Faster Deliveries and Smarter Order Assignments for an On-Demand Meal Delivery Platform

Wenzheng Mao[1] Liu Ming [2] Ying Rong [3] Christopher S. Tang [4] Huan Zheng[5]

Oct 2019

Abstract

Problem definition: This paper investigates the impact of delivery performance on future customer orders for an on-demand meal delivery service platform. We also identify factors (e.g., delivery driver's local area knowledge and experience) that can affect delivery performance. Using our results, we illustrate how one can develop an "order assignment policy" that can help a platform to increase future customer orders.

Academic/Practical Relevance: Our intent is to identify the underlying factors and develop an order assignment policy that can help an on-demand meal delivery service platform to grow.

Methodology: By analyzing transactional data obtained from an online meal delivery platform in Hangzhou (China) over a two-month period in 2015, we examine the impact of meal delivery performance on a customer's future orders. Through a simulation study, we illustrate the importance of incorporating our empirical results into the development of a smarter "order assignment policy".

Results: We find empirical evidence that an "early delivery" is positively correlated with customer retention: a 10-minute earlier delivery is associated with an increase of one order per month from each customer. However, we find that the negative effect on future orders associated with "late

[1] Faculty of Business and Economics, The University of Hong Kong, Hong Kong, China.

[2] School of Management and Economics, The Chinese University of Hong Kong (Shenzhen), China

[3] Antai College of Economics and Management, Shanghai Jiao Tong University, Shanghai, China.

[4] UCLA Anderson School of Management, University of California, Los Angeles, U.S.A.

[5] Antai College of Economics and Management, Shanghai Jiao Tong University, Shanghai, China.

deliveries" is much stronger than the positive effect associated with "early deliveries". Moreover, we show empirically that a driver's individual local area knowledge and prior delivery experience can reduce late deliveries significantly. Finally, through a simulation study, we illustrate how one can incorporate our empirical results in the development of an order assignment policy that can help a platform to grow its business through customer retention.

Managerial Implications: Our empirical results and our simulation study suggest that to increase future customer orders, an on-demand service platform should address the issues arising from both the supply side (i.e., driver's local area knowledge and delivery experience) and the demand side (i.e., asymmetric impacts of early and late deliveries on future customer orders) into their operations.

Keywords: Startup Operations, Order Assignment, Delivery Performance, Operations Efficiency.

---

1. Introduction

As more people migrate from rural to urban areas to seek better job opportunities and living conditions, the demand for convenient services continues to rise. At the same time, the advent of real time location information systems and mobile payment systems has spawned various types of online matching platforms[1] such as Uber and Lyft that "enable individuals and/or entities as buyers and sellers to *transact* (i.e., search and match) effectively and efficiently by employing various internet-connected digital communication devices" (European Commission, 2016). Despite their exponential growth, most online platforms are not yet profitable. For example, though Uber launched its IPO in 2019 with over 91 million users, the platform has incurred a profit loss in the billions (Franklin 2019). However, many investors support various innovative online platforms because they care more about long-term growth than short-term profits (Clark 2019).

[1] The reader is referred to Chen et al. (2018) for a comprehensive discussion about different types of on-line platforms.

Outside of ride-hailing services, on-demand meal delivery platforms are experiencing an annual growth rate of 9.3%. To get more online customers, many restaurants are partnering with food-ordering and delivery platforms such as Grubhub (US), Deliveroo (UK), and Ele.me (China) to capture office workers and consumers who are too busy to eat out or take out.[2] By posting menus on these platforms, participating local restaurants enable customers to place meal orders (to be picked up and delivered by drivers). These platforms charge restaurants a commission fee based on the value of each order. According to a Statista Report (2019), worldwide revenue of online food delivery amounts to US$94,385 million, and China accounts for more than 42% of the worldwide revenue. For example, Meituan (China), a publicly traded online food ordering and delivery platform, has over 300 million registered members, 3.6 million registered restaurants, and 2.7 million delivery drivers delivering 24 million meals in China on a daily basis.[3]

As more meal delivery platforms enter the market, they must compete on customer satisfaction (e.g., meal delivery service) in order to retain existing customers and acquire new customers, especially when the customer's cost of switching is very low. While customer retention is caused by higher customer satisfaction, which is affected by better service quality (Anderson et al. 2004, Bolton and Lemon 1999, Bolton et al. 2006, Court and Vetvik 2009, Gomez et al. 2004, Morgan and Rego 2006), it is unclear if these results hold in the context of online food delivery platforms especially because, unlike most service operations that involve only two parties (customers and servers), online meal delivery service platforms coordinate "three different parties": on-demand customers who place orders from restaurants with listings on the platform, independent restaurants (who prepare meals for those online customers as well as their own dine-in customers), and delivery

[2] In 2018, McDonald's partnered with Uber Eats, and Dunkin' partnered with DoorDash.

[3] According to Statista Report (2019) `https://www.statista.com/outlook/374/100/online-food-delivery/worldwide`, China is the biggest market (US$39,888m), followed by the United States (US$19,472m), India (US$7,092m), the United Kingdom (US$3,810m) and Germany (US$2,083m).

drivers with different local area knowledge and experience (who need to travel to restaurants to pick up the orders and then deliver the orders to customers assigned by the platform). These observations, along with the importance of future revenue growth for online platforms (as discussed earlier), have motivated us to study the following two research questions:

1. *What is the impact of an early (or a late) meal delivery on a customer's future orders?*

2. *What is the impact of the driver's local area knowledge (i.e., delivery location familiarity) and experience (i.e., meal delivery experience) on the earliness (or lateness) of a meal delivery?*

We also learned from industry leaders that different meal delivery platforms use in-house algorithms to "assign" orders to drivers to minimize average delivery time by incorporating different factors.[4] However, these platforms have not yet taken the drivers' experience and local area knowledge or the asymmetric effect of early/late delivery into consideration when assigning orders to drivers. This observation motivates us to examine the third research question:

3. *How should a platform assign customer orders (associated with different restaurant locations and customer locations) to drivers (with different local area knowledge, delivery experience, and physical locations) in order to effectively increase a customer's future orders?*

We examine the first two questions by analyzing transactional data from 40,786 meal orders provided by a Chinese online meal delivery platform that occurred between July 1 and August 31, 2015 in Hangzhou, China.[5] To examine our first research question, we develop an "additive hazard model" to estimate each customer's future order; this, however, can be unobservable because such an order may occur after our observation period.[6] Our empirical analysis reveals that early deliveries

[4] Factors include the physical distance between the restaurant and the customer's location, the popularity of a restaurant, peak hours, and weather conditions.

[5] Our data includes information about customer order placement time, required (or promised) delivery time for each order, driver assignment time, meal pick-up time, meal delivery time, customer location, driver characteristics (local area knowledge and delivery experience) and restaurant characteristics (price, location, etc.).

[6] We also implement a method (akin to the two-stage least square method) to deal with potential endogeneity issues through multiple instrumental variables (IVs). We also control for variables including weather, traffic congestion,

can increase a customer's future orders: delivering a meal 10 minutes earlier than expected can increase a customer's future demand by 1.03 orders per month. We also discover an "asymmetric effect" of early versus late deliveries. An earlier delivery can boost a customer's future orders slightly: 10 more minutes earlier on an early delivery can increase a customer's future orders by 0.74 orders per month. However, a late delivery can reduce a customer's future orders significantly: being 10 more minutes late on a late delivery can reduce a customer's future orders by 2.70 orders per month.

To examine the second question, we develop a regression model that incorporates various control variables. We find that a driver's local area knowledge and delivery experience can have significant impacts on delivery time. Specifically, we find that a driver who has 30 days additional work experience can reduce delivery time by 5.10 minutes per order, and a driver with local area knowledge can decrease delivery time by 3.33 minutes per order.

Using the results associated with the first two questions (i.e., the asymmetric effect of early/late deliveries and the effect of driver's local area knowledge and experience), we examine the third question by developing an "order assignment algorithm" that can help the platform to increase its future customer orders. Specifically, we first develop three different order assignment policies. The first policy is a benchmark that focuses on *minimizing delivery distance*, but it does not take the driver's knowledge and experience or the asymmetric impact of early/late deliveries into consideration. The second policy *minimizes delivery time* by utilizing the result of our second question (which takes the impact of the driver's local knowledge and experience into consideration regarding delivery time). The third policy utilizes our results from both the first and the second questions to *maximize a customer's future orders (as a function of delivery performance)*, where

time periods, restaurants, and driver-specific characteristics. Our robustness checks and other tests such as weak-IV (Stock and Yogo 2002) and the effective F-value (Olea and Pflueger 2013) provide support for the appropriateness of our approach.

the delivery performance depends on the driver's knowledge and experience. We evaluate the performance of these three policies via a simulation study.[7] Our simulation study suggests that the third policy can increase orders by 0.024 per customer per day (or 0.72 per month) than the second policy and by 0.0369 more orders per customer per day (or 1.107 per month) than the first policy.

In summary, our analysis yields three key findings. First, we find empirical evidence that faster meal delivery time can enable an online meal delivery service platform to generate more future orders from existing customers. Second, we find an asymmetric effect: early deliveries can increase future orders, but late deliveries can severely reduce future orders. Third, we show that an efficient way to assign orders to drivers is to take delivery drivers' knowledge and experience into account along with the asymmetric impacts of early and late deliveries.

Our empirical and simulation findings have several managerial implications. First, to increase a customer's future orders, a platform's order assignment algorithm should take the driver's local area knowledge and delivery experience into consideration. Second, as the platform grows its business by hiring many new drivers, there will be an inherent heterogeneity among drivers in terms of knowledge (familiarity with different locations) and experience (familiarity with the pick-up and delivery operations). However, the platform can leverage various emerging information technologies including *indoor position and navigation systems* that can suggest the most efficient indoor route for drivers to deliver their orders quickly.[8] By using these smart apps that facilitate knowledge-

[7] Our simulation study samples random orders during peak hours from our data set. Following the common practice that orders are assigned intermittently (e.g., every 5 minutes), we first generate a batch of random samples of customer orders and a random sample of driver locations across the 5-minute time window from our data set. We then assign these orders to different drivers according to those three aforementioned assignment policies.

[8] Some indoor locating and positioning technologies require smart sensors to be installed within some commercial buildings. For example, Here.com is one of the leading companies that develops indoor tracking and positioning smart apps. However, the development of an indoor position system for customers who are located in different buildings

and experience-sharing among different drivers, the platform can reduce their meal delivery time even further. Third, by noting that the impacts of early versus late deliveries on future customer orders are asymmetric, a smarter order assignment should strike the right balance between early and late deliveries (instead of focusing purely on shortest meal delivery time).

This paper is organized as follows. Section 2 reviews relevant literature and Section 3 describes our data. Section 4 analyzes our first and second research questions. Specifically, we examine how an early (or late) delivery affects future customer orders and how a driver's local area knowledge and delivery experience affect delivery time. We then develop and compare three order assignment policies though a simulation study in Section 5. Section 6 concludes the paper.

2. Literature Review

The operational issues arising from online platforms have drawn significant interests from operations management researchers (Chen et al. 2018, Hu 2019, Benjaafar and Hu 2019). There are two main mechanisms with which a platform can coordinate supply and demand. Bai et al. (2018), Cachon et al. (2017), Taylor (2018), Chen and Hu (2019) explore different pricing mechanisms. In addition to pricing, an assignment mechanism (or dispatch mechanism) is an alternative way to improve the efficiency of a platform. The classical dispatch policy is to match customers and drivers based on the closest distance (Lyu et al. 2019, Ozkan and Ward 2019). Recently, researchers have investigated the dispatch policy from various perspectives. Using data from Didi Chuxing, Wang et al. (2017) consider the stable matching problem for order dispatching. Chu et al. (2018) provide

can be extremely costly. As an alternative, the platform can capture different (actual) indoor routes from their own drivers by collecting successive snapshots from the drivers' mobile device cameras. By using these indoor snapshots, one can build a database of images that is suitable for estimating a location in a venue. Once the database is built, a mobile device moving through the venue can take snapshots that can be interpolated into the venue's database, yielding location coordinates. These coordinates can be used in conjunction with other location techniques for higher accuracy. By using various learning algorithms, it is possible for the platform to develop smart apps that can help drivers with limited knowledge and experience to deliver orders to indoor locations faster.

a novel approach to the order announcement mechanism to reduce the drivers' gaming behavior. Hu and Zhou (2019) enumerate sufficient conditions for the optimal dispatching policy to possess a priority hierarchy structure. Ozkan and Ward (2019) propose a continuous linear programming approach and prove its asymptotical optimality in a large market. Lyu et al. (2019) use online convex optimization by incorporating different metrics for order dispatching. We note that Xu et al. (2018), Zhang et al. (2017) use machine learning techniques to develop different large scale order dispatch policies to improve different performance metrics (profit, order acceptance rate, customer wait time, etc.). More broadly, our work also belongs to the field studying last mile delivery, which is within the scope of smart city operations (Mak 2018a, Qi and Shen 2019). Qi et al. (2018) investigate the scalability issues of applying shared mobility to solve last mile delivery problems. Mak (2018b) explores the benefit of utilizing in-store customers to deliver orders for online customers for last mile urban delivery.

Our work is related to empirical studies of delivery's impact on demand. In the B2C environment, delivery speed is one of the most important service measures. Fisher et al. (2019) identify a four percent demand increase from affected consumers after a leading U.S. apparel retailer opened a new distribution center to shorten its delivery time. Cui et al. (2019) show that the removal of a high-quality logistics provider, namely, SF Express, would result in a 14.56% sales reduction for the Alibaba platform. Luo et al. (2019) show that logistics information provided through different means, such as word of mouth, claimed inventory availability and store location, have different impacts on the sales of cameras and mobile phones on the Alibaba platform. In addition, delivery performance also plays a critical role in the B2B environment. Peng and Lu (2017) show that different types of customers, such as trade customers and original equipment manufacturers, value delivery performance metrics differently.

Our paper is different from the above literature in the following areas. First, we study the impact of delivery performance on future customer orders in the context of online meal delivery services.

Second, we investigate the impact of driver's knowledge and experience on delivery performance, and we find empirical evidence that shows early versus late deliveries have asymmetric effects on future orders. Third, our simulation study indicates that incorporating this asymmetric effect into the development of an order assignment policy that matches customer orders (restaurant locations and customer locations) and drivers (driver locations) can help a platform to increase customer retention.

3. Data Description

We collected our data from a large Chinese online meal delivery platform based in Hangzhou (China) over a two-month period (from July 1 to August 31, 2015). This platform used full-time employees as delivery drivers to perform pick-up and delivery operations in Hangzhou throughout our sample time period. The raw data captures the chronicle of each customer order, which begins at the moment a customer places an order and ends at the moment the assigned driver delivers the order to the customer's location. Specifically, we have real-time information about customer order placement time, platform dispatch time (the time at which a meal order is assigned to a driver), driver receipt time (the time at which a driver confirms the order assignment), meal pick-up time (including the time when the driver arrives at the restaurant and the time when he leaves the restaurant with the assigned meal order), and order delivery time. Figure 1 depicts the distribution of orders placed during different hours on a typical day, indicating that the "noon period" (10 a.m.-1 p.m.) is the peak period.

3.1. Dependent Variable: Future Order

The dependent variable of interest is the customer's future order (which is a measure of customer retention), and it is "estimated" from the time elapsed between the current order and the next order. Specifically, we define $Duration_i$ as the estimated time elapsed (measured in days) between the current order i and the next order placed on the platform by the *same* customer. However, due to the limitations of the time span of our dataset, we cannot observe the actual time of the

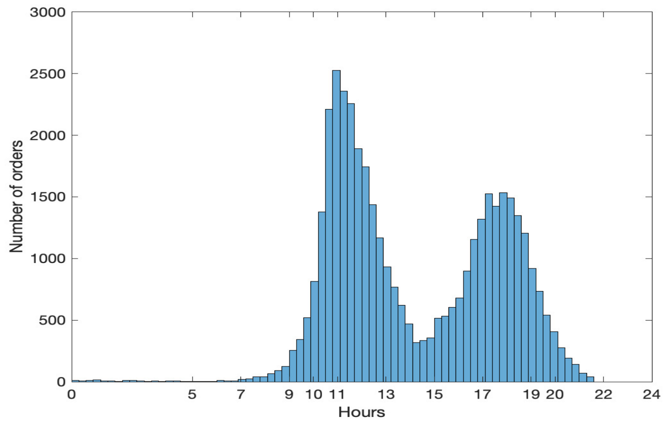


Figure 1 **Number of orders received across different hours**

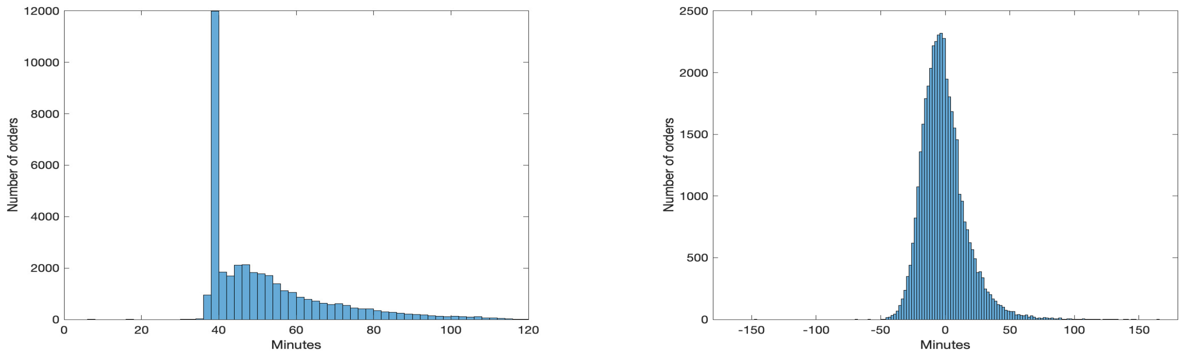
next order for some orders. For example, for those orders placed close to the end of our observation period, the next order would likely occur after our observation period. As a result, to construct the duration time, we define $Duration_i$ as the interorder time and define δ_i as the censoring indicator. Specifically, we define $Duration_i = t_{next(i)} - t_i$, where t_i stands for the time that order i was placed, and $t_{next(i)}$ denotes the time of the next order placed by the same customer. If no order is observed by the end of the observation period T, we set $t_{next(i)} = T$. Additionally, we define $\delta_i = 1$ if the next order is not observed before T, otherwise $\delta_i = 0$.

3.2. Independent Variable: Time Gap

We are interested in a key independent variable, $TimeGap_i$, which measures the time difference (in minutes) between the delivery time (the time at which order i is delivered to a customer) and the "required time" (the time at which this customer expects to receive the order i). There are two types of required time, depending on the order type. First, for a "reserved" order, the required time is the expected delivery time specified by the customer when she placed the order. Second, for an on-demand order (i.e., a "nonreserved" order), the required time is the "estimated" delivery time estimated by the platform, taking order placement time, restaurant (e.g., meal preparation time) and travel distance into consideration.[9] Based on our discussion with the management of the platform, the required time is reported to the customers *truthfully* according to the estimated delivery time. This is because if the reported required time is shorter than the estimated delivery time, the platform will face a higher risk of having a late delivery that can negatively affect customer retention. Additionally, if the reported required time is longer than the estimated delivery time, the platform will face a higher risk of losing an order (order cancellation), which can negatively affect customer acquisition.

Figure 2(a) displays the distribution of the required time interval, i.e., the time interval between the required time and the order placement time. Except for reserved orders (which are placed at least one hour before the platform dispatches orders to drivers), 77.12% customers (in our two-month period sample data) were promised they would receive their meal within one hour after placing an order on the platform. Moreover, by calculating the time difference (between the order finishing time and required time), it can be observed that a nonpositive value of $TimeGap_i$ represents an on-time or early delivery, in which case the customer receives the order before her required time. A positive value of $TimeGap_i$ reflects a late delivery, in which case the customer receives the order later than her required time. Figure 2(b) shows the distribution of the time gap of orders across the whole sample, in which 50.17% are on-time or early deliveries, and the other 49.83% are late deliveries. This observation verifies that the platform reports the required time by using the mean (i.e., the estimated delivery time), as explained earlier.

[9] Since our dataset did not record which orders were reserved, we define a reserved order as an order that is placed at least one hour before the platform dispatches this order to a driver. For example, if a customer places an order at 8 a.m. with a required receipt time of 11 a.m., the platform will only begin to dispatch this order (to a driver) at 10 a.m., which is two hours later than the time at which the order was placed (8 a.m.). We shall classify this order as a reserved order. The measurement, which defines 5.11% reserved orders, is based on dispatch time because it reveals how the platform considers whether an order is urgent or not.



(a) Distribution of required time interval

(b) Distribution of time gap

Figure 2 Distribution of Required Time Interval and Time Gap

3.3. Control Variables

Our model includes various control variables that captures different characteristics of different orders and restaurants together with traffic congestion and weather conditions (temperature). Order characteristics include price of order, distance (between restaurant and customer locations), and whether the order is a reserved order. Additionally, we include a variable $OutstandingOrder_i$ to denote the number of on-hand orders processed by a driver (i.e., driver's utilization) when order i is assigned to the driver. In addition, because we do not have the exact information about the travel distance between restaurant and customer locations, we use the longitudes and latitudes associated with the addresses of restaurants and customers recorded in our data to estimate the travel distance between the restaurant and the customer location associated with order i and call it $Distance_i$, where:

$$Distance_i = 110.574 \times |CLA_i - RLA_i| + 111.320 \times |CLO_i - RLO_i| \times \cos\left(\frac{\pi(CLA_i + RLA_i)}{360}\right), \quad (1)$$

CLA_i and RLA_i are order i's customer address and restaurant address latitudes, and CLO_i and RLO_i are order i's longitudes for customer address and restaurant address, respectively.[10]

[10] The formula we adopt is the spherical earth surface formula, which is commonly used in empirical studies. The parameters for this calculation are available from the U.S. National Geospatial Intelligence Agency at https://msi.nga.mil/MSISiteContent/Staticiles/Calculators/degree.html.

Our control variables include the restaurant's characteristics. Specifically, we incorporate the stay time $ResStay_i$, which captures the time a driver spent at the restaurant waiting for order i. The stay time captures the restaurant's efficiency; a short stay time suggests the restaurant is efficient and the driver does not need to wait long to pick up her order. We also include the coefficient of variation of the number of historical orders for the restaurant, i.e., $CV\_ResOrders$. The coefficient of variation characterizes the variability of the number of orders that the restaurant faces; a smaller coefficient of variation indicates a more stable number of orders that the restaurant receives everyday. In addition, the controls vector includes the traffic congestion level[11] at the moment when the driver picks up her order, outside temperature, two dummy variables associated with peak periods (i.e., 10 a.m.-1 p.m. and 5 p.m.-7 p.m.), and weekday dummy variables.

3.4. Driver's Local Area Knowledge and Delivery Experience

While we are interested in examining the impact of a driver's delivery experience and local area knowledge on early and late deliveries (relative to the required time), the driver's job start date is not available in our database. To overcome this shortcoming, we use the date (5 July 2015) as a cutoff date so that, if a driver appears in our database before 5 July, 2015[12], we define this driver as an "existing driver"; otherwise, we classify this driver as a "new driver". To avoid potential bias, we use only orders delivered by new drivers in our analysis. Specifically, by defining the first day when each new driver appears in our database as his "start" date, the experience of a new driver associated with order i; i.e., $Experience_i$, is measured according to the number of days elapsed

[11] The congestion level (0 to 10) is provided by the Government of Hangzhou, and it is based on weighted travel time associated with different segmented distances under current traffic conditions. More details can be found at `http://jtw.beijing.gov.cn/ysj/xxgk/flfg/r764/201111/P020141111625117643012.pdf`.

[12] The time between two orders delivered by the same drivers is less likely to exceed 5 days because they are full-time employees. Thus, if our data shows no record for the driver before 5 July, 2015, this driver is considered as a new driver. Robustness checks of other cutoff dates such as 3 July, 4 July and 6 July for the following analysis lead to the same results.

between his start date and the date when order i took place. Associated with each order i, we define each driver's local area knowledge; i.e., $Local_i$, as a dummy indicator variable. We set the value to 1 if the driver has visited the neighborhood within 200x200 square meters of the location associated with order i during prior meal deliveries as shown in our database; otherwise, we set the value to 0. Eventually, we arrive at the final data set for our analysis, which is based on $40,786$ meal orders delivered by 161 new drivers over a two-month observation period.

3.5. Summary Statistics

Table 1 reports the summary statistics of our data, and Table A1 (in the Appendix) provides the definition of variables.

4. Empirical Analysis

We begin by examining our first research question: What is the impact of an early (or late) meal delivery on a customer's future orders? However, due to various omitted or unobserved factors (e.g., promotion activities conducted by restaurants/platforms), we cannot directly establish a causal relationship between early (or late) deliveries and future customer orders. To address the potential issue of endogeneity, we identify two instrumental variables (IVs), i.e., driver's local area knowledge (i.e., delivery location familiarity) and experience (i.e., meal delivery experience), and we conduct a two-stage analysis by incorporating these two IVs.

Our first stage analysis, to be presented in Section 4.2, entails the validation of instrumental variables and their roles in our second research question regarding the impact of the driver's knowledge (i.e., delivery location familiarity) and experience (i.e., meal delivery experience) on the earliness (or lateness) of meal deliveries. By studying the second research question, we can understand the impact of driver heterogeneity on time gap.

In addition, the result of the second research question will serve as our first-stage estimation of Time Gap with IVs (for the first research question). Armed with this estimated value of Time Gap, we proceed to the second-stage estimation, which involves an additive hazard model, in Section 4.3. Through this hazard model, we can investigate our first research question by establishing the causality between early (or late) meal deliveries and future customer orders in Section 4.5.

Table 1 Summary Statistics

Variables	Mean	Standard deviation	Min	Max
Duration (days)	3.02	5.03	0	51
Indicator of censoring	0.71	0.45	0	1
Time gap between finished time and required time (minutes)	-1.21	17.49	-187.57	230.47
Stay time in restaurant (minutes)	5.79	7.42	0	94.52
Number of Outstanding Orders	0.70	0.84	0	4
Average number of daily orders in each restaurant	16.76	24	0.02	104.34
Standard deviation of the number of daily orders from each restaurant	5.43	5.84	0.13	27.39
Coefficient of variation of the number of orders from each restaurant	0.57	0.42	0.20	7.87
Price of an order (Yuan)	60.50	47.30	0	1505
Distance between restaurant and customer (kms)	1.63	1.14	0	12.59
Congestion level	2.6	1.13	0.42	8.44
Temperature (Celsius)	30.23	3.83	18	39
Driver's delivery experience (days)	17.14	12.05	1	55
Driver's local area knowledge	0.48	0.5	0	1

	Number of Observations
Customers	20,591
Drivers	161
Restaurants	1,439
Total Orders	40,786
Reserved orders	2,338
On-time/early-delivery orders	20,463

4.1. The Impact of Delivery Performance on Future Orders

4.1.1. A Preliminary Analysis Before we present the aforementioned two-stage analysis, let us first conduct a "preliminary analysis" of delivery performance and future customer demand. To address the concern about those unobserved future orders that may occur after our two-month observation period, we use a survival analysis associated with the hazard rate of time duration for customers to place the next order on the platform as our response variable. Specifically, we apply an additive hazard model as our base model (Aalen 1989), which is known to be more flexible than the proportional hazard model, to incorporate transformations as well as time-dependent covariates. Specifically, our additive hazard model can be expressed as:

$$h(t|TimeGap_i, \boldsymbol{X_i}) = \alpha(t) + \beta(t) \cdot TimeGap_i + \boldsymbol{\gamma(t) \cdot X_i}. \tag{2}$$

Here, $h(t|TimeGap_i, \boldsymbol{X}_i)$ is the hazard rate, which characterizes the "instantaneous probability" of placing another order at time t. $TimeGap_i$ is the time gap between order i and the next order placed by the same customer. The vector X_i are control variables associated with order i, as described in Section 3.3 (e.g., price, distance, congestion level, temperature, categorical indicator (whether it is a reserved order), time and weekday dummy variables). Notice that for ease of exposition, we do not include unobserved variables in the above hazard model. However, the complete model addressing unobserved variables is described in Appendix A2.

Table 2 displays the estimated value of our coefficients associated with $\beta(t)$ and $\gamma(t)$. As indicated, $TimeGap_i$ has a significantly negative effect on a customer's tendency to place a future order. In other words, the earlier the order arrives, the more frequently a customer will place the next order. Specifically, since the time gap is measured in minutes, if the delivery time is shortened by 10 minutes, the "instantaneous probability" (of the next order) increases by $0.00026 * 10 = 0.0026$ per customer per day, which corresponds to an increase of 0.078 orders per customer per month. In addition, we apply MacFadden's Pseudo R^2, proposed by McFadden et al. (1973), to measure the overall goodness of fit for the additive hazard model. This method employs an iterative maximum likelihood estimation process (unlike OLS). In addition, the Pseudo R^2 gives a value of 0.27, which indicates a good fit.[13]

[13] Louviere et al. (2000) propose a goodness-of-fit using McFadden's Pseudo R^2 for fitting the overall model. Because MacFadden's Pseudo R^2 is considerately lower than the traditional R^2 index, McFadden suggests values between 0.2 and 0.4 should be taken to represent a good fit of the model.

Table 2 Survival Analysis of Customer Future Orders (without IVs)

	coefficient	z	p
Intercept	0.247***	15.086	<0.01
TimeGap	**−0.00026*** **	-3.548	<0.01
Num_Orders	-0.005**	-1.961	0.050
ResStay	-0.0002	-1.011	0.312
Distance	0.029***	13.974	<0.01
Price	1.16E-05***	-33.463	<0.01
Congestion	0.005**	2.335	0.020
Temperature	0.002**	2.218	0.027
ReservedTrue	0.022	0.669	0.503
CV_ResOrders	0.109***	17.112	<0.01
Monday	0.017*	1.709	0.087
Tuesday	0.013	1.406	0.160
Wednesday	0.017*	1.931	0.054
Thursday	0.003	0.285	0.776
Saturday	-0.018***	-3.169	<0.01
Sunday	-0.049***	-7.296	<0.01
10 A.M.	0.028***	4.158	<0.01
11 A.M.	0.017***	2.916	<0.01
12 A.M.	0.008	1.167	0.243
1 P.M.	0.012	1.429	0.153
5 P.M.	-0.005	-0.214	0.830
6 P.M.	-0.002	-0.175	0.861
7 P.M.	0.031***	3.331	<0.01
Observations	40,786		
McFadden Pseudo R^2	0.27		

*: $p < 0.1$; **: $p < 0.05$; ***: $p < 0.01$

4.1.2. Endogeneity Issue and Instrumental Variables Although we observe a significant effect of $TimeGap_i$ on a consumer's future orders, as estimated in Equation (2), we cannot infer a causal relationship between these two variables. Additionally, the estimated coefficient of $TimeGap_i$

can be biased. This is because there are some omitted or unobserved factors that may affect a customer's future orders and the time gap simultaneously. For example, we do not have information about promotional activities launched by the platform or restaurants (e.g., special, limited-time discounts; discount coupons for placing next orders; and new restaurant menus) that can largely stimulate consumers to place more future orders. As more orders are triggered by these unobservable factors, driver utilization can increase significantly, which can increase the time gap.

To tackle this endogeneity issue, we incorporate two instrumental variables in the survival analysis, as illustrated in Equation (2), and construct a two-stage survival analysis (akin to the two-stage least square regression model). In the first stage, we regress IVs and other controls on $TimeGap_i$ by using a simple OLS model. In doing so, we obtain the estimated $TimeGap_i$ for each order i. In the second stage, we regress the "estimated" $TimeGap_i$ (obtained from the first stage) and other controls on $Duration_i$ applying a similar additive hazard model to Equation (2).[14]

We use driver's experience ($Experience_i$) and local area knowledge ($Local_i$) as our instrumental variables (IVs). It is known that a good IV should be correlated with the instrumented independent variable (i.e., $TimeGap_i$), and it should be uncorrelated with the dependent variable (i.e., future customer order) with all current covariates being controlled. For the first sign of a good IV, also known as the "relevant condition", we can show a significant correlation between the performance of a driver, measured by delivery time gap, and his local knowledge as well as experience in our first-stage estimation (in section 4.2). It is intuitive that local knowledge and delivery experience can help drivers to become more familiar with the delivery process and navigating, which can effectively shorten the time gap. Second, we argue that $Experience_i$ and $Local_i$ satisfy the "exclusive condition", resolving the second condition of a good IV. Customers cannot observe a driver's $Experience_i$ and $Local_i$, especially when the driver is assigned to an order by a platform that

[14] Appendix A2 provides a proof about unbiased estimation of the survival model when instrumental variables are incorporated.

currently does not take a driver's $Experience_i$ and $Local_i$ into consideration in assigning orders and providing the required time estimation. Therefore, consumers' belief on both delivery time and required time will not be updated according to a driver's $Experience_i$ and $Local_i$.[15] Specifically, the correlation coefficient between future orders and IVs are 0.11 and 0.09, respectively (for $Experience_i$ and $Local_i$). Hence, we can conclude that the exclusive condition is also satisfied for these two variables (i.e., local area knowledge and experience) to be selected as valid IVs.

4.2. First Stage Estimation: The Impact of Driver's Characteristics on Time Gap

As explained earlier, our analysis entails a two-stage estimation process. Our first stage estimates the time gap by including two instrumental variables (i.e., driver's local area knowledge and delivery experience) and other control variables via a simple linear regression model as follows.

$$\begin{aligned} TimeGap_i = {} & \beta_0 + \beta_1 Experience_i + \beta_2 Local_i + \beta_3 Distance_i + \beta_4 Price_i \\ & + \beta_5 Reserved_i + \beta_6 ResStay_i + \beta_7 CV\_ResOrders_i + \boldsymbol{\theta \cdot Control_i} + \varepsilon_i. \end{aligned} \tag{3}$$

Table 3 presents our results for five different variants of the model described in Equation (3), where different variants are associated with different control variables. Specifically, in Model 1, we observe that the coefficients of both delivery experience and local area knowledge are significantly negative; an extra 10 days of work experience leads to a 1.5 minute reduction in time gap. In addition, a driver with local area knowledge results in a decrease of 3.25 minutes in time gap. For the other variables, distance and outstanding orders have significantly positive effects, and reserved order has a negative effect on time gap, while price has little effect (though it is significant) on time gap. Model 2 extends Model 1 by incorporating restaurant characteristics, and the results are consistent. In addition, the result shows that waiting time spent in a restaurant and the coefficient of variation

[15] Some variables including outstanding orders and waiting time at the restaurant are also not observed by customers. However, the value of these variables tends to increase when the demand exceeds the supply. Therefore, if the value of these variables changes, a customer's expectation about required time will be affected because the customer can anticipate the system load.

Table 3 OLS Regression of Driver's Delivery Time Gap with embedded IVs

	Model 1	Model 2	Model 3	Model 4	Model 5
(Intercept)	-6.72 (0.21)***	-9.59 (0.24)***	-9.96 (0.29)***	-3.70 (0.76)***	-4.02 (0.78)***
Experience	**-0.15 (0.01)*****	**-0.16 (0.01)*****	**-0.16 (0.01)*****	**-0.17 (0.01)*****	**-0.17 (0.01)*****
Local	**-3.25 (0.17)*****	**-3.34 (0.17)*****	**-3.37 (0.17)*****	**-3.34 (0.17)*****	**-3.33 (0.17)*****
OutstandingOrders	8.32 (0.09)***	8.97 (0.09)***	8.70 (0.11)***	8.75 (0.11)***	8.70 (0.11)***
Distance	1.33 (0.08)***	1.31 (0.08)***	1.30 (0.08)***	1.31 (0.08)***	1.30 (0.08)***
Price	0.00 (0.00)***	0.00 (0.00)***	0.00 (0.00)***	0.00 (0.00)***	0.00 (0.00)***
ReservedOrder	-4.30 (1.71)*	-4.40 (1.68)**	-4.00 (1.68)*	-3.92 (1.68)*	-3.81 (1.68)*
ResStay		0.39 (0.01)***	0.40 (0.01)***	0.40 (0.01)***	0.40 (0.01)***
CV_ResOrders		1.93 (0.18)***	1.83 (0.18)***	1.81 (0.18)***	1.80 (0.18)***
10 A.M.			-1.53 (0.28)***	-1.61 (0.28)***	-1.55 (0.28)***
11 A.M.			1.14 (0.28)***	0.94 (0.29)**	0.98 (0.29)***
12 P.M.			1.10 (0.30)***	1.06 (0.31)***	1.08 (0.31)***
1 P.M.			-0.91 (0.36)*	-0.72 (0.36)*	-0.71 (0.36)*
5 P.M.			1.08 (0.29)***	1.27 (0.30)***	1.44 (0.31)***
6 P.M.			2.39 (0.30)***	2.44 (0.32)***	2.66 (0.32)***
7 P.M.			2.21 (0.37)***	1.79 (0.37)***	1.84 (0.37)***
Congestion				-0.26 (0.09)**	-0.35 (0.09)***
Temperature				-0.18 (0.02)***	-0.16 (0.02)***
Weekday FE	No	No	No	No	Yes
Observations	40,786	40,786	40,786	40,786	40,786
Adj. R^2	0.24	0.35	0.40	0.42	0.42

Note: Numbers in brackets are standard deviations for the corresponding parameter estimates.

*: $p < 0.1$; **: $p < 0.05$; ***: $p < 0.01$

(i.e., variable of $CV\_ResOrders$ – variability of the number of orders that the restaurant has to handle) can increase time gap.

For Models 3, 4 and 5, we examine the hour effect, because orders tend to surge during peak hours (10 a.m.-1 p.m. and 5 p.m.-7 p.m.). We can observe from Table 3 that the results associated with Models 3, 4 and 5 are consistent with Models 1 and 2. In addition to that, we find that the time gap tends to be higher during peak hours.

In summary, by examining the results associated with all five models shown in Table 3, we can conclude that a driver's delivery experience as well as his local area knowledge can reduce the time gap through faster delivery services even if meal delivery service is considerably simpler than other services or activities (see the analysis of experience for banks (Staats and Gino 2012) and project management (Calvo et al. 2019). These results answer our second research question about the impact of driver's local area knowledge and experience on the earliness of a meal delivery.

4.3. Second-Stage Estimation: The Impact of "Estimated" Time Gap on Future Customer Orders

While all models in Table 3 are based on the linear regression model as stated in (3), we use Model 5 as our first stage estimation of the time gap measure for our second stage estimation of a customer's future orders via the additive hazard model, as previously explained. To do so, let us denote order i's "estimated" time gap $TimeGap\_Fit_i$ as estimated by Model 5. Then, we replace the independent variable $TimeGap_i$ with $TimeGap\_Fit_i$ in our hazard model, as stated in Equation (2), and re-estimate the corresponding coefficients. Denote the instrumental variable vector as Z, and we can conduct this estimation by applying the following additive hazard model conditional on Z, which was proposed in a survival context by Tchetgen et al. (2015), i.e.,

$$\tilde{h}(t|\boldsymbol{X_i}, Z) = \tilde{\alpha}(t) + \tilde{\beta}(t) \cdot TimeGap\_Fit_i + \tilde{\boldsymbol{\gamma}}\boldsymbol{(t)} \cdot \boldsymbol{X_i}. \tag{4}$$

Because of the incorporation of instrumental variables, coefficient estimations of $\tilde{\beta}(t)$ and $\tilde{\boldsymbol{\gamma}}\boldsymbol{(t)}$ are different from those associated with our base model as stated in Equation (2) (without the inclusion of any IVs). This second-stage estimation enables us to measure how the "estimated" time gap based on Model 5 affects a customer's future orders; this is intended to answer our first research question. Notice that in the estimation due to the additional uncertainty generated from the estimation of $TimeGap\_Fit$ in the first-stage regression of Equation (3), we apply nonparametric bootstrap (Efron and Tibshirani 1994) to obtain more accurate estimates of standard errors together with p-value in the second-stage Equation (4).

By using "full sample" data that includes both early and late deliveries (labeled as Full-Sample), Table 4 (Full Sample) shows the result of our estimation of $\tilde{\beta}(t)$ and $\tilde{\boldsymbol{\gamma}}(\boldsymbol{t})$. Akin to our base model (without IVs) as shown in Table 2, Table 4 (Full Sample) implies that the "estimated" time gap *TimeGap_Fit* (based on Model 5 with two embedded IVs) continues to have a significantly negative effect on future customer orders. More importantly, we find a much stronger effect than before: delivering a meal order 10 minutes before the required time can create an increase of $0.0344(0.00344 * 10)$ orders per customer per day, which corresponds to an increase of 1.03 orders per customer per month.

To put our results into perspective, let us consider Meituan's meal delivery service. In the second quarter of 2019, Meituan processed 2.085 billion orders[16] for 44.19 million monthly active consumers[17] and generated a revenue of \$1.86 billion. By using these figures, we can estimate the value created by reducing time gap by 10 minutes through faster delivery services thereby leading to one extra order per customer per month. The associated increase in future orders can be translated into a \$121.81 $(44.19 * (1.86/2.085) * 1.03 * 3)$ million increase in revenue on a quarterly basis.

4.4. Validity of Instrument Variables

To justify the validity of our choice of IVs, we conduct three different tests. The first Hausman test (Hausman 1978) gives a p-value less than 0.1%, indicating the existence of endogeneity. This finding also indicates that we should not apply our base model as stated in Equation (2) as the coefficient estimates are biased. Second, we test for weak instruments. The first-stage F-value is 517.1, much greater than the basic rule-of-thumb value of 10 or 19.93, which is the 5% critical value proposed by Stock and Yogo (2002). Thus, we can reject the null hypothesis that the instruments of local experience and work experience are weak; this then favors the alternative hypothesis that these IVs are strong. In addition, the robustness check using the effective F-value proposed by

[16] `http://meituan.todayir.com/attachment/201908231717191783443873_en.pdf`

[17] `http://boyue.analysys.cn/view/article.html?articleId=20019271&columnId=8`

Table 4 Survival Analysis of Future Customer Orders with embedded IVs

	Full − Sample		*Early − Delivery*		*Late − Delivery*	
	coefficient	z	coefficient	z	coefficient	z
Intercept	0.205***	12.300	0.213***	6.800	0.179***	3.610
TimeGap_Fit	**−0.00344***	-4.540	**−0.00246***	-2.880	**−0.00978***	-2.540
Num_Orders	0.019***	2.600	2.28E-04***	-0.234	0.026***	1.960
ResStay	0.001**	2.000	3.84E-04**	-0.131	0.003**	2.830
Distance	0.034***	14.300	0.029***	8.990	0.036***	10.900
Price	0.000***	-29.900	-9.45E-06***	-19.400	-1.43E-05***	-26.400
Congestion	0.004*	1.840	0.001*	0.499	0.006*	1.740
Temperature	0.002**	2.270	0.003**	3.260	0.001**	0.516
ReservedTrue	0.014	0.533	0.133	1.120	-0.055	-0.377
CV_ResOrders	0.114***	17.600	0.116***	13.200	0.114***	11.100
Monday	0.017*	1.680	0.031*	2.300	0.007*	0.340
Tuesday	0.011	-3.290	0.019	1.440	0.005	0.350
Wednesday	0.014	-6.740	0.032	2.650	0.001	-0.090
Thursday	0.006	0.618	0.027	2.240	-0.01	-1.090
Saturday	-0.018***	1.240	-0.011***	-1.660	-0.023***	-2.710
Sunday	-0.046***	1.630	-0.030***	-3.050	-0.058***	-6.040
10 A.M.	0.026***	3.770	0.034***	2.650	0.015***	1.980
11 A.M.	0.022***	3.500	0.030***	2.760	0.010***	1.540
12 A.M.	0.013*	1.740	0.018*	1.540	0.007*	0.728
1 P.M.	0.011	1.320	0.025	1.550	0.002	0.307
5 P.M.	0.001	0.500	0.016	1.370	-0.003	0.041
6 P.M.	0.007	0.862	0.028	2.030	-0.006	-0.308
7 P.M.	0.038***	3.870	0.041***	2.130	0.037***	3.190
Observations	40,786		24,063		16,723	
McFadden Pseudo R^2	0.30		0.23		0.26	

*: $p < 0.1$; **: $p < 0.05$; ***: $p < 0.01$

Olea and Pflueger (2013) for the first stage gives the same result. Since we have more instruments (two) than the number of endogenous variables (one), an additional Sargan test examines the overidentification problem. The p-value of 0.576 is large enough that we cannot reject the null hypothesis that the over-identification restrictions are valid. Consequently, we can conclude that our IVs are valid and that the model as estimated is not misspecified.

4.5. Asymmetric Impact of Time Gap on Future Customer Orders

To further investigate how early and late deliveries affect a customer's future orders, we divide our data into two subsamples: early deliveries with $TimeGap_i$ being nonpositive in one sample (labeled as Early-Delivery) and late deliveries with $TimeGap_i$ being positive in a second sample (labeled as Late-Delivery). Then, we conduct the same two-stage estimation with IVs for these two subsamples separately. Table 4 (Early-Delivery) and Table 4 (Late-Delivery) summarize our results. Observe from Table 4 (Late-Delivery) that a late delivery has a much stronger effect on future customer orders. Specifically, if the delivery time of a late order decreases by 10 minutes (i.e., less late by 10 minutes) given an unchanged required time, future customer orders per customer per day increase by 0.0978 ($0.00978 * 10$), or an increase of 2.7 orders per customer per month. However, the effect of early deliveries is relatively mild. The estimated coefficient as reported in Table 4 (Early-Delivery) is around 0.0246 ($0.00246 * 10$), i.e., an increase of 0.74 orders per customer per month is associated with a 10-minute earlier delivery. Overall, the effect of a later delivery on future customer orders is around four times stronger than that of an early delivery. The asymmetric effect implies that, to increase a customer's future orders, the platform should assign orders to drivers more wisely in order to avoid late deliveries. In addtion to the effect of time gap, the coefficients of other variables and control variables are consistent with the results based on the Full Sample as shown in Table 4.

4.6. Robustness Checks

4.6.1. Reserved Orders The required time for those "on demand" orders is set by the platform according to the estimated delivery time, which is based on customer location, order and restaurant characteristics, driver utilization, congestion, etc. However, we do not observe order cancellations due to customers finding the required time set by the platform to be unacceptably long. As a robustness check, we can focus our analysis on reserved orders alone, because they are unlikely to be cancelled by the customers, especially as the required time is selected by the customers when they place their reserved orders.[18] We perform the same two-stage analysis by focusing on those reserved orders alone. The result is summarized in Table 5, which exhibits similar coefficient of $TimeGap\_Fit$ as shown in Table 4 (i.e., -0.00347 versus -0.00344). Additionally, the ratio of the effect of time gap for early to late deliveries for the subsample of reserved orders (i.e., $0.00943/0.00267 = 3.53$) is also close to the ratio for the whole sample without isolation of reserved orders (i.e., $0.00978/0.00246 = 3.97$). Thus, we can conclude that the effect of order cancellations (those who do not place orders on the platform due to a concern about a long required time) is negligible.

4.6.2. Repeat Customers Although we use a survival analysis to address the issue of unobserved future orders that occur after our observation period, there is a concern that some customers may never order again from the platform after placing their first orders. Thus, the coefficient of $TimeGap\_Fit$ in our previous analysis (Table 4) may be over or under estimated. To address this concern, we use the sample of repeat customers who made purchases at least twice within our observation period. We repeat the same two-stage analysis by focusing on repeat customers. Table 6 shows our results based on repeat customers. Comparing Table 4 with Table 6, we can see the coefficients of $TimeGap\_Fit$ in two tables, i.e., -0.00344 (in Table 4) and -0.00348 (in Table 6), indicating that an increase of 1.03 orders and 1.04 orders per month, respectively, are not significantly different in full sample regressions. Moreover, the impact of late delivery is always around four times greater than that of early delivery on future customer orders in both analyses. Therefore, we can conclude that the effect of time gap on a customer's future orders is robust no matter whether the customer is a repeat customer or not.

[18] Recall that reserved orders are defined in section 3, and are orders whose dispatch times (from platform) are at least one hour later than the time they were placed (by the customer).

Table 5 Robustness Check: Reserved Orders Only

	Full − Sample		*Early − Delivery*		*Late − Delivery*	
	coefficient	z	coefficient	z	coefficient	z
Intercept	0.067***	2.108	0.050***	1.250	0.235***	1.510
TimeGap_Fit	**−0.00347*****	-2.593	**−0.00267*****	-2.100	**−0.00943*****	-1.930
Num_Orders	0.032***	1.981	0.016***	1.850	0.005***	1.940
ResStay	0.001**	1.399	0.001**	-0.053	4.75E-05**	0.080
Distance	0.010***	2.861	0.014***	1.880	0.009***	2.260
Price	-3.97E-06***	-5.991	-3.70E-06***	-6.020	-8.43E-06***	-4.010
Congestion	0.009*	1.604	0.010*	1.130	0.002*	0.333
Temperature	-0.001**	-1.563	0.002**	0.848	-0.003**	-2.400
ReservedTrue	0.048	5.184	0.047	3.670	0.042	3.190
Monday	0.021***	1.305	0.039***	1.310	0.024***	1.040
Tuesday	0.007	0.584	0.016	0.916	0.000	-0.296
Wednesday	-0.007	-0.462	-0.008	-0.350	0.001	-1.260
Thursday	0.011***	0.699	0.015***	0.705	0.007***	0.231
Saturday	-0.005*	-0.363	0.000*	0.215	-0.002*	-0.004
Sunday	-0.008***	-0.620	0.005***	0.531	-0.016***	-0.005
10 A.M.	0.043	4.413	0.026	1.560	0.042	3.530
11 A.M.	0.086***	5.776	0.064***	2.980	0.073***	3.460
12 A.M.	0.129***	4.841	0.141***	3.200	0.111***	3.080
1 P.M.	0.038*	1.093	0.084*	1.040	0.032*	0.592
5 P.M.	-0.011	-0.788	0.009	0.566	-0.017	-1.230
6 P.M.	0.049	1.639	0.113	2.040	0.008	0.067
7 P.M.	0.149	2.508	0.305	2.730	-0.044	-0.787
Observations	2,338		1,442		896	
McFadden Pseudo R^2	0.23		0.18		0.20	

*: $p < 0.1$; **: $p < 0.05$; ***: $p < 0.01$

5. Order Assignment Policies: A Simulation Study

In the last section, we established empirical evidence indicating that late deliveries have a stronger negative impact on a customer's future orders than the positive impact generated by early deliv-

Table 6 Robustness Check: Survival Analysis for Demand with Repeat Customers

	Full − Sample		***Early − Delivery***		***Late − Delivery***	
	coefficient	z	coefficient	z	coefficient	z
Intercept	0.486***	10.200	0.304***	3.840	0.289***	1.633
TimeGap_Fit	**−0.00348***	-2.490	**−0.00220***	-2.010	**−0.01050***	-1.870
Num_Orders	0.029***	2.540	0.037***	2.900	0.026***	1.400
ResStay	0.001**	1.590	0.002**	2.210	0.000**	0.452
Distance	0.035***	5.780	0.033***	3.850	0.036***	4.590
Price	-1.89E-05***	-18.000	-1.53E-05***	-10.600	-2.14E-05***	-15.400
Congestion	0.015*	3.020	0.010*	1.560	0.022*	2.980
Temperature	0.008**	6.100	0.010**	5.230	0.006**	2.450
ReservedTrue	-0.053	-0.139	0.343	1.310	-0.06	1.320
CV_ResOrders	0.101***	6.900	0.074***	4.190	0.115***	4.900
Monday	0.034*	1.260	0.047*	1.270	0.019*	0.120
Tuesday	0.011	0.299	0.047	1.520	-0.015	-0.916
Wednesday	0.030	1.540	0.081	2.930	-0.001	-0.315
Thursday	-0.014	-1.110	0.022	0.654	-0.043	-1.980
Saturday	-0.003***	-0.809	0.016***	0.282	-0.018***	-1.330
Sunday	-0.068***	-4.050	-0.039***	-1.410	-0.091***	-4.140
10 A.M.	0.040***	2.500	0.032***	1.240	0.049***	2.490
11 A.M.	0.006***	0.826	0.021***	1.280	-0.002***	0.174
12 A.M.	0.000*	0.601	0.010*	0.907	-0.005*	0.204
1 P.M.	-0.013	-0.148	-0.035	-0.552	-0.003	0.286
5 P.M.	-0.024	-0.947	-0.038	-0.801	-0.02	-0.768
6 P.M.	-0.031	-1.610	0.005	0.209	-0.07	-2.570
7 P.M.	0.031***	1.260	0.022***	0.408	0.042***	3.880
Observations	27624		16825		10799	
McFadden Pseudo R^2	0.28		0.27		0.22	

*: $p < 0.1$; **: $p < 0.05$; ***: $p < 0.01$

eries. Additionally, a driver with local area knowledge and more delivery experience can perform the delivery service faster. These two results suggest that, to increase future customer orders, the platform should take (a) the impact of a driver's knowledge and experience on time gap (or delivery time), and (b) the asymmetric impact of early and late deliveries on future orders into consideration when assigning orders to drivers. After establishing these two empirical results, we are interested in exploring our third research question: how should one incorporate factors (a) and (b) to develop an order assignment policy that can effectively increase a customer's future orders? Upon discussing with the management, we learned that their order assignment policy did not take driver knowledge and experience or the asymmetric impact of early/late deliveries into consideration (primarily because the effect of factors (a) and (b) was not known to the management until our empirical study). This observation motivates us to develop an effective order assignment policy that incorporates both factors (a) and (b). Additionally, we examine the benefit of our order assignment policy over two benchmark policies via a simulation study.

To evaluate the performance of our policy that takes factors (a) and (b) into consideration, we first establish two benchmark policies. The first policy (Policy 1) is a benchmark policy that assigns orders to drivers by minimizing total delivery distance. This policy takes neither factor (a) nor (b) into consideration, which mimics the policy used by the platform.[19] The second policy (Policy 2) assigns orders to drivers by minimizing the total estimated delivery time. This policy takes factor (a) but not factor (b) into consideration. Finally, our policy (Policy 3) takes both factors into consideration, and it assigns orders to drivers by maximizing the estimated number of future orders (which depends on each driver's knowledge and experience and the asymmetric impact of early/late deliveries).

5.1. Simulating Orders and Driver Locations

Before we define these three policies formally, let us first describe how we evaluate these three policies via a simulation study. Note that our data set contains information about each driver's

[19] The actual policy adopted by the platform is not available to us even though we know that the adopted policy does not take either factor into consideration. Therefore, Policy 1 captures the spirit of the policy adopted by the platform.

location at the moment he arrives at the designated restaurant's (or customer's) location, but it does not contain information about each driver's location in real time. Hence, we cannot simulate order assignments dynamically in real-time. To overcome this challenge, we take a different tack. Instead of assigning each incoming order to a driver in real time, we consider a discrete time approach in which the platform assigns orders accumulated over a short-time period (e.g., every 5 minutes) to drivers. Per our discussion with the management, dispatching orders using discrete time (e.g., every 5 minutes) appears to be consistent with actual practice.

In our simulation study, we consider a specific instant when the platform needs to assign N orders (accumulated within a short-time period, say, 5 minutes) to N drivers. We first generate these N "random orders" (consisting of customer and restaurant locations) randomly from a pool orders received during the noon peak hours (10 a.m. - 1 p.m.) in our data set. Then, to develop a proxy of the locations of those N drivers, we assume that all N drivers are completing their deliveries at different customer locations at that specific instant. To do so, we generate a set of N "random locations" for the drivers by selecting another N customer order locations randomly from our database during the same time period. As a result, we have N randomly generated "driver locations" and N "new orders" (with both restaurant and customer locations) for that specific instant. By using this information, we can compute the performance metrics associated with each policy when we assign order i to driver j, where $i, j = 1, \cdots, N$ associated with different policies.

5.2. Three Order Assignment Policies

All three assignment policies are based on the following standard assignment constraints:

$$\sum_{i=1}^{N} x_{ij} = 1, \quad \forall j = 1 \dots N, \sum_{j=1}^{N} x_{ij} = 1 \quad \forall i = 1 \dots N, x_{ij} = \{0, 1\}, \quad \forall i, j, \tag{5}$$

where $x_{ij} = 1$ if order i is assigned to driver j and equals 0, otherwise. However, different assignment policies are established according to different objective functions.

Specifically, Policy 1 focuses on minimizing the total distance, where the travel distance d_{ij} when we assign order i to driver j is computed according to Equation (1). In this case, the order assignment policy under Policy 1, x_{ij}, is the optimal solution to the following assignment problem:

$$\text{(Policy 1)} \ \min \ \sum_{i=1}^{N}\sum_{j=1}^{N} d_{ij}x_{ij} \ \text{ subject to (5).}$$

Next, Policy 2 aims to minimize the total delivery time, where the delivery time t_{ij} when driver j is assigned to order i is given as:

$$t_{ij} = \frac{d_{ij}}{s} + \hat{\beta}_{exp} Exp_j + \hat{\beta}_{local} Local_{ij} + outstanding_j. \tag{6}$$

Here, s is the travel speed (km/hour) which is assumed to be the same among all drivers. In the simulation, we draw s from Uniform[10, 20] so that the average travel speed is 15 km/hour; this resembles the actual travel speed in Hangzhou during noon hours. Exp_j is driver j's delivery experience (in days). $Local_{ij}$ is a dummy variable indicating whether driver j had previously visited the neighborhood of order i's customer location.[20] Additionally, because it is possible that the driver will have outstanding orders to finish when he gets this new assignment, we use $outstanding_j$ to denote the additional delivery time to complete outstanding orders by driver j.[21] Hence, the order assignment policy under Policy 2, x_{ij}, is the solution to the following assignment problem:

$$\text{(Policy 2)} \ \min \ \sum_{i=1}^{N}\sum_{j=1}^{N} t_{ij}x_{ij} \ \text{ subject to (5).}$$

[20] Because our data set does not contain delivery experience for drivers appearing before July 5, 2015, we generate Exp_j from Uniform[0, 60] and $Local_{ij}$ from a Bernoulli distribution with a success rate 0.2. $\hat{\beta}_{exp}$ and $\hat{\beta}_{local}$ are set according to Table 3 (Model 5), i.e., $\hat{\beta}_{exp} = -0.17$ and $\hat{\beta}_{local} = -3.33$.

[21] The number of outstanding orders is randomly drawn from the data. If driver j has n outstanding orders on hand, $outstanding_j$ is randomly generated from a normal distribution $N(20n, 9\sqrt{n})$ (in minutes), because the mean and standard deviation of service time are approximately 20 minutes and 9 minutes, respectively. Note that the service time of order j is calculated as $finishTime_j - \max(dispatchTime_j, previousOrderFinishTime_j)$.

Finally, Policy 3 focuses on maximizing total future customer orders. To determine future customer orders for different order assignments, we need to utilize the empirical results that were established earlier. First, for each randomly selected order i from our database, we have information about the corresponding restaurant and customer locations as well as the required time r_i for the order. Recall from Table 4 that an early delivery of $(r_i - t_{ij})^+$ minutes is associated with an estimated coefficient 0.00246 (denoted by $\hat{\beta}_e$) and a late delivery of $(t_{ij} - r_i)^+$ minutes is associated with an estimated coefficient 0.00978 (denoted by $\hat{\beta}_l$), where the delivery time t_{ij} when driver j is assigned to order i is given in (6). By noting that early (late) deliveries can increase (decrease) a customer's future orders, the order assignment policy under Policy 3, x_{ij}, is the solution to the following assignment problem:

$$\text{(Policy 3)} \;\; \max \;\; \sum_{i=1}^{N} \sum_{j=1}^{N} \left[\hat{\beta}_e (r_i - t_{ij})^+ - \hat{\beta}_l (t_{ij} - r_i)^+ \right] x_{ij} \;\; \text{subject to (5).}$$

5.3. Simulation Results

We conduct our simulation study to evaluate the aforementioned three policies as follows. We randomly generate 500 scenarios by using the data set associated with orders that occur during the noon time period (10 a.m. - 1 p.m.). Each of the 500 scenarios is based on $N = 150$ randomly generated driver locations and $N = 150$ randomly generated orders (restaurant and customer locations as well as the required time for each order). For each scenario, we determine the optimal assignment for each policy by solving the three corresponding (assignment) problems described above. In addition to the optimal order assignment for each policy, we compute different performance metrics: total delivery distance (in kms), total delivery time (in minutes), on-time/early delivery percentage (i.e., the proportion of those 150 randomly generated orders completed before the required time), and the increase in a customer's future orders (i.e., the increase of a customer's reorders per day as predicted by our additive hazard model previously presented).

By computing the average performance associated with the three order assignment policies across all 500 scenarios, we summarize our results in Table 7. Observe from the table that all three

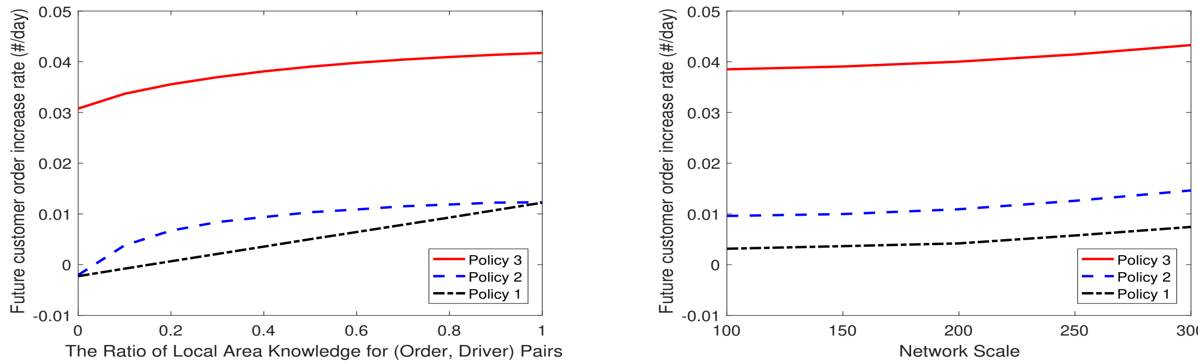
Table 7 Performance Comparison across Three Order Assignment Policies

	Policy 1 (Min Distance)	Policy 2 (Min Delivery Time)	Policy 3 (Max Future Orders)
Avg. Delivery Distance (Km)	3.56	3.66	3.87
Avg. Delivery Time (Minute)	37.44	36.12	37.38
Avg. On-Time/Early Delivery (%)	73.3%	75.2%	88.7%
Avg. Future Order Increase (order/day)	0.0010	0.0095	0.0379

order assignment policies have a similar performance in terms of travel distance and delivery time. However, Policy 3 outperforms the other two benchmark policies in terms of average on-time/early deliveries and future order increase. The reason why Policy 3 dominates the other two policies in these two dimensions is because Policy 3 takes the asymmetric effect of early and late deliveries on future customer orders into consideration when assigning orders to drivers. In addition, the consideration of the impact of a driver's knowledge and experience on time gap alone can generate more future orders as indicated by the comparison between Policy 2 and Policy 1.

Next, we compare performance across all three order assignment policies in terms of the rate of future orders increases as we vary the local area knowledge of drivers. To do so, we generate 500 scenarios by following the same process as described earlier. However, to vary the local area knowledge of drivers, we define a parameter $l \in [0, 1]$ and we vary this parameter l from 0 to 1. For each given value of l and for each of the 500 scenarios, we simulate the local area knowledge of each driver and order pair $Local_{ij}$ as a Bernoulli trial so that $Local_{ij} = 1$ with probability l and equals 0 with probability $(1 - l)$. For each randomly generated value of $Local_{ij}$, we can use Equation (6) to compute the delivery time t_{ij} so that we can determine Policy 2 and Policy 3 by solving the corresponding assignment problems. Figure 3(a) depicts our results as we vary l from 0 to 1.

Based on Figure 3(a), we observe that Policy 3 bests the other two policies in terms of future order increase rate because it takes the asymmetric effect of early/late deliveries on future orders into consideration. For Policy 1, the future order increase rate grows linearly as the local area knowledge



(a) Impact of driver's local area knowledge

(b) Impact of network scale

Figure 3 **Performance of three order assignment policies**

level l increases because the underlying order assignment does not take local area knowledge into consideration. Therefore, future order increase is gained purely from the reduction in delivery time due to the increase in driver's local experience level l. By comparing Policy 1 and Policy 2, it is evident that the benefit of incorporating the impact of a driver's knowledge and experience on time gap is high when the driver's local area knowledge l is modest. In addition, Figure 3(a) also suggests that, under policies 2 and 3, the future order increase rate exhibits a decreasing marginal return as the driver's local area knowledge level l increases. Hence, it is sufficient for the platform to have 20% of driver and order pairs with local area knowledge because the marginal benefit to increasing local area knowledge beyond 20% is limited. If one treats delivery drivers with local area knowledge as a flexible resource, our study of delivery services echoes the classical finding about process flexibility. That is, a little flexibility can generate a performance that is close to full flexibility (Jordan and Graves 1995, Chou et al. 2010, Wang and Zhang 2015, Désir et al. 2016).

Finally, we compare performance across all three order assignment policies when we vary the number of orders N between 100 and 300. As the number of orders within the same geographical region N grows, the chance of finding a driver who is located near a designated restaurant/customer location will increase. (This is akin to the "scaling" effect exhibited in ride-hailing services: a passenger's waiting time and a driver's idle time decrease as the number of passengers and drivers increases.) By using the same approach as before and keeping the driver's local area knowledge level l at 0.2, we obtain the results shown in Figure 3(b). Observe from Figure 3(b) that all three assignment policies can generate more future orders as N increases due to the scaling effect. However, Policy 3 continues to dominate the other two policies.

6. Conclusion and Discussion

We have investigated how delivery performance (Time Gap) affects the revenue growth of an on-demand meal delivery platform by analyzing transactional data from 40,786 orders delivered by 161 drivers over a two-month period in Hangzhou. We found empirical evidence that early deliveries can increase future customer orders, while late deliveries can strongly decrease future customer orders. Additionally, we established empirical evidence indicating that a driver's local area knowledge and delivery experience can significantly affect a driver's delivery time.

By leveraging our empirical results, we developed two benchmark order assignment policies and established a policy that aims to maximize total future customer orders by taking the asymmetric impact of late and early delivery on future orders and each driver's local area knowledge and delivery experience into consideration when assigning orders to drivers. Through our simulation study of 500 scenarios (with 150 random orders in each scenario), we illustrated that, by taking both the asymmetric impact of delivery time gap and each driver's local area knowledge and delivery experience into consideration when assigning orders to drivers, the platform can significantly increase a customer's future orders.

Beyond our assignment policy that takes each driver's local area knowledge and delivery experience into consideration, some platforms have developed incentives to discourage drivers from delivering their orders late. For instance, Meituan and Ele.me in China penalize drivers for late deliveries. However, this penalty system can backfire. For instance, to avoid being penalized, many meal delivery drivers in China prioritize speed over safety, which has caused traffic accidents and deaths in China.[22] This observation raises an interesting question to be examined in the future: What is the right (reward and/or penalty) incentive mechanism that can encourage more on-time deliveries without sacrificing safety?

Additionally, our data revealed that drivers wait for 5.79 minutes (on average) to pick up their orders in restaurants. Therefore, it is of interest to examine the following question: Is there a way for the platform to help restaurants to plan their operations better so that drivers do not need to waste too much time waiting for orders? Should the platform also develop mechanisms to encourage customers to place more "reserved" orders so that the platform can better coordinate with restaurants and drivers to improve delivery performance?

Akin to other smart city initiatives (Mak 2018a, Qi and Shen 2019), the on-demand nature and the huge volume of data related to meal delivery services in a city creates opportunities to develop innovative solutions. Ultimately, the OM research community has more novel planning and coordinating problems to explore.

References

Aalen, O. O. (1989). A linear regression model for the analysis of life times. *Statistics in medicine*, 8(8):907–925.

Anderson, E. W., Fornell, C., and Mazvancheryl, S. K. (2004). Customer satisfaction and shareholder value. *Journal of marketing*, 68(4):172–185.

Bai, J., So, K. C., Tang, C. S., Chen, X., and Wang, H. (2018). Coordinating supply and demand on an on-demand service platform with impatient customers. *Manufacturing & Service Operations Management*.

Benjaafar, S. and Hu, M. (2019). Operations management in the age of the sharing economy: What is old and what is new? *Forthcoming, Manufacturing and Service Operations Management*.

Bolton, R. N. and Lemon, K. N. (1999). A dynamic model of customers' usage of services: Usage as an antecedent and consequence of satisfaction. *Journal of marketing research*, 36(2):171–186.

[22] "Speed over safety? China's food delivery industry warned over accidents," *Reuters*, 28 September, 2017.

Bolton, R. N., Lemon, K. N., and Bramlett, M. D. (2006). The effect of service experiences over time on a supplier's retention of business customers. *Management Science*, 52(12):1811–1823.

Cachon, G. P., Daniels, K. M., and Lobel, R. (2017). The role of surge pricing on a service platform with self-scheduling capacity. *Manufacturing & Service Operations Management*, 19(3):368–384.

Calvo, E., Cui, R., and Serpa, J. C. (2019). Oversight and efficiency in public projects: A regression discontinuity analysis. *Management Science*.

Chen, Y. and Hu, M. (2019). Pricing and matching with forward-looking buyers and sellers. *Manufacturing & Service Operations Management*.

Chen, Y.-J., Dai, T., Korpeoglu, C. G., Körpeoğlu, E., Sahin, O., Tang, C. S., and Xiao, S. (2018). Innovative online platforms: Research opportunities. *Manufacturing & Service Operations Management, Forthcoming*.

Chou, M. C., Chua, G. A., Teo, C.-P., and Zheng, H. (2010). Design for process flexibility: Efficiency of the long chain and sparse structure. *Operations research*, 58(1):43–58.

Chu, L. Y., Wan, Z., and Zhan, D. (2018). Harnessing the double-edged sword via routing: Information provision on ride-hailing platforms. *Working Paper, University of Southern California*.

Clark, K. (2019). Unicorns aren't profitable, and wall street does not care. *TechCrunch*.

Court, David, D. E. S. M. and Vetvik, O. J. (2009). The consumer decision journey. *McKinsey Quarterly*, (3):96—-107.

Cui, R., Li, M., and Li, Q. (2019). Value of high-quality logistics: Evidence from a clash between sf express and alibaba. *Management Science*, Forthcoming.

Désir, A., Goyal, V., Wei, Y., and Zhang, J. (2016). Sparse process flexibility designs: is the long chain really optimal? *Operations Research*, 64(2):416–431.

Efron, B. and Tibshirani, R. J. (1994). *An introduction to the bootstrap*. CRC press.

Fisher, M., Gallino, S., and Xu, J. (2019). The value of rapid delivery in online retailing. *Journal of Marketing Research*, Forthcoming.

Franklin, J. (2019). Uber unveils ipo with warning it may never make a profit. *Reuters*.

Gomez, M. I., McLaughlin, E. W., and Wittink, D. R. (2004). Customer satisfaction and retail sales performance: an empirical investigation. *Journal of retailing*, 80(4):265–278.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica: Journal of the econometric society*, pages 1251–1271.

Hu, M. (2019). From the classics to new tunes: A neoclassical view on sharing economy and innovative marketplaces. *Working Paper, University of Toronto*.

Hu, M. and Zhou, Y. (2019). Dynamic type matching. *Rotman School of Management Working Paper*, (2592622).

Jordan, W. C. and Graves, S. C. (1995). Principles on the benefits of manufacturing process flexibility. *Management Science*, 41(4):577–594.

Louviere, J. J., Hensher, D. A., and Swait, J. D. (2000). *Stated choice methods: analysis and applications*. Cambridge university press.

Luo, J., Rong, Y., and Zheng, H. (2019). Impacts of logistics information on sales: Evidence from an online marketplace. *Working Paper*.

Lyu, G., Cheung, W. C., Teo, C.-P., and Wang, H. (2019). Multi-objective online ride-matching. *Available at SSRN 3356823*.

Mak, H.-Y. (2018a). Enabling smarter cities with operations management. *Available at SSRN*.

Mak, H.-Y. (2018b). Peer-to-peer crowdshipping as an omnichannel retail strategy. *Available at SSRN 3119687*.

McFadden, D. et al. (1973). Conditional logit analysis of qualitative choice behavior.

Morgan, N. A. and Rego, L. L. (2006). The value of different customer satisfaction and loyalty metrics in predicting business performance. *Marketing Science*, 25(5):426–439.

Olea, J. L. M. and Pflueger, C. (2013). A robust test for weak instruments. *Journal of Business & Economic Statistics*, 31(3):358–369.

Ozkan, E. and Ward, A. (2019). Dynamic matching for real-time ridesharing. *Available at SSRN 2844451*.

Peng, D. X. and Lu, G. (2017). Exploring the impact of delivery performance on customer transaction volume and unit price: evidence from an assembly manufacturing supply chain. *Production and Operations Management*, 26(5):880–902.

Qi, W., Li, L., Liu, S., and Shen, Z.-J. M. (2018). Shared mobility for last-mile delivery: Design, operational prescriptions, and environmental impact. *Manufacturing & Service Operations Management*, 20(4):737–751.

Qi, W. and Shen, Z.-J. M. (2019). A smart-city scope of operations management. *Production and Operations Management*, 28(2):393–406.

Staats, B. R. and Gino, F. (2012). Specialization and variety in repetitive tasks: Evidence from a japanese bank. *Management science*, 58(6):1141–1159.

Stock, J. H. and Yogo, M. (2002). Testing for weak instruments in linear iv regression.

Taylor, T. A. (2018). On-demand service platforms. *Manufacturing & Service Operations Management*, 20(4):704–720.

Tchetgen, E. J. T., Walter, S., Vansteelandt, S., Martinussen, T., and Glymour, M. (2015). Instrumental variable estimation in a survival context. *Epidemiology (Cambridge, Mass.)*, 26(3):402.

Wang, X., Agatz, N., and Erera, A. (2017). Stable matching for dynamic ride-sharing systems. *Transportation Science*, 52(4):850–867.

Wang, X. and Zhang, J. (2015). Process flexibility: A distribution-free bound on the performance of k-chain. *Operations Research*, 63(3):555–571.

Xu, Z., Li, Z., Guan, Q., Zhang, D., Li, Q., Nan, J., Liu, C., Bian, W., and Ye, J. (2018). Large-scale order dispatch in on-demand ride-hailing platforms: A learning and planning approach. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 905–913. ACM.

Zhang, L., Hu, T., Min, Y., Wu, G., Zhang, J., Feng, P., Gong, P., and Ye, J. (2017). A taxi order dispatch model based on combinatorial optimization. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 2151–2159. ACM.

A1. Variable Summary

Table A1 Summary of Variables

Variable	Abbreviation	Definition
Time interval between order i and the next order placed by the same customer	$Duration_i = t_{next(i)} - t_i$	t_i corresponds to the day order i was placed, and $t_{next(i)}$ denotes the time the next order was placed by the *same* customer. We set $t_{next(i)} = T$ if no next order was placed by the end of the observation period T.
Censored indicator of $Duration_i$	δ_i	$\delta_i = 1$ when there is no next order being placed before T, otherwise $\delta_i = 0$
Delivery time gap	$TimeGap_i$	Actual delivery time of order i - the platform's estimated delivery time of order i
Delivery distance	$Distance_i$	Travel distance between restaurant and customer for order i
Waiting time in a restaurant for order i	$ResStay_i$	Amount of time that the designated driver has to wait at the restaurant to pick up order i
Coefficient of variation of restaurant orders	$CV\_ResOrder_i$	CoV (standard deviation / mean) of daily orders associated with the restaurant that processes order i
Meal price	$Price_i$	Price of order i
Traffic congestion level	$Congestion_i$	Congestion level when order i is placed.
Week day dummy	$Monday, Tuesday, \ldots$	$Monday_i = 1$ if order i is placed on Monday. $Tuesday_i = 1$ if order i is placed on Tuesday. ...
Peak hours dummy	10A.M., 11A.M.,...	Dummies for peak hours from 10 A.M. to 1 P.M., and 5 P.M. to 7 P.M.
Driver's experience	$Experience_i$	Driver's working days between 5 July 2015 and the day when order i is assigned
Driver's local area knowledge	$Local_i$	$Local_i = 1$ if order i is delivered by a driver who had previously visited the neighborhood (i.e., $200m \times 200m$ region)

A2. Additive Hazard Model with Instrumental Variables

We now extend our additive hazard model to incorporate instrument variables (IVs) and show that applying an IV using a similar-to-2SLS method can address the endogeneity problem and give us unbiased coefficient estimates. Primarily under an additive hazards model, Tchetgen et al. (2015) described two methods including the two-stage regression analysis in a survival context. By following the similar logic, we here demonstrate how IVs work in our model. By recalling from Section 4.1.1 that $TimeGap_i$ represents the time gap between the required time and the actual delivery time of order i and from Section 3.3 that our control variables $\boldsymbol{X}_i$ associated with each order i include price, distance and delivery time, etc., we try to estimate the effect of these variables on the time to customer's next order, denoted by t_{next}. Recall that we first implement the following model without IVs in our estimation in Section 4.1.1,

$$h(t|X^*, \boldsymbol{X}) = \alpha(t) + \beta(t)X^* + \boldsymbol{X}\boldsymbol{\gamma(t)}, \tag{A1}$$

where $TimeGap_i$ is replaced by X^*, $\boldsymbol{X}_i$ is replaced by $\boldsymbol{X}$. In doing so with unobserved variables ignored, we inevitably encounter an endogeneity problem and the coefficient estimates, i.e., $\beta(t)$ and $\boldsymbol{\gamma(t)}$, are biased. Thus, we cannot exclude unobserved variables without dealing with the underlying endogeneity issue in the estimation, and we need to develop a different approach. First, we rewrite a complete form of Equation (2) by including unobserved variables as follows:

$$h(t|X^*, \boldsymbol{X}, U, Z) = \tilde{\alpha}(t) + \tilde{\beta}(t)X^* + \boldsymbol{X}\boldsymbol{\tilde{\gamma}(t)} + \zeta(t)U, \tag{A2}$$

where U is an unobserved variable and Z are instrumental variables. Notice that although this is the complete formation, we cannot directly estimate Equation (A2) because of the unobserved variable U.

Next, we need to show that, by incorporating instrumental variables Z and applying the similar-to-2SLS method in the second-stage estimation, we can remove the effect of unobserved U while keeping the coefficients of covariates unbiased as in Equation (A2). We derive the survival function (i.e., the tail distribution of survival time) associated with the model as stated in (A2), where:

$$S(t|X^*, \boldsymbol{X}, U, Z) = exp\{-\int_0^t [\tilde{\alpha}(s) + \tilde{\beta}(s)X^* + \boldsymbol{X}\tilde{\boldsymbol{\gamma}}(\boldsymbol{s}) + \zeta(s)U]ds\}. \tag{A3}$$

Utilizing the instrumental variable of Z, we can delve into the correlation between Z and the endogenous variable X^* by using the following linear regression model (as our standard first-stage estimation of $TimeGap_i$):

$$X^* = \alpha_0 + \beta_0 Z + \boldsymbol{X}\boldsymbol{\gamma_0} + \epsilon, \tag{A4}$$

where the error term ϵ is assumed to have a mean of 0 and be independent of Z. By assuming that there is a statistically significant correlation between X^* and Z (as verified in our empirical analysis in Section 4.1.2 and Section 4.4), we do not further analyze the strength of such a correlation or the goodness of fit for this model as it requires additional tests. For Z to be a valid IV, time to the next order, i.e., t_{next}, and instrumental variables Z must be conditionally independent given (X^*, $\boldsymbol{X}$, U)[23], i.e., the exclusion restriction condition.

For ease of exposition, let $D = d(Z, \boldsymbol{X}) = \alpha_0 + \beta_0 Z + \boldsymbol{X}\boldsymbol{\gamma_0}$, where $D = E(X^*|Z, \boldsymbol{X})$. Then, we can rewrite the survival function given in (A3) as:

$$S(t|X^*, \boldsymbol{X}, U, Z) = exp\{-\int_0^t [\tilde{\alpha}(s) + \tilde{\beta}(s)D + \boldsymbol{X}\tilde{\boldsymbol{\gamma}}(\boldsymbol{s})]ds\} \cdot exp\{-\int_0^t [\tilde{\beta}(s)\epsilon + \zeta(s)U]ds\} \tag{A5}$$

Our goal is to deal with unobserved variables so that the estimation of the survival function does not depend on U. Then, we have:

$$\begin{aligned} S(t|\boldsymbol{X}, Z) &= E(S(t|X^*, \boldsymbol{X}, U, Z)|\boldsymbol{X}, Z) \\ &= exp\{-\int_0^t [\tilde{\alpha}(s) + \tilde{\beta}(s)D + \boldsymbol{X}\tilde{\boldsymbol{\gamma}}(\boldsymbol{s})]ds\} \cdot E[exp\{-\int_0^t [\tilde{\beta}(s)\epsilon + \zeta(s)U]ds\}|\boldsymbol{X}, Z]. \end{aligned} \tag{A6}$$

[23] We allow that the controlled X^* given Z can be conditionally correlated with the unobserved variable U, i.e., $cov(\epsilon, U|Z) \neq 0$.

First, by assuming that U is independent of Z in addition to the previous independence assumption between ϵ and Z (in order to ensure Z to be a valid IV), we get:

$$E[exp\{-\int_0^t [\tilde{\beta}(s)\epsilon + \zeta(s)U]ds\}|\boldsymbol{X}, Z] = E[exp\{-\int_0^t [\tilde{\beta}(s)\epsilon + \zeta(s)U]ds\}|\boldsymbol{X}] \tag{A7}$$

Second, because $\boldsymbol{X}$ is the covariate vector without an endogeneity problem and is assumed to have no correlation with either U or ϵ, we have:

$$S(t|\boldsymbol{X}, Z) = exp\{-\int_0^t [\tilde{\alpha}(s) + \tilde{\beta}(s)D + \boldsymbol{X}\tilde{\boldsymbol{\gamma}}(\boldsymbol{s})]ds\} \cdot E[exp\{-\int_0^t [\tilde{\beta}(s)\epsilon + \zeta(s)U]ds\}] \tag{A8}$$

Combining these two observations as stated in (A7) and (A8), we can write our additive hazard model from the survival function $S(t|\boldsymbol{X}, Z)$ as:

$$\begin{aligned}\tilde{h}(t|\boldsymbol{X}, Z) &= \tilde{\alpha}(t) + \tilde{\beta}(t)D + \boldsymbol{X}\tilde{\boldsymbol{\gamma}}(\boldsymbol{t}) - \frac{\partial logE[exp\{-\int_0^t [\tilde{\beta}(s)\epsilon + \zeta(s)U]ds\}]}{\partial t} \\ &= \tilde{\tilde{\alpha}}(t) + \tilde{\beta}(t)D + \boldsymbol{X}\tilde{\boldsymbol{\gamma}}(\boldsymbol{t}),\end{aligned} \tag{A9}$$

where $\tilde{\tilde{\alpha}}(t) = \tilde{\alpha}(t) - \frac{\partial logE[exp\{-\int_0^t [\tilde{\beta}(s)\epsilon + \zeta(s)U]ds\}]}{\partial t}$.

In summary, we have shown that by incorporating instrumental variables Z and following the similar-to-2SLS procedure, our estimation function in the second stage manages to remove the effect of unobserved variables and still maintain its coefficients of $\tilde{\beta}(s)$ and $\tilde{\boldsymbol{\gamma}}(\boldsymbol{t})$ unbiased, which is the same as in the complete hazard model of Equation (A2). ■