TOWARD A SYSTEMATIC APPROACH TO DRIVERS' BEHAVIORAL STUDY ON MOBILITY-ON-DEMAND RIDE SHARING SYSTEMS

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Master of Science

By

Lan Liu

August 2017

ABSTRACT

With continuous growth of urban populations, transportation system faces many challenges related to increasing demand of real time services, limited government investment and sustainability of environment. Ride sharing, as a mobile-internet-based transportation service mode, has gained wide popularity across the world. Instead of operating a fixed fleet, a ride sharing platform consolidates supplies from independent drivers with dynamic and flexible schedules. A ride sharing system provides drivers with a flexible working method and also improves passengers' trip experiences in respect of easy reservation and convenient access to trip information. Unlike traditional taxi business, where supply is constant, ride sharing systems interact with a dynamic fleet. Therefore, adequate study for drivers' behaviors is of great research interest, as it not only enriches behavioral study for suppliers in economic activities but also supports design of operation strategy for ride sharing system. This research proposes a comprehensive and data driven method that implements behavioral study based on Multinomial Logit Model (MNL) and Mixed Logit Model, from the family of Random Utility Maximization models. Furthermore, in order to explore operation strategies, a simulation framework of ride sharing system is developed. Operation strategies that involve consideration of drivers' behaviors are proposed and simulated, which attempt to improve the ride sharing system's operation efficiency.

Keywords: Ride sharing system; Drivers' operation behavior; Discrete Choice; Multinomial Logit Model; Mixed Multinomial Logit Model; Simulation; Maximum weighted matching; Operation efficiency.

BIOGRAPHICAL SKETCH

Lan was born in NeiJiang, a tiny city in Sichuan Province, China. Lan started her school year at the age of two, attending a kindergarten located on a historical street. As a two-year-old little girl, Lan loved going to school only because her dearly grandfather would buy her favorite pie on the way home each single day. Lan had a wonderful childhood in a beautiful small village with her grandparents.

When Lan was five, she was sent to a boarding school to start primary education. On the very first day of school, she even had no idea about how to make up bed nor comb hair. It was extremely difficult to be away from home and grow up alone. Fortunately, she found the best company from all kinds of books. Lan had numerous splendid dusks holding a book and waiting for dinner.

Hesitating on what to do with life in the first year of college, Lan decided to go aboard and explore for more possibilities. Two and a half year in Oklahoma State University helped her build up the enthusiasm for study again and also encouraged her to continuously discover for potential.

Two years at Cornell as a graduate student in transportation system engineering has been an incredible experience for Lan. She had chance to study courses that were totally fresh to her existing knowledge, which further inspired her curiosity and found things she desired to explore. Days at Cornell have been an amazing journey that not only extends her knowledge

but also strengthens her up to face any type of challenges with confidence and courage. After graduation, Lan will start working as an algorithm researcher at a ride sharing company, where she hopes she can continuously build up and realize her dream.

ACKNOWLEDGMENTS

First, I would like to thank my advisor, Dr. Ricardo Daziano for his instruction and help. His course CEE 6640 has been one of the most inspiring courses I ever took, which opened my mind and found a new and amazing research area. Then, I would also like to thank my advisor Dr. James Renegar, who has inspired me to explore the field of Operational Research and always been so supportive.

Also, I would like to thank my family and friends, who have given me constant encouragement. Thank you to Tong, who has encouraged me to explore the field of computer science and been be my ever best teammate. Thank you to PeiPei OuYang, who gave me the sincerest care and support to get through difficulties. Thank you to YiQin Qiu, who has encouraged and helped me so much from the first year of college. I also want to thank my parents who have tried their best to support me to discover and fulfill my dream.

Table of Contents

ABSTRACT 3

BIOGRAPHICAL SKETCH 4

ACKNOWLEDGMENTS 6

CHAPTER 1: INTRODUCTION 8

CHAPTER 2: DEVELOPMENT OF MULTINOMIAL LOGIT MODEL (MNL) 13

FOR DRIVERS' OPERATION BEHAVIORS 13

2.1 DATA DESCRIPTION 14
2.2 MATHEMATICAL ILLUSTRATION FOR MULTINOMIAL LOGIT MODEL 15
2.3 DRIVER CLUSTERING 17
2.4 SPECIFICATION FOR MNL'S CHOICE AND ATTRIBUTE SET 19
2.5 RESULTS ANALYSIS FOR MULTINOMIAL LOGIT MODEL (MNL) 27
2.5.1 RESULTS ANALYSIS FOR INDEPENDENT SAMPLE 28
2.5.2 RESULTS ANALYSIS FOR SAMPLES ACROSS DATES AND GROUPS 33
2.6 MIXED LOGIT MODEL (MIXL) 53
2.6.1 DEVELOPMENT FOR MIXL 53
2.6.2 NUMERICAL EXAMPLE ANALYSIS 54

CHAPTER 3: RIDE-SHARING SYSTEM SIMULATION 57

3.1 IMPLEMENTATION OF RIDE-SHARING SYSTEM SIMULATION 58
3.1.1 SIMULATION DESIGN FOR SUPPLY FROM DYNAMIC DRIVER FLEET 58
3.1.2 SIMULATION DESIGN FOR DEMAND 60
3.2 RIDE-SHARING PLATFORM STRATEGY 61
3.3 DETAILED SIMULATION STEPS 63
3.4 RIDE-SHARING SIMULATION AND CASE STUDY 64
3.4.1 DATA DESCRIPTION 64
3.4.2 NUMERICAL CASE STUDY 65

CHAPTER 4: CONCLUSION 76

REFERENCE 79

Chapter 1: Introduction

Ride sharing, as an internet-based transportation service mode, has experienced tremendous growth worldwide during recent years. Along with the fast development of mobile technology, ride sharing has evolved from a naval transportation concept to a much more widely used transportation service. There are numerous ride-sharing platforms operating across the world, such as Uber, based in the United States, and DiDiChuXing, based in China. These platforms have built large market shares in transportation service. For example, according to one industrial report (1), DiDiChuXing processes approximately 10 million trip requests daily and has around 1.5 million registered drivers. DiDiChuXing provides ride-sharing services for over 400 million customers across 400 cities in China. The scale of this business has also inspired enriched service types such as carpooling, luxury taxi services, and transportation for hitchhiking across cities.

Essentially, a ride-sharing platform consolidates real-time information about both demand and supply, while also performing instant driver-order matching and notification. Unlike traditionally hailing a taxi on the street, ride sharing provides passengers with real-time trip service and updates through mobile Application (APP). With ride-sharing platforms' mobile APP, passengers are able to make instant or advanced trip reservations and track the real-time location of their reserved vehicle conveniently. Generally, one trip request contains the following information: desired pickup time, desired pickup location, and destination. After receiving these types of trip requests, a ride-sharing platform takes charge of matching feasible driver-order pairs and sends matched trip information to both the driver and the passenger, accordingly.

Ride-sharing services are believed to create positive experiences for both passengers and drivers (2). For passengers, ride sharing systems improve their travel experience by saving them from searching for a taxi on the street and providing them with a better-planned trip. Drivers also have a more positive experience, because they are assigned orders instead of looking for passengers by themselves. Additionally, drivers operate with a flexible and independent schedule on a ride-sharing platform, as they do not have any employment relationship with the ride sharing company. The participating drivers have the right to decide when and where to enter or leave the system, which leaves them free to consider their own schedule availability and operation preferences. As a result, a ride-sharing platform interacts with a dynamic fleet rather than a relative fixed one, which differs significantly from traditional taxi operators.

The population of drivers on a ride-sharing platform is composed of individuals with different schedule flexibilities and revenue objectives (3). The dynamic nature of this fleet brings with it substantial efficiency challenges in dispatching supply to meet demand (4). For an operator that interacts with a fixed fleet, the major uncertainty about operation status comes from stochasticity within the demand distribution along time and geographical areas. With a constant supply, an appropriate prejudgment for future demand distribution is adequate to design an efficient dispatch scheme. However, when interacting with a dynamic fleet, the stochastic distributions of both demand and supply should be closely studied to design the optimal operation scheme. As the distribution of supply directly relates to drivers' operation decisions, studying driver behavior is essential to improve matching efficiency for ride-sharing systems.

Additionally, an appropriate supply boosting scheme is necessary for a ride-sharing system, which often faces a demand–supply gap issue. According to operation records released by

DiDiChuXing, order requests that do not receive a response due primarily to lack of supply account for approximate 12–15% of total received requests. Boosting supply in an operation context where more drivers are desired can improve passengers' experience by reducing the probability of a non-response, and can also increase revenue for both drivers and the ride-sharing platform. Without establishing an employment relationship between drivers and ride-sharing systems, boosting supply often requires providing subsidies to drivers. An improved understating of drivers' operation behaviors can support the design of a subsidy allocation scheme, which can lead to higher increase of supply.

While many researches have focused on passengers' characteristics and corresponding behaviors in the context of transportation mode choice (5,6,7), few have studied trip service suppliers' choice behaviors during operation. This gap in research is due to the fact that little motivation exists to explore the issue of supplier behavior in the traditional context where supply is constant, and also due to limited public data resources. Meanwhile, a ride-sharing platform based on mobile internet technology can accumulate massive, detailed operation records that reflect the interaction between a stochastic demand and a dynamic supply. On the supply side, the dynamic distribution results from drivers' independent decisions during the operation process. Therefore, operation records from ride sharing-platforms offer a promising data resource for the study of suppliers' behaviors, which will not only enrich understanding of suppliers' behaviors in business activities, but will also support more efficient, effective designs for the operation strategies of a ride-sharing system. However, ride-sharing trip data that are publicly available, such as Uber's operation data in TLC trip record of New York City[1] and DiDiChuXing's released public dataset[2], usually provide records only for processed orders.

[1] http://www.nyc.gov/html/tlc/html/about/trip_record_data.shtml, accessed January, 2017.

[2] http://research.xiaojukeji.com/, accessed January 2017

In order to study drivers' behaviors, a comprehensive method for extracting behavioral information from limited data sources needs to be established.

On a ride-sharing platform, drivers' operation decisions are based on independently evaluating their current operation statuses. Specifically, operation decisions can be categorized as continuous operation, taking a rest, and exiting the system. Different operation decisions can be understood as drivers' discrete choices when faced with operation statuses that constantly change. After determining drivers' operation statuses, the Multinomial Logit Model (MNL) (8)—taken from family of Random Utility Maximization models which are widely adopted for research in fields such as urban economics, transportation and marketing (9)—is used to analyze drivers' decision processes. By incorporating MNL, one major assumption is that drivers, in response to evaluating their current operation status, make decisions about whether or not to provide another service after completing their previous order. Accordingly, a driver's operation status can be described according to three factors: the system environment, the newly assigned order's characteristics, and the driver's own operation history during the day. Drivers with distinct operation statuses are assumed to be different behavioral respondents, considering their observed heterogeneities. The proposed description for operation status builds a method to establish MNL choice attributes sets from limited operation records that only contain information about successfully matched driver-order pairs.

Furthermore, in order to better capture heterogeneity within driver population, drivers are clustered into different groups based on operation performance. Driver clustering makes up for MNL's lack of flexibility in describing heterogeneity to a certain extent and provides straightforward comparison scenarios. Moreover, the commonly adopted extension for MNL, Mixed Logit Model (MIXL), is also proposed to allow for random heterogeneity within specific

attributes and to further validate results from MNL. The purpose of incorporating both MNL and MIXL is to deliver quantitative estimates of drivers' preferences regarding proposed attributes. The correlation and tradeoff between different attributes are explored, as well.

Estimates from MNL and MIXL provide a more comprehensive understanding of drivers' operation behaviors in a ride-sharing system, thus not only enriching related behavioral studies but also offering improvements for ride-sharing systems' operation performance. When a ride-sharing platform matches queuing orders and available drivers in unique pairs, the value of each pair can be evaluated from multiple angles, such as revenue contribution from the order, the driver's operation experience, and the passenger's trip experience. Using these components, ride-sharing developers can improve the system's matching strategy and improve its operation efficiency. If using weight to describe the value of a unique driver–order pair, different weighting strategies can be proposed to achieve different operation objectives.

In order to comprehensively explore operation strategies, one simulation framework for ride-sharing systems is developed in this research. Given stochasticity in both demand and supply distribution, operation optimization strategies that consider drivers' behaviors are developed. The proposed strategies make use of estimates from MNL to improve matching efficiency in respect to matched order volume and drivers' operation experience. While applying different operation strategies, the simulation framework delivers criteria for system performance in detail. Thus, comprehensive analysis of various strategies' effects can be conducted, which further supports operation improvement.

Chapter 2: Development of Multinomial Logit Model (MNL) for Drivers' Operation Behaviors

Study of individual choices has been an important subject in various research fields, such as urban economics, marketing and transportation. A comprehensive analysis for participants' choices can enrich understanding toward human behaviors as well as provide insights for numerous operation contexts, including commodity pricing, product design and resource allocation. Multinomial Logit Model (MNL), developed by McFadden in 1974 (8), has been the most widely applied method for the study of individual choices. The theory basis for MNL is that individual makes choices according to Random Utility Maximization (RUM). MNL's advantages are mainly exhibited by efficient estimation and simple interpretation for model results. With a closed-form choice probability and globally concave likelihood function, MNL can be estimated by Maximum Likelihood Estimator (MLE) efficiently (9).

In the context of ride sharing, a comprehensive analysis for drivers' operation behaviors is important as it supports operation strategies from multiple aspects, such as subsidy allocation for supply boosting and efficiency improvement for driver-order matching scheme. With operation records in format as introduced in Section 2.1, alternatives and attributes set can be established based on reasoning presented in Section 2.4.

On the other hand, MNL assumes uniform preferences across individuals, which can be unrealistic especially for essentially heterogeneous behavioral respondents. In order to make up for MNL's lack of flexibility, drivers on a specific sample date are clustered into 6 groups according to their operation performance. MNL is then integrated to observations associated with each driver cluster separately. When conducting drivers' behavioral study, both

independent analysis based on estimates from single driver cluster and cross-group comparison analysis are implemented. The multi-level analysis provides behavioral insights from different views and validation for the proposed framework of MNL implementation from various perspectives as well. Additionally, Mixed Logit Model (MIXL), which is a popular extension for MNL (9), is also implemented with sample dataset in order to explore individual-specific preference.

2.1 Data Description

In this research, the accessed data resource is operation record from DiDiChuXing[3], a major ride sharing company in China. The released dataset contains order requests of DiDiChuXing's ride sharing service in an anonymous city during January 1st to January 21st in 2016. Both order requests that successfully matched with available driver and those not responded due to lack of supply are included in the the dataset. The volume of order requests is 406,553 on daily average, with the highest at 535,823 and the lowest at 322,284. Each record describes basic information for a trip request, which is organized as shown in Table 1. As stated above, the operation records are based on order requests, which are not direct behavioral information. In order to perform behavioral study with MNL, a comprehensive method for extracting information of drivers' operation decisions and operation statues is proposed in Section 2.4.

Table 1 Data description for operation record

Order_ID	Order_ID is unique for each request sent; However, if customer sent request repeatedly due to failure of responds, the order_id would stay same. This leads to two potential insights: repeating requests should be considered as one demand when performing demand prediction; with numerous records, the customers' waiting budget could be approximated from data.

[3] http://research.xiaojukeji.com/, accessed January 2017. The dataset was released for DiDi Algorithm Contest and used for this research through request.

Driver_ID	Driver_ID is unique as well as consistent across the dataset. For an order record, if driver_id =='NULL', that means the order is not matched. (due to lack of supply).
Passenger_ID	Unique and consistent across the dataset
Start_District	Hashing code; 66 areas. Size for each area is relatively small, with average price of inner trip around CNY 8 (2km)
Dest_District	Hashing code; There are approximate 360 areas and consistent with start_district code. Trips within marked starting areas count form over 75% of total operation records.
Price	CNY
Date	2016-01-01 – 2016-01-31; around 400,000 orders per day
Time	i.e. 10:00:00

2.2 Mathematical illustration for Multinomial Logit Model

A full review for MNL can be found at (8,9). If there are J alternatives in choice alternative set and J>= 2, MNL summarizes random utility for individual *n* to choose alternative *j* as:

$$U_{jn} = V_{jn} + \varepsilon_{jn}, \varepsilon_{jn} \sim EV1(0, \lambda)$$

$$= X_{jn}\beta + \beta_j + \varepsilon_{jn}$$

Where, V_{jn} describes deterministic utility and ε_{jn} stands for independent random error term. Specifically, MNL assumes independent and identically distributed extreme value type 1 for error term. The model estimation can be derived from proposed distribution for ε_{jn} accordinly. X_{jn} stands for attributes set that perceived by individual *n* if choosing alternative *j*.

Following equation gives estimated probability for individual *n* to choose alternative *j*:

$$P_{jn} = \Pr(j|C_n) = \Pr\left(U_{jn} \geq U_{in}, \forall i \in C_n, i \neq j\right)$$

$$= \Pr\left(\varepsilon_{in} - \varepsilon_{jn} \leq V_{jn} - V_{in}, \forall i \in C_n, i \neq j\right)$$

$$= \int_{\varepsilon n} I\left(\varepsilon_{in} - \varepsilon_{jn} \leq V_{jn} - V_{in}, \forall j \in C_n, i \neq j\right) f(\varepsilon_n) d\varepsilon_n$$

Where, C_n is the choice alternative set for individual n.

According to inference above, only utility difference between alternatives can be identified. Estimation for choice probability is then based on utility difference. If applying Gumbel or Fisher-Tippet extreme value type I (EVI) as assumed distribution for random error term, ε_{jn}, choice probability estimation can be derived as following:

$$P_{jn} = \frac{\exp\left(\lambda^{-1} V_{jn}\right)}{\sum_{i \epsilon C_n} \exp\left(\lambda^{-1} V_{in}\right)}$$

Where, λ is the scale parameter and scaled to 1.

Maximum Likelihood Estimator (MLE) is widely adopted for model estimation, especially for large scale sample dataset(9). MLE attempts to maximize joint probability for observing the dataset. Coefficients' estimates are generated from derivation of likelihood function presented as follows. The global concave feature of likelihood function makes it convenient to be optimized.

$$\ell(\beta; y|x) = \prod\prod {P_{jn}}^{y_{jn}}$$

Where, y_{jn} equals to 1 if individual n choose alternative j and equals to 0 if not.

Additionally, Relative Risk Ratio (RRR) of MNL is applied for results analysis. Relative Risk Ratio in MNL represents that when a specific independent variable changes by one unit, the ratio value of choice probabilities for two alternatives. Then, RRR of conditional logit can be formulized as below.

$$RRR = \frac{P_{mt'}}{P_{nt'}} \Big/ \frac{P_{mt}}{P_{nt}}$$

$$\frac{P_{mt}}{P_{nt}} = \frac{\exp\,(X_{mt}\beta)}{\exp\,(X_{nt}\beta)} = \exp\,(X_{mt} - X_{nt})\beta$$

$$RRR^{u} = exp^{\beta_u}$$

Where, X_{mt} denotes regressor that corresponds to individual *t* choosing alternative *m* and $X_{mt'}$ represents regressor that changes by one unit in specific attribute u from X_{mt}.

2.3 Driver Clustering

Drivers participating in the ride-sharing system are essentially different from each other in respect of revenue objective, schedule flexibility, and preference for geographic area. Meanwhile, drivers also have independency for operation decision, which results in a dynamic fleet with greater heterogeneity regarding operation behavior as compared with a fixed driver fleet. Figure 1 illustrates percentage of total supply that provided by drivers with different volume of orders taken on January 2nd, 2016.

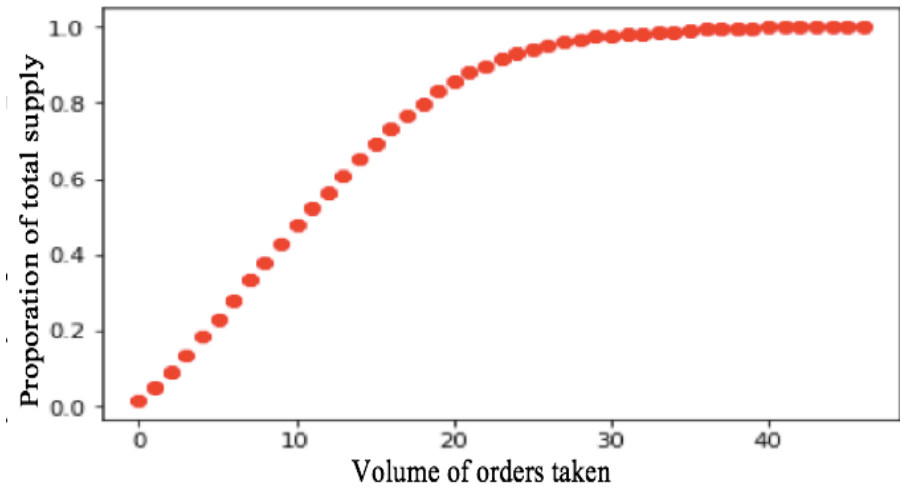


Figure 1 Curve of supply population

According to Figure 1, the volume of orders taken for an individual driver can vary from 1 to 52 within the driver fleet. Drivers with volume of orders taken lower than 10 provide approximate 49% of total trip services. The observation suits the intuition that a ride sharing platform attracts participation of many part-time drivers. These part-time drivers operate with a more flexible schedule compared with full-time drivers, and tend to be affected more by

operation statuses. Meanwhile, the curves for supply population on different dates can be well fitted by similar sigmoid functions, suggesting a stable pattern of population component.

In order to study drivers' behaviors from their discrete choices, investigating existence of different driver groups can be helpful for better capturing heterogeneity within drivers' characteristics. Drivers' revenue performance is of major importance as it directly reflects drivers' revenue objective.

Then, three attributes, drivers' total volume of orders taken, average revenue per order and total revenue, are proposed to constitute driver clustering criteria. Specifically, drivers are clustered into 6 groups based on optimization of within-variance. Besides, analysis for drivers' operation process requires an adequate number of observations per driver to reflect driver's choice situations. Therefore, when performing driver clustering, drivers with volume of orders taken lower than 5 are exclude from driver sample.

The popular unsupervised algorithm, K-means clustering, is applied for driver grouping in this research (11). A brief mathematical formulation for K-means clustering is recalled as follows.

Given a dataset that contains n points that described by x;

$k\ is\ the\ number\ of\ clusters\ chosen\ by\ reseracher$;

$C_1, C_2 \dots C_k\ denote\ the\ partition\ of\ points$;

$n_p\ represents\ number\ of\ points\ in\ C_p$;

$M\ is\ the\ clustering\ result$;

Various optimization objectives can be used for K-means clustering. In this research, minimizing within-cluster variance is adopted. Then the objective formulization of K-means clustering can be illustrated below (11).

$$M = \min \sum_{p=1}^{k} \sum_{i \in C_p} \left\| x_i - r_p \right\|_2^2$$

$$where, r_p = \frac{1}{n_p} \sum_{x \in C_p} x \text{ , representing average position of } C_p$$

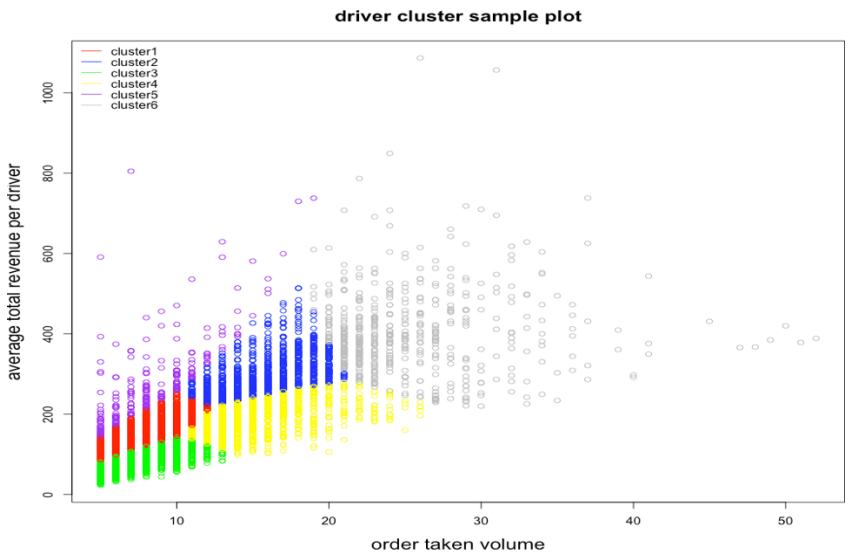


Figure 2 Driver clustering on sample date January 5th, 2016

Figure 2 above illustrates clustering results on a sample date of January 5th, 2016.

According to Figure 2, except for driver cluster marked by purple, the other four clusters all concentrate on relative small domain on both volume of orders taken and total revenue. On contrast, the cluster marked by purple distributes sparsely along both axes and contains the fewest observations. Therefore, the other five clusters are supposed to more promising samples for MNL integration.

2.4 Specification for MNL's choice and attribute set

On a ride sharing platform, where participating drivers have full control for when and where to enter or leave the system, drivers' operation behaviors can be analyzed as a series of discrete choices. After independent evaluation of current operation status, an individual driver

continuously makes operation decisions about whether to provide another trip service. The outcome choices are assumed to be those associate with the largest random utility.

Based on empirical context, drivers' operation decisions can be summarized into three alternatives, continuous operation, taking a break and exiting the system. Continuous operation represents that the driver is ready to provide another service after completing previous order. Break means that the driver will take a break after completion and enter the system again after break. Exit describes situation where driver exiting the system without entering again during this calendar date.

However, drivers' operation status is more complicated to describe as it can involve numerous factors that are difficult to identify. Besides, the factor sets considered by different individuals can be heterogeneous. Therefore, an appropriate description method for operation status, based on data resources that commonly available, is essential for conducting drivers' behavioral study. Despite for differences across individuals, drivers make operation decisions mostly based on tradeoff among revenue objective, level of fatigue and schedule flexibility. While schedule flexibility for each individual driver is hard to identify, revenue objective and level of fatigue can be observed from operation records. Further, drivers' operation experiences regarding speed of reward accumulation, efficiency of order matching and one specific order's characteristic also correlates to drivers' operation decisions. Then, considering availability of data resources as well, three major components are involved into establishment of description method for operation status, which are system environment, newly assigned order's characteristic and drivers' own operation history during the day.

According to reasoning above, MNL's attribute set is proposed as summarized in Table 2. Specifically, there are three alternative-specific attributes and 16 individual-specific attributes. Detailed explanations for each attribute are presented as well.

Table 2 Attribute set for MNL

	Notation	**Simplified notation**[4]
Alternative Specific Attributes		
Expected value of next revenue	$\mathrm{E}(revenue_k + revenue_{k+1})$	*expm*
Expected resting time length	$\mathrm{E}(rest_t)$	*rest*
Expected working time	$\mathrm{E}(work_k + work_{k+1})$	*e_w*
Expected revenue per time unit for next order	$\mathrm{E}(rpt_{k^5} + rpt_{k+1})$	*rpt*
Individual Specific Attributes		
(Drivers' individual record)		
Revenue from previous order	$revenue_{k-1}$	*pervious_o*
Average reward for the day	$\overline{revenue}_k$	*ave_r*
Accumulative reward for the day	$\sum_{i=1}^{k-1}(revenue_i)$	*acc_r*
Accumulative reward for the period	$\sum_{i=b}^{k-1}(revenue_i)$	*acc_p*
Accumulative working time for the day	$\sum_{i=0}^{t}(work_i)$	*acc_t*

[4] Simplified notations are exchangeable with full notations in following discussion.
[5] The "rpt" is abbreviation for "reward per unit of time".

Accumulative working time for the period	$\sum_{i=p}^{t}(work_i)$	*acc_d*
Volume of order taken for the day	*order_vol*	*order_vol*
Volume of order taken for the period	*order_p*	*order_p*
Time difference between two consecutive orders	$timestamp_k - timestamp_{k-1}$	*diff_t*
Weighted moving average for revenue per order	$\overline{revenue_{w,k}}$	*wma*
Percentage of current order's reward accounted for current total reward	$(revenue_k / \sum_{i=1}^{k}(revenue_i)) * 100\%$	*percentage_c*
(Current order's time characteristic and related system condition)		
Whether during morning rush hour: 08:00-09:00	*Drush1*	*Drush1*
Whether during afternoon rush hour: 17:30-18:30	*Drush2*	*Drush2*
Whether during nap time: 14:00-15:00	*Dnap*	*Dnap*
Percentage of current time slice's order volume accounted for daily total	*vol_p*	*vol_p*

Explanation for alternative specific attributes

The four alternative specific attributes give estimates for drivers' expected values in respect of monetary reward, working time and rest time if taking one newly assigned order. These attributes change with drivers' operation decision accordingly. Detailed explanations for each alternative specific attribute are listed as following.

$\mathrm{E}(revenue_k + revenue_{k+1})$: It represents driver's expectation for next revenue, including revenue from current order k and next order, $k+1$. Notice that when operating on DiDiChuXing, drivers do not know exact revenue for current order until completion. Therefore, current order's revenue is also included into drivers' expectation for next revenue. For each alternative in the choice set, corresponding expected revenue is estimated with major assumption that drivers expect next revenue according to historical revenue information. In this research, mathematical description for driver's expectation of next revenue when deciding to operate continuously is proposed as following.

$$\mathrm{E}(revenue_k + revenue_{k+1})_{continous\ opertion}$$

$$= \boldsymbol{\beta_1}(\text{previous order's revenue} + current\ order's\ revenue$$

$$+ \ \text{weighted moving average revenue} + \text{average revenue})$$

where, $\boldsymbol{\beta_1}$ *represent a constant weighting vector*

When estimating expected revenue for taking a break and exiting the system, time of value is applied to approximate discount rate. That means, when driver takes a break or exits the system, expectation for next revenue is still included as one component but discounted when transformed from future value to present value. Naturally, the value is discounted along the time interval between two consecutive operating period. Equation below presents the formulation of expected revenue for operating choices of taking a break and exiting the system.

$$\mathrm{E}(revenue_k + revenue_{k+1})_{taking\ break/exit}$$

$$= \frac{\mathrm{E}(revenue_k + revenue_{k+1})_{continous\ opertion}}{\left(1 + r_{exp}\right)^n}$$

where,

r_{exp} is the discount rate and n is the number of time periods in between .

$\mathrm{E}(rest_t)$*:* $\mathrm{E}(rest_t)$ stands for expected time length of non-operating at choice situation t, which varies according to operation choice. In this research, drivers are assumed to estimate resting length based on historical time interval between orders.

$\mathrm{E}(work_k + work_{k+1})$: $\mathrm{E}(work_k + work_{k+1})$ represents expected working time in hour, including estimated working time for current order k and next order $k + 1$.

$$\mathrm{E}(work_k + work_{k+1})_{continous\ operation}$$

$$= \boldsymbol{\beta}_2(current\ order's\ estimated\ time\ length$$

$$+ previous\ order's\ time\ length)$$

where,

$$\boldsymbol{\beta}_2\ represents\ a\ constant\ weighting\ vector$$

Similarly, $\mathrm{E}(work_k + work_{k+1})$ for taking a break and exiting the system are applied discount rate to reflect drivers' perception for future working time.

$$\mathrm{E}(work_k + work_{k+1})_{taking\ a\ break\ |\ exiting\ system}$$

$$= \frac{\mathrm{E}(work_k + work_{k+1})_{continous\ operation}}{\left(1 + r_{e\_w}\right)^{\mathrm{n}}}$$

where,

r_{e_w} is the discount rate and n is the number of time periods in between .

$\mathrm{E}(rpt_k + rpt_{k+1})$*: rpt* describes drivers' expected reward per time unit. As neither expected reward nor expected working time alone can describe drivers' tradeoff in between,

$\mathrm{E}(rpt_k + rpt_{k+1})$ is included. The $\mathrm{E}(rpt_k + rpt_{k+1})$ for each alternative is obtained from $\mathrm{E}(revenue_k + revenue_{k+1})$ divided by $\mathrm{E}(work_k + work_{k+1})$ accordingly.

Explanation for individual specific attributes

Drivers under different operation status are considered to be different behavioral respondents. There are 16 individual specific attributes in total, which describe drivers' heterogeneous operation status from three aspects, system environment, newly assigned order's characteristic and drivers' own operation history during the day.

$revenue_{k-1}$: $revenue_{k-1}$ is the revenue from previous order , $k-1$;

$\overline{revenue}_k$: $\overline{revenue}_k$ is a driver's average reward per order until current order k;

$\sum_{i=1}^{k-1}(revenue_i)$: $\sum_{i=1}^{k-1}(revenue_i)$ is a driver's accumulated reward until previous order $k-1$;

$\sum_{i=0}^{t}(work_i)$: $\sum_{i=0}^{t}(work_i)$ is a driver's accumulated working time in the system until choice situation t;

$\sum_{i=p}^{t}(work_i)$: $\sum_{i=p}^{t}(work_i)$ is a driver's accumulated working time length for current working period, from choice situation p to current choice situation t;

order_vol: "order_vol" is the volume of order taken from the beginning of this working day until current order k;

order_p: "order_p" is the volume of order taken from the beginning of current working period until current order k;

$(revenue_k / \sum_{i=1}^{k}(revenue_i)) * 100\%$: $(revenue_k / \sum_{i=1}^{k}(revenue_i)) * 100\%$ is the percentage that revenue from current order k accounted for accumulated revenue.

$timestamp_k - timestamp_{k-1}$: $(timestamp_k - timestamp_{k-1})$ is the time difference between current order k and previous order $k-1$;

$\overline{revenue_{w,k}}$: $\overline{revenue_{w,k}}$ is the weighted moving average revenue per order for individual driver until current order k;

$$\overline{revenue_{w,k}} = \frac{k * revenue_k + (k-1) * revenue_{k-1} + \cdots + 1 * revenue_1}{k + (k-1) + \cdots + 1}$$

vol_p: "vol_p" represents percentage of order volume during current order's time slice accounted for total order volume during the day. Specifically, there are 72 consecutive time slices during one day, with each time slice covering 20 minutes.

Dnap: Dnap is a binary variable that represents whether order's time stamp falls in common nap period, 14:00 – 15:00.

Drush1: Drush1 is a binary variable indicating whether order's time stamp within morning rush hour, 08:00 – 09:00.

Drush2: Drush2 is a binary variable indicating whether order's time stamp within evening rush hour, 17:30 – 18:00.

Notice that for the three dummy variables for special periods, Dnap, Drush1 and Drush2, the time bound can be adjusted according to real time demand and supply condition. However, by using constant time bound here, a general analysis for special periods can be conducted. The robustness of corresponding time bounds can be analyzed by MNL estimates from various samples across driver groups and operation dates.

2.5 Results Analysis for Multinomial Logit Model (MNL)

With MNL specified in Section 2.4, drivers' operation behaviors can be analyzed based on dataset illustrated in Section 3.1. In this section, estimates from MNL are presented along with interpretations in detail. As stated in Section 2.1 and 2.3, the dataset contains operation records from January 1st to January 21st in 2016, and 6 driver clusters are sampled for each date. For each driver cluster, corresponding operation records constitute one sample dataset for MNL implementation.

From results generated by samples across dates, MNL proposed in Section 4 delivers promising estimates that suit intuition and have statistical significance for most of the coefficients. Meanwhile, comprehensive insights can be summarized through results analysis of independent samples as well as comparisons between different samples.

The two subsections demonstrate results analysis for a single sample on specific date and also samples across dates and groups. The multilevel analysis aims to present independent interpretations of coefficients and correlation within estimates as well. Besides, the change

pattern of coefficients across different driver groups and dates can be identified through comparison analysis. The consistency of MNL estimates across various samples also validate reasoning and methodology proposed.

2.5.1 Results analysis for independent sample

Table 3 MNL Results Sample for one driver cluster on January 5th, 2016

Attributes	Estimates	Relative risk ratio	Individual specific attributes with alternative specific coefficients	Estimates	Relative risk ratio
Alternative specific constants			1:wma	0.074	1.077
1:(intercept)	3.328	27.873	2:wma	0.038	1.039
2:(intercept)	0.782	2.186	1:previous_o	0.027	1.028
Alternative specific attributes with generic coefficients			2:previous_o	0.030	1.030
rest	0.024	1.024	1:acc_d	0.094	1.099
rpt	0.458	1.581	2:acc_d	0.153	1.165
Alternative specific attributes with alternative specific coefficients			1:acc_t	-0.023	0.977
3:e_w	-1.723		2:acc_t	-0.022	0.978
1:e_w	-1.944	0.802	1:acc_r	-0.009	0.991
2:e_w	-1.570	1.165	2:acc_r	-0.010	0.990
3:expm	-0.144		1:order_vol	-0.207	0.813
1:expm	-0.147	0.997	2:order_vol	-0.278	0.757
2:expm	-0.132	1.012	1:order_p	-0.173	0.841
			2:order_p	-0.129	0.879
			1:diff_t	0.680	1.974
			2:diff_t	0.117	1.124
			1:ratio	-0.650	0.522
			2:ratio	-1.105	0.331
			1:rush1	-0.864	0.422
			2:rush1	-1.120	0.326
			1:rush2	-0.133	0.875
			2:rush2	0.205	1.228
			1:vol_p	0.399	1.490
			2:vol_p	0.315	1.370
			1:nap	0.522	1.686
			2:nap	0.882	2.416

Table 3 presents MNL estimates for a single sample of driver cluster on January 7th, 2016[6]. Exiting the system is used as reference level within the three choice alternatives. Estimates with statistical significance at 95% confidence level are highlighted with blue. Detailed

[6] Table 5.1 uses simplified notations of attributes. Comparison table is included in Appendix.

interpretations for specific attributes are illustrated as follows. Besides, inferences in the analyses are certeris pribus.

Table 4 Descriptive statistics for sample cluster

Number of drivers	Average of total revenue(2)	Standard deviation of (2)	Average of order taken volume (3)	Standard deviation of (3)	Average of average revenue per	Standard deviation of (4)
719	307.3	54.9	14.4	2.3	22.5	3.2

Interpretations for alternative-specific attributes

- Alternative Specific Constants (ASC): Compared to exiting the system, ASC for continuous operation and taking a break are both positive and statistically significant, with value of 3.328 and 0.782 respectively. Corresponding risk ratios are 27.873 and 2.186, which indicate that continuous operation and taking a break are both preferred than exiting the system, with probability increase of 268.7% for continuous operation and 118.6% for taking a break.

- $\mathrm{E}(work_k + work_{k+1})$*:* As $\mathrm{E}(work_k + work_{k+1})$ represents expected working time in hour, it's in accordance with intuition to obtain negative estimates with -1.944, -1.570 and -1.723 for all three alternatives respectively, implying that longer expectation for continuous operation time relates to lower drivers' utility. Relative risk ratios for continuous operation and taking a break are 0.802 and 1.165 respectively. When expected working time increases by one hour, drivers' probability of continuous operation is 0.802 times of exiting the system, a 19.8% decrease. On the other hand, drivers' probability for taking a break is 1.165 times of exiting the system, a 16.5% increase.

- $\mathrm{E}(rest_t)$*:* The positive estimate, 0.024, suggest that drivers prefer schedule flexibility for choice decision.

- *exp:* The negative estimates, -0.147, -0.132 and -0.144, for *exp* seem to be contrast to intuition, as *exp* describing expectation for reward. However, this can be reasonably explained by more comprehensive understanding for drivers' decision process. Expectation for reward can not be analyzed alone as each order must associate with certain operation time length, which necessarily leads to tradeoff between time and reward. The unexpected sign for *exp* can be caused by the attribute's lack of independent explaining capability.

- $\mathrm{E}(rpt_k + rpt_{k+1})$*:* Embedding of rpt within attributes set aims to capture tradeoff between time and reward more comprehensively. MNL delivers positive estimate ,0.458, for coefficient of rpt as expected, indicating that drivers prefer higher reward per time unit.

Interpretations for individual-specific attributes

- $revenue_{k-1}$: This alternative-specific attribute has positive estimates, 0.027 and 0.03, for both continuous operating and break, as well as with statistical significance.

  If revenue from pervious order increases by one Yuan in CNY, driver's probability for continuous operation is 1.028 times of exiting the system, representing 2.8% increase. Driver's probability for taking a break is 1.030 times of exiting the system, a 3.0% increase as well. The increases in probabilities imply that higher reward from pervious order can encourage drivers to stay in the system, delivering insights for how historical reward can serve as incentive for boosting supply.

- $\sum_{i=1}^{k-1}(revenue_i)$: The relative risk ratios for accumulated reward are 0.991 and 0.990 for continuous operation and taking a break respectively. If accumulated reward increases by one Yuan in CNY, drivers' probability of continuous operation is 0.991 times of exiting the system, indicating a 9% decrease.

- *order_vol*: When volume of order taken increases, it's intuitive that drivers' level of fatigue increases and in different way from tiredness accumulated by working time length. Except for operation time length, volume of order taken also includes drivers' working process for picking up and dropping off passengers. The estimates are negative and also with statistical significance for continuous operating or just taking a break, with value of -0.207 and -0.268 respectively. After taking one more order, driver's probability of continuous operation will decrease by 18.7% as compared to exiting the system. On the other hand, driver's probability of taking a break is 0.931 times of continuous operation, a 6.9% decrease.

- *order_p*: Negative estimates are generated with order_p for both continuous operation and taking a break, with value of -0.173 and -0.129 accordingly. It suits with intuition that higher order volume taken for current working period relates to lower probability for staying in the system. While completing one more order during current working period, driver's probability of continuous operation is 0.841 times of exiting the system, a 15.9% decrease.

- $(revenue_k / \sum_{i=1}^{k}(\square evenue_i)) * 100\%$: The attribute relates to both accumulative reward and expected reward and has a general decreasing trend along time. When the percentage vibrates within a relatively low value interval, a high price order might lead the driver to exit the system. This can be explained that with tiredness and reward being accumulated, drivers expect a high price order to complete the work for the day.

- $timestamp_k - timestamp_{k-1}$: During continuous operating period, longer time difference between orders often relates to longer travel distance. The positive estimates for

diff_t, 0.680 for continuous operation suggest that drivers are motivated to operate continuously with longer-distance order.

- *wma*: For appropriate description for drivers' sensibility of their revenue along operating period, weighted moving average reward is applied with higher weight given to closer period. The positive estimates, 0.074 and 0.038, reflect incentive from higher wma for driving behavior. Furthermore, when weighted moving average reward increases by one Yuan in CNY, driver's probability of continuous operation 1.077 times of exiting the system and 1.037 times of taking a break, equaling to 7.7% increase and 3.7% increase respectively.

  Additionally, when incorporating drivers' average revenue per order into attributes set, it does not have statistical significance across samples. It can be concluded that average revenue per order is not an appropriate description for drivers' perception toward their revenue performance. Drivers tend to focus more on revenue from orders closer to current operation time.

- vol_p: vol_p describes system operation status instead of individual driver's. The discussed sample delivers positive estimates of vol_p for both continuous operation and taking a break, with value of 0.399 and 0.315 respectively. During operation period when order volume accounts for one more percentage for total volume of orders, driver's probability to continuously operate is 1.490 times of exiting the system and 1.088 times of taking a break. The increased probability for continuous operation suggests drivers' improved operation experiences during time period with higher order volume.

- Drush1: During morning rush hour, drivers' probability of continuous operation is 0.422 times of exiting the system, even with high order volume during morning rush hour. Also, drivers' probability of taking a break also decreases by 67.4% as compared to exiting the system. The disutility associated with morning rush hour might correlate to drivers' limited schedule flexibility during the period.

- Drush2: Unlike rush 1, rush 2 do not have estimates with statistical significance at 95% confidence level. The estimate magnitude for continuous operation is -0.133 while for taking a break is 0.205. The non uniform estimates imply drivers' heterogeneous schedule flexibility during afternoon rush period.

- Dnap: During 14:00 to 15:00, drivers' probability of taking a rest is 2.416 times of exiting the system and 1.432 times of continuous operation, suggesting 141.6% and 43.2% increase respectively. Then, during specified nap hour, drivers show the highest preference for taking a break.

2.5.2 Results analysis for samples across dates and groups

Section 2.5.1 gives detailed model interpretation for one driver cluster sample. The explanation for results can be supported by both intuition and rational inference. In order to further validate proposed MNL incorporation, Section 2.5.2 delivers analysis for results across driver groups and dates. Most of the samples generate results that can be well explained by consistent reasoning discussed in Section 2.5.1. Also, most attributes are observed to be with statistical significance across samples.

Except for further validating MNL incorporation from model interpretation of independent samples, the analysis across samples can deliver enriched insights regarding heterogeneity among different driver groups and operation dates. From cross group and date comparison scenarios, drivers' heterogeneity and its quantitative effects on drivers' decision process can be analyzed.

Sample Data Description

Datasets for 21 consecutive dates are adopted for numerical example, which covers 2016-01-01 to 2016-01-21.

The driver clustering method discussed in Section 2.2 is adopted for each day, which results in 6*21=126 separate samples. Table 5 illustrates estimates for all 6 driver cluster samples on January 7th, 2016. Dataset from each sample date is processed in the same way. It can be observed that MNL incorporation have well performed consistency regarding estimates' statistical significance as highlighted with blue. Also, for individual-specific attributes that directly describe drivers' individual operation record, estimates from different samples are consistent for whether being positive or negative.

Further, in order to find similar groups across dates, all the samples are clustered into 6 groups again via the same clustering criteria, average revenue per order, total revenue and order taken volume. Table 6 presents statistics for the 6 driver groups and number of samples contained.

Table 5 MNL estimates across groups on January 7th, 2016

Coefficients	***Estimates***					
	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
1:(intercept)	5.325	3.336	3.994	2.172	3.630	3.523
2:(intercept)	2.044	1.552	-0.159	-2.197	0.983	0.403
rest	0.217	0.008	0.059	-0.044	0.000	0.053
rpt	0.068	1.045	0.348	0.314	0.393	0.606
1:wma	0.024	0.080	0.103	0.138	0.058	0.145
2:wma	0.005	0.028	0.062	0.076	0.027	0.048
1:previous_o	0.014	0.040	0.037	0.048	0.032	0.053
2:previous_o	0.012	0.015	0.030	0.042	0.033	0.046
1:acc_d	0.139	-0.064	-0.068	-0.087	0.040	-0.017
2:acc_d	0.062	0.149	-0.104	-0.064	0.041	0.102
1:acc_t	-0.011	-0.042	0.022	-0.120	-0.053	-0.016
2:acc_t	-0.008	0.038	0.068	-0.214	-0.037	-0.053
1:acc_r	0.002	-0.015	-0.004	-0.003	-0.009	-0.014
2:acc_r	0.005	-0.008	-0.005	-0.001	-0.010	-0.010
1:order_vol	-0.304	-0.055	-0.167	-0.057	-0.200	-0.068
2:order_vol	-0.410	-0.192	-0.143	-0.114	-0.182	-0.189
1:order_p	-0.262	-0.211	-0.065	0.034	-0.207	-0.022
2:order_p	-0.181	-0.279	-0.055	0.098	-0.211	0.008
1:diff_t	0.691	0.760	1.028	0.951	0.666	0.905
2:diff_t	-0.076	0.015	0.393	0.556	0.060	0.242
1:ratio	-0.561	-0.383	-1.355	-0.201	-0.905	-1.091
2:ratio	-0.768	-0.200	-1.715	-1.098	-1.304	-1.945
1:rush1	-0.753	-0.490	-1.254	-1.680	-0.870	-2.299
2:rush1	-1.244	-0.364	-1.675	-2.473	-1.209	-2.873
1:rush2	-0.552	0.256	0.078	-0.042	-0.053	-0.405
2:rush2	-0.350	0.761	0.133	0.214	0.053	-0.317
1:vol_p	0.165	0.093	1.151	1.587	0.324	1.092
2:vol_p	0.629	-0.189	1.763	2.219	0.362	1.409
1:nap	-0.507	0.663	0.391	2.275	0.429	0.646
2:nap	-1.078	0.784	0.741	2.498	0.933	0.909
3:e_w	-1.461	-1.476	-3.599	-3.147	-1.509	-3.220
1:e_w	-1.781	-1.879	-3.443	-2.859	-1.780	-3.047
2:e_w	-1.354	-1.315	-3.269	-2.670	-1.372	-2.681
3:expm	-0.068	-0.187	-0.223	-0.282	-0.126	-0.301
1:expm	-0.073	-0.190	-0.228	-0.294	-0.132	-0.313
2:expm	-0.065	-0.194	-0.215	-0.277	-0.122	-0.295

Table 6 Driver groups' statistics

	Group1	Group2	Group3	Group4	Group6
order_vol	7.42	7.93	25.73	15.36	16
ave	12.33	21.55	16.71	14.17	21.25
revenue	91.49	170.89	430.01	217.65	340.00
num of samples	15	11	15	17	14

Specifically, 22 samples belong to Group5. However, these samples suffer from lack of observation and result in poor model performance. Besides, 32 samples that not have record completeness or not result in statistical significance for over 70% of proposed attributes are excluded from this analysis. Table 6 illustrated basic information for driver groups. Observation number for each sample is approximate 15,000. All analyses are using reference level of exiting the system and certeris pribus unless otherwise specified. Following Table 7 presents brief summary for results across 72 samples, with positive estimates highlighted by green. It can be observed that most attributes have consistent estimated signs across different samples.

Table 7 Brief summary of results across samples

Sample ID	X1..intercept.	X2..intercept.	rest	rpt	X1.wma	X2.wma	X1.previous_c	X2.previous_c	X1.acc_d	X2.acc_d	X1.acc_t	X2.acc_t	X1.acc_r	X2.acc_r	X1.order_yo	X2.order_yo	X1.order_p	X2.order_p	X1.diff_t	X2.diff_t	X1.percentage_c	X2.percentage_c	X1.rush1	X2.rush1	X1.rush2	X2.rush2	X1.vol_p	X2.vol_p	X1.nap	X2.npa	X3.e_w	X1.e_w	X2.e_w	X3.exp	X1.exp	X2.exp	Date	Group ID
9	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	2	1
14	1	0	0	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	3	1
29	1	1	1	1	1	1	1	1	1	1	0	1	0	0	0	0	0	0	1	1	0	0	0	0	1	1	1	0	1	1	0	0	0	0	0	0	5	1
36	1	1	0	1	1	1	1	1	1	1	0	1	0	0	0	0	0	0	1	1	0	0	0	0	1	1	1	0	1	1	0	0	0	0	0	0	6	1
38	1	1	1	1	1	1	1	1	0	1	0	1	0	0	0	0	0	0	1	0	0	0	0	0	1	1	1	0	1	1	0	0	0	0	0	0	7	1
44	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	0	8	1
75	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	1	1	0	1	1	0	0	0	0	0	0	13	1
82	1	1	1	1	1	1	1	1	1	1	0	1	0	0	0	0	0	0	1	1	0	0	0	0	1	1	1	0	1	1	0	0	0	0	0	0	14	1
87	1	1	1	1	1	1	1	1	0	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	1	1	0	1	1	0	0	0	0	0	0	15	1
93	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	16	1
101	1	1	1	1	1	1	1	1	0	1	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1	0	0	0	0	0	0	17	1
104	1	1	1	1	1	1	1	1	1	1	0	1	0	0	0	0	0	0	1	1	0	0	0	0	1	1	1	0	1	1	0	0	0	0	0	0	18	1
109	1	1	1	1	1	1	1	1	0	1	0	1	0	0	0	0	0	0	1	0	0	0	0	0	1	1	1	0	1	1	0	0	0	0	0	0	19	1
120	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	0	20	1
122	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	1	1	0	1	1	0	0	0	0	0	0	21	1
8	1	0	1	1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	2	2
18	1	0	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	3	2
19	1	0	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	4	2
27	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	0	5	2
41	1	1	0	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	0	7	2
50	1	0	0	1	1	1	1	1	1	1	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	9	2
72	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0	0	12	2
80	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	0	14	2
91	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	16	2
97	1	0	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	1	1	0	0	1	1	0	0	1	1	1	1	0	0	0	0	0	0	17	2
106	1	0	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	1	1	0	0	0	0	1	0	1	1	1	1	0	0	0	0	0	0	18	2
23	1	0	1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	1	1	1	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0	0	4	3
26	1	0	0	1	1	1	1	1	0	0	0	1	0	0	0	0	0	1	1	1	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0	0	5	3
31	1	0	1	1	1	1	1	1	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	0	6	3
40	1	0	1	1	1	1	1	1	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0	0	7	3
46	1	0	0	1	1	1	1	1	0	0	0	0	0	1	0	0	1	1	1	1	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	8	3
51	1	0	1	1	1	1	1	1	0	0	0	0	0	1	0	0	1	1	1	1	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0	0	9	3
57	1	0	0	1	1	1	1	1	0	0	0	1	0	1	0	0	1	1	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	0	10	3
63	1	0	1	1	1	1	1	1	0	0	0	0	0	1	0	0	0	1	1	1	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0	0	11	3
67	1	0	1	1	1	1	1	1	0	0	0	0	0	1	0	0	1	1	1	1	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	12	3
77	1	0	0	1	1	1	1	1	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	0	13	3
81	1	0	1	1	1	1	1	1	0	0	0	0	0	1	0	0	1	1	1	1	1	0	0	0	0	1	1	1	1	1	0	0	0	0	0	0	14	3
100	1	0	0	1	1	1	1	1	0	0	0	0	0	0	0	0	1	1	1	1	0	0	1	1	1	1	1	1	1	1	0	0	0	0	0	0	17	3
105	1	0	0	1	1	1	1	1	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	18	3
112	1	0	0	1	1	1	1	1	0	1	1	0	0	0	0	0	1	1	1	1	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0	0	19	3
126	1	0	0	1	1	1	1	1	0	0	0	1	0	0	0	0	1	0	1	1	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0	0	21	3
12	1	0	1	1	1	1	1	1	0	0	0	1	0	0	0	0	1	0	1	1	0	0	1	1	0	1	1	1	1	1	0	0	0	0	0	0	2	4
17	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	1	0	1	1	1	0	0	0	0	0	0	0	0	3	4
22	1	0	0	1	1	1	1	1	0	1	1	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	4	4
25	1	1	1	1	1	1	1	1	0	1	1	1	0	0	1	0	0	0	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	0	5	4
32	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	6	4
42	1	0	0	1	1	1	1	1	0	1	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	7	4
52	1	0	1	1	1	1	1	1	0	1	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	9	4
59	1	0	0	1	1	1	1	1	0	0	1	1	0	0	0	0	0	0	1	1	0	0	1	1	0	0	1	1	1	1	0	0	0	0	0	0	10	4
62	1	1	1	1	1	1	1	1	0	1	1	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	11	4
71	1	0	1	1	1	1	1	1	0	1	1	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	12	4
78	1	1	1	1	1	1	1	1	0	0	1	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	13	4
83	1	0	0	1	1	1	1	1	0	0	1	1	0	0	0	0	0	0	1	1	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0	0	14	4
95	1	0	0	1	1	1	1	1	0	1	1	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	16	4
102	1	0	0	1	1	1	1	1	1	1	0	1	0	0	0	0	1	0	1	1	0	0	1	1	0	0	1	1	0	0	0	0	0	0	0	0	17	4
107	1	0	0	1	1	1	1	1	0	1	1	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	18	4
110	1	0	1	1	1	1	1	1	0	1	1	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	19	4
124	1	1	1	1	1	1	1	1	0	0	1	1	0	0	1	0	0	0	1	1	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	21	4
24	1	0	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	4	6
28	1	0	1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	5	6
34	1	1	1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	6	6
39	1	0	1	1	1	1	1	1	0	0	0	1	0	0	0	0	0	0	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	0	7	6
47	1	0	1	1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	8	6
53	1	0	0	1	1	1	1	1	1	1	0	1	0	0	0	0	0	0	1	1	0	0	1	1	0	1	1	1	1	1	0	0	0	0	0	0	9	6
55	1	0	1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	0	1	1	0	0	0	1	0	1	1	1	1	1	0	0	0	0	0	0	10	6
65	1	1	1	1	1	1	1	1	0	0	1	1	0	0	0	0	0	1	1	1	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0	0	11	6
79	1	0	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	14	6
90	1	0	1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	15	6
94	1	0	0	1	1	1	1	1	0	0	0	0	0	0	0	0	1	1	1	1	0	0	1	1	0	0	1	1	1	1	0	0	0	0	0	0	16	6
98	1	0	1	1	1	1	1	1	0	0	1	1	0	0	0	0	0	0	1	1	0	0	1	1	1	1	1	1	1	1	0	0	0	0	0	0	17	6
108	1	0	1	1	1	1	1	1	0	0	1	1	0	0	0	0	0	1	1	1	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0	0	18	6
125	1	1	1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	21	6

Coefficients analysis for time sensitive attributes

Within the attributes set proposed as in Table 2, vol_p, Drush1, Drush2 and Dnap are four individual specific attributes that related to order's time characteristic. Instead of describing individual driver's operation record during the day, these four attributes capture time sensitivity within drivers' operation status. Time sensitive attributes not only reflect consideration for drivers' schedule flexibility, but also the change of system's operation environment along time.

- **Analysis for *vol_p***

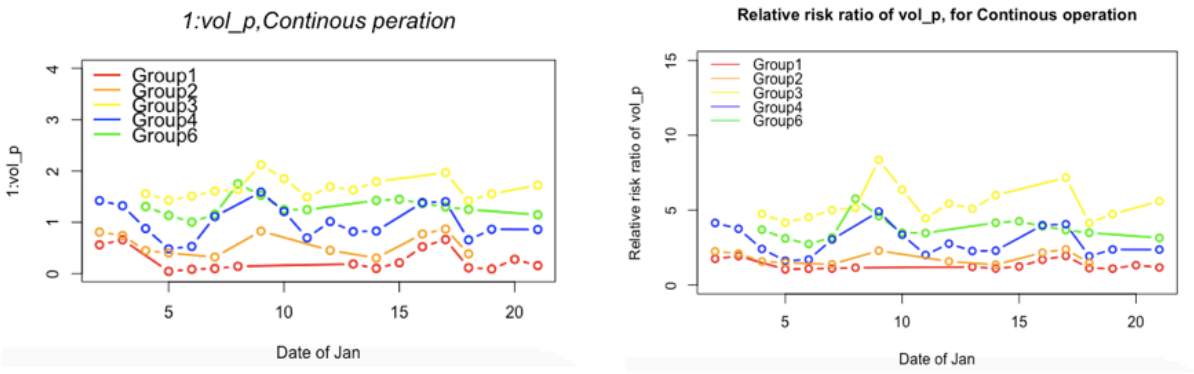


Figure 3 Coefficients and RRR for vol_p across dates and groups for continuous operation

Table 8 Average of vol_p estimates for continuous operation across groups

	Average of vol_p Estiamtes for continous operation				
Group_ID	Group1	Group2	Group3	Group4	Group6
vol_p	0.261	0.576	1.667	1.004	1.308
RRR	1.331	1.821	5.399	2.88	3.764

According to Figure 3, *vol_p* has positive coefficients for continuous operation across all the samples. The estimates are also with statistical significance at 95% confidence level or above. When drivers interact with a ride-sharing platform with higher order volume density, their operating experience can be improved from shortened idling period. Generally, drivers can expect higher operation efficiency during time period with high order volume and increased probability for achieving higher revenue goal. Intuitively, drivers can be motivated for

continuous operating by higher order volume density, with all other condition hold. The positive estimates for *vol_p's* coefficient not only validates consistency with general intuition but also illustrates pronounced pattern across groups.

It can be observed that coefficients for different groups correspond to different value domain and have significant magnitude relationship when compared with each other. Actually, the order of groups regarding associated *vol_p's* coefficient value is in correspondence with order regarding groups' average total revenue per driver. For example, with highest revenue value of 430.01 CNY among groups, Group3 also has largest estimated coefficients for *vol_p* with average value of 1.667. Drivers' revenue performance reflects their unique operation strategy and business object. Larger coefficient for *vol_p* indicates drivers' higher sensitivity toward order volume density and tendency for continuous operating. A better optimized strategy that results in higher revenue performance often relates to better judgment of order volume density fluctuation within the whole day.

Besides, *vol_p's* coefficients also show magnitude difference between weekdays and weekends. For all five groups from date of January 1st , 2016 to January 21st,2016, vol_p's coefficients' often increase during weekends and stay relatively flat during weekdays. The two peak values, 2.12 and 2.01, in Figure 5.5 correspond to two weekends (January 9th, January 16th , January 17th) respectively. During weekends, drivers are less affected by their other daily routine and have higher level of schedule flexibility toward ride-sharing business. Increased operation efficiency resulted from higher order volume density will be valued more by a more schedule-flexible driver population.

- **Analysis for Drush1 and Drush2**

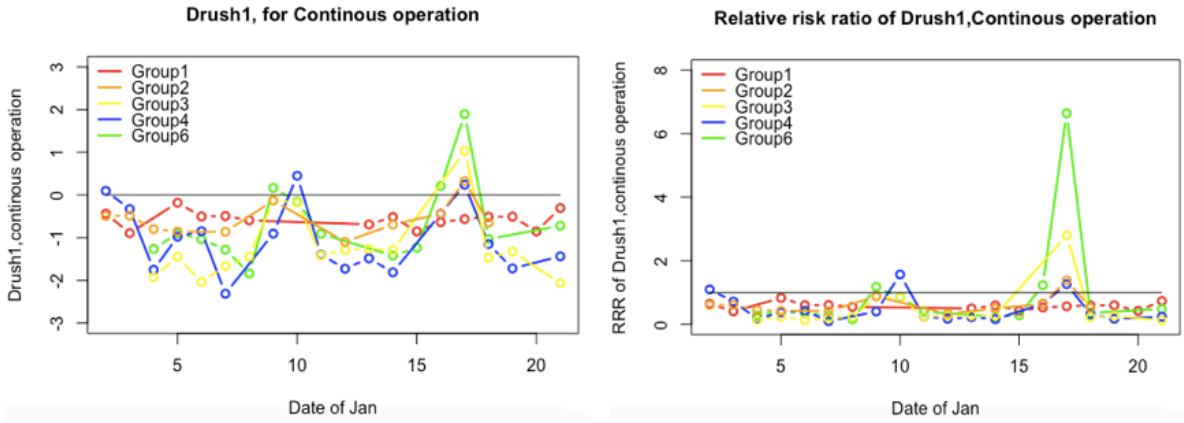


Figure 4 Coefficients and RRR of Drush1 across dates and groups for continuous operation

Attribute *Drush1* describes whether current order belongs to morning rush hour, which is 8:00am to 9:00am. Figure 4 above illustrates coefficients for Drush1 across dates and groups. The coefficients have negative value for most of samples as well as statistical significance. Although weekdays and weekends show different trip demand pattern along the day, the morning rush hour always correspond to high order volume density. Despite for positive motivation generated by high order volume density, drivers have lower probability for continuous operation during morning rush hour, which can correspond to drivers' limited schedule flexibility during this time. Negative estimates with smaller absolute values and a few positive estimates all occur during weekends, which can be explained by drivers' increased schedule flexibility on Saturday and further validates previous reasoning.

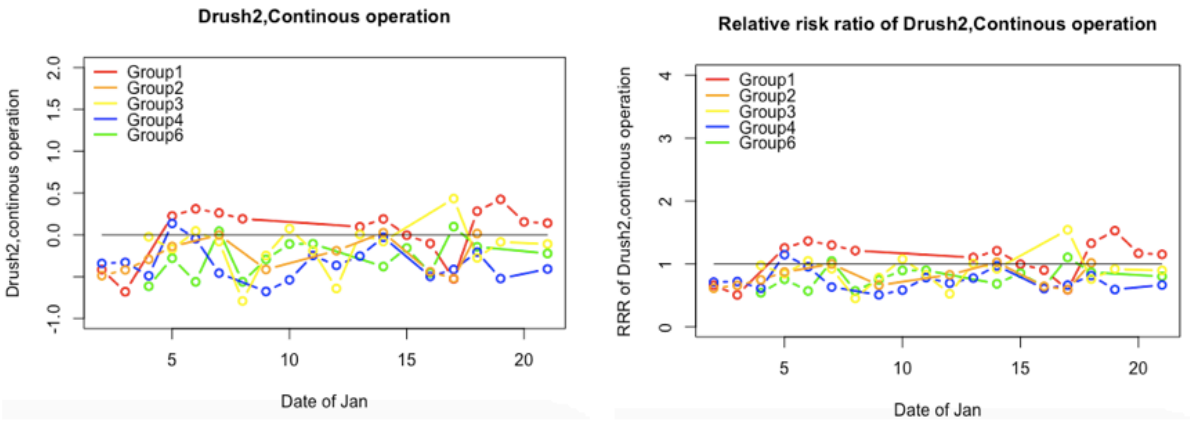


Figure 5 Coefficients and RRR of Drush2 across dates and groups for continuous operation

Attribute Drush2 represents whether current order occurs within afternoon rush hour, which corresponds to 17:30pm to 18:30pm. Unlike Drush1, Drush2 does not have uniform negativity nor positivity for its coefficient estimates across samples, which indicates more significant heterogeneity within drivers' preference toward afternoon rush period. Group1, which has the lowest revenue among groups, has positive estimates for most of its samples. The observation suggests that Group1 contains drivers who tend to operate part-timely after work.

- **Analysis for Dnap estimates**

Based on data observation, 14:00 pm to 15:00pm does not have pronounced high order volume density but often relates to significant demand and supply gap. This might relate to common Chinese life style for taking noon nap. The individual specific attribute Dnap describes whether current order happens during noon nap time.

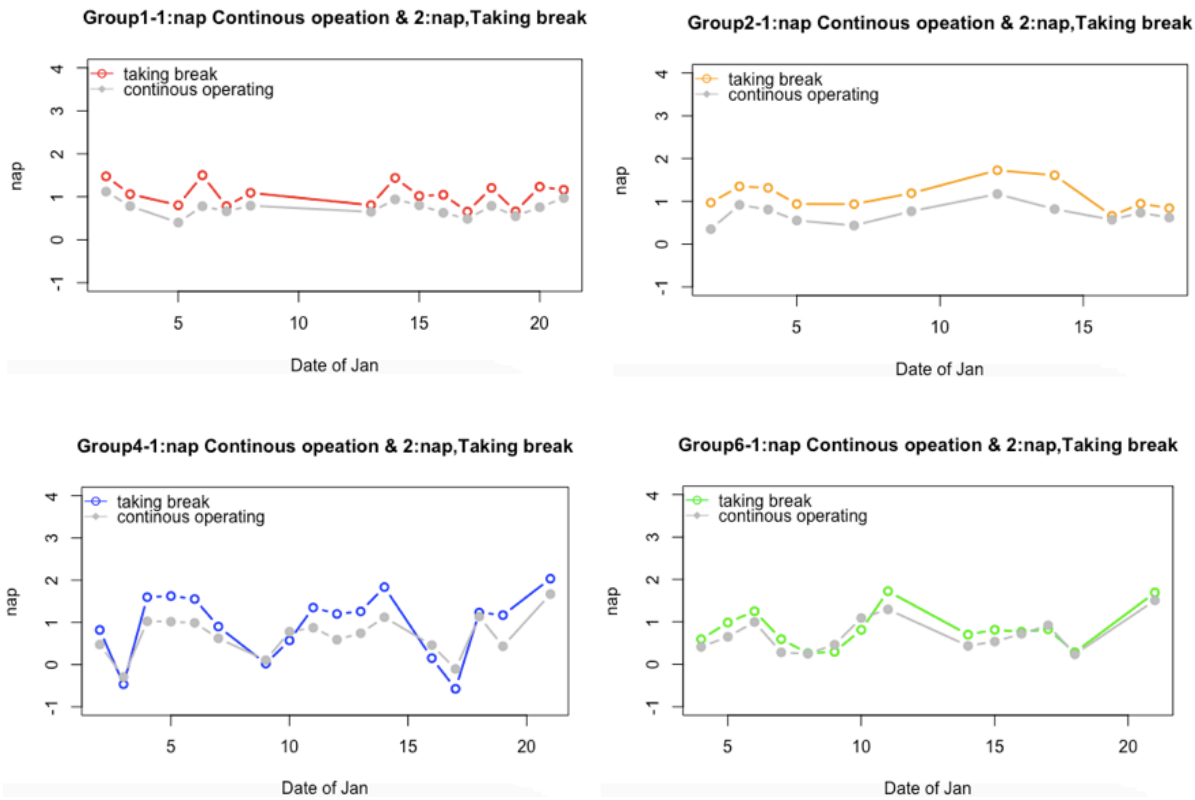


Figure 6 Coefficients for nap across dates and groups

Table 9 Average of nap estimates across groups

Average of nap Estimates across groups			
Group ID	**1:nap**	**2:nap**	**RRR_nap**
Group1	0.741	1.063	1.380
Group2	0.705	1.135	1.538
Group3	2.135	2.368	1.263
Group4	0.685	0.958	1.314
Group6	0.699	0.827	1.137

Compared with continuous operating, drivers show larger preference toward taking break during 14:00pm to 15:00pm across both dates and groups. Specifically, compared with continuous operation, average value of nap estimates for taking a break is 43.4%, 61.1%, 10.9%, 39.9% and 18.4% higher for each driver group respectively. Moreover, during specified nap hour, drivers' probability for taking a break is 1.380, 1.538, 1.263, 1.314 and 1.137 times of continuous operation for each group respectively. The uniform observation indicates a common

life style's effects on drivers' operation process and supports explanation for demand and supply gap during this period.

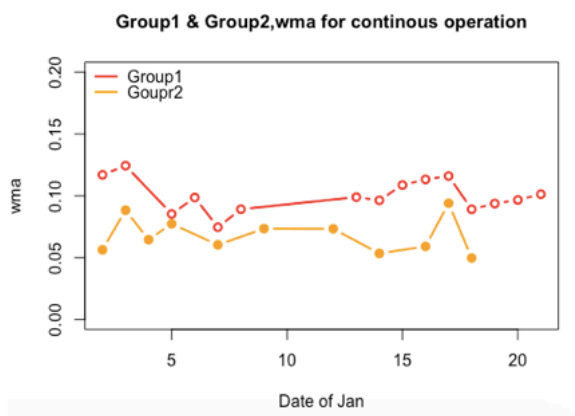
Monetary estimates analysis and comparison across groups and dates

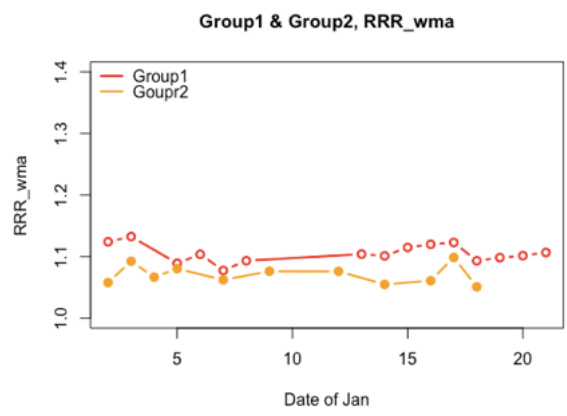
Among individual-specific attributes that describe individual driver's operation record, three attributes are directly related to monetary reward, wma, previous_o and acc_r. Unlike time sensitive attributes that mainly reflect driver's schedule flexibility, monetary reward related attributes correlate closer with drivers' revenue objective. Besides, while acc_r represents drivers' accumulative monetary reward, order_vol records for drivers' accumulated tiredness during operation. Comprehensive analysis for these four attributes can shed light on understanding of drivers' heterogeneous revenue objectives and tradeoffs between reward and level of fatigue.

Table 10 Average of monetary attributes' estimates and RRR across groups[7]

	Average of Estimates for continous operaiton							
Group ID	**wma**	**RRR_wma**	**previous_o**	**RRR_previouso_o**	**acc_r**	**RRR_acc_r**	**order_vol**	**RRR_order_vol**
Group1	0.100	1.105	0.0401	1.041	-0.016	0.984	-0.087	0.916
Group2	0.068	1.071	0.0338	1.034	-0.011	0.989	-0.177	0.838
Group3	0.158	1.171	0.0443	1.045	-0.002	0.998	-0.046	0.955
Group4	0.154	1.167	0.0513	1.053	-0.013	0.987	-0.047	0.954
Group6	0.100	1.105	0.0397	1.040	-0.005	0.995	-0.157	0.855

- **Analysis of *wma* and previous_o**





[7] RRR_μ represents relative risk ratio for attribute μ.

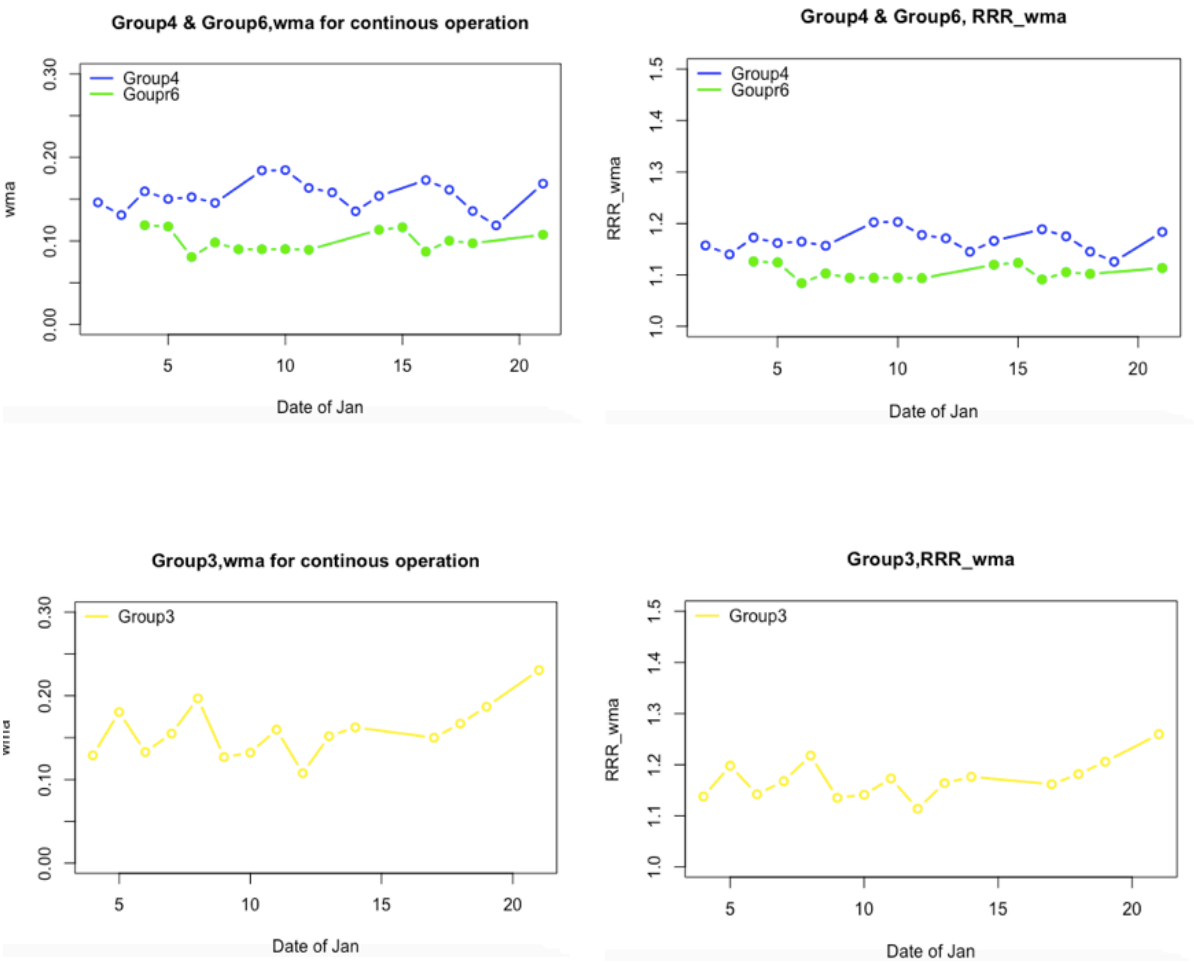


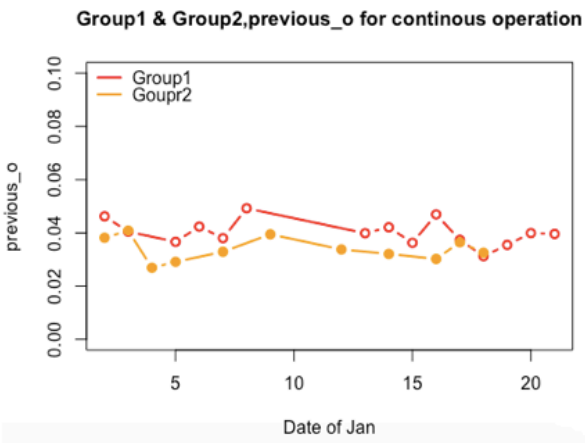
Figure 7 Coefficients and RRR of wma across dates and groups for continuous operation

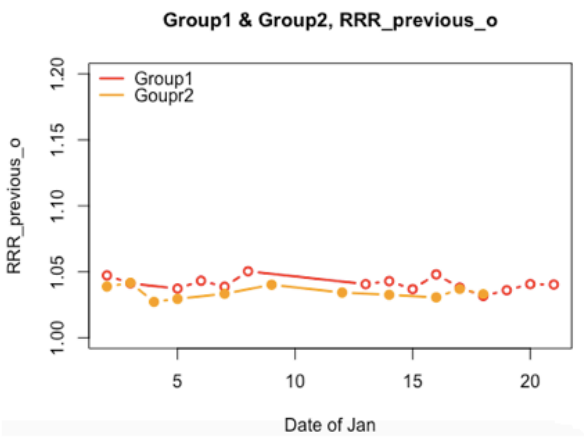
The two individual specific attributes, *wma* and *previous_o*, both relate to operation efficiency in respect of monetary reward per order. It suits with intuition that both *wma* and *previous_o* has positive coefficients for continuous operation across all samples, with statistical significance as well. Driver groups with similar order taken volume per driver but much different revenue show pronounced different valuations for *wma and previous_o*. In general, driver group with higher revenue corresponds to lower preference toward both wma and previous_o, as shown in Figure 7 and Figure 8. Specifically, compared with Group1 and Group4, Group2 and Group6 have consistent lower estimates for wma respectively across all samples. According to Table 10, the average value of Group1's wma estimates is 46.9% higher than average value of Group2's wma estimates, and the average value of Group4's wma

estimates is 54.6% higher than Group6's average value of wma estimates. Similar pattern can be observed for previous_o as well. On average, according to Table 10 with one Yuan increase in wma, drivers in Group1 have 10.5% increase in probability for continuous operation while Group2 have 7.1% increase, as compared to exiting the system. Similarly, drivers in Group4 have probability for continuous operation that 1.167 times of probability for exiting the system, higher than the relative risk ratio of 1.105 for Group6.

Besides, for groups with close revenue value but different average revenue per order, such as Group2 and Group4, group with lower average reward per order corresponds to higher valuation toward wma for continuous operation. Specifically, with one Yuan increase in wma, drivers in Group2 have 7.1% increase in probability for continuous operation compared to exiting the system, while drivers in Group4 have 16.7% increase. The observation again supports the inference that drivers devoting less comparable work efforts usually focus more on operation efficiency instead of accumulative performance.

Additionally, groups with lower volume of orders taken, Group1 and Group2, have more obvious timing pattern regarding weekends. There exists increasing trend for wma coefficient during weekends.





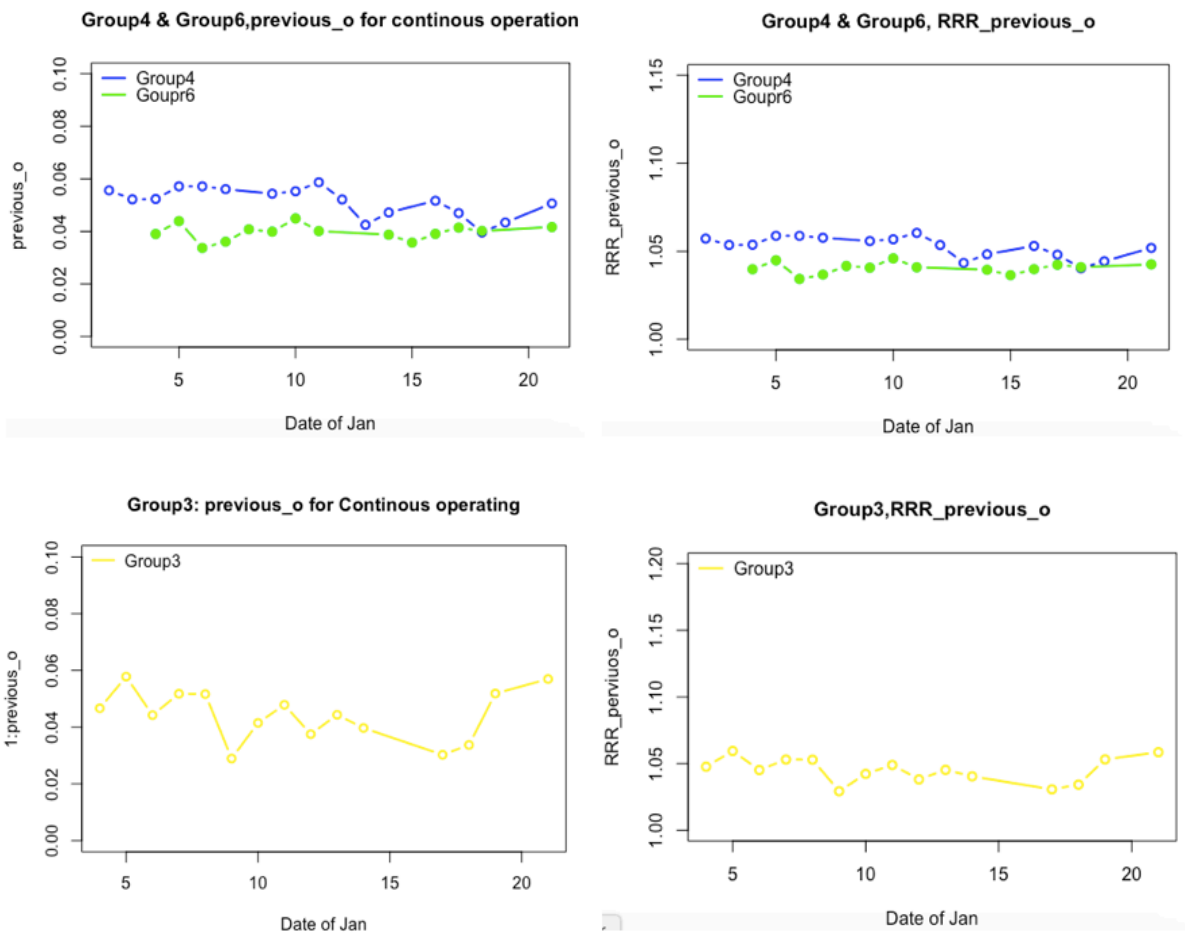


Figure 8 Coefficients and RRR of previous_o across dates

On the other hand, when compared coefficients for *wma* and *previous_o*, although both positive and statistic significant, their value domain varies from each other. Across all comparable samples, *wma* has 177.3% larger estimates than *previous_o* on average, which sheds light on describing drivers' common rationality during evaluation process. On average, one Yuan increase in wma results in 12.4% increase in probability of continuous operation compared to 4.3% increase resulted form one Yuan increase in previous_o. Drivers tend to put a time decreasing weight on his or her order reward record and consider the weighted average a more important factor than previous reward for operation decision.

- **Analysis of order_vol and acc_r**

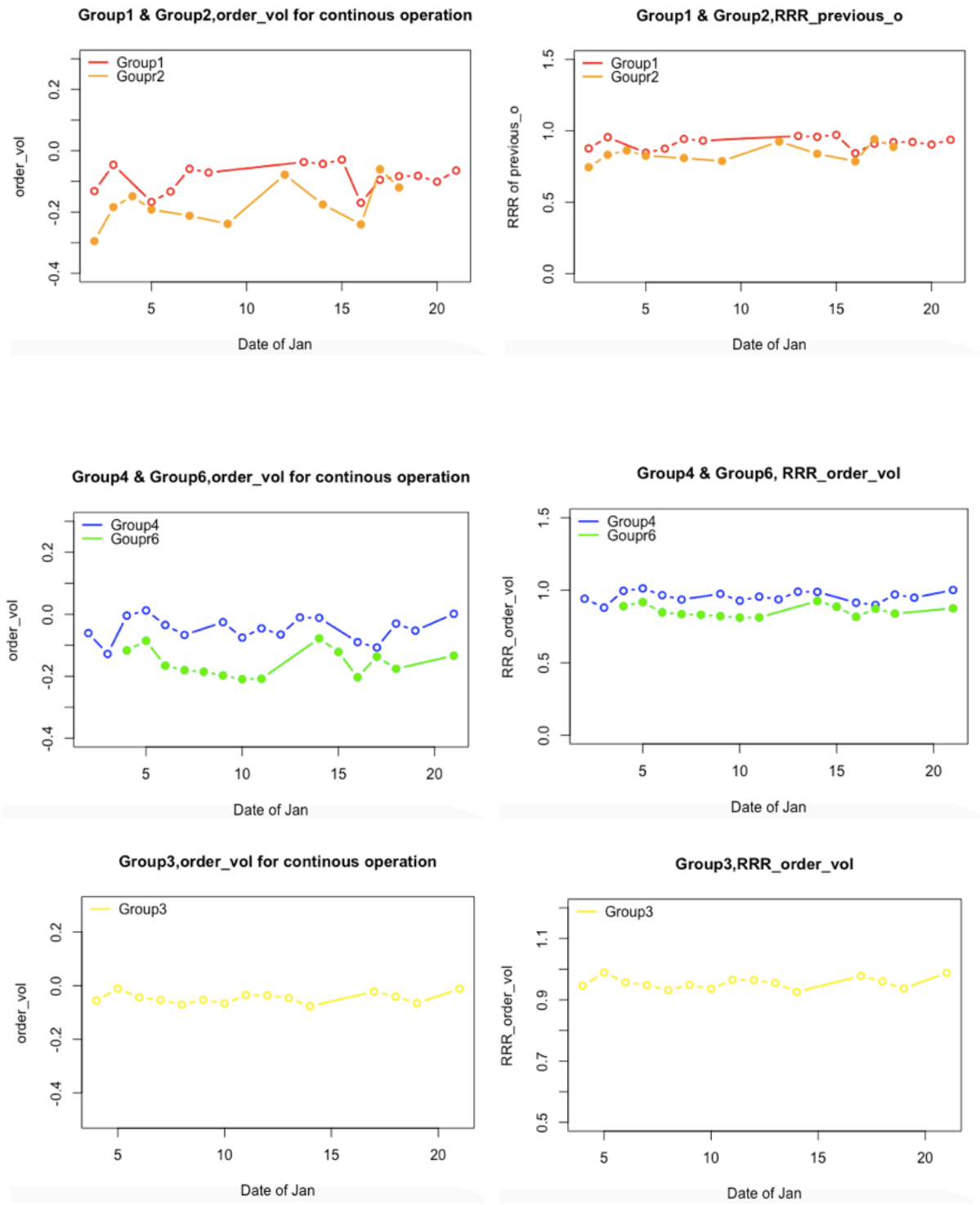


Figure 9 Coefficients and RRR of order_vol across dates and groups for continuous operation

As an attribute indicating drivers' accumulative tiredness during operation, *order_vol* has negative coefficient estimates across all samples as shown in Figure 9. With increasing order taken volume, drivers are less likely to continuously operate. Besides, compared with *acc_t,*

which describes accumulative working time length directly, *order_vol* has better performance as an attribute. *Order_vol* has estimates with statistical significance across samples while *acc_t* does not. Unlike *acc_t* simply representing working time length, *order_vol* also contains information for tiredness accumulated by passenger pickup and drop off, which often involves with additional searching and parking process.

Meanwhile, the two comparison scenario composed with Group 1&2 and Group 4&6 respectively also show similar difference regarding coefficient magnitude within groups. With similar order taken volume, *order_vol* has smaller absolute value for group with lower average revenue per order. Specifically, the average of order_vol estimates' absolute value of Group1 is 50.5% lower than Group2's, and the average of order_vol estimates' absolute value of Group4 is 70.1% lower than Group6's.

It is in line with intuition because higher reward order often relates to longer travel time and distance, both positively correlated with tiredness.

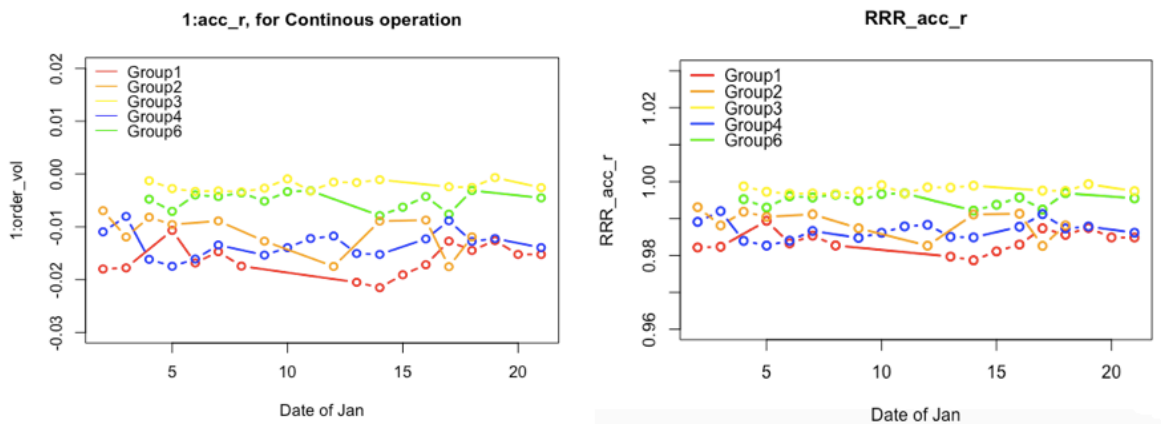


Figure 10 Coefficients and RRR of acc_r for continuous operation across dates and groups

The negative and statistic significant estimates for *acc_r* are in correspondence with intuition that drivers are less likely to continuous operating as approaching revenue goal. Besides, the order of value domain for groups are almost the same order of groups' revenue value. Unit increase in accumulative monetary reward associates with different valuation among groups, with the largest influence on Group1, which has the lowest total revenue among 5 groups.

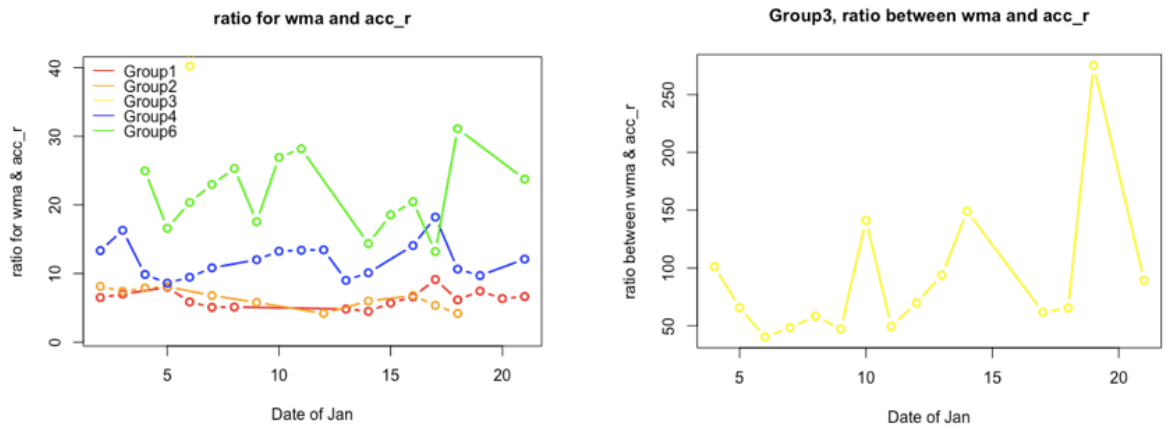


Figure 11 Ratio for wma and acc_r across dates and groups

Considering different scale of monetary attributes acc_r and wma, absolute ratio between their coefficients are illustrated in Figure 11. The ratio describes drivers' tradeoff between operation efficiency and accumulated reward, which inspires a more comprehensive understanding for drivers' revenue goal. Group's average revenue per order represents average reward accumulation speed within this group. From Figure 11 it can be observed that the higher the group's revenue, the higher the ratio between wma and acc_r. Meanwhile, the ratio generally has value higher than corresponding group's average revenue per order. Therefore, it can be inferred that drivers always put higher valuation on recent operation efficiency, wma, than accumulated reward. With higher speed of reward accumulation, drivers are more likely to continuously operating and earning more, instead of exiting the system with approaching some fixed reward goal. Specifically, Group 3 has significant larger comparison ratio between wma

and acc_r than other groups and its own average revenue per order. This suggests that drivers in Group3 have the largest flexibility in their monetary reward goal. After taking orders with high weighted average reward, they are more likely to continuously operating and adjust to a higher reward goal accordingly.

Willingness to work analysis

Willingness to pay is a commonly adopted measure in customers' choice behavior modeling, which describes the amount of how many a customer is ready to pay in exchange for one unit change in corresponding attribute (9). Similarly, based on attributes set proposed as in Table 4, willingness to work can be defined as a driver's propensity to accept an hour of work in exchange for a certain amount of monetary reward. The estimation for a driver's willingness to work is illustrated as following.

$$Willingess\ to\ work = u(X(s_0 + work_0) - u(X(s_0 + work_1)$$
$$= \beta_{rpt} * \beta_{e\_w} * 60$$

where,

u stands for utility function of operation status X, which composed by $s_0\ and\ work_0$; β_{rpt} and β_{e_w} represent estimates for *rpt* (in CNY per minute) and *e_w* (in hour) from MNL respectively. The constant multiplier 60 transfers minute to hour.

Consistent with reasoning in Section 2.5.2, different driver groups have different values for willingness to work, from which specific patterns can be identified.

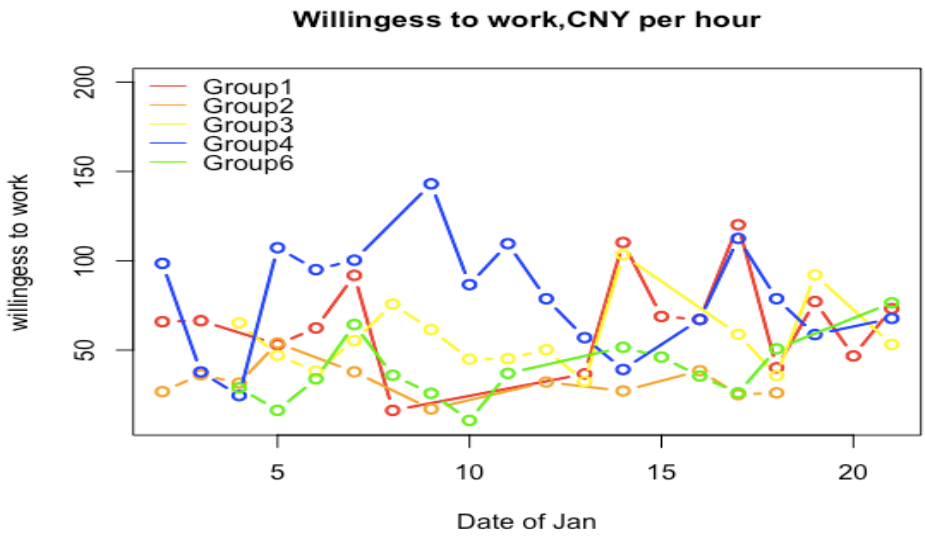


Figure 12 Willingness to work estimation across dates and groups

Figure 12 illustrates willingness to work across dates and groups. Unlike estimates discussed in Section 2.5.2, such as wma, previous_o, acc_r and order_vol, each driver group has willingness to work estimates that fluctuate significantly along different dates. Besides, differences between groups can not be observed directly from Figure 12. The instable estimates reflect MNL's limitation in capturing heterogeneity by adopting generic estimates for *rpt* and *e_w*.

In order to offset the instability from limited heterogeneity consideration, each driver group takes average value for willingness to work across dates as a generic estimate. Table 11 below summarizes corresponding estimate for each driver group.

Table 11 Willingness to work estimates for each driver group, in CNY

Average values	Group1	Group2	Group3	Group4	Group6
rpt	0.66	0.33	0.25	0.46	0.21
e_w (abosolute value)	1.65	1.63	3.96	2.96	2.95
willingess to work	66.43	32.04	57.27	80.15	38.51

- **Analysis for willingness to work from comparison between groups**

According to Table 11, Group1 and Group2 have similar average order taken volume, 7.42 and 7.93 respectively, but different average revenue values. Group2 has average total revenue of 170.89 CNY, which is 86.8% higher than Group1's average total revenue of 91.49 CNY. According to Table 11, Group1 has estimated willingness to work at 66.43 CNY, 107% higher as compared to Group2's estimate of 32.04 CNY. Although Group1 and Group2 have similar order taken volume, Group2 receives higher total revenue mainly from taking orders with longer travel distance and time. Difference between Group1 and Group2's estimated willingness to work indicate that drivers operating longer in the system are willing to operate at lower monetary reward per hour.

The conclusion can be further supported by comparison between Group4 and Group6, as well as comparison between Group3 and Group4.

Group4 and Group6 both have order taken volume of approximate 15, but Group6 has average total revenue of 340 while Group4 has 217.65. Similarly, Group6's estimate for willingness to work is much lower than Group4, with value of 38.51 CNY compared to Group4's 80.15 CNY.

Meanwhile, compared with Group4, Group3 has similar mean value for average revenue per order at 16.71 CNY. However, Group3 has average order taken volume of 25.7 while Group4 has 15.4. From the comparison, it can be inferred that drivers in Group3 operate much longer than drivers in Group4. When comparing Group3 and Group4's estimated willingness to work, Group4 shows 40.0% higher value than Group3.

These three comparison scenario can support inference that drivers operating longer in the system are willing to operate at lower monetary reward per hour. The conclusion reflects

drivers' heterogeneous expectation regarding monetary reward in unit operation time. Besides, the insights can support incentive scheme design for boosting supply with limited subsidiary resource.

2.6 Mixed Logit Model (MIXL)

2.6.1 Development for MIXL

While MNL allows for straightforward estimation and easy interpretation, it is lack of flexibility in capturing heterogeneous preferences across individuals. The commonly adopted extension for MNL, Mixed Logit Model (MIXL), improves model flexibility by allowing taste preferences to vary randomly across individuals based on prior assumed distribution (9). Instead of generating fixed parameters, MIXL estimates for random parameters that accommodate unobserved heterogeneity across individuals. Unobserved heterogeneity can be assumed to vary in population following a proposed density function.

In MIXL, the random utility for individual n to choose alternative j is formulized as following.

$$U_{jn} = V_{jn} + \varepsilon_{jn}, \varepsilon_{jn} \sim EV1(0, \lambda)$$

$$= X_{jn}\beta_n + \beta_j + \varepsilon_{jn}$$

where,

β_n represents individual taste heterogeneity within population and follows continuous density $f(\beta_i|\theta)$ that with parameters θ.

Estimation for MIXL can be achieved by Maximum simulated likelihood (9).

From MNL established in Section 4, abundant insights regarding drivers' operation behaviors have been obtained. However, the underlying assumption that individuals perceive attributes relative to expected working time or reward, such as rpt and e_w, uniformly limits associated understanding from being more realistic. Therefore, MIXL is implemented with the same attributes set in Table 2 but allowing randomization for rpt and e_w.

2.6.2 Numerical example analysis

One sample cluster is chosen randomly to be used for MIXL numerical case study. Following Table 12 illustrates descriptive statistics of the sample cluster, which contains 1,287 drivers and 19,379 associated observations.

Table 12 Descriptive statistics for sample cluster on 01-02-2016

Number of drivers	Average of total revenue (2)	Standard devition of (2)	Average of order taken volume (3)	Standaard deviation of (3)	Average of average revenue per order (4)	Standard deviation of (4)
1287	223.4	44.3	15.1	3.1	15.1	2.9

Specifically, the alternative specific attributes, rpt and e_w, vary across behavioral respondents according to:

$$\beta_{rpt,i} = \beta_1 + \pi_{11} order\_vol + \sigma_1 \eta_{1i}$$

$$\beta_{e\_w,i} = \beta_2 + \pi_{21} orde_{vol} + \pi_{22} acc_r + \sigma_2 \eta_{2i}$$

where,

η_{1i} and η_{2i} all belong to N (0,1).

Table 13 below demonstrates corresponding MIXL estimates sample cluster. Estimates from MNL is included in Table 13 for comparison purpose. Highlighted cells in both tables represent estimates with statistical significance.

Table 13 MIXL results for sample driver cluster on 01-02-2016

Attributes	**Estimates**	**Individual specific attributes with atlernative specific**	**Estimates**
Alterantive specific constants		1:order_p	-0.025
1:(intercept)	2.259	2:order_p	-0.067
2:(intercept)	-2.135	1:ratio	-0.607
Alternative specific attributes with generic coefficients		2:ratio	-0.682
exp	-0.331	1:acc_t	0.030
Alternative specific attributes with randomized coefficient		2:acc_t	0.080
e_w	-3.773	1:acc_d	-0.187
rpt	1.626	2:acc_d	-0.157
rpt.order_vol	-0.070	1:previous_o	0.056
e_w.order_vol	-0.007	2:previous_o	0.049
e_w.acc_r	0.002	1:diff_t	1.209
sd.e_w	0.908	2:diff_t	0.669
sd.rpt	0.988	1:rush1	0.076
		2:rush1	0.852
		1:rush2	-0.337
		2:rush2	0.015
		1:vol_p	1.533
		2:vol_p	2.112
		1:nap	0.568
		2:nap	0.890
		1:wma	0.159
		2:wma	0.104
		1:acc_r	-0.014
		2:acc_r	-0.011

Table 14 MNL results for sample driver cluster on 01-02-2016

Attributes	Estimates	Individual specific attributes with alternative specific coefficients	Estimates
Alternative specific constants		1:wma	0.142
1:(intercept)	2.857	2:wma	0.065
2:(intercept)	-1.528	1:previous_o	0.057
Alternative specific attributes with generic coefficient		2:previous_o	0.044
rest	0.019	1:acc_d	-0.132
rpt	0.497	2:acc_d	-0.035
Alternative specific attributes with alternative specific coefficients		1:acc_t	-0.006
3:e_w	-3.697	2:acc_t	0.008
1:e_w	-3.417	1:acc_r	-0.011
2:e_w	-3.268	2:acc_r	-0.005
3:exp	-0.297	1:order_vol	-0.063
1:exp	-0.314	2:order_vol	-0.158
2:exp	-0.294	1:order_p	-0.003
		2:order_p	-0.017
		1:diff_t	0.940
		2:diff_t	0.371
		1:ratio	-0.996
		2:ratio	-1.160
		1:rush1	0.090
		2:rush1	0.843
		1:rush2	-0.332
		2:rush2	0.046
		1:vol_p	1.419
		2:vol_p	1.991
		1:nap	0.494
		2:nap	0.804

According to Table 13 and Table 14, MIXL obtains consistent estimates with MNL for fixed parameters with expected signs and similar magnitudes. The two randomized coefficients, rpt and e_w, both generate mean values and standard deviations with statistical significance. The results imply that there exists significant variation in how e_w and rpt are perceived across behavioral respondents. Detailed interpretations for rpt and e_w are illustrated as below.

rpt: In MIXL defined above, rpt is assumed to have a normal distribution. The mean value estimate is 1.626 with standard deviation of 0.988. The negative and significant rpt. order_vol,

-0.070, indicates that drivers taking more orders already are less sensitive toward expected reward in unit time, which is in accordance with expectation. As tiredness accumulated, continuous operation incentive from reward will decrease.

e_w: The mean value estimate for e_w from MIXL, -3.733, has expected sign and similar magnitude with MNL. Both mean value and standard deviation are with statistical significance as well. However, neither e_w. order_vol nor e_w.acc_r is statistic significant, which might due to observed heterogeneity being explained by order_vol and acc_r.

Chapter 3: Ride-Sharing System Simulation

Multinomial Logit Model (MNL) for drivers' operation behavior description not only sheds light on understanding drivers' decision process comprehensively, but also delivers insights for ride-sharing platform's operation efficiency improvement. The challenge brought by demand and supply gap is always critical for transportation system, including ride-sharing platform. Based on operation record, orders that not been matched successfully due to lack of supply account for approximate 12% - 15% of total order requests. For a ride-sharing platform, where drivers have full decision right of whether entering or exiting system, understanding for fleet's dynamic characteristics from MNL can help optimizing both supply incentive scheme and real-time order-driver matching strategy.

For example, drivers belonging to different groups and being under different operation situation has heterogeneous valuation for monetary reward increment. If targeting for supply boosting, the limited subsidy resources can be allocated more efficiently by being assigned to drivers in operation situation where monetary reward increment resulting in higher marginal utility. Besides, if orders having destination at area with significant demand and supply gap

can be paired with drivers who have high possibility for continuous operation, valid supply can be better centered to area in need by fulfilling current request and help resolving future demand and supply gap issue.

In order to investigate feasibility and potential effects of above optimization suggestions, simulation framework for ride-sharing operation is proposed and implemented in this research. In each matching window, available drivers and queuing orders are assigned weights based on associated operation strategy. Maximum weighted matching algorithm for complete bipartite graph (10) is then incorporated to realize one to one matching between driver and order. The simulation framework allows for illustrating potential effects of different operation strategies, from both systematic performance and individual driver's reaction.

3.1 Implementation of Ride-Sharing System Simulation

The objective for the proposed simulation is to illustrate how ride-sharing systems respond to different operation strategies. Performance evaluation criteria involve the strategy's effects on all major participants in the system, including the drivers, platform operator, and passengers. Furthermore, in order to better reflect a real operation situation, the simulation framework uses observations from real operation records to propose assumptions whenever possible. The following subsections illustrate critical assumptions and describe in detail the design of the simulation.

3.1.1 Simulation design for supply from dynamic driver fleet

Simulating drivers' interactions with the system is critical due to their dynamic characteristics. On a ride-sharing platform, drivers participate with full control over their operation schedules. They decide when and where to enter the system, as well as when and where to exit. They also

decide whether to stay in the same area or move to another when idling. These behaviors all have profound effects on the system's supply distribution regarding both time and area.

For drivers' initialization, real operation records are used to determine the entering area and time. On the sample date for simulation, the driver fleet will start at the same time and area as the original record, thus reflecting drivers' initial choice behaviors for that operation.

Rebalance of supply is a popular research problem in transportation (2) and can be even more challenging when considering a dynamic fleet. Efficient rebalancing, from an operator's point of view, often aims to reduce available drivers' idling periods and meet more demand requests from passengers. There also exists rebalancing behavior in the ride-sharing system based on drivers' own decisions instead of following operators' guidance. Current researches conducted on ride sharing often often assume that drivers will stay in the same area after arriving at a previous passenger's destination (2). However, from data observation, this assumption is unrealistic and will affect system performance. Specifically, in approximately 44.3% of total dropping-off cases, drivers choose to rebalance themselves to another area and then continue operating. The reasons behind drivers' rebalancing behavior relates primarily to their judgment of future demand distribution, as well as personal preference. Therefore, following statistical probabilities from operation records, drivers in the proposed simulation framework will rebalance themselves after arriving at one area.

From statistics regarding drivers' rebalancing probabilities, two main conclusions can be drawn. First, drivers arriving at a lower-demand area have a higher probability of rebalancing, and vice versa. Second, drivers arriving at different areas will target different areas for rebalancing. These two conclusions imply that both drivers' judgments about future demand

and areas' relative geographic positions influence their rebalancing activity. The statistical probability will be applied to simulate where the drivers, if they decide to rebalance, will rebalance themselves after dropping off a passenger. In the simulation, drivers will choose to rebalance if they keep idling for multiple time steps in the same area. The threshold for idling tolerance is set to be a random number drawn from uniform distribution between 7 and 20 in minutes, in order to reflect drivers' heterogeneity. Drivers' idling period length is simulated by their consecutive unsuccessful matching times in the same area. In addition, due to the limitation of incomplete records, there are 290 arrival areas, which are different from the 66 coded start areas. Drivers dropping off at these areas are assumed to immediately rebalance themselves to one of the 66 start areas following probability distribution, if they do not leave the system.

MNL defines drivers' operation behavior as either operating continuously, taking a break, or exiting. However, in the simulation, if the assumption is that drivers will leave the system whenever the probability for taking a break or exiting is the highest among the three alternatives, then this assumes that MNL has perfect prediction power. The assumption can be inappropriate for simulation, as it ignores drivers' possible changes in behavior under a different matching scheme. Also, it can be conflicting to use prediction as ground truth for simulation. Therefore, a random draw from the choice set following estimated alternative probabilities is used as the simulated driver decision.

3.1.2 Simulation design for demand

Real-time order requests are used as demands in the simulation. The waiting time tolerance for each order is assumed to be a random number drawn from uniform distribution between 3 and

7 in minutes. If the order is not successfully matched to a driver after the tolerance period, this order is assumed to be dropped by the passenger.

3.2 Ride-sharing platform strategy

The purpose of the ride-sharing platform is to queue passenger pickup orders and match then with available drivers in unique pairs. While the classic transportation queuing model applies the first-come, first-served rule and then matches demand and supply indifferently, a ride-sharing system might lose system-wide efficiency by following this model. As ride-sharing system efficiency often suffers from inconsistent distributions of supply and queuing demand, assigning appropriate weight to each order and driver pair offers a promising solution to the problem of distributing available supply to areas in need.

However, on a ride-sharing platform, the distribution of available supply along time and area closely relates to the effects of processed orders and drivers' own choices, both of which lead to available supply distribution's stochastic feature. Therefore, there exist significant challenges in formulating an accurate mathematic description for available supply's dynamic distribution during the operation period. When one area's forecasted supply need can be described as the difference between future demand and available supply, the amount of available supply and associated drivers' individual operation statuses can be difficult to estimate, even compared with a promising demand volume prediction. On the other hand, if using predicted demand volume to evaluate whether a specific area is need of supply, it's highly possible that excess demand will accumulate in any area with high predicted demand volume. Unlike the difference between amount of available supply and real-time demand which involves dynamic information from both supply and demand, predicted demand volume only uses partial information of the entire operation situation.

Considering the factors that have just been described, an area's future supply need is proposed to be evaluated based on estimated future available supply distribution and the predicted demand volume. Further, the estimation for available supply at a given time within the particular area of interest is composed of available supply at the current time step, an exact measurement which can be obtained from a current operation record. One area's available supply at a given time consists of three parts: the newly entering drivers, existing drivers who have dropped off passengers, and idling drivers. The number of newly entering drivers can be efficiently predicted from historical records, and the time step when drivers are supposed to drop off passengers can also be estimated via travel distance and traffic conditions. However, one major challenge arises from the status of idling drivers, whose availability is influenced by the dynamic effects of orders that will be processed between the current time step and the future time step of interest.

Two estimation methods are proposed for calculating future available supply, resulting in two slightly different order and driver pair-weighting strategies. The first future available supply estimation gathers information on current differences between supply and demand within an area at the time of interest as well as total availability of future supply obtained from the current time step. The second method divides the availability of future supply that can be obtained from the current time step into three parts as described above: newly entering drivers, drivers who are supposed to drop off passengers, and idling drivers. The interaction between idling drivers and orders to be processed at the current time step and the future time of interest will be further estimated according to historical orders' destination distribution. For orders that originate and terminate within the same area, the driver will be available at the same area after dropping off passengers if the driver does not leave the system. In contrast, a driver taking an

order with the origin and destination in different areas will change area availability if the driver operates continuously. The historical statistics of orders' destination distribution within a specific area, whether the destination falls within the same area as the origin or not, is then applied to describe idling drivers' changes in availability. Tradeoff between supply components from different sources can be reflected using constant weighting parameters. Detailed formulas for estimating future supply availability are included in numerical example analysis presented in Section 3.3.

3.3 Detailed Simulation steps

- Each time step is measured in units of one minute. At each time step and within each of the 66 areas, order requests that are sent and drivers who are available at the time step are used to construct a complete bipartite matching graph. The two node groups in the graph represent driver and order, respectively. The edge between order and driver indicates matching feasibility. Using the complete bipartite matching graph, it is assumed that orders and drivers within the same area are all allowed to pair with each other. Also, orders still within the holding period in this area will be added to the queue of orders yet to be matched.

- For each possible order-driver pair in the pool, weight is assigned according to different weighting strategies. The probability for continuous operation is defined as the probability that drivers will continue to operate after completing a specific queuing order. Drivers' operation statuses will be updated temporarily with each feasible order in the queue to estimate this probability.

- Next, the maximum weighted algorithm will be applied to perform order-driver matching. Orders that receive no response will be held and added into the next round until they exceed

the tolerance period. Drivers who are not matched successfully will be marked as idle. When idle rounds exceed the drivers' tolerance, drivers will start to self-rebalance.

- Travel time between each origin and destination pair is estimated using operation records with corresponding origins and destination. The average travel time between each origin and destination pair, plus a randomized component, compose the simulated order completion time.

- Drivers' operation status in the form of MNL attributes set, order-taken history, and location at each time step are updated accordingly.

3.4 Ride-sharing simulation and case study

3.4.1 Data Description

The sample dataset for this simulation example is generated from a January 5th, 2016 operation record which includes all 5,689 drivers belonging to 5 driver groups analyzed in Section 5, as well as 68,290 associated orders. Of all of the order requests made on 01-05-2016, 15.35% of the requests were not responded to, mainly due to lack of available supply. Of these non-responded orders, 10,579 are randomly selected for inclusion in the simulation dataset. Further, 42.4% of non-responded orders correspond to a time step during morning rush hour, 8:00am to 9:00am. Therefore, the simulation is performed for first 800 time steps in the day, from 00:00:00 to 13:20:00, aiming to cover the morning peak period when major demand and supply gaps occur, as well as the following period which experiences low gap volume. Figure 13 presents frequency plots for both matched order volume and non-responded order volume during the first 800 time steps in the original operation record. In the real operation record, a

total of 28,477 order requests are successfully matched, while 7,641 orders do not receive a response.

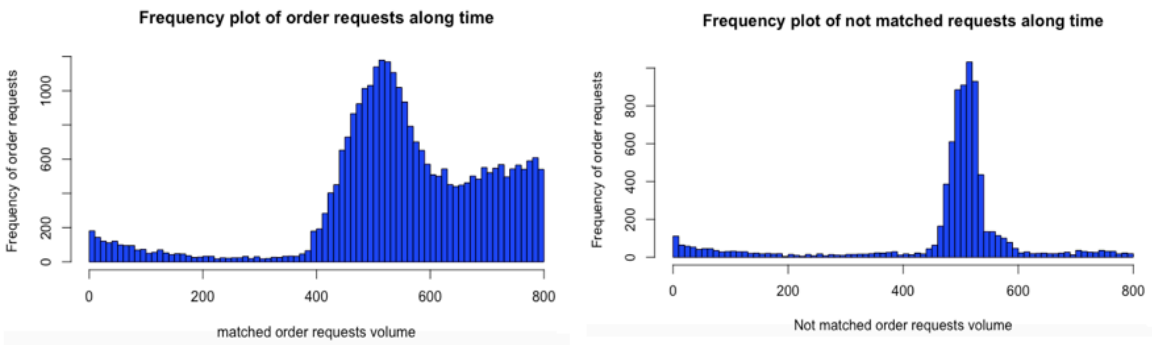


Figure 13 Frequency plot for order requests along time

As the sample driver fleet is composed of 5 driver clusters, each driver cluster's corresponding MNL estimates are applied for choice probability fitting during simulation. Notice that the simulation results are not suitable for direct comparison with original operation results, mainly due to incomplete or unknown operation information regarding detailed geographic data, exact order completion time, and original matching strategy. Furthermore, simulation results are only based on a partial driver fleet and order requests from the real system, while original operation results were generated from a real system that is significantly different regarding driver fleet size and order volume. The results comparison is consequently limited to different scenarios proposed and simulated under the same framework.

3.4.2 Numerical Case Study

In a dynamic fleet, each specific driver and order pair's potential effect on the performance of the entire ride-sharing platform is unique in various aspects, including revenue contribution, the centering of available supply, and participants' travel or working experience. An optimization strategy that assigns weight to each driver and order pair dynamically can be designed to target different objectives accordingly. The numerical example in this section illustrates the weighting strategy discussed in Section 3.2, which aims improve available supply

centering and corresponding simulation results analysis. Table 15 demonstrates the weighting formulation and key parameters' values applied in a related optimization strategy.

Table 15 Order and driver pair weighting scenarios

Weighting Scenarios	**Driver-order weighting formulation, *w(driver, order)***	**Estimated demand and supply gap at *$area_i$, $timestep_j$,and gap_{ij} , from current timestep, T***
Scenario 1	equal weight	Not applied
Scenario 2	w = driver's probability of continuous operation * estimated demand and supply gap at order's destination area	$gap_{ij} = (demand_{iT,T} - supply_{iT,T})*\beta_1 + (demand_{ij,T} - \text{supply}_{ij,T}) * \beta_2$ *where* β_1 and β_2 are constant parameters, 0.3 and 0.7 are adopted respectively
Scenario 3	w = driver's probability of continuous operation * estimated demand and supply gap at order's destination area	$gap_{ij} = (demand_{iT,T} - supply_{iT,T})*\beta_1 + (demand_{ij,T} - (\text{newly entering supply}_{ij,T} + \text{supply from order completion}_{ij,T} + \text{estimated idling drivers' supply}_{ij,T}) * \beta_2$ estimated idling drivers' supply $_{ij,T}$ = total idling supply $_{ij,T}$ * inside order ratio $_i$ * α *where* β_1 and β_2 are constant parameters, 0.3 and 0.7 are adopted respectively; α is constant parameter describing ratio of available supply from drivers completing inside order during $timestep_T$ and $timestep_j$; inside order refers to orders with both origin and destination in the same $area_i$

The proposed ride-sharing simulation framework delivers results for three weighting scenarios, described in Table 15, from which detailed operation information can be analyzed for strategy comparison and insight generation. Scenario 1, where no weighting strategy is applied for the duration of the whole simulation, is used as a control scenario to check Scenarios 2 and 3.

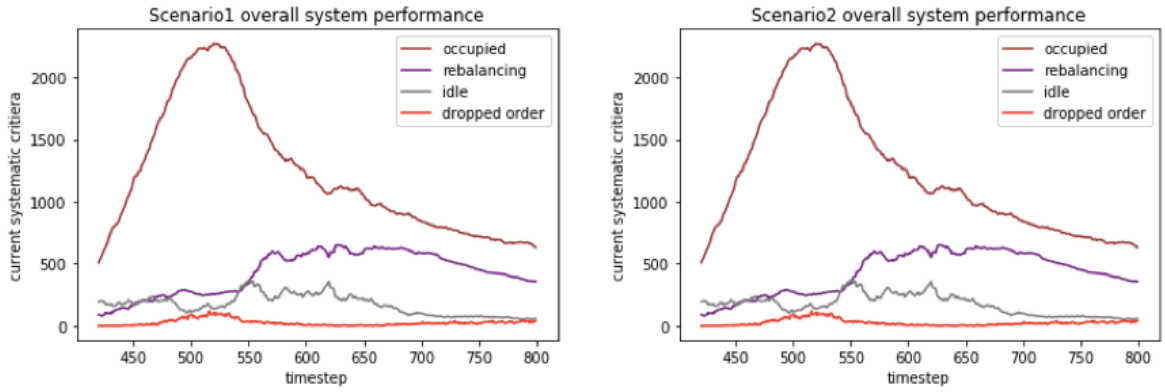
Table 16 Simulation results for three weighting strategies

Weighting Scenarios	**Scenario 1**	**Scenario 2**	**Scenario 3**
Number of participating drivers with matched orders	4,395	4,427	4,425
Simulation timestep coverage	Timestep 0 to 800	Timestep 0 to 800	Timestep 0 to 800
Start timestep for weighting strategy	NA	Timestep 420	Timestep 420
Drivers' average total revenue, in CNY	86.3	88.9	89.3
Standard deviation for drivers' average total revenue	37.8	36.3	36.2
Total volume of matched orders	25,475	26,730	26,828
Total operation time in system, in minutes	523,042	538,029	543,263
Total idling time in system	128,855	129,924	131,943
Total rebalancing time in system	176,816	135,922	121,130
Ratio of rebalancing time to total system time	21.3%	16.9%	15.2%

Ratio of idling and rebalancing time in total system time	37.0%	33.1%	31.7%

According to Table 16, both scenarios where weighting strategies are applied show improved system performance in respect to increased matched order volume, decreased idling time, and decreased rebalancing time accumulating in the system. Specifically, the idling time represents a period when drivers are available for order assignments but do not receive any and stay in the same area as their last completed order's destination. Rebalancing time corresponds to the period when drivers' idling time exceeds the tolerance interval and they choose to rebalance themselves to another area.

Figure 14 demonstrates overall system operation performance at each time step during the simulation period for all three scenarios. It can be observed that in respect to time step, excessive demand occurs from time step 500 to time step 550, while frequent idling and rebalancing occur from time step 600 to time step 650.



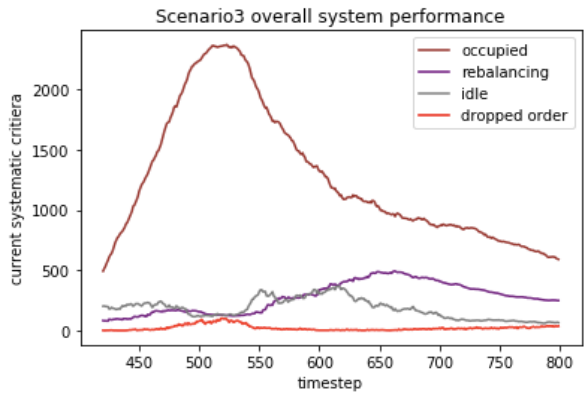


Figure 14 Overall system performance for three weighting strategies

Compared with Scenario 1, where each order and driver pair has equal weight, Scenarios 2 and 3 attempt to resolve the demand and supply gap issue by dynamically targeting valid supply to areas in need. Higher weight is assigned to driver and order pairs in which the driver has a higher continuous operation probability and the order destination has a larger estimated demand and supply gap at the approximate time of arrival. When drivers who are choosing to operate continuously are dispatched to areas in need, the occurrence of drivers voluntarily rebalancing due to long idling time in an area decreases in general. Specifically, total matched order volume increases by 4.93% in Scenario 2, when compared with Scenario 1. The idling and rebalancing time, as a percentage of total system time, decreases from 37.0% in Scenario 1 to 33.1% in Scenario 2. This improvement is primarily due to decreasing the amount of rebalancing time, which decreases from 21.3% in Scenario 1 to 16.9% in Scenario 2. As rebalancing to other areas generally correlates to higher time and monetary cost compared with idling in the same area, the significant decrease in rebalancing time improves drivers' operation experiences by reducing destination-less driving periods and related fuel consumption. Drivers who perform less "empty" driving have improved operation efficiency in that they reduce tiredness and increase net profit.

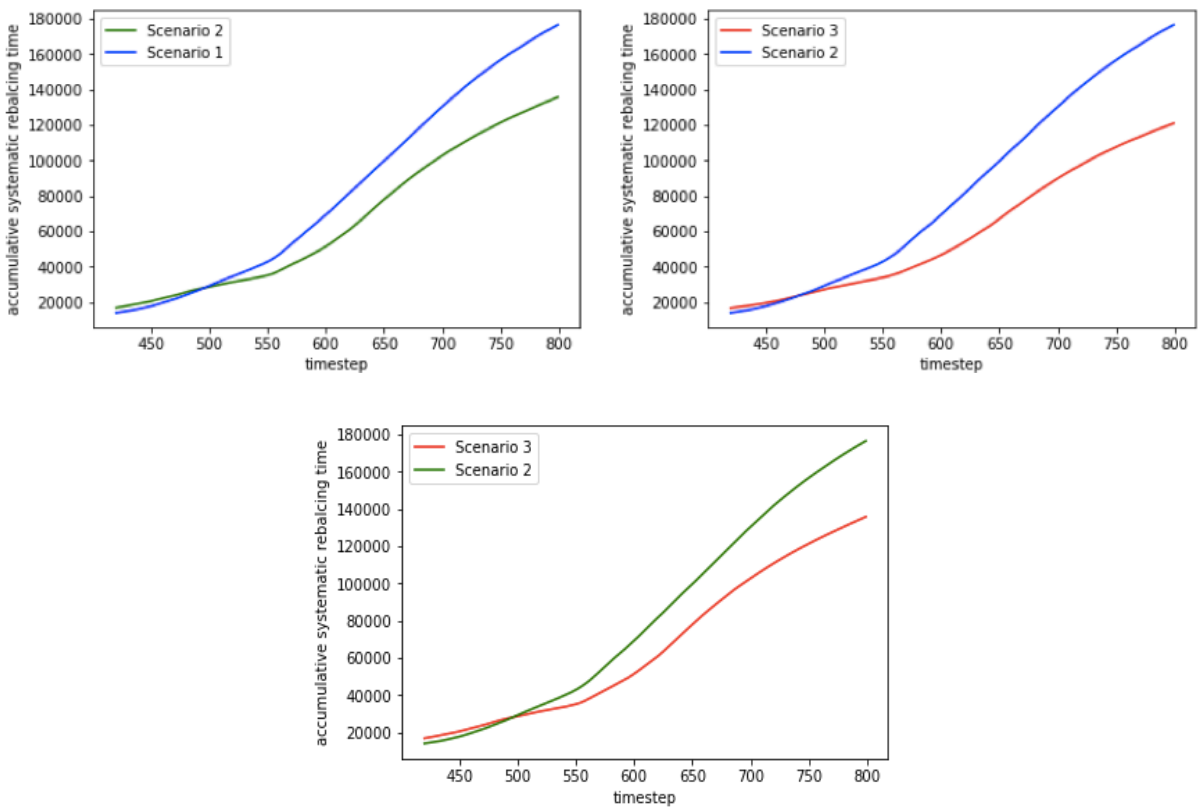


Figure 15 Accumulated rebalancing time in system along simulation time steps

Similarly, compared with Scenario 1, Scenario 3 results in better system performance: a 5.3 percentage point reduction in idling and rebalancing time, a 6.1 percentage point reduction for rebalancing time, a 5.3% higher order matched volume, and 3.5% higher average driver revenue. When compared with Scenario 2, although Scenario 3 does not show a significant increase regarding total matched orders, it further decreases the percent of rebalancing time from 16.9% (in Scenario 2) to 15.2%. This reduction in rebalancing time reflects improved supply matching efficiency. Figure 15 demonstrates that, compared with Scenario 1, Scenario 3 reduces rebalancing time in the system both stably and significantly after applying a corresponding weighting strategy. Additionally, Scenario 3 also shows obvious rebalancing time reduction after time step 540, when compared with Scenario 2.

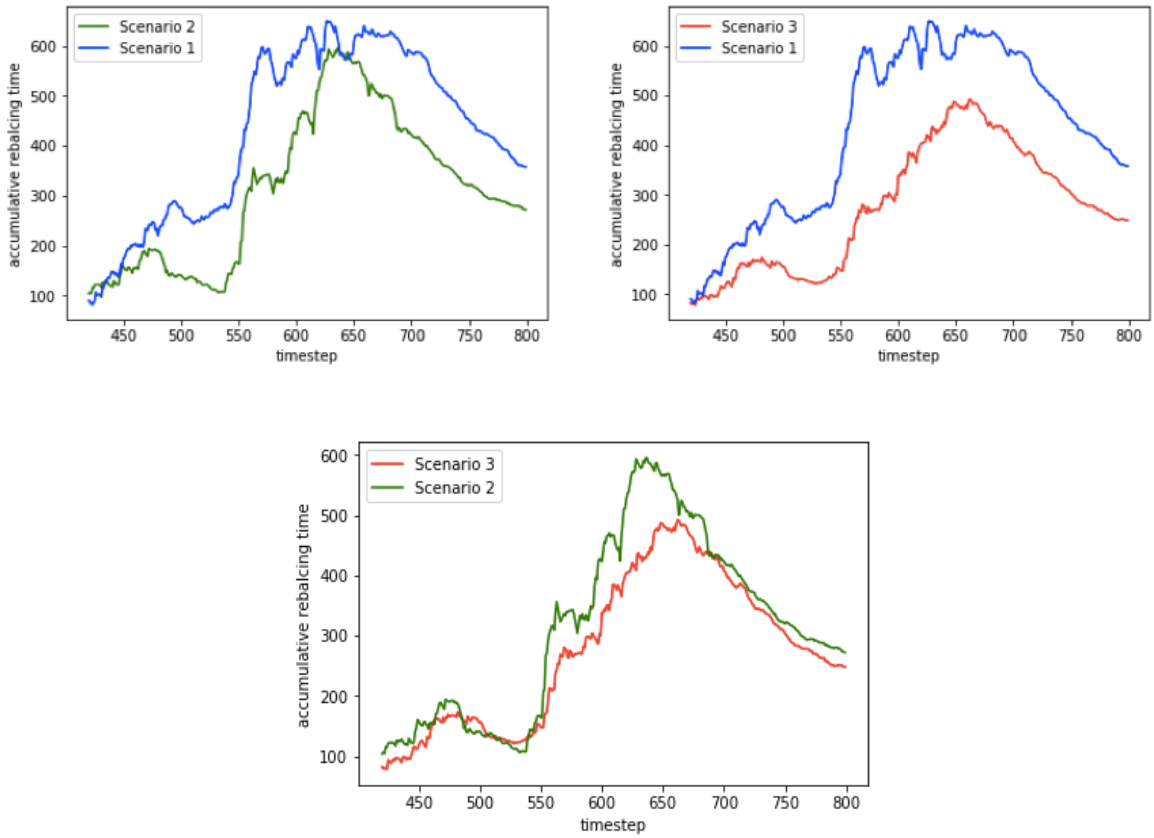


Figure 16 Rebalancing time at each time step in system during simulation period

Figure 16 illustrates systematic rebalancing time at each time step. Except for the obvious rebalancing time reduction obtained by Scenarios 2 and 3 along the entire simulation period (as compared to Scenario 1), a time pattern for the system's rebalancing time can also be observed. When no weighting strategy is applied, systematic rebalancing time consistently increases until time step 650. The increasing trend positively correlates with the total request volume over time and also reflects inefficiency in the order and driver matching process. In Scenarios 2 and 3, except for the decreasing trend that results from reduced order request volume after time step 650, there also exists one period where systematic rebalancing time decreased, between time step 480 and time step 540. The decreasing trend validates the optimization strategy's positive effect on reducing rebalancing during morning rush hour—the point in the day when the demand and supply gap is the most obvious.

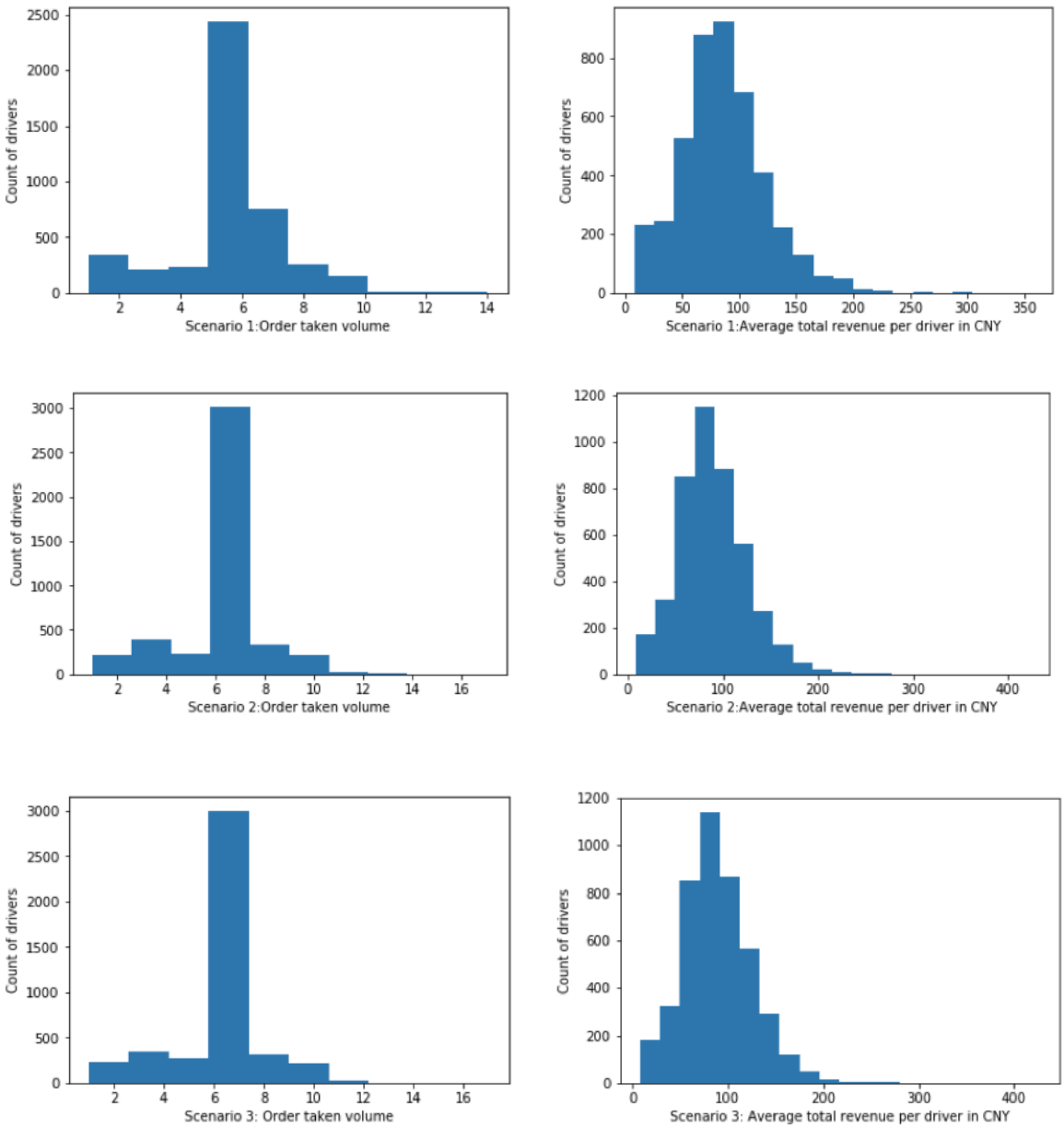


Figure 17 Frequency plots for order taken volume and average total revenue per driver

As shown in Table 16 and Figure 17, Scenarios 2 and 3 produced not only an increase in drivers' average total revenue when compared with Scenario 1, but also a lower standard deviation of drivers' average total revenue. The lower standard deviation for drivers' average accumulated revenue reflects a more even distribution of orders taken among drivers.

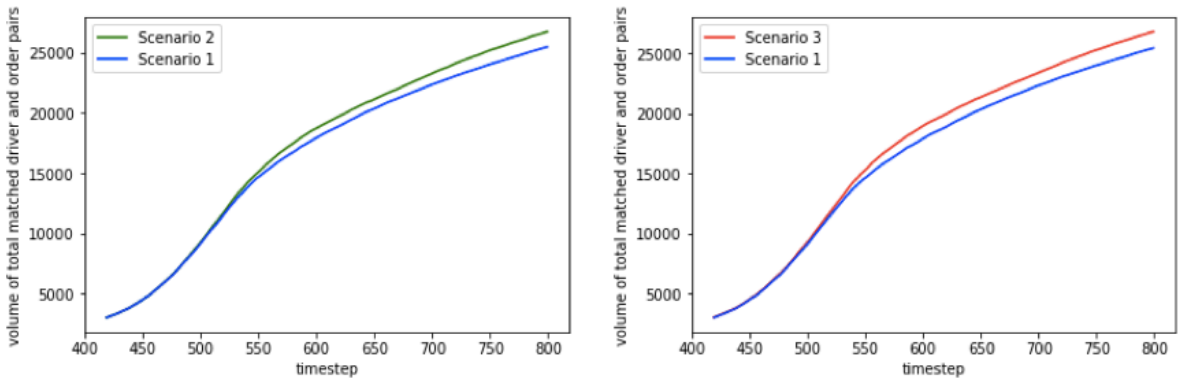


Figure 18 Volume of total matched driver-order pairs along time step

The application of a weighting strategy during matching starts from time step 420 in both Scenarios 2 and 3. According to Figure 18, total matched order volume increases incrementally in Scenarios 2 and 3 after time step 540, which corresponds to 09:00:00. It can be concluded that the weighting strategy based on the estimated gap between demand and supply performs better during the relatively flat hump than during morning peak hour. The significant demand and supply gap during morning rush hour limits the optimization effect of the weighting strategy, as most available drivers are occupied rather than idling.

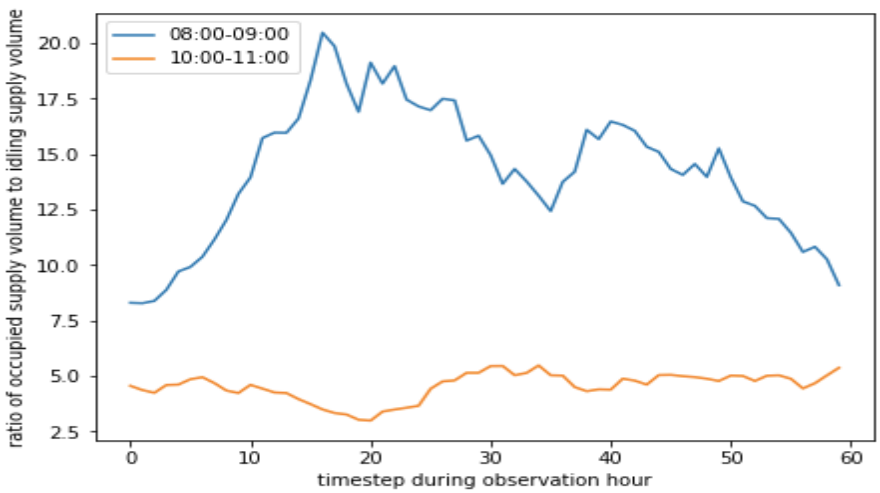


Figure 19 Ratio of occupied supply volume to idling supply volume

Figure 19 shows that the ratio of occupied supply volume to idling supply volume during 08:00:00 to 09:00:00 is substantially higher than the ratio during 10:00:00 to 11:00:00, with an

average ratio value of 14.3 and 4.6, respectively. Therefore, an optimization strategy that distributes valid supplies to areas in need can improve system performance more effectively when there actually exists excessive supply within the overall system. The strategy's optimization potential lies in assigning higher priority to dispatching available drivers to areas in need rather than allowing them to idle. However, when overall demand greatly surpasses supply, the optimization potential is accordingly limited. Although the total matched order volume from Scenarios 2 and 3 does not significantly exceed the volume from Scenario 1 during rush hour, substantial reduction in rebalancing time can already be observed, as shown in Figure 16.

Except for system performance evaluation criteria, the simulation framework also delivers detailed information for individual drivers' operation statuses and corresponding estimates for continuous operation probability. The heterogeneity of drivers' operation statuses results in different change pattern of continuous operation probability during the operation process.

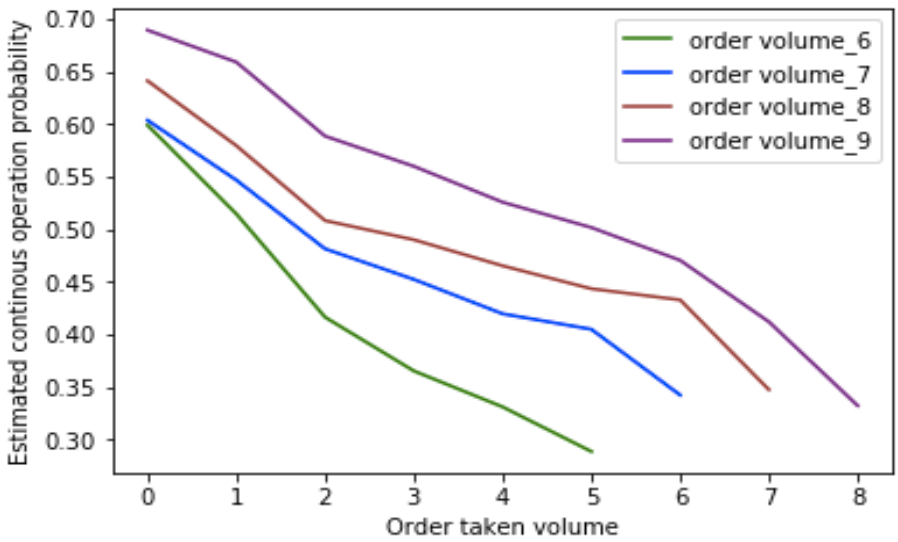
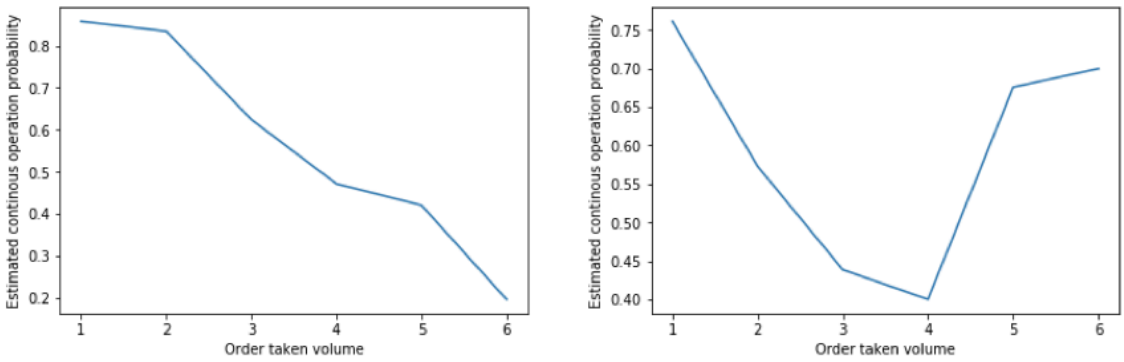


Figure 20 Estimated continuous operation probability across driver groups

By grouping drivers according to the total volume of orders they took in the simulation, Figure 20 illustrates how different groups' average estimated continuous operation probability changes as the volume of orders the drivers take accumulates. The graph shows that drivers' average probability for continuous operation generally decreases as the volume of orders taken increases, a pattern that is consistent across driver groups. Additionally, the driver group with a larger order taken volume starts from a higher value and decreases more slowly in respect to average continuous operation probability. These observations further support the interference of different driver groups' heterogeneous reactions during the operation process.

In addition to analyzing drivers in groups, individual drivers' operation statuses can also be generated from simulation results. Although drivers can be grouped by their operation results (such as order taken volume, total revenue and average revenue per order), drivers also have significant heterogeneity within the operation process on individual level, even when they have similar operation results. Figure 21 presents four typical change patterns for individual drivers' continuous operation probability estimates throughout the operation process.



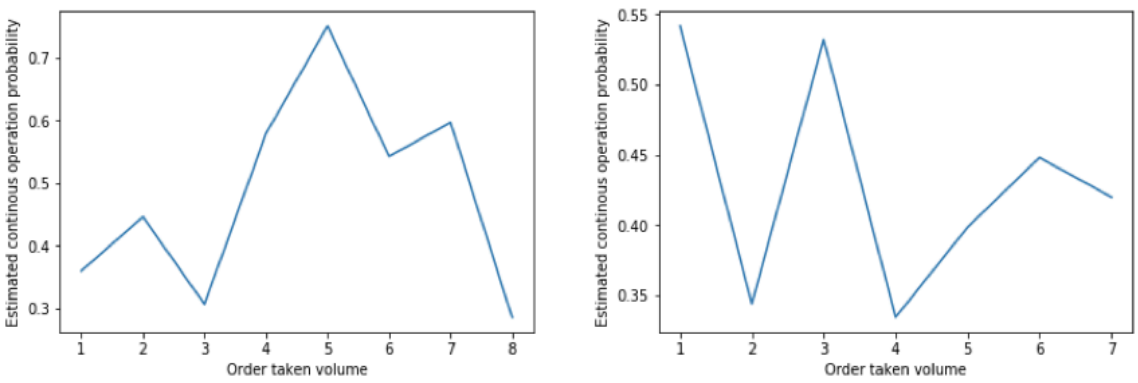


Figure 21 Four typical change patterns for continuous operation probability

The changes in drivers' patterns of continuous operation probability reflect heterogeneity across different drivers and their corresponding order sequence. The most common pattern is the decreasing trend at the top left. The other three patterns can be described as "V" shaped, flipped "V" shape, and constant fluctuation. The exploration of individual drivers' operation processes can be meaningful for analyzing the correlation between order sequence and a driver's operation status.

Chapter 4: Conclusion

In this thesis, a data driven method that studies drivers' operation behaviors is developed and implemented. Drivers' operation decisions on a ride sharing platform are summarized to three alternatives, continuous operation, taking a break and exiting the system. Then, drivers' operation behaviors can be analyzed as discrete choices corresponding to different operation statuses. The widely accepted Multinomial Logit Model (MNL) is adopted for behavioral study of drivers' operation. With operation records released by DiDiChuXing, which composed by basic information of trip requests, a comprehensive approach that describes drivers' operation statues is proposed. The description method extracts behavioral information from indirect data resource and builds up the choice attributes set for MNL.

As MNL suffers from lack of flexibility for individual-specific preference, driver clustering that aims at capturing heterogeneity across driver groups and Mixed Logit Model (MIXL) which allows for random parameters for individuals are both performed. Driver clustering based on K-means clustering algorithm prepares samples of operation records separately for each cluster on sampled date. Results analyses based on independent sample and comparison scenarios across samples are conducted, which deliver insights from multiple angles and further validate proposed methodology. Across 72 samples, the implementation of MNL delivers promising estimates, which suit with intuition and have statistical significance for the most of attributes. Time patterns within drivers' behavior and differences between driver groups are analyzed in detail as well. MIXL which designed to allow for randomization in specific attributes generates promising results that generally in line with MNL. Besides, the randomized coefficients give more information about behavioral respondents' unobserved heterogeneity as well as correlations within different attributes.

The behavioral study for drivers' operation not only enriches understanding of suppliers' behaviors in business activities, but also support operation strategy design for ride sharing systems. On a ride sharing platform, where supply distributes stochastically based on drivers' dynamic decisions, an adequate understanding and prejudgment of drivers' behaviors are essential for the design of optimal schemes. In order to investigate different strategies' effects in a comprehensive way, a simulation framework of ride sharing system is developed.

With the behavioral study method proposed before, drivers' operation statuses can be tracked conveniently and prejudgments for drivers' operation decisions can be obtained from corresponding MNL estimates. In order to make the simulation more realistic, major assumptions regarding drivers' first occurrence in system and self-rebalance behaviors are all

simulated according to real operation records. Then, involving consideration for both demand and supply in future time steps, two operation strategies are designed and simulated with sampled operation records. Both strategies aim at improving driver-order matching efficiency by assigning appropriate weight to each unique driver-order pair in queue. Drivers with higher probabilities for continuous operation are assigned priority in pairing with orders that have destinations in areas which predicted to have gap between demand and supply. The two proposed strategies differ slightly from each other in evaluating for future supply. Compared with reference strategy that no weight assigned to driver-order pair, these two strategies achieve better operation performance in respect of increased volume of matched orders, increased revenue and decreased rebalancing time for drivers. The improvements support the potential for behavioral study of drivers' operation, which can be of great importance in ride sharing systems' design of operation strategies.

Except for proposed strategies, the simulation framework is also capable of exploring various other strategies. Additionally, insights from behavioral study can be applied from multiple angles as well. For example, while the proposed strategies aim at improving matching efficiency, optimal schemes of subsidiary allocation for boosting supply can also be developed from drivers' behavioral study and explored through the simulation framework. Current research can inspire future exploration for both behavioral insights and optimization strategies.

REFERENCE

1. Retrieved July 30, 2017, from http://www.didichuxing.com/en/

2. Spieser, K., Samaranayake, S., Gruel, W., & Frazzoli, E. (2016, January). Shared-vehicle mobility-on-demand systems: a fleet operator's guide to rebalancing empty vehicles. In *Transportation Research Board 95th Annual Meeting* (No. 16-5987).

3. Banerjee, S., Gollapudi, S., Kollias, K., & Munagala, K. (2017, April). Segmenting Two-Sided Markets. In *Proceedings of the 26th International Conference on World Wide Web* (pp. 63-72). International World Wide Web Conferences Steering Committee.

4. Agatz, N., Erera, A., Savelsbergh, M., & Wang, X. (2012). Optimization for dynamic ride-sharing: A review. *European Journal of Operational Research*, *223*(2), 295-303.

5. Xiong, C., Hetrakul, P., & Zhang, L. (2014). On ride-sharing: a departure time choice analysis with latent carpooling preference. *Journal of Transportation Engineering*, *140*(8), 04014033.

6. Daziano, R. A. (2015). Inference on mode preferences, vehicle purchases, and the energy paradox using a Bayesian structural choice model. *Transportation Research Part B: Methodological*, *76*, 1-26.

7. Muñoz, B., Monzon, A., & Daziano, R. A. (2016). The Increasing Role of Latent Variables in Modelling Bicycle Mode Choice. *Transport Reviews*, *36*(6), 737-771.

8. McFadden D (1974).\Conditional Logit Analysis of Qualitative Choice Behavior." In P Zarem-
bka (ed.), Frontiers in Econometrics, pp. 105{142. Academic Press, New York.

9. Train K (2009). Discrete Choice Methods with Simulation. 2nd edition. Cambridge University
Press.

10. Evans, J. R., & Minieka, E. (1992). Optimization algorithms for networks and graphs. New York: Dekker.

11. James, G., Witten, D., Hastie, T., & Tibshirani, R. (2015). *An introduction to statistical learning with applications in R*. New York: Springer.