drivers to be on the road.

On December 17th (a weekend), all four areas had similar patterns. The surge pricing was more dynamic throughout the day. It is puzzling that the surge price peaked around 6-7am in all four areas, when there should not be a high demand for Uber service. Such abnormality is hard to explain, and one may suspect that there might be errors in obtaining the correct weekend mornings' surge price information in this dataset.

6 ANALYSIS

6.1 Demand Forecast

Table 6.1 below provides the RMSEs using testing data under 10-fold cross-validation sets. We see that the ensemble methods do not necessarily work better than the penalized regression. Among different statistical learning models, LASSO has outperformed the others. In fact, the regularized linear regression models (LASSO, ridge, and elastic net) have close RMSEs that are significantly lower than other models such as trees and neural networks. The univariate ARIMA has the worst predictive power, which seems to be surprising since in Table A.3 (XGBoost variable importance table) *MA7* and *ts* have the most importance.

Moreover, LASSO outperforms the more complicated ensemble regressors, such as XGBoost and GBDT. LASSO usually perform better when the linearity assumption is close to valid, when there are no strong interaction effects, or when there are only small numbers of "meaningful" variables." On the other hand, the ensemble models here might have overfitting issues that reduce its extrapolability.

Table 6.1: MSE/RMSE of all results

Model	RMSE
LASSO	46,050
Ridge	46,179
Elastic Net	46,106
XGBoost	49,116
GBDT	100,103
Neural Network	49,116
ARIMA	104,309

Additionally, from XGBoost Importance table in Table A.3, compared to temporal information (*ts* and *MA7*), one could find that weather factors have very low explanatory power. Among them, precipitation and average wind speed are the most important, followed by temperature and snow. This meets with common sense, where rainfall and wind may influence one's decision whether to take an FHV or not.

6.2 Wait time, Surge price, and Socio-economics

The importance vector reported by random forest algorithm in Table 6.2 below indicates that the *percent_POC* has much higher importance than the rest of the variables, followed by the geographic information (*BASE*) and the proportion of surge price over 1. The socio-economic information such as percentage of poverty or median household income does not seem to have very significant influence.

Table 6.2: Random Forest Importance Vector

Importance	Feature
0.4921	percent_POC
0.1523	BASE
0.1455	proportion
0.0303	percent_poverty
0.0269	POP100
0.0266	medianHouseholdIncome
0.0141	isWeekend

Then, what about the directions of influence? Referring to Table 6.3, Table 6.4 and Table 6.5, the estimate for *percent_POC* is consistently positive and significant across all three products after controlling other factors, and the estimate for percent_poverty is positive and highly significant in the case of UberX. These suggest that higher ethnic minority rates or poverty rates in a neighborhood are correlated with longer wait time, whereas the surge is likely to have an association with less wait time: racial and economic compositions in a neighborhood seem to matter with service quality.

It is also interesting to see a positive direction for both *percent_poverty* and *medianHousholdIncome* in the case of Uber, which is likely due to the multicollinearity between the two variables. However, since the estimated parameter of poverty rates is much higher than that of *medianHouseholdIncome*, the general direction for poverty status is likely to be more indicative, meaning poorer neighborhoods tend to expect longer wait time.

Table 6.3: GLM Statistics for UberX

	Estimate	Std. Error	t value	Pr(>|t|)
proportion	-1.597e-01	1.758e-02	-9.080	< 2e-16 ***
percent_poverty	3.700e-02	1.239e-02	2.986	0.00303 **
medianHouseholdIncome	8.920e-06	4.912e-06	1.816	0.07024 .
percent_POC	4.502e-02	6.479e-03	6.949	1.85e-11 ***
isWeekend	8.756e-02	2.373e-01	0.369	0.71238
BASE	-1.027e-02	4.466e-03	-2.301	0.02201 *

Significance Level : ‘***’ 0.001; ‘**’ 0.01; ‘*’ 0.05; ‘.’ 0.1; ‘ ’ 1

Table 6.4: GLM Statistics for UberXL

	Estimate	Std. Error	t value	Pr(>|t|)
proportion	-4.457e-01	3.548e-02	-12.562	< 2e-16
percent_poverty	-1.355e-02	1.195e-02	-1.134	0.2575
medianHouseholdIncome	9.801e-06	4.562e-06	2.149	0.0324 *
percent_POC	6.024e-02	5.939e-03	10.143	< 2e-16 ***
isWeekend	1.016e+00	2.281e-01	4.455	1.14e-05 ***
BASE	-2.471e-03	3.966e-03	-0.623	0.5337

Significance Level : ‘***’ 0.001; ‘**’ 0.01; ‘*’ 0.05; ‘.’ 0.1; ‘ ’ 1

Table 6.5: GLM Statistics for UberBLACK

	Estimate	Std. Error	t value	Pr(>|t|)
proportion	-8.166e+00	1.616e+00	-5.053	7.08e-07 ***
percent_poverty	-1.867e-03	1.528e-02	-0.122	0.902848
medianHouseholdIncome	2.009e-05	5.954e-06	3.374	0.000824 ***
percent_POC	8.511e-02	7.175e-03	11.862	< 2e-16 ***
isWeekend	-5.615e-01	2.700e-01	-2.080	0.038294 *
BASE	2.253e-02	5.285e-03	4.263	2.60e-05 ***

Significance Level : '***' 0.001; '**' 0.01; '*' 0.05; '.' 0.1; ' ' 1

Moreover, the estimated parameter for proportion of surge price over 1 is consistently negative and statistically significant in all three products (UberX, UberXL, Uber BLACK), after controlling for socio-economic, temporal, and geographical factors. Therefore, a higher proportion of surge prices is likely to be associated with less wait time. Though this may be the general trend, would it be possible that surge prices in one neighborhood have interactive effects on its adjacent neighborhoods? In that case, surge price might not perform well in redistributing Uber traffics into areas with high demands. In order to confirm the consistency of the surge price effect, I have subsequently estimated the direction of influence with a Bayesian network with the UberX data, along with other covariates in this dataset.

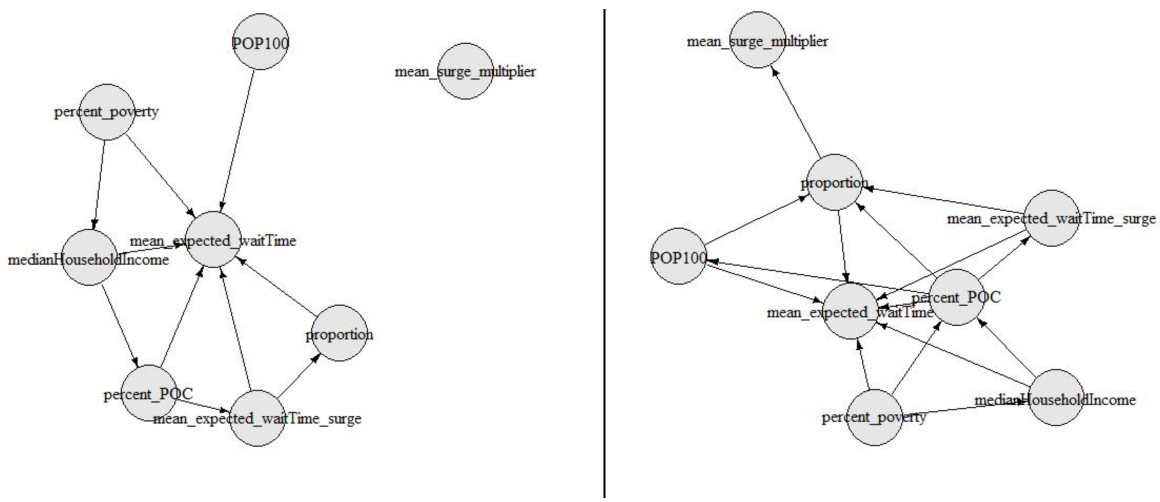
Table 6.6: Coefficients of the Edges Directing to *mean_surge_price*

mean_surge_price	percent_POC	percent_poverty
.	-5.37×10^{-2}	3.52×10^{-2}
medianHouseholdIncome	POP100	proportion
-6.27×10^{-6}	-1.07×10^{-4}	3.76×10^{-1}

The signs of influences on mean wait time can be referred to in Table 6.6 above (with the tuning parameter λ = 7.1157), and the estimates here are conceptually similar to the coefficients in GLM. One can find that poverty rates and proportion of surge price are positively correlated to waiting time, whereas population, median household income, and ethnic minority rates are negatively correlated to waiting time. The opposite directions of poverty rates and income meet with expectation, but the coefficient sign for ethnic minority rates is different from the sign in the GLM output.

Moreover, Figure 6.1 on the next page illustrates the structure graphs of the network, which correspond to lambda = 7.1157 and lambda = 2.1179 after parameter tuning. For lambda = 7.1157, the Bayesian network structure has 10 edges, where mean surge price (*mean_surge_multiplier*) has been isolated from the rest of the variables. When lambda = 2.1179, surge pricing and percentage of surge pricing within a specific region is connected. Here, the Bayesian network has properly captured the connection between the two variables that are both a function of the surge pricing. In addition, the connection among *percent_poverty*, *medianHouseholdIncome*, and *percent_POC* suggests the correlation among demographic and economic features.

On the other hand, when the variable *proportion* is taken away from the network, the connection between surge price and waiting time is established, as one may discover in Figure 6.2 below. The isolation in Figure 6.1 (a) and the connectivity in Figure 6.2 suggest that the surge price may have inconsistent causal attributions to the wait time. Unlike the consistent negative correlation in GLM modelling, when proportion of surge price has been included in the Bayesian network structure, surge pricing seems to have a very weak impact on wait time. Therefore, the surge price - waiting time relationship should be further examined under the scope of local areas and time windows.



(a) λ = 7.1157

(b) λ = 2.1179

Figure 6.1: Estimated DAGs under different tuning parameters λ

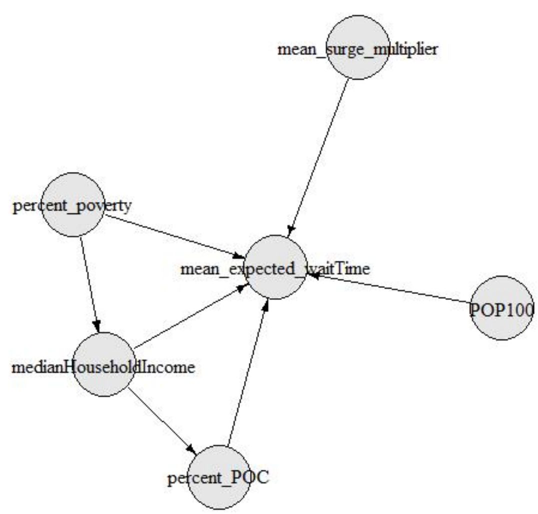


Figure 6.2: Estimated DAG excluding variable *proportion*, λ = 7.1157

As a result, I have created plots of information on wait time and surge price difference, in four areas at 7-9am and 5-7pm separately. The graphs can be referenced from Figure A.1 to Figure A.6 in the Appendix section. Recall that Areas 1, 2, and 3 are adjacent areas. In areas with statistically significant regression slopes, the vast majority has a positive slope, which means that one area does sap another area's surge price effect. For instance, the positive slope in Area2-Area3 AM pair means that, when the Dupont Circle (residential and business area) has lower surge price than the adjacent federal district, the wait time in Dupont Circle will climb up. This agrees with the expectation that the comparative incentive of higher surge price in Federal Government District saps away Uber drivers in Dupont Circle, regardless whether the price in Dupont Circle is surging or not.

On the contrary, in the cases of Area1 - Area2 PM (Figure A.1) and Area3 - Area1 PM (Figure A.5), the statistically significant slope is negative, meaning that even if Area 2 has higher surge price than Area 1, Area 1 tends to have less wait time. This might provide evidence that surge price in these areas works in a more static way instead of interacting with adjacent areas, although they only constitute 2 out of the 12 statistically significant cases.

However, the dominating positive trend above can also be attributed to 1) unusually high demand in Area A at afternoon rush hours or 2) severe traffic congestion that drives up the wait time; both reasons could result in longer wait time regardless of the increase in surge price in that particular area. Nonetheless, these possibilities still suggest that a higher surge price itself does not necessarily improve service quality calibrated as waiting time, as many other factors come into play. Thus, even if an area has a high surge price, there seems to lack systematically positive evidence that it is related to less waiting time when the neighboring area has even higher surge prices, at least in a majority of local areas and time periods we have examined here.

7 CONCLUSION AND DISCUSSION

7.1 Conclusion

In predicting daily pickups in NYC, the temporal information (such as 7-day moving average and weekday/weekend separation) provides more predictive power, while in general weather features do not serve as important influential factors. Nonetheless, time series models such as ARIMA does not quite capture the future trend, whereas more parsimonious models, such as LASSO in this empirical study, serve better in the forecast. These empirical results could be taken into consideration in ride-hailing services' future incentive designs to the drivers.

Meanwhile, this paper has studied relationships between service quality (wait time as the proxy) and factors such as dynamic pricing, urban demographics, and socio-economics. Similar to Jennifer and Diakopoulos (2016)'s findings, the empirical GLM results on the Washington D.C. dataset suggest that ethnic minority rates is positively correlated with longer wait time for UberX service. In fact, the ethnic minority rates has been a consistent and statistically significant factor in correlation with wait time across UberX, UberXL, and Uber BLACK services. Surge price is also another possible factor in influencing wait time, yet the direction of effect is uncertain: although higher surge pricing seems to be correlated to less wait time in general, there is insufficient evidence for the same pattern when we look into Bayesian network analysis, as well as further examination in particular neighborhoods and time windows. The reasons might be: 1) drivers do not rely heavily on the surge multiplier since it changes rather abruptly even in several minutes; 2) surge price does not work in a static way, but rather interacts dynamically with other neighborhood factors. However, a more comprehensive review of the existing data with rigorous experiment design is needed before such supposition could be substantiated.

7.2 Limitations and Future Works

One limitation in this thesis is the scope of the response variables. The pickup demand in this thesis is in fact the fulfilled demand, which is only a point on the actual demand curve, whereas one needs more information about the whole demand and supply curve to render insights in solving the redistribution issue. Similarly, waiting time has been used as a proxy of service quality for explanatory analysis. However, service quality clearly spans over a wide spectrum of criteria; for instance, drivers and riders' ratings and comments could be included in the study to offer a more complete picture in assessing service quality.

In the future demand forecast, as the geographic information of pickups is further grouped into 5 boroughs, the prediction and analysis could be further granulated into more refined levels for comparison. Additionally, in 2017, the NYCTLC required ride-hailing services to provide information such as drop-off locations. Having both pickup and drop-off information, independent researchers could probe into Uber's redistribution problems and supply-demand dynamics with a more comprehensive understanding.

Meanwhile, there are many more statistical learning methods that could be adopted to improve predictive and explanatory power. In the demand forecast, the fluctuating nature in the series may be better captured by generalized autoregressive conditional heteroskedasticity (GARCH) model. Additionally, as surveyed in the literature review section, in recent years many more finely-tuned deep learning methods, such as LASSO-STAR and LSTM, have been developed to address the shortcomings that one single category of modelling has.

Finally, the positive correlation between waiting time and racial composition in Washington D.C. neighborhoods, after controlling other factors such as poverty rates and surge pricing, has also posed important questions. As Jennifer and Diakopoulos (2016) pointed out,

> while any sort of racially biased agenda by Uber is extremely unlikely, [some empirical studies] suggest that race does play a role in predicting the service quality of uberX in different neighborhoods. [...] In light of the recent Supreme Court ruling that discriminatory intention is not a precondition to be held accountable for discrimination in the housing market, what are the implications of unintentional discriminatory side-effects being uncovered in other domains? (para. 14)

The answer to address the fairness in service delivery and tools for quantification will help to reinforce the working ethics in the rise of the gig economy.

A Appendix

Table A.1: Variable Description for NYC Pickup Data (n = 1000)

Variable Name	**Variable Description (unit if possible)**
Continuous Variable	
Orders	Number of realized pickups (i.e. fulfilled demand) from Uber X
AWND	Average daily wind speed (tenths of meters per second)
PRCP	Precipitation (tenths of mm)
SNOW	Snowfall (mm)
SNWD	Snow depth (mm)
TAVG	Average temperature (tenths of degrees C)
TMAX	Maximum temperature (tenths of degrees C)
TMIN	Minimum temperature (tenths of degrees C)
WDF2	Direction of fastest 2-minute wind (degrees)
WDF5	Direction of fastest5 2-second wind (degrees)
WSF2	Fastest 2-minute wind speed (tenths of meters per second)
WSF5	Fastest 5-second wind speed (tenths of meters per second)
MA7	7-day moving average of Number of realized pickups by Uber X
MA3	3-day moving average of Number of realized pickups by Uber X
STD	standard deviation of the pickup number up to date
Discrete Variable	
ts	observation index
weekday	whether it is a weekday (1=Yes, 0=No)
Borough	Borough (1 = Manhattan, 2 = Brooklyn, 3 = Queens, 4 = The Bronx, 5 =Staten Island)

Table A.2: Variable Description for Washington D.C. Uber Waiting Time Data (n = 2506)

Variable Name	**Variable Description (unit if possible)**
Continuous Variable	
mean_expected_waitTime	average wait time (secs)
mean_expected_waitTime_surge	average wait time when price surge is present (secs)
mean_surge	average surge multiplier (available for Uber X, Uber XL, and Uber Black)
proportion_waitTime	percentage of wait time over 360 seconds (%)
proportion	percentage of surge pricing (%)
percent_poverty	poverty rate (%)
medianHouseholdIncome	median household income (in dollars)
percent_POC	percentage of residents who are People of Color (%)
Discrete Variable	
BASE	census tract (179 tracts)
POP100	population count
product	Uber car product types (UberX, uberXL, UberBLACK, UberSUV, uberX + Car Seat, BLACK CAR + Car Seat, SUV + Car Seat)
isWeekend	whether it is a weekend (1=Yes, 0=No)

Table A.3: XGBoost Importance Table

Importance	Feature
49.05%	MA7
44.48%	ts
3.67%	weekday
0.88%	STD
0.38%	PRCP
0.32%	AWND
0.19%	TMIN
0.19%	WSF5
0.17%	WDF2
0.17%	WDF5
0.16%	WSF2
0.13%	TMAX
0.10%	TAVG
0.07%	SNWD
0.04%	SNOW

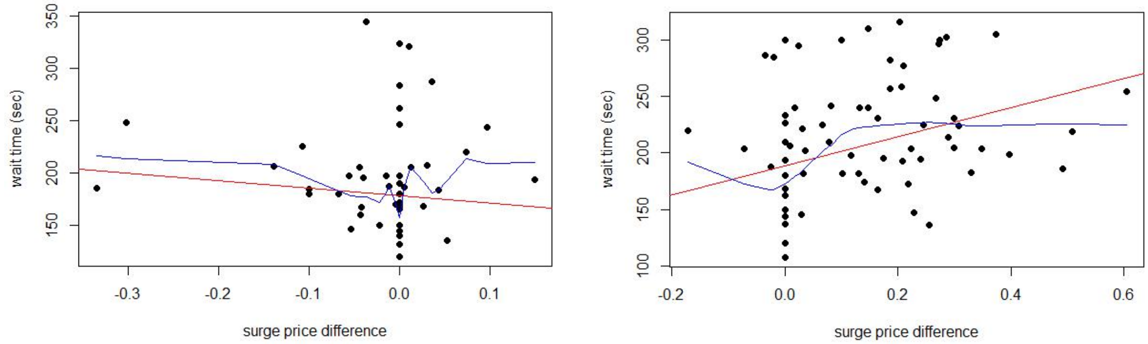


Figure A.1: Wait Time at Area1 (vs. Area2) and Surge Price Differences, 7-9 AM (Not Significant) and 5-7 PM

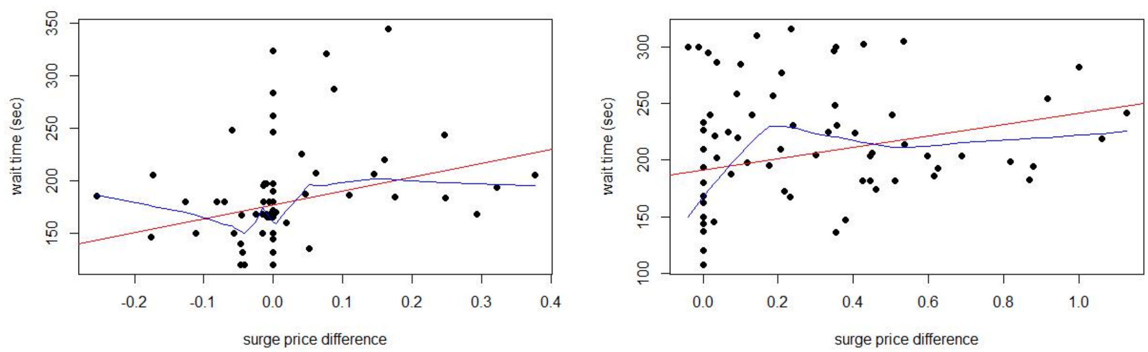


Figure A.2: Wait Time at Area1 (vs. Area3) and Surge Price Differences, 7-9 AM and 5-7 PM

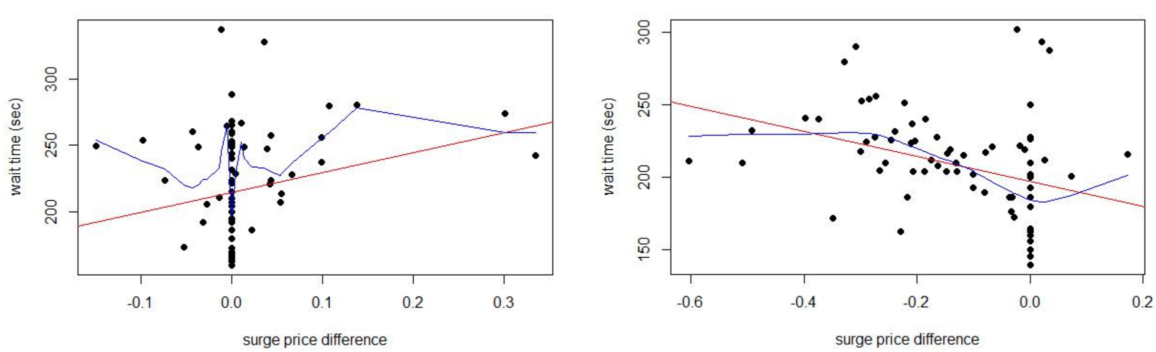


Figure A.3: Wait Time at Area2 (vs. Area1) and Surge Price Differences, 7-9 AM and 5-7 PM

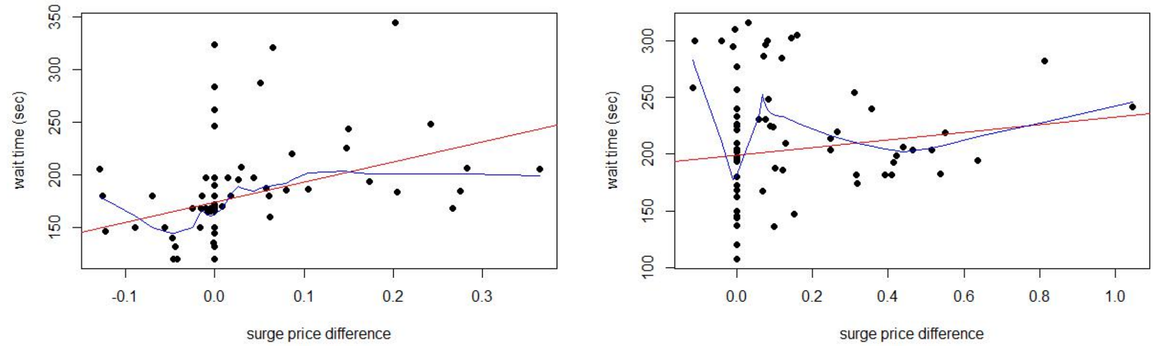


Figure A.4: Wait Time at Area2 (vs. Area3) and Surge Price Differences, 7-9 AM and 5-7 PM (Not Significant)

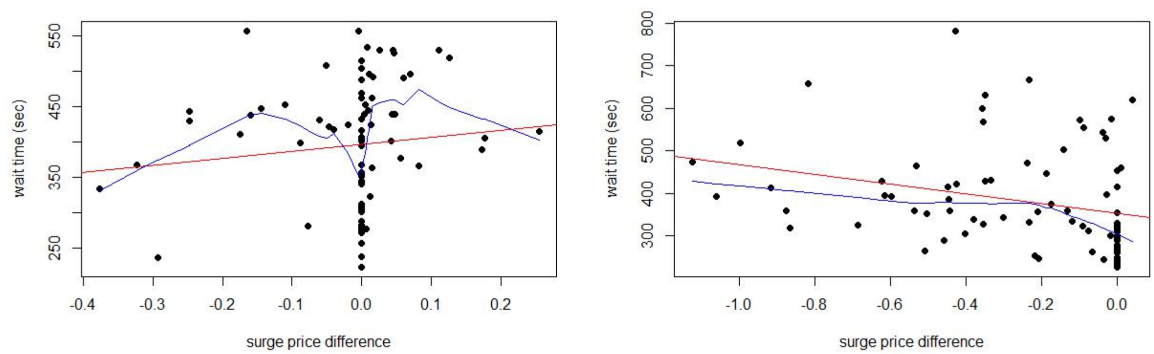


Figure A.5: Wait Time at Area3 (vs. Area1) and Surge Price Differences, 7-9 AM and 5-7 PM

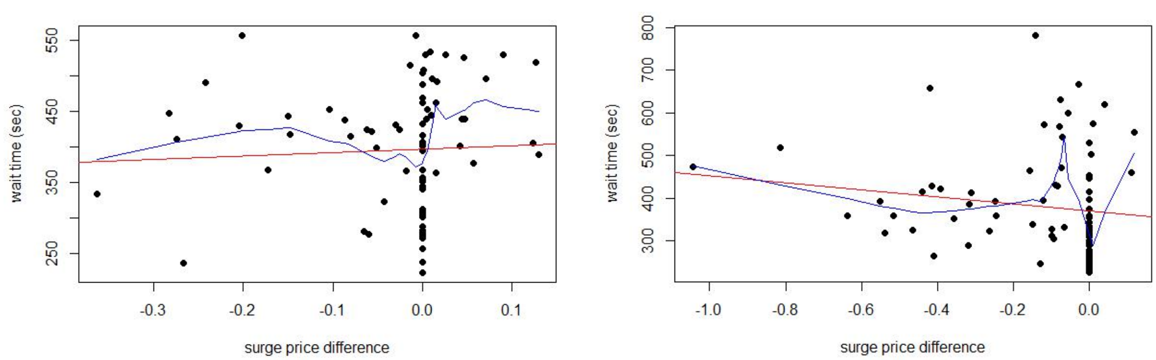


Figure A.6: Wait Time at Area3 (vs. Area2) and Surge Price Differences, 7-9 AM (Not Significant) and 5-7 PM (Not Significant)

REFERENCES

Aragam, B., Gu, J., & Zhou, Q. (2017). Learning large-scale Bayesian networks with the sparsebn package. arXiv preprint arXiv:1703.04025.

Bell, F., & Smyl, S. (2018, September 6). *Forecasting at Uber: An Introduction*. Retrieved April 10, 2019, from https://eng.uber.com/forecasting-introduction/

Brodeur, A., & Nield, K. (2018). An empirical analysis of taxi, Lyft and Uber rides: Evidence from weather shocks in NYC. *Journal of Economic Behavior & Organization*, *152*, 1-16.

Castiglione, J., Chang, T., Cooper, D., Hobson, J., Logan, W., Young, E., Charlton, B., Wilson, C., Mislove, A., Chen, L. & Jiang, S. (2016). TNCs today: a profile of San Francisco transportation network company activity. *San Francisco County Transportation Authority (June 2016)*.

Charlton, B. (2017, June 13). Visualizing Uber and Lyft trips in San Francisco: More than 200,000 trips a day. Retrieved March 25, 2019, from https://hackernoon.com/visualizing-uber-and-lyft-usage-in-san-francisco-928208b1978a

Chen L. Ampountolas, K. & Thakuriah, P. (2019). Predicting Uber Demand in NYC with Wavenet. In *The Fourth International Conference on Universal Accessibility in the Internet of Things and Smart Environments*. Smart Accessibility 2019.

Chen, L., Mislove, A., & Wilson, C. (2015, October). Peeking beneath the hood of uber. In *Proceedings of the 2015 Internet Measurement Conference* (pp. 495-508). ACM.

Chen, M. K., & Sheldon, M. (2016, July). Dynamic Pricing in a Labor Market: Surge Pricing and Flexible Work on the Uber Platform. In *Ec* (p. 455).

Cohen, P., Hahn, R., Hall, J., Levitt, S., & Metcalfe, R. (2016). *Using big data to estimate consumer surplus: The case of uber* (No. w22627). National Bureau of Economic Research.

Diakopoulos, N. (2015). How Uber surge pricing really works. *The Washington Post*, *17*.

Faghih, S. S., Safikhani, A., Moghimi, B., & Kamga, C. (2019). Predicting Short-Term Uber Demand in New York City Using Spatiotemporal Modeling. *Journal of Computing in Civil Engineering*, *33*(3), 05019002.

Ge, Y., Knittel, C. R., MacKenzie, D., & Zoepf, S. (2016). *Racial and gender discrimination in transportation network companies* (No. w22776). National Bureau of Economic Research.

Jiang, S., Chen, L., Mislove, A., & Wilson, C. (2018, April). On Ridesharing Competition and Accessibility: Evidence from Uber, Lyft, and Taxi. In *Proceedings of the 2018 World Wide Web Conference on World Wide Web* (pp. 863-872). International World Wide Web Conferences

Steering Committee.

Laptev, N., Yosinski, J., Li, L. E., & Smyl, S. (2017). Time-series extreme event forecasting with neural networks at uber. In *International Conference on Machine Learning* (No. 34, pp. 1-5).

Schneider, T. (2015). Analyzing 1.1 billion NYC taxi and Uber trips, with a vengeance.

Schmidt, S. (2018, October 1). *Examining Changes in New York's Taxi and For-Hire Vehicle Industry, One Hour at a Time*. New York City Taxi and Limousine Commission. Retrieved April 3, 2019, from https://medium.com/@NYCTLC/factbook-deep-dive-1-e4dd57075945

Stark, J., & Diakopoulos, N. (2016). Uber seems to offer better service in areas with more white people. That raises some tough questions. *The Washington Post*.

Turakhia, C. (2017, November 10). *Engineering More Reliable Transportation with Machine Learning and AI at Uber*. Retrieved April 15, 2019, from https://eng.uber.com/machine-learning/

Wang, C., Hao, P., Wu, G., Qi, X., & Barth, M. (2018). *Predicting the Number of Uber Pickups by Deep Learning* (No. 18-06738).

Wang, D., Cao, W., Li, J., & Ye, J. (2017, April). DeepSD: supply-demand prediction for online car-hailing services using deep neural networks. In *2017 IEEE 33rd International Conference on Data Engineering (ICDE)* (pp. 243-254). IEEE.

Zander, G. (2017). *Predicting taxi passenger demand using artificial neural networks* (Master's thesis). Retrieved from: http://www.nada.kth.se/~ann/exjobb/gustav_zander.pdf