Hindawi
Journal of Advanced Transportation
Volume 2022, Article ID 9120129, 22 pages
https://doi.org/10.1155/2022/9120129

Research Article

A Bi-Level Optimization Model for Ride-Sourcing Platform's Spatial Pricing Strategy

Wei Tang[1,2,3] **Heng Wang**[4] **Yang Wang**[1] **Chuqiao Chen**[5] **and Xiqun (Michael) Chen**[3]

[1]*Platform of Transport Technology Thinktank (Research Institute of Highway, Ministry of Transport), Beijing 100088, China*
[2]*Center for Balance Architecture, Zhejiang University, Hangzhou 310007, China*
[3]*College of Civil Engineering and Architecture, Zhejiang University, Hangzhou 310007, China*
[4]*The Architectural Design & Research Institute of Zhejiang University Co., Ltd. (UAD), Hangzhou 310007, China*
[5]*School of Cyberspace Security, HangZhou Dianzi University, Hangzhou 310018, China*

Correspondence should be addressed to Xiqun (Michael) Chen; chenxiqun@zju.edu.cn

Received 20 December 2021; Revised 19 January 2022; Accepted 26 January 2022; Published 25 February 2022

Academic Editor: Yajie Zou

This article investigates a long-term optimal spatial pricing strategy for a ride-sourcing platform that serves a particular (possibly populated) area with profit-driven service providers (i.e., drivers) and time- and price-sensitive customers (i.e., passengers). By observing that oftentimes, the price strategy is anisotropic and spatial-dependent, both the supply and request are endogenous, and we build an analytical bi-level optimization mode. In the upper-level formulation, the ride-sourcing platform aims at setting up the spatially heterogeneous pricing strategy to maximize its total profit. However, in the lower level, we solve the trip distribution model that characterizes the flow rates among zones given the travel demand rate at each zone. We prove that when the platform seeks to expand its business, the optimal number of participating drivers and their optimal wages will be influenced not only by the pricing strategy but also by the level of service of the entire platform. Our further investigation shows that the profit at a particular zone can be influenced by the potential customers' service requests from other zones. Finally, we use the real-world data provided by DiDi Chuxing to numerically illustrate our model and theoretical results.

1. Introduction

The emergence of ride-sourcing companies such as Uber, Lyft, and Didi Chuxing has changed daily mobility modes of transportation in many cities. These transportation network companies provide internet-based on-demand ride services platforms that intelligently match ride requests for customers (i.e., passengers) at a particular location to the nearby affiliated service providers (i.e., drivers). Different from traditional street-hailing taxi, ride-sourcing services are fully capable of positioning idle vehicles to satisfy unmatched customers' requests, making it possible to effectively influence the behavior of participants (e.g., drivers and passengers) via economic means (e.g., pricing). The online matching mechanism of e-hailing platforms significantly decreases the searching friction between potential customers and drivers and adaptively adjusts the searching radius for a ride-hailing market. The on-demand ride services help satisfy travel demand of impatient customers. Consequently, a large number of cities where street hails were previously unsustainable are now experiencing more efficient ride services through ride-sourcing platforms [1].

Since customers' demand is nonstationary and could vary significantly from time to time; therefore, it is more beneficial for the on-demand ride-sourcing companies to hire independent and flexible drivers to satisfy customers' requests. However, compared with traditional taxi drivers, ride-sourcing drivers do not have fixed renting cost to pay (of course, there is the commission rate of each ride order charged by the platform); thus, their work participation is primarily driven by their earnings. In particular, since ride-sourcing drivers only get compensations by fulling the requests, their earnings are almost proportional to the utilization of their vehicles, which is highly related to the demand of customers [2]. On the other hand, the customers' demand is driven by the fare and their waiting time (we consider the passenger-driver matching time as a part of waiting time), and in turn, the waiting time is primarily dependent on the number of vacant vehicles. Therefore, the numbers of participating drivers and customers are endogenously dependent on both the wage rate and price rate decided by the on-demand ride services platform.

In spite of the interplay of driver and customers, the interaction between zones is also an important issue that needs to be carefully considered in the ride-hailing market. Thus, this article considers a spatially dependent wage rate. In fact, suppose that the wage rate stays the same across all the different zones. Then if the number of active drivers at a particular zone drops dramatically, we hope that the utilization of the rest of the active drivers in this zone will increase due to fewer participating drivers. However, the waiting time of customers will also increase, causing both a reduction in demand and a decrease in the utilization of the active drivers. Therefore, in order to maximize the profit, the on-demand ride services platform needs to have a fundamental understanding of supply (i.e., drivers) and demand (e.g., customers) to determine the optimal wage and price. Therefore, it must frequently coordinate the supply with the demand in different zones by (a) deciding the proper spatially heterogeneous wage rates based on the potential zonal supply (i.e., the total number of active drivers) and (b) setting up a proper spatially dependent pricing scheme based on potential demand (i.e., the total number of customers).

To address the complex relationship between spatial supply and demand, this article develops an analytical framework for an on-demand ride services platform to determine its price rate and wage rate for customers and drivers, respectively. In this framework, we make the following assumptions:

(a) We divide a large-scale transportation network into zones, while the exogenous parameters vary from zone to zone. Drivers can deliver customers from one zone to another, and then, they can be used in the zone at the end of their trip. Recall Bimpikis et al. [3]; we also assume that in a long term, the average number of drivers in a particular zone cannot exceed the number of registered drivers in the zone.

(b) We consider a long-term equilibrium service network, where each zone's inflow rate and outflow rate are balanced; that is to say, after a long-term process, the market clears up the imbalance of supply, so that each zone's total inflow rate is equal to its total outflow rate. As a result, the driver moves completely with the order without the need for additional empty dispatch or cruise. Importantly, the long-term demand characteristics (i.e., flow rate between different zones) are endogenously determined through a trip distribution model, which is also related to interzone distance and speed.

(c) The price rate and wage rate are zone-dependent in our analytical framework and are determined by the platform before requests are sent by customers.

(d) At each zone, since the customers' waiting time is related to both supply (i.e., the number of participating drivers) and demand (i.e., the customers' arrival rate), we propose a double queuing model to deductively replicate the matching process within each zone.

(e) Finally, we assume that each driver is not willing to participate unless her/his wage rate is larger than the reservation rate [4], that is, opportunity cost; similarly, customers will choose the ride-sourcing service only if their utilities are above a certain level.

With these aforementioned assumptions, we formulate a bi-level optimization model, where the upper level is to maximize the total profit of an on-demand ride service platform, and the lower level is trip distribution model. Our results show that at each zone, its optimal price rate is highly related to the service quality, which are summarized below.

We first consider the situation where an emerging on-demand ride service platform aims to expand its market penetration at the market entry stage and ensure its levels of service (i.e., the ratio of served requests to potential ones) for each zone to be homogeneous. That is a simple case when the levels of service are meant to be equal across all zones. Under this setting, we are able to characterize the optimal pricing rate for each zone.

We also conduct a series of sensitivity analyses on the optimal price rate, wage rate, and other implicit variables (i.e., the number of participating drivers, and customers' service requests). Our derived analytical results show that the increase in the driving speed at a given zone will reduce the number of its participating drivers and also shorten the average waiting time. The increase in the number of potential drivers at a given zone will result in an increase in the number of its participating drivers but a decrease in the average waiting time.

However, the increase in either the driving speed or number of potential drivers at a given zone will reduce its optimal wage rate but improve the optimal level of service. Consequently, the increased level of service will further increase the number of participating drivers and wage rates of other zones. It suggests that if the platform adopts the spatial price rate, they should set a lower wage rate when the supply increases at a particular zone and raise the wage rate in other zones to keep the similar level of service for the entire ride-hailing market.

On the other hand, the increase in the customers' value of waiting time at a zone will lead to a raise in its corresponding optimal wage rate but a reduction in the total profit and level of service. As a result, the reduction in the level of service reduces the number of participating drivers and the wage rate in other zones. It suggests that the platform should raise the wage rate to attract more drivers to mitigate the longer waiting time. The increase in travel distances will also increase the wage and price rates as well as the waiting time. Therefore, for instance, the platform should increase the wage rate if some arterial roads are under construction in a zone. Our numerical experiments using the real-world data provided by Didi Chuxing justify these theoretical results.

Next, we extend our model to a more flexible setting under which the levels of service may differ across different zones to maximize total profit. We prove that the optimal number of drivers is concave in the level of service. But the optimal level of service may not be unique. Besides, the local profit of zone A decreases in the level of service of zone B if the travel distance between them is larger than the average travel distance in zone A and the average travel time between them is shorter than the average travel time in zone A.

The remaining of the article is organized as follows. Section 2 presents a brief review of the related literature. In Section 3, we present our modeling framework, including the behavior modeling of drivers and customers, double queuing model, and trip distribution model. In Section 4, we develop the spatial pricing model for analyzing a situation in which the on-demand ride services platform expects to maximize its profit with an identical level of service from zone to zone. We analyze the sensitivity of the optimal price and wage rates. We further extend our model to a more general setting where the levels of service can be spatially heterogeneous. In Section 5, the numerical experiments based on real-world data from Didi Chuxing (the largest on-demand ride service platform in China) have been conducted to illustrate our theoretical derivations. Section 6 concludes the article. Appendix A provides notation, and Appendix B presents the mathematical proofs for all propositions.

2. Literature Review

The emerging ride-sourcing platforms (e.g., Uber, Lyft in the US, and DiDi in China) have a profound impact on urban mobility and transportation sectors by matching the supply of drivers with the travel demand in a real-time manner. These on-demand ride services platforms can be regarded as a special form of emerging sharing economy enterprises. In the literature, there are many studies investigating the operational strategies on the sharing economy from various industries, for example, resource, item, bicycle, car, and parking spot sharing. Wang et al. [5] proposed a two-phase optimization model in resource sharing pickup and delivery process. Wang et al. [6] analyzed the collaboration among service providers in a logistics network. Behrend and Meisel [7] analyzed a setting under which a platform operated both item-sharing and crowd-shipping. Çelebi et al. [8] presented a bicycle sharing system design, which incorporated both capacity allocation and location decisions. Lu et al. [9] built a two-stage stochastic integer programming model, which was to maximize the profit of the car sharing platform. Hara and Hato [10] proposed a Vickrey Clarke Groves mechanism for a car sharing system. They found that the negative price (or benefit) could encourage the relocation of vehicles under the auction mechanism. Xiao et al. [11] solved a shared parking problem by designing double-auction mechanisms under the parking spot allocation rule and the transaction payment rule. Instead of public parking spot sharing problem, Xu et al. [12] investigated the sharing problem for the private parking spots and showed the mechanism would significantly increase the social welfare and also the budget balance of the company. However, on-demand ride services are different from most of the other sharing services. In a ride-sharing system, the requests need to be served in time to prevent customers from opting in the other transportation modes. Other on-demand services usually reserve the service in advance. This feature results in a very different customer decision-making process and also a different timing of service requests. Next, we will present a review on the relevant literature of on-demand ride services.

The development of on-demand ride services platforms has motivated researchers to discuss various operational issues. Dong et al. [13] studied the dual sourcing problem of drivers in the ride-sourcing market, where some platforms recruit contractual drivers to cope with the uncertainty of labor supply from freelance/self-scheduling drivers. Ke et al. [14] investigated the regulatory outcomes of various representative government regulations, including price-cap regulation, vehicle fleet size control, wage regulation, income regulation, car utilization rate regulation, commission charge regulation, etc. Yu et al. [15] investigated three scenarios of Chinese government policies on on-demand ride services companies by introducing a two-period game theoretical model. By assuming that both the customers' valuation for the service and the providers' reservation wage rate followed a Bernoulli distribution, Taylor [2] discovered that agent independence would lower the price, while delay sensitivity increased the optimal price even if customers' valuation uncertainty was moderate. Bai et al. [4] expanded the work of Taylor [2] to a continuous distribution and presented a new mathematical model based on the M/M/k queuing theory. Their study suggested the on-demand ride services platform increased price and wage rates. Furthermore, when customers are more sensitive to the waiting time, the platform should increase the wage rate. Recently, Ma and Zhang [16] have studied the traffic flow rate patterns of a dynamic ride-sharing service in a single bottleneck corridor. Masoud and Jayakrishnan [17] investigated the ride-matching problem and solved it with the "Ellipsoid Spatiotemporal Accessibility Method." Qian et al. [18] examined two essential problems (i.e., optimal assignment and behaviors of participants) on the operation strategy and policy-making of the ride-sharing service by formulating the problem as an integer linear programming problem. Stiglic et al. [19] studied the benefits of introducing the meeting points in a ride-sharing platform. Their simulation results indicated that meeting points could significantly increase the number of matched participants.

The impact of dynamic pricing (also known as surge pricing) of on-demand ride services has drawn much attention recently. Some studies investigated this issue by analyzing historical data of a large-scale road network. For example, Chen and Sheldon [20] analyzed a subset of Uber data (about 25 million trips). Their study showed that in contrast to the income-target behavior, more drivers tended to participate when the surge pricing was high. Chen et al. [21] analyzed data from Uber App and found that surge pricing would motivate a new supply of drivers, but suppress the demand of passengers. However, its impact on existing supply (i.e., existing drivers) is still unknown.

Some studies focus on theoretical analyses on dynamic pricing. Hu and Zhou [22] assumed that the demand and supply were only determined by price and wage rates, respectively. They showed that the optimal price is a U-shape function of the exogenous wage. Riquelme et al. [23] characterized the waiting time using an M/M/k queuing model. Their study concluded that the performance of any dynamic pricing strategies could not exceed that of the optimal static pricing policy; however, dynamic pricing was more robust when system parameters changed. Cachon et al. [24] studied several pricing schemes and found that dynamic pricing usually achieved the near-optimal profit compared to the fixed price or fixed wage. Chen and Wang [25] studied the pricing problem for a last-mile transportation system with various customer types, which was numerically tested by the real data in Singapore. Lei et al. [26] presented a multiperiod model that solved the problems of dynamic pricing and vehicle dispatching in the on-demand ride-sharing system. They formulated a dynamic mathematical programming with equilibrium constraints and applied an approximate dynamic programming (ADP) algorithm to solve the overall problem. Xu et al. [27] formulated a double-ended queuing model to prove that the supply curve in a ridesharing system with limited matching radius was always backward bending, while a smaller matching radius caused a weaker bend.

This article aims at the spatial pricing problem for on-demand ride-sourcing service. There are several studies focusing on the spatial pricing in traditional modes of transportation, for example, Chao and Friesz [28] and Fisk [29]. To the best of our knowledge, there are only a few studies related to the topic. Chen et al. [30] proposed a reinforcement learning enhanced agent-based modeling and simulation framework to solve the spatial-temporal pricing problem for a ride-sourcing platform. In theoretical analysis, Bimpikis et al. [3] investigated how the spatial demand pattern affected the optimal spatial pricing, profits, and consumer surplus in a network. By assuming that (a) the arrivals of idle vehicles followed the spatial Poisson point process; and (b) spatial pricing was not allowed, Zha et al. [31] found that the on-demand platform might increase the price rate to avoid the inefficiency of supply. Furthermore, the platform might set a price rate higher than the level of price rate, which guaranteed market clearing. In spite of the insightful conclusion of this stream of studies, they assumed that the customer demand was irrelevant to the queuing time and supply was irrelevant to the system utilization. Different from all the previous works, our model based on the queuing theory captures the change of the total waiting time of customers and the overall system's utilization related to the ratio between supply and demand. Moreover, we describe the dynamic demand characteristics by using the Wilson entropy model [32, 33].

Finally, note that our work is closely related to Bai et al. [4], where the authors only considered a single-zone pricing problem. In their model, the waiting time was approximated by the M/M/k queuing model, and the earning rate was relevant to the supply and demand ratio. Differently, this article studies the spatial pricing problem, which is much more complicated and flexible than that studied in Bai et al. [4]. In addition, our work differs from Bai et al. [4] in the following aspects: (a) we use a more realistic double-ended queue to estimate the waiting time, which has a better characterization of the interaction between drivers and passengers; (b) we extend the original single-zone model into a spatial pricing model by introducing the trip distribution model; and (c) our model enables that every zone differs in exogenous parameters as well as decisions of customers and drivers.

3. Modeling Framework

This section presents the modeling framework about individual drivers' and customers' behavior, queuing model, and trip distribution model for the strategic analysis of the on-demand ride services platform. The requests can only be served by drivers in the same zone, and drivers and customers are both assumed to be heterogeneous within the zone. The queuing model captures the interaction between the number of participating drivers and the arrival rate of effective ride requests. The interaction between zones is captured by the trip distribution model. The model aims at maximizing the platform profit under different scenarios, while the demand and supply reach the spatial equilibrium in a representative ride-sourcing market.

In this section, we consider a network $\mathcal{G}(\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ denotes a set of nodes (i.e., regions) and $\mathcal{E}$ denotes a set of links. A node here means a zone rather than a specific location. For simplicity, we assume that $\mathcal{G}$ is a complete graph, which includes all the self-loops, that is,, $\mathcal{E} = \mathcal{V} \times \mathcal{V}$. For each link $(i, j) \in \mathcal{E}$, let L_{ij} denote the average unit of the service distance from region i to region j, V_{ij} denotes the average service speed from region i to region j, and q_{ij}^* denotes the customers' demand rate (i.e., flow rate) from region i to region j. We also suppose that the flow rate variables $\mathbf{q}^*$ will be determined by the trip distribution model.

The rest of notation will be introduced in the subsequent subsections. For ease of reference, Appendix A lists the basic notation used in this section.

3.1. Demand Side: Utility of Passengers.

In this section, we present a customer behavior model in which the customer is sensitive to the price and waiting time. The model integrates the price and waiting time with the service rate. Considering a long-term equilibrium service network, inflow rate and outflow rate of zone $i \in \mathcal{V}$ remain equivalent.

Arbitrary zone i bears a maximum potential customer demand rate $\overline{\lambda}_i, i \in \mathcal{V}$ (e.g., the number of customers' ride requests per minute). It is assumed the platform sets price rate $p_i, i \in \mathcal{V}$ per service unit (e.g., dollar per kilometer) in zone i. The average travel distance per trip $d_i(q^*)$ is related to the travel flow rate distribution among zones, which is formulated as

$$d_i(q^*) = \frac{\sum_{j\in\mathcal{V}} L_{ij} q_{ij}^*}{\sum_{j\in\mathcal{V}} q_{ij}^*}, \quad \forall i \in \mathcal{V}, \tag{1}$$

where q_{ij} is the flow rate from zone i to j, and L_{ij} is the distance between zone i and j. To model heterogeneous customers, we assume that the service valuation per service unit, that is, v_i, varies among customers. According to the notation in the appendix, the expected customer surplus without waiting equals to $(v_i - p_i)d_i(q^*)$. To capture the waiting time sensitivity, we define the expected utility function of a customer in zone i as

$$U_i(v_i) = (v_i - p_i)d_i(q^*) - c_i W_i, \tag{2}$$

where c_i denotes the cost of waiting per unit time, W_i represents the expected waiting time, and $U_i(v_i)$ represents the surplus or utility function for customers in zone i. We assume that customers will choose other transportation modes if their demands are not satisfied by ride-sharing services, which means that customers with utility function $U_i(v_i)$ lower than 0 will choose other transportation modes, and with $U_i(v_i)$ higher than 0 will choose ride-sharing service. Thus, the effective demand rate (i.e., the realized customer request rate) λ_i is determined by the distribution of $U_i(v_i)$. It assumes that a customer will choose the service only if their surplus (utility function) satisfies $U_i(v_i) \geq 0$, that is,

$$\lambda_i = \text{Prob}\{U_i(v_i) \geq 0\}\overline{\lambda}_i = \text{Prob}\left\{v_i \geq p_i + \frac{c_i W_i}{d_i(q^*)}\right\}\overline{\lambda}_i. \tag{3}$$

Let us define the proportion of served requests to the total number of requests as the level of service, that is, $s_i = \text{Prob}\{v_i \geq p_i + c_i W_i / d_i(\mathbf{q}^*)\}$. Then, the effective customer request rate λ_i is given by

$$\lambda_i = s_i \overline{\lambda}_i. \tag{4}$$

According to Bai et al. [4], we assume the the service valuation per service unit v is uniformly distributed from 0 to 1. From equation (3), we can derive that price rate p_i satisfies

$$p_i = 1 - s_i - \frac{c_i W_i}{d_i(q^*)}. \tag{5}$$

3.2. Supply Side: Ride-Sourcing Drivers. Let K_i be the number of potential drivers who may provide ride services in zone i. Generally, K_i represents the number of registered providers on the platform. Assume that the platform sets the wage rate in zone i as w_i (e.g., dollars per hour). Given zone price p_i, denote k_i as the corresponding number of effective service providers who are willing to participate the platform, in which $k_i \leq K_i$. Denote $\mu_i(q^*)$ as the average service speed (e.g., km/h) in zone i, which is related to the travel flow rate distribution among zones. It can be formulated by

$$\mu_i(q^*) = \frac{\sum_{j\in\mathcal{V}} V_{ij} q_{ij}^*}{\sum_{j\in\mathcal{V}} q_{ij}^*}, \quad \forall i \in \mathcal{V}. \tag{6}$$

Thus, $\mu_i(q^*)/d_i(q^*)$ represents the expected service rate of drivers in zone i (i.e., the number of requests served per minute). Given the effective customer request rate λ_i and the number of participating drivers k_i mentioned above, the utilization of these k_i participating providers equals to $\lambda_i d_i(q^*)/k_i\mu_i(q^*)$. We know that k_i must follow the constraint $k_i\mu_i(q^*)/d_i(q^*) > \lambda_i$ as the realized request rate cannot exceed the service rate. The expected wage per unit time for a participating driver equals to the wage per service unit w_i multiplied by the average service speed $\mu_i(q^*)$. Hence, the wage per unit time of a participating provider equals to

$$w_i\mu_i(q^*)\frac{\lambda_i d_i(q^*)}{k_i\mu_i(q^*)} = \frac{w_i\lambda_i d_i(q^*)}{k_i}. \tag{7}$$

To model drivers' heterogeneity, drivers are assumed to have a random reservation earning rate r per unit time, corresponding to their outside options. We denote the cumulative distribution function of reservation rate r as $G(\cdot)$. A registered driver with reservation rate r will only offer service if the average participating earning rate $w_i\lambda_i d_i(q^*)/k_i$ is larger than or at least equal to r. Let β_i denote the ratio of providers who are willing to offer ride service with the wage per unit time $w_i\lambda_i d_i(q^*)/k_i$. Thus, $\beta_i = \text{Prob}\{r \leq w_i\lambda_i d_i(q^*)/k_i\} = G(w_i\lambda_i d_i(q^*)/k_i)$, the corresponding number of participating drivers k_i can be calculated by

$$k_i = \beta_i K_i \tag{8}$$

According to Bai et al. [4], we assume the reservation rate r is uniformly distributed from 0 to 1. Thus, we have

$$\beta_i = \frac{w_i\lambda_i d_i(q^*)}{k_i}. \tag{9}$$

Combining equations (8) and (9), we can express the wage rate w_i as a function of the number of participating providers k_i as follows:

$$w_i = \beta_i\frac{k_i}{\lambda_i d_i(q^*)} = \frac{k_i^2}{K_i\lambda_i d_i(q^*)}. \tag{10}$$

3.3. Waiting Time Estimation: A Double Queuing Model. In practice, ride-sourcing platforms, such as Didi Chuxing, segment space into bordered matching blocks. Within each block, the platform organizes requests into a virtual queue and then matches them sequentially with idle drivers arriving in the block. Inspired by Matsushima and Kobayashi [34], Shi and Lian [35, 36], and Xu et al. [27], we formulate the waiting mechanism between customers and drivers as a double queuing model, in which we assume that ride requests arrive following a Poisson process with rate λ_i (i.e., requests per unit time). Each request may consist of one to four passengers, who can be taken by a vehicle. Vacant vehicles (drivers) also arrive following a Poisson process with rate $k_i\mu_i(q^*)/d_i(q^*)$. The online matching time between drivers and ride requests can be negligible compared with the waiting time. Thus, we assume that the matching is instantaneous. The requests and drivers are arranged based on a first-come-first-served discipline.

With the assumption that the matching process between a driver and a ride request is instantaneous, at any time, there are either ride requests or drivers waiting to be matched in the queue. We define a Markov process $\{N_i(t); t > 0\}$ to model the system, where $\{N_i(t) > 0\}$ represents the number of ride requests waiting to be matched in the queue at time t, $\{N_i(t) > 0\}$ represents the number of drivers waiting at time t, and $N_i(t) = 0$ represents that all drivers have been matched with all requests at time t.

Denote the state space $\mathcal{S}_i = \{-k_i, -k_i + 1, \ldots, -1, 0, 1, \ldots, +\infty\}$. Therefore, $\{N_i(t) > 0\}$ is a one-dimensional birth-and-death process with state space $\mathcal{S}_i$. Figure 1 illustrates the waiting process where row represents the waiting process in a zone and each circle represents a certain state.

According to Little's law [37], the average queuing length L_i, L_i^v, and expected waiting time W_i, W_i^v of passengers and drivers are as follows:

$$L_i = \frac{k_i\mu_i(q^*)\left(\lambda_i d_i(q^*)/k_i\mu_i(q^*)\right)^{k_i+1}}{k_i\mu_i(q^*) - \lambda_i d_i(q^*)}. \tag{11a}$$

$$W_i = \frac{d_i(q^*)\left(\lambda_i d_i(q^*)/k_i\mu_i(q^*)\right)^{k_i}}{k_i\mu_i(q^*) - \lambda_i d_i(q^*)}. \tag{11b}$$

$$\begin{aligned} L_i^v &= k_i + \lambda_i d_i(q^*)\left(\frac{\lambda_i d_i(q^*)}{k_i\mu_i(q^*)}\right)^{k_i} \\ &\quad \cdot \left(\frac{1 - \left(k_i\mu_i(q^*)/\lambda_i d_i(q^*)\right)^{k_i}}{k_i\mu_i(q^*) - \lambda_i d_i(q^*)}\right). \end{aligned} \tag{11c}$$

$$\begin{aligned} W_i^v &= \frac{d_i(q^*)}{\mu_i(q^*)} + d_i(q^*)\left(\frac{\lambda_i d_i(q^*)}{k_i\mu_i(q^*)}\right)^{k_i} \\ &\quad \cdot \left(\frac{1 - \left(k_i\mu_i(q^*)/\lambda_i d_i(q^*)\right)^{k_i}}{k_i\mu_i(q^*) - \lambda_i d_i(q^*)}\right). \end{aligned} \tag{11d}$$

3.4. Bi-Level Optimization Model. The flow rate among different zones affects the spatial price and wage set by the platform, which in turn influences the travel demand at each zone, as well as the flow rates among different zones. To model the systematic dynamics, we formulate a bi-level optimization model, where in the upper level, the platform aims at maximizing its profit given the flow rates, and the lower level is the trip distribution model that characterizes the optimal flow rates among zones given the travel demand rate at each zone.

We present our bi-level optimization model as follows:

$$\begin{aligned} \max_{k,s} \quad \pi(k,s) &= \sum_{i\in\mathcal{V}} (p_i - w_i)\lambda_i d_i(q^*) \\ &= \sum_{i\in\mathcal{V}} \Big[\overline{\lambda}_i d_i(q^*) s_i(1 - s_i) - c_i\overline{\lambda}_i s_i d_i(q^*) \\ &\quad \cdot \frac{\left(\overline{\lambda}_i s_i d_i(q^*)/\mu_i(q^*)k_i\right)^{k_i}}{\mu_i(q^*)k_i - \overline{\lambda}_i s_i d_i(q^*)} - \frac{k_i^2}{K_i}\Big]. \end{aligned} \tag{12a}$$

$$\text{s.t.} \quad K_i \geq k_i \geq \frac{s_i\overline{\lambda}_i d_i(\mathbf{q}^*)}{\mu_i(\mathbf{q}^*)}, \quad \forall i \in \mathcal{V}, \tag{12b}$$

$$1 \geq s_i \geq 0, \quad \forall i \in \mathcal{V}, \tag{12c}$$

where the trip flow rate matrix $\mathbf{q}^*$ equals the optimal solution to the lower-level trip distribution optimization model, formulated by

$$\mathbf{q}^* \in \arg\min_{\mathbf{q}} \sum_{i\in\mathcal{V}}\sum_{j\in\mathcal{V}} \left(q_{ij}\ln q_{ij} - q_{ij}\right). \tag{13a}$$

$$\text{s.t.} \quad s_i\overline{\lambda}_i = \sum_{j\in\mathcal{V}} q_{ij}, \quad \forall i \in \mathcal{V}, \tag{13b}$$

$$\sum_{m\in\mathcal{V}} q_{jm} = \sum_{i\in\mathcal{V}} q_{ij}, \quad \forall j \in \mathcal{V}, \tag{13c}$$

$$q_{ij} \geq 0, \quad \forall i, j \in \mathcal{V}. \tag{13d}$$

For the upper-level model, the objective of the platform in equation (12a) is to maximize the overall profit, where at zone $i \in \mathcal{V}$, $p_i - w_i$ represents the net profit (i.e., the difference between price and wage) with p_i defined in equation (5) and w_i is defined in equation (10), and $\lambda_i d_i(\mathbf{q}^*)$ represents the total travel distance per unit time with λ_i defined in equation (4) and $d_i(\mathbf{q}^*)$ defined in equation (1). The constraints in equation (12b) indicate that (a) the number of participating drivers cannot exceed the number of registered drivers; and (b) the service rate must be no smaller than the realized request rate. The constraints in equation (12c) ensure that the service rate must be between 0 and 1.

For the convenience of derivation, we formulate the lower-level model () according to the Wilson entropy model [32], which has an explicit form of the optimal solution. In the lower-level model, flow rate q_{ij} from zone i to zone j should satisfy constraints in equations (13b) and (13c), which show that (a) the outflow rate of zone i equals to the summation of all the flow rates leaving from it; and (b) the summation of all flow rate out from zone j equals to the summation of all flow rates into j. We use the second constraint since we only consider long-term optimal optimization rather than short-term; thus, the outflow rate of a zone is equal to its inflow rate.

The following proposition shows that we can obtain the closed-form solution of the lower-level model ().

Proposition 1. *Let $\mathbf{q}^*$ be the optimal solution to lower-level model in Eq.() for a given service vector $\mathbf{s}$. Then,*

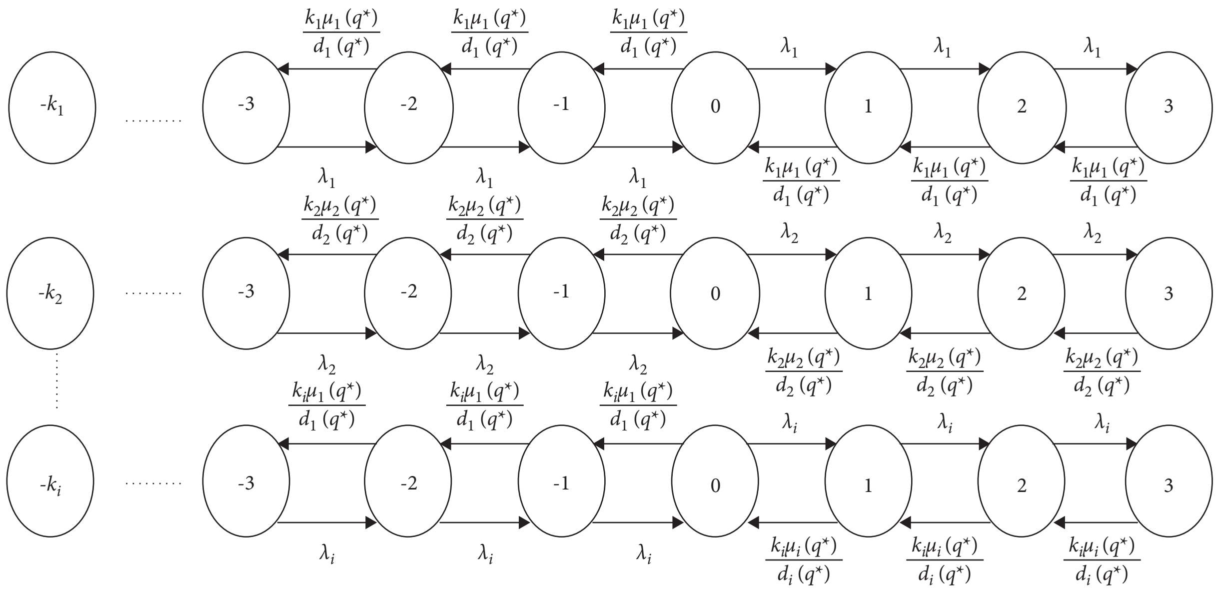


FIGURE 1: The state transition process of a double queuing model.

$$q_{ij}^* = \frac{s_i s_j \bar{\lambda}_i \bar{\lambda}_j}{\sum_{\tau\in\mathcal{V}} s_\tau \bar{\lambda}_\tau}, \quad \forall i, j \in \mathcal{V}. \tag{14}$$

By subscribe flow rate matrix $\mathbf{q}^*$ in (14) into the upper-level model (12a)–(12c), we can simplify it as follows:

$$\max_{\mathbf{k},\mathbf{s}} \quad \pi(\mathbf{k},\mathbf{s}) = \sum_{i\in\mathcal{V}} \left[\frac{\sum_{j\in\mathcal{V}} L_{ij}\bar{\lambda}_j s_j}{\sum_{j\in\mathcal{V}} \bar{\lambda}_j s_j} \bar{\lambda}_i s_i (1-s_i) - c_i \frac{\left(\bar{\lambda}_i s_i / k_i \sum_{j\in\mathcal{V}} L_{ij}\bar{\lambda}_j s_j / \sum_{j\in\mathcal{V}} V_{ij}\bar{\lambda}_j s_j\right)^{k_i+1}}{1 - \bar{\lambda}_i s_i / k_i \sum_{j\in\mathcal{V}} L_{ij}\bar{\lambda}_j s_j / \sum_{j\in\mathcal{V}} V_{ij}\bar{\lambda}_j s_j} - \frac{k_i^2}{K_i} \right], \tag{15a}$$

$$\text{s.t.} \quad \bar{\lambda}_i s_i \frac{\sum_{j\in\mathcal{V}} L_{ij}\bar{\lambda}_j s_j}{\sum_{j\in\mathcal{V}} V_{ij}\bar{\lambda}_j s_j} \le k_i \le K_i, \quad \forall i \in \mathcal{V}. \tag{15b}$$

$$0 \le s_i \le 1, \quad \forall i \in \mathcal{V}. \tag{15c}$$

4. Analytical Results

We analyze the bi-level optimization model under two special cases. In the first case, we consider the situation where an emerging on-demand ride service platform aims to expand its market penetration at the market entry stage and ensure its level of service for each zone to be homogeneous. In the other case, we study that the platform has been at a stable stage or has dominated the ride-sourcing market and targets at the highest profit regardless of heterogeneous levels of service. Throughout this article, we use superscript $*$ to denote the optimal solutions and optimal profit.

4.1. Case 1: Spatial Homogeneity—Identical Level of Service. In this subsection, we suppose that the on-demand ride services platform is at the entry stage of a competitive ride-sourcing market environment and is devoted to expanding its penetration by maintaining an identical level of service across all zones. More formally, we make the following assumption.

Assumption 1. In upper-level models (15a)–(15c), we assume that $s_i = s$ for all $i \in \mathcal{V}$; that is, the platform would like to choose a homogeneous level of service in Case 1.

Under Assumption 1, upper-level models (15a)–(15c) become

$$\max_{\mathbf{k},s} \quad \pi(\mathbf{k},s) = \sum_{i\in\mathcal{V}}\left[\frac{\sum_{j\in\mathcal{V}} L_{ij}\overline{\lambda}_j}{\sum_{j\in\mathcal{V}}\overline{\lambda}_j}\overline{\lambda}_i s(1-s) - c_i\frac{\left(\overline{\lambda}_i s/k_i\sum_{j\in\mathcal{V}} L_{ij}\overline{\lambda}_j/\sum_{j\in\mathcal{V}} V_{ij}\overline{\lambda}_j\right)^{k_i+1}}{1-\overline{\lambda}_i s/k_i\sum_{j\in\mathcal{V}} L_{ij}\overline{\lambda}_j/\sum_{j\in\mathcal{V}} V_{ij}\overline{\lambda}_j} - \frac{k_i^2}{K_i}\right], \tag{16a}$$

$$\text{s.t.} \quad \overline{\lambda}_i s\frac{\sum_{j\in v} L_{ij}\overline{\lambda}_j}{\sum_{j\in v} V_{ij}\overline{\lambda}_j} \le k_i \le K_i, \quad \forall i\in v \tag{16b}$$

$$0\le s_i\le 1, \quad \forall i\in\mathcal{V}. \tag{16c}$$

We can derive the following mathematical properties of models (16a)–(16c).

Proposition 2. *Under Assumption 1, models (16a)–(16c) exhibit the following properties:*

(1) *The objective function $\pi(\mathbf{k},s)$ is biconcave in $\mathbf{k}$ and s; that is, it is concave in s for any fixed $\mathbf{k}$ and concave in $\mathbf{k}$ for any fixed s;*

(2) *Let $(\mathbf{k}^*, s^*)$ denote the optimal solution of model (). Then, we have $0 < s^* < 1$ and $s\overline{\lambda}_i\sum_{j\in\mathcal{V}} L_{ij}\overline{\lambda}_j/\sum_{j\in\mathcal{V}} V_{ij}\overline{\lambda}_j < k_i^* \le K_i$, $\forall i\in\mathcal{V}$.*

Proposition 3. *Under Assumption 1, let the payout ratio, that is, the proportion of wage to price that the platform sets, be flexible among different zones. Denote ρ_i^* as the optimal proportion of the realized demand rate to service rate $s\overline{\lambda}_i\sum_{j\in\mathcal{V}} L_{ij}\overline{\lambda}_j/\sum_{j\in\mathcal{V}} V_{ij}\overline{\lambda}_j < k_i^* \le K_i$ at zone i; then,*

(1) *When K_i increases, k_i^*, s^*, and π^* increase; if k_i^* is a boundary solution (i.e., $k_i^* = K_i$), then w_i^* increases in K_i; if $-1/\rho_n^* + 2 + \log(1/\rho_n^*) \ge 0\,(1\ge\rho_n^*\ge 0.3178)$, $\forall n\in\mathcal{V}$, then both k_n^* and w_n^* increase;*

(2) *When V_{ij} increases, both k_i^* and w_i^* decrease, and both s^* and π^* increase; if $-1/\rho_n^* + 2 + \log(1/\rho_n^*) \ge 0\,(1\ge\rho_n^*\ge 0.3178)$, $\forall n\in\mathcal{V}, n\ne i$, then both k_n^* and w_n^* increase;*

(3) *When c_i increases, both k_i^* and w_i^* increase, and both s^* and π^* decrease; if $-1/\rho_n^* + 2 + \log(1/\rho_n^*) \ge 0\,(1\ge\rho_n^*\ge 0.3178)$, $\forall n\in\mathcal{V}, n\ne i$, then both k_n^* and w_n^* decrease;*

(4) *When L_{ij} increases, k_i^* increases and s^* decreases; if $-1/\rho_n^* + 2 + \log(1/\rho_n^*) \ge 0\,(1\ge\rho_n^*\ge 0.3178)$, then k_n^*, w_n^*, $\forall n\in\mathcal{V}, n\ne i$ decrease.*

All proofs are relegated to the appendix. Based on the analytical results of Proposition 3, we summarize these monotonicity properties in Table 1. The main insights for the spatial model in the scenario of spatial homogeneity under Assumption 1 are as follows.

First, the platform should reduce the local wage rate w_i^* as the number of available service providers K_i or service speed V_{ij} increases. The reason is that with an increase in registered drivers or travel speed, the service rate could be raised even if the wage rate stays constant. Thus, the platform could relatively reduce wage rate w_i to achieve greater profit without reducing the service rate. It indicates that both a greater number of registered drivers and better road infrastructure help the platform achieve their higher profit. However, the optimal price p_i^* is not monotonic in K_i. The behavior occurs because the expected waiting time increases convexly in the system utilization.

When the number of potentially participating drivers K_i or travel speed V_{ij} is not large or fast enough that drivers serve with high utilization, the bottleneck for the profit growth is mainly on the supply side. The platform operates in high utilization, and waiting time W_i accounts for a large proportion in the custom utility function. In this case, an increase in supply from a higher K_i can significantly increase the number of operating drivers k_i and reduce waiting time W_i. Thus, the platform can afford to rise the optimal price p_i^* and decrease the optimal wage rate w_i^*, while still maintaining a higher realized customer request rate λ_i and achieving a higher profit π^*. On the other hand, when the number of registered driver K_i is too large that drivers serve with low utilization, the bottleneck is mainly on the demand side. The system operates in low utilization. In this case, an increase in K_i would slightly reduce waiting time W_i since the waiting time is already short. As W_i only accounts for a slight part of the customer utility function, a decrease in the optimal wage rate w_i^* will not influence much of the waiting time, while a decrease in price p_i^* would significantly increase the real customer demand rate λ_i. The platform should reduce the optimal price p_i^* to incentivize a higher λ_i and reduce w_i^* to achieve a higher profit π^*. Overall, we explain that the queuing effect results in the nonmonotony of the optimal price p_i^*. The nonmonotonic property of p_i^* is obviously due to the nonlinear effect of utilization on waiting time. If the effect of waiting time cost on customer demand is not captured (i.e., $c_i = 0$), apparently both p_i^* and w_i^* decrease in K_i.

The increase in the overall level of service s^* will affect the parameters in all zones. In order to keep the identical level of service, the increase/decrease in s^* needs the increase/decrease in the driver number in the rest of zones. Thus, $k_n^*, \forall n\in\mathcal{V}$, always changes in the same direction as s^*. Similarly, in order to raise/reduce the driver number, the

Table 1: Summary of sensitivity analysis in Proposition 3.

Increasing parameter	Wage w_i^*	Drivers k_i^*	Service level s_i^*	Drivers, wage k_n^*, w_n^*	Profit π^*
Potential providers K_i	Not monotone	↓	↑	↑	↑
Service speed V_{ij}	Not monotone	↓	↓	↑	↑
Service distance L_{ij}	Not monotone	↓	↑	↓	↑
Waiting cost c_i	Not monotone	↑	↑	↓	↓

Here, ↓ = decrease and ↑ = increase.

wage rate w_n^*, $\forall n \in \mathcal{V}$ should be set higher/lower, which explains the result that w_n^* always changes in the same direction as s^*, too.

Second, we find that when the waiting cost c_i increases, the platform should offer a higher wage rate w_i^*. This strategy attracts more drivers k_i^* to join the platform and reduces the optimal profit of the platform π^*. An increase in c_i would increase the waiting time cost if the demand rate is constant. Consequently, the platform should increase the optimal wage w_i^* to reduce the demand rate. Therefore, it helps achieve a lower utilization and reduce the waiting time.

Finally, the increase in L_{ij} will cause an increase in the optimal number of drivers k_i^* and decrease in the level of service s^*. The increase in L_{ij} increases the waiting time, if k_i and s stay constant. Thus, the optimal number of drivers k_i^* should be raised to save the waiting time, but s^* would decrease.

4.2. Spatial Heterogeneity: Flexible Level of Service

Assumption 2. The on-demand ride services platform stays on the stable stage in a monopoly ride-sourcing market environment and aims at maximizing its profit by allowing spatially heterogeneous or flexible levels of service $\mathbf{s}$ among all zones.

Under this assumption, the optimal levels of services s^* might not be unique, but the optimal $\mathbf{k}^*$ is unique for any fixed s^*.

Proposition 4. *Under Assumption 2, the objective function π is concave in $\mathbf{k}$ for any fixed $\mathbf{s}$, but is not necessarily concave in $\mathbf{s}$ for any fixed $\mathbf{k}$. Given $\mathbf{s}^*$, the optimal solution $\mathbf{k}^*$ uniquely exists, and k_i^*, $\forall i \in \mathcal{V}$, is either the solution of $\partial\pi/\partial k_i = 0$ or equals K_i. The sensitivity analyses in terms of k_i in Proposition 3 still hold.*

Proposition 5. *Under Assumption 2, let the payout ratio, that is, the proportion of wage to price that the platform sets, be flexible among different zones. Given any $\mathbf{s}$, then*

(1) The profit π_i decreases in s_i, if $s_i > 1/2$, $\forall i \in \mathcal{V}$

(2) The profit π_j decreases in s_i, if $L_{ji} > d_j$ and $L_{ji}\mu_j < V_{ji}d_j$, and increases in s_i, if $L_{ji} < d_j$ and $L_{ji}\mu_j > V_{ji}d_j$

Based on the analytical results of Propositions 4 and 5, we summarize the main insights for the spatial model in the scenario of spatial heterogeneity under Assumption 2 as follows:

First, the partial derivative of $\mathbf{k}$ is consistent with Proposition 3 under Assumption 1, and the objective function is still concave in $\mathbf{k}$ for any fixed $\mathbf{s}$. However, the model is no longer concave in $\mathbf{s}$ for any fixed $\mathbf{k}$, which complicates the analytical results. Thus, it is difficult to reach more insightful analytical results under Assumption 2. To numerically solve this problem, we use alternating projection to reach a near-optimal solution, which satisfies the well-known KKT conditions.

Second, under the assumption that the valuation per service unit v is uniformly distributed following $[0, 1]$, we suggest the platform not pursuit a level of service higher than 1/2. It might be intuitive for us that a higher level of service usually results in the higher profit in the multi-enterprise environment. But in a monopoly ride-sourcing market, a high level of service requires the low price as well as short waiting time, which requires the high wage rate to attract an enough number of drivers. Thus, a high level of service might not help the platform gain the maximum profit. In some circumstances, the platform keeps the service under a not very high level to maximize its profit.

Finally, travel speed as well as travel distance is dependent on the travel flow rate distribution among zones; that is, a zone's state would be affected by the states of other zones. We can interpret Proposition 5 as follows. Once a zone's demand rate increases, the proportion of its flow rate to the overall flow rate is enlarged. Thus, a zone's average distance d_i tends to be enlarged/lessen if the distance between two zones is greater/less than the original average distance d_i. Thus, the increase of s_i would increase the average local travel distance d_i. Analogously, if travel time L_{ji}/V_{ji} between the two zones j and i is less than the original average travel time, then the increase of s_i would reduce local travel time d_i/μ_i. For instance, there might be an expressway connecting two regions, which are far from each other. Overall, if the increase in the local request arrival rate in zone i helps increase distance d_j and reduce travel time d_i/μ_i in zone j, the local profit π_j will be enlarged. On the contrary, if the increase in the local request arrival rate in zone i decreases distance d_j and increases travel time d_i/μ_i in zone j, the optimal local profit π_j will decrease.

5. Numerical Illustrations

5.1. Parameters. In this section, we present some numerical experiments with parameters calibrated by using the DiDi Chuxing data in Hangzhou, China. In the experiments, we divide Hangzhou into 7 × 7 zones. Each zone is a square with

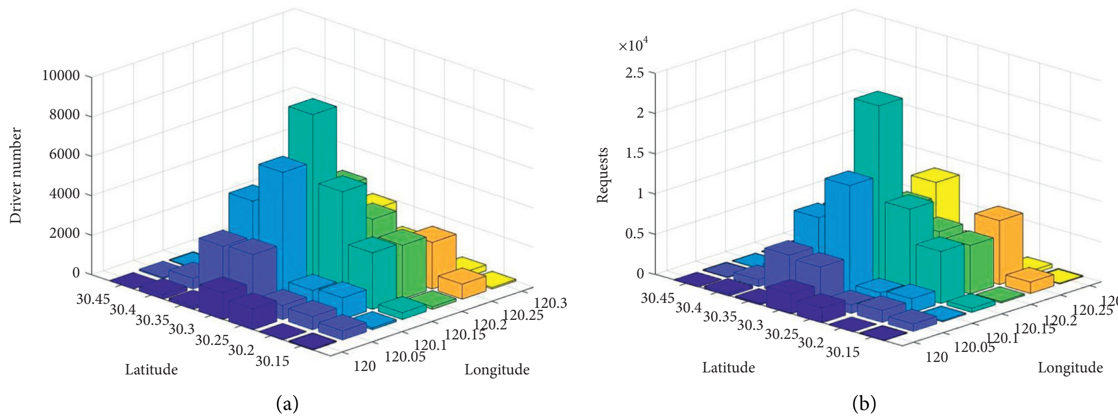


Figure 2: The spatial distribution of drivers and passengers' ride requests in zones. (a) The number of registered drivers (K_i) in Hangzhou. (b) The number of requests in Hangzhou.

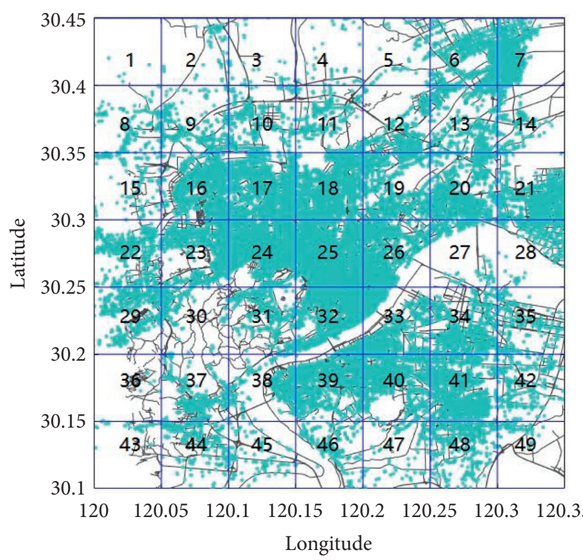


Figure 3: Spatial distribution of on-demand ride requests in Hangzhou, China.

the side length of about 5 km, which is approximately the margin of a driver's service radius. We analyze the one-week sampling order data between March 6, 2017 and March 12, 2017 from DiDi Chuxing. The average demand and number of available drivers per day are shown in Figures 2 and 3, respectively. As can be seen, both the numbers of drivers and ride requests are much more active in the urban center than suburbs, and the ride request calling even drops to 38 times per day in the margin zones. It is worthy to note that it is the sampling dataset that we analyze, and the sample data consist of 50% of total requests, so the real requests exceed the values shown in the figures, but the spatial distribution pattern maintains the same.

We examine the average income for citizens in Hangzhou in 2018. The majority of citizens earn 2,000 to 8,000 CNY (US $100 is approximate CNY 670) per month, including 18.3% earn 2,000 CNY to 3,000 CNY, 17.4% earn 3,000 CNY to 4,500 CNY, 18.6% earn 4,500 CNY to 6,000 CNY, 15.9% earn 6,000 CNY to 8,000 CNY, and 8.7% earn 8000 CNY to 10,000 CNY. Considering that the drivers need bear expenses of car insurance, fuel cost, and car maintenance, we estimate that a hourly wage rate of 35 CNY is required for a ride-sourcing driver to provide service. Thus, the hourly wage reservation r is assumed to be distributed uniformly between 35 CNY and 45 CNY.

Table 2: Summary of parameters based on empirical data statistics.

Parameter	Max	Min	Mean	Source
K_i	622	2	106	Data
λ_i	$930\,h^{-1}$	$1.58\,h^{-1}$	$134\,h^{-1}$	Data
V_{ij}	68.29 km/h	9.66 km/h	34.38 km/h	Data
L_{ij}	64.20 km	1.70 km	22.07 km	Data
c_i	—	—	200 CNY/h	Assumption
v	4 CNY/km	1 CNY/km	2.5 CNY/km	Assumption
r	45 CNY/h	35 CNY/h	40 CNY/h	Assumption

We also investigate the price that DiDi charges from the customers in 2017. Although the dynamic pricing strategy

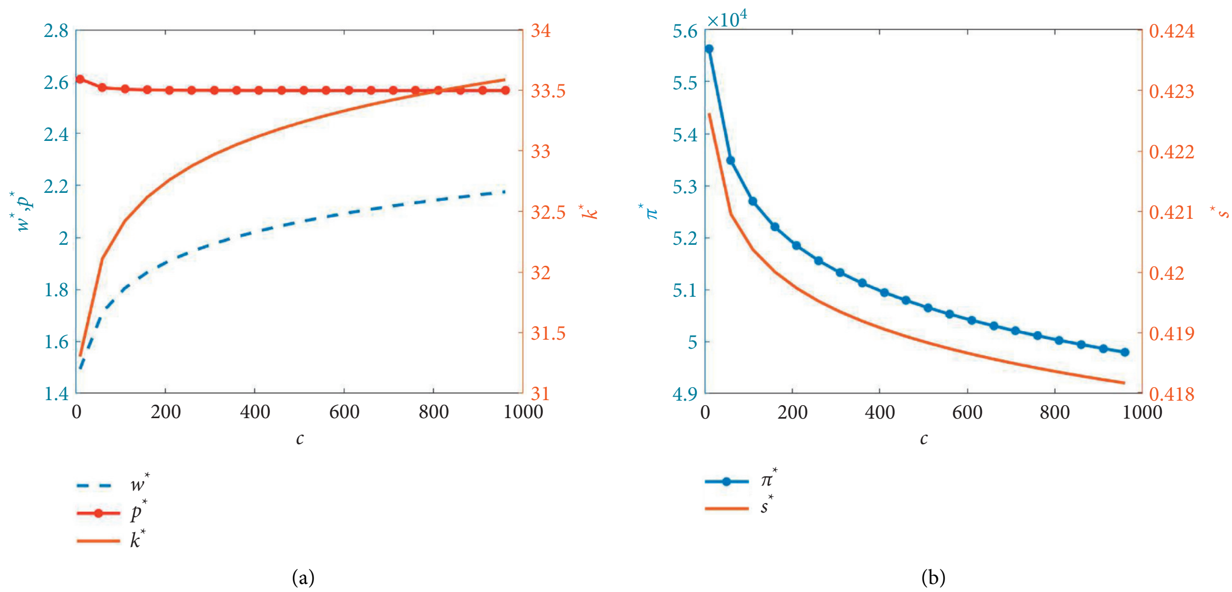


FIGURE 4: Sensitivity analysis of endogenous variables in terms of **c**. (a) p^*, w^*, and k^*. (b) s^*.

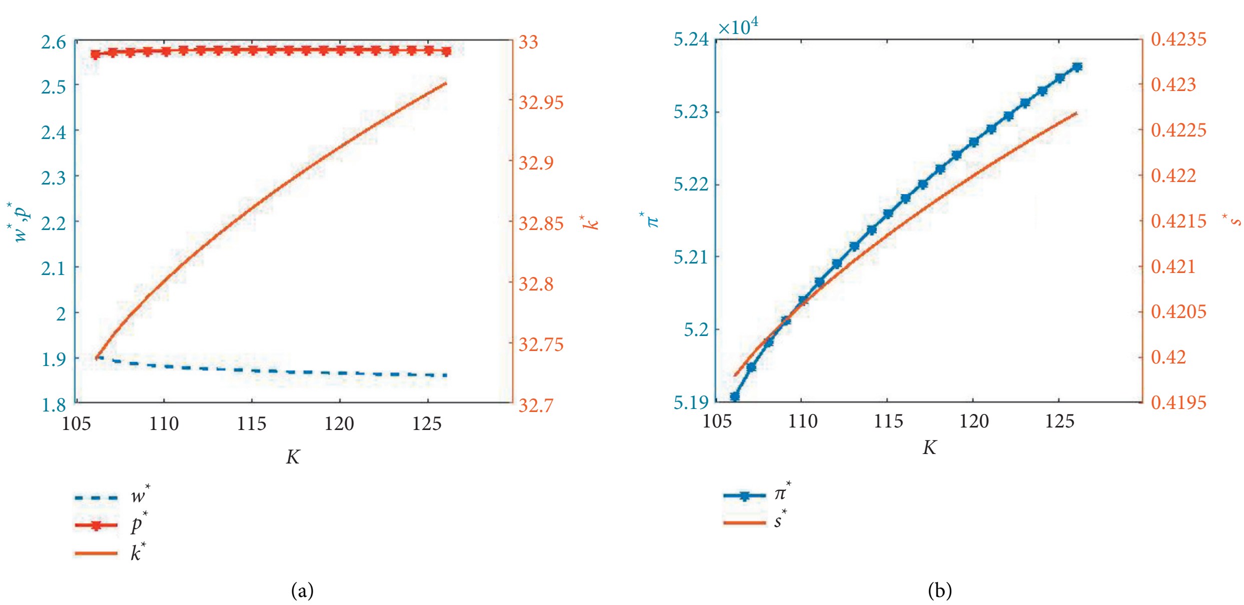
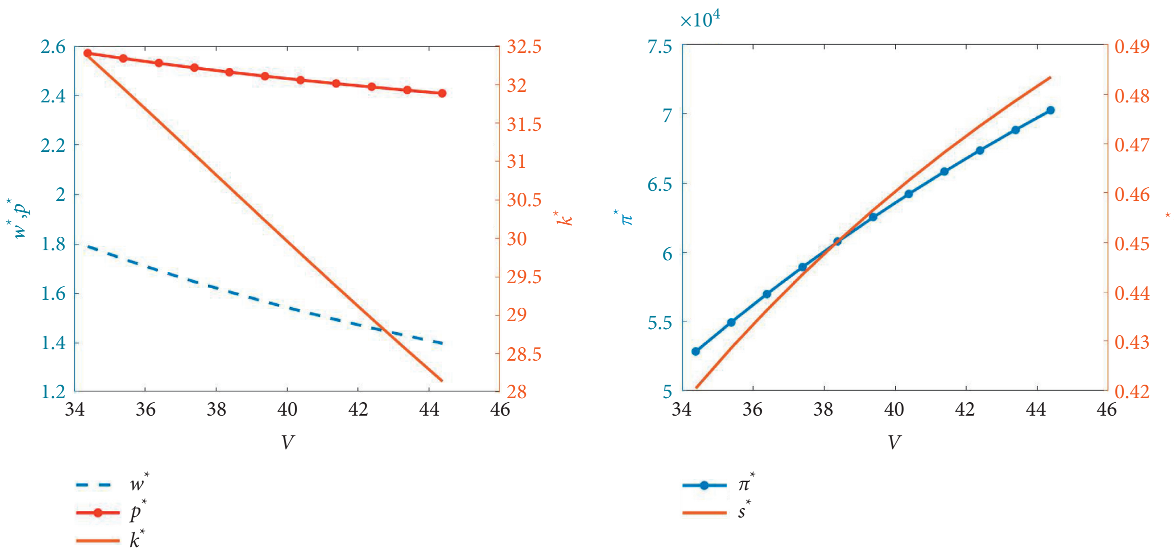


FIGURE 5: Sensitivity analysis of endogenous variables in terms of **K**. (a) p^*, w^*, and k^*. (b) s^*.

DiDi takes in the rush hour, the price is 3.2 CNY/km from 11 PM to 6 AM in the midnight or early morning, 2.4 CNY/km from 6 AM to 7 AM, 2.5 CNY/km from 7 AM to 9:30 AM and from 4 PM to 7 PM, and 2.3 CNY/km for the rest of time periods. Since most orders focus on the daytime, we set the customer valuation per kilometer v be uniformly distributed between 2 CNY and 4 CNY. The value of time c is a difficult parameter to accurately estimate. Thus, we roughly set c_i as 200 CNY for all zones.

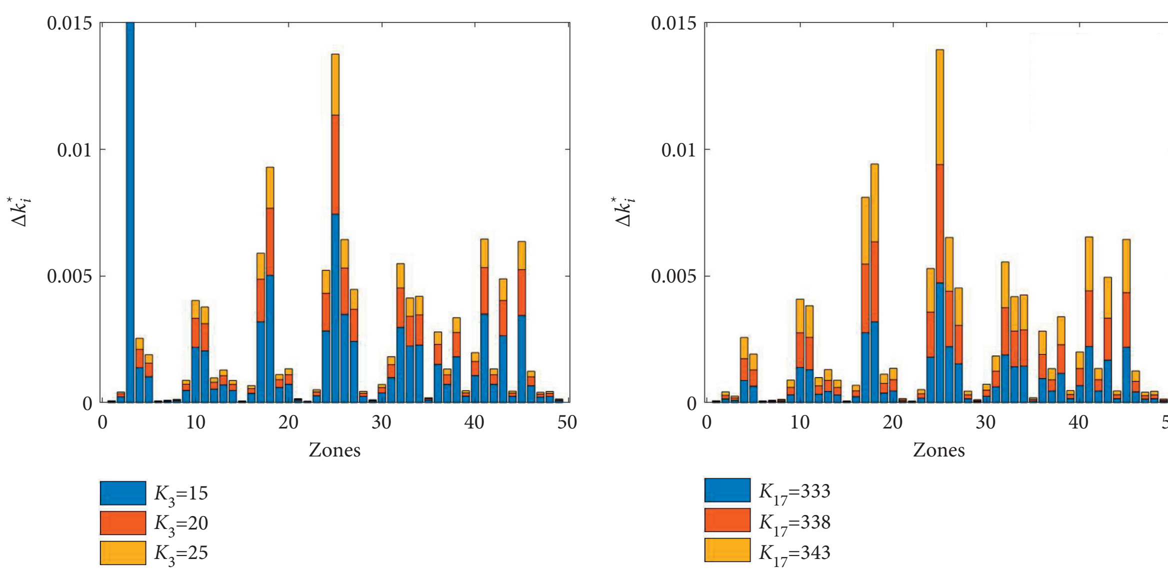
The parameters of customer demand rate $\overline{\lambda_i}$ and the number of potential drivers K_i are estimated from the real data. We set demand rate $\overline{\lambda_i}$ as the average demand rate per hour from the dataset and K_i as the average number of different driver IDs that appear in the dataset. Travel distance L_{ij} and speed V_{ij} between zone i and j are estimated from the average value of travel distance and speed from zone i to j in the real dataset. Table 2 summarizes the parameters and corresponding sources used in the numerical experiments.

5.2. Illustration Based on Real Data. We numerically solve the model under Assumption 1 with the interior point method in MATLAB 2017b. Let us denote the average price as p, average wage as w, and average number of drivers as k. Figure 4 illustrates the changes in the average optimal price p^*, wage w^*, number of drivers k^*, and level of service s^*, as the waiting cost c increases from 10 CNY to 1,000 CNY. As



(a) (b)

Figure 6: Sensitivity analysis of endogenous variables in terms of **V**. (a) p^*, w^*, and k^*. (b) s^*.



(a) (b)

Figure 7: The change of k_i^*, $\forall i \in \mathscr{V}$, as K_3 and K_{17} increase. (a) The change of k_i^*, $\forall i \in \mathscr{V}$, as K_3 increases. (b) The change of k_i^*, $\forall i \in \mathscr{V}$, as K_{17} increases.

can be observed, both the average w^* and k^* increase as c increases, the optimal level of service s^* (and the effective service rate λ^*) decreases, and the optimal price p^* slightly decreases. The result is caused for the waiting time cost increases as the waiting cost c increases, which decreases the customer demand rate. The platform should raise the wage rate w^* to attract more drivers and then reduce waiting time. The optimal price p^* should also be slightly reduced to attract more customers.

Figure 5 illustrates the sensitivity of optimal variables as total K_i increases, which we denote as the increase in K. Intuitively, the average wage w^* decreases and average participating drivers k^* increases as K increases, both the optimal level of service s^* and effective service rate λ^* increase. As shown in Figure 6, the sensitivity of those optimal endogenous variables as V_{ij} increases, which we denote as the increase of V, is similar to Figure 5. Both the average wage w^* and k^* decrease as V increases, and both the

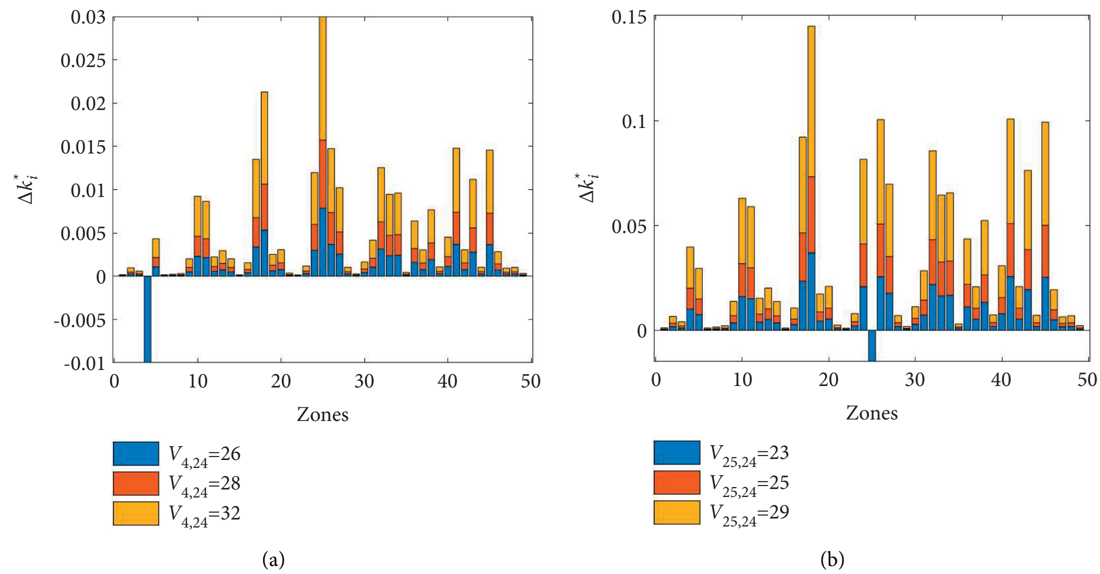


Figure 8: The change of k_i^*, $\forall i \in \mathcal{V}$, as $V_{4,24}$ and $V_{25,24}$ increase. (a) The change of k_i^*, $\forall i \in \mathcal{V}$, as $V_{4,24}$ increases. (b) The change of k_i^*, $\forall i \in \mathcal{V}$, as $V_{25,24}$ increases.

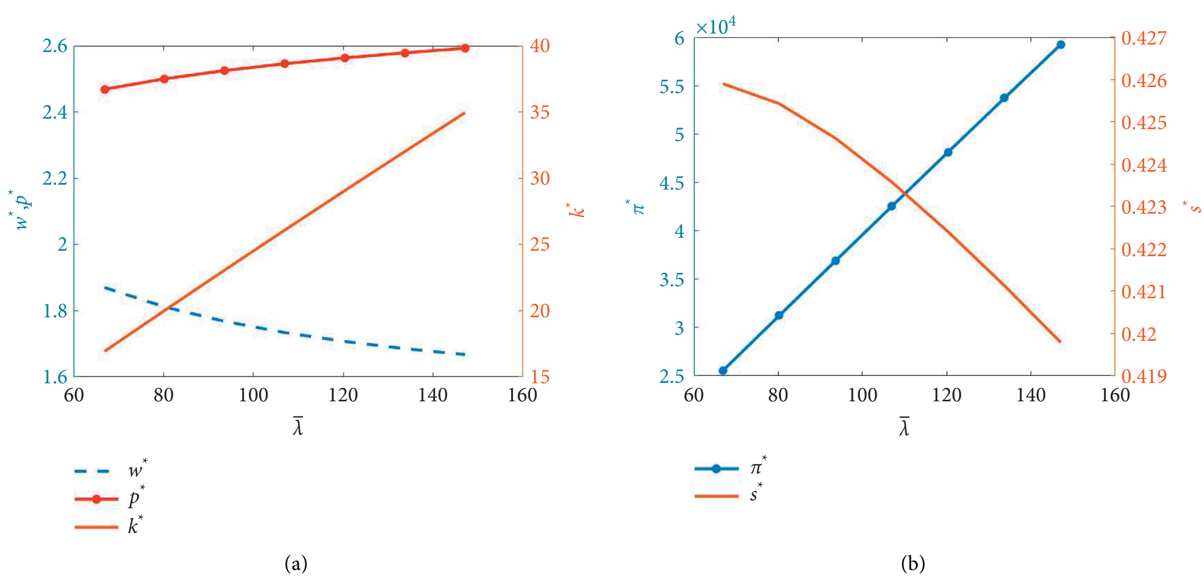


Figure 9: Sensitivity analysis of endogenous variables in terms of $\bar{\lambda}$. (a) $p^*, w^*,$ and k^*. (b) s^*.

optimal level of service s^* and effective service rate λ^* increase. The increase of either K or V causes the increase in the service rate, decrease in the waiting time, and attracts more customers. Thus, to some extent, both either the increase of K or V results in the analogous effect.

Figure 7 shows the change of local k_i^* among all zones with the increase of K_3 in which Δk^* represents the difference of k_i^* between the original one when $K_3 = 10$. The results show that with the increase in the level of service, k_n^*, $\forall n \in \mathcal{V}, n \neq i$, increases. The trend is similar to the increase of K_{17}. We choose regions 3 and 17 because they are the suburb area and the center area, respectively. Figure 8 shows the change of local k_i^* among all zones with the increase of a specific $V_{4,24}$. Similarly, the increase in travel speed helps increase s^*, which increases k_n^*, $\forall n \in \mathcal{V}, n \neq i$. The trend is similar to the increase of $V_{25,24}$.We choose regions 4 and 25 because they are the suburb area and the center area, respectively.

Figures 9 and 10 show the change of parameters with the increase of total $\bar{\lambda}_i$ and L_{ij}, which we denote as the increase of $\bar{\lambda}$ and L. The increase of either $\bar{\lambda}$ or L increases the demand rate. With the increased request rate, the number of effective drivers k_i^* increases. The level of service s^* decreases to inhibit the excessive demand. And the platform could correspondingly reduce the wage rate w_i^*. Figure 11 illustrates the change of local k_i^* among all zones with the

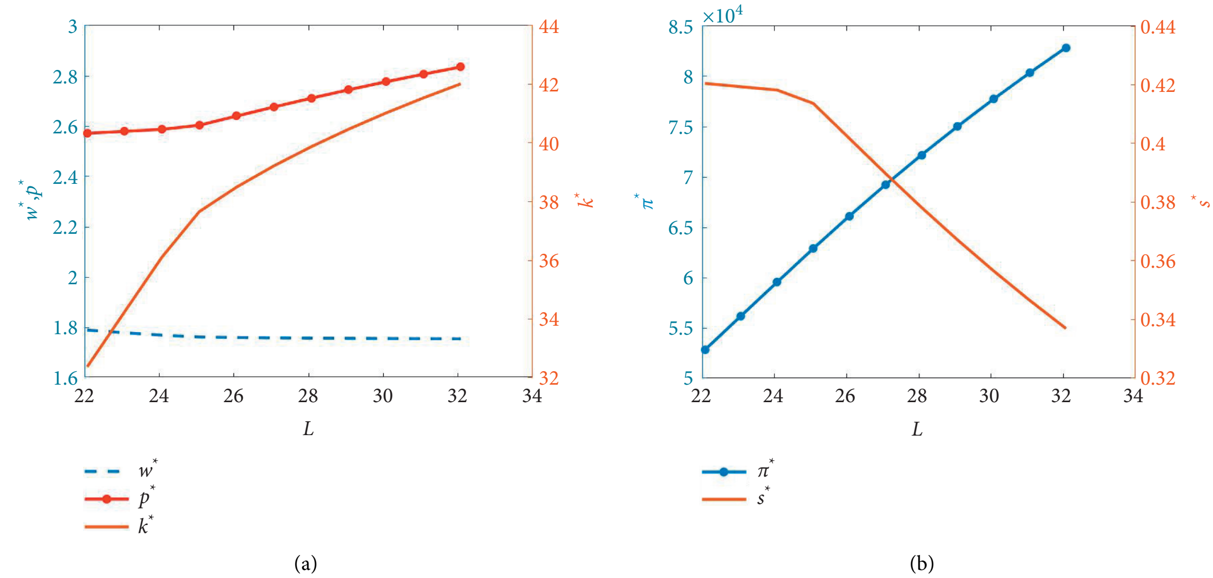


FIGURE 10: Sensitivity analysis of endogenous variables in terms of L. (a) p^*, w^*, and k^*. (b) s^*.

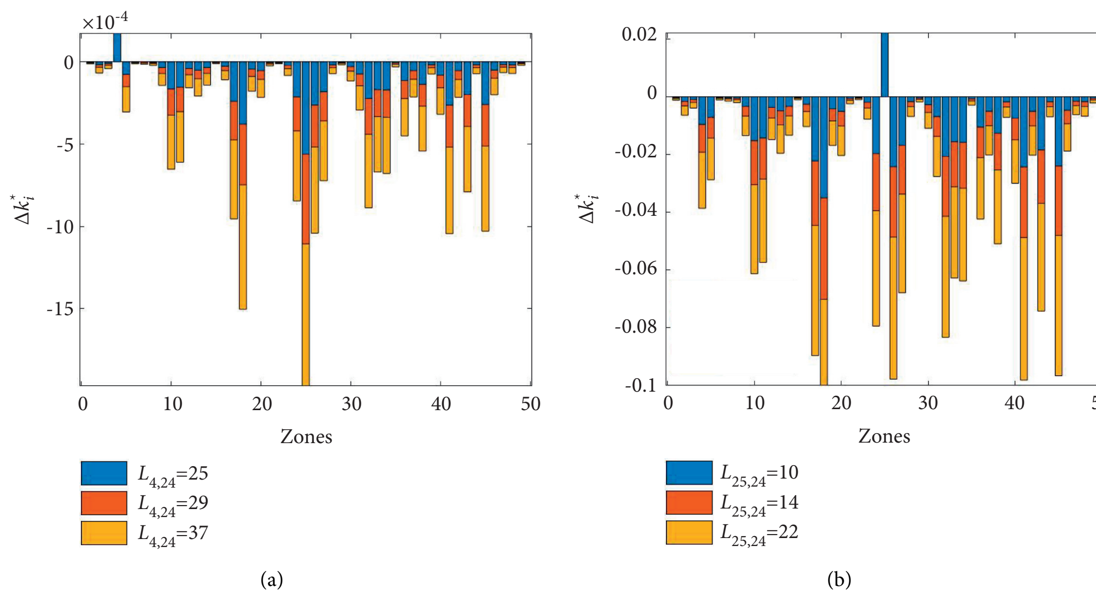


FIGURE 11: The change of k_i^* as $L_{4,24}$ and $L_{25,24}$ increases. (a) The change of k_i^* as $L_{4,24}$ increases. (b) The change of k_i^* as $L_{25,24}$ increases.

increase of $L_{4,24}$ in which Δk^* represents the difference of k_i^* between the original one when $L_{4,24} = 22$ km. As can be seen, in spite of zone 4, k_i^* increases among the rest of zones. The trend is similar to the increase of $L_{25,24}$. We choose regions 4 and 25 because they are the suburb area and the center area, respectively.

Finally, we compared the platform profit under three different scenarios, that is, the flexible level of service and flexible payout ratio, identical level of service and flexible payout ratio, and identical level of service and identical payout ratio in Figure 12. The model is no longer a biconcave optimization problem when the on-demand ride services platform aims at maximizing its profit by allowing spatial heterogeneity or flexible levels of service among all zones. To numerically solve this problem, we use alternating projection to reach the optimal solution. Obviously, the most flexible case reaches the greatest profit, while the identical payout ratio significantly reduces platform revenue.

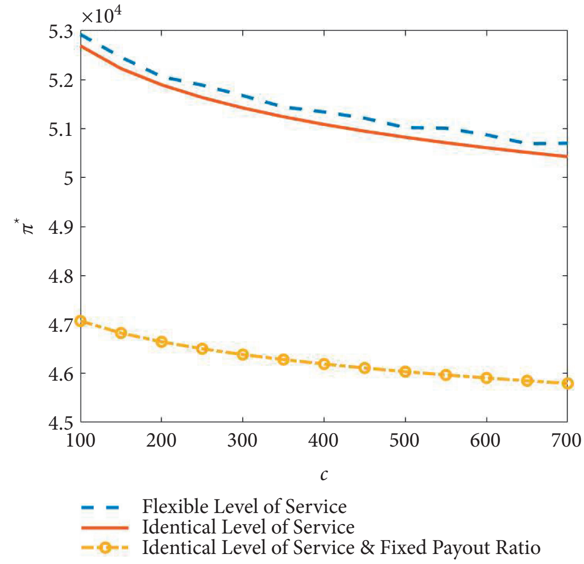


FIGURE 12: Sensitivity analysis of π^* in terms of c.

Conversely, the negative effect of identical level of service on profits is weak, and the platform guarantees service level in areas with low demand at the expense of profit.

6. Conclusions

This article investigates the long-term effects of spatial pricing and its regulation for a ride-sourcing platform. In our model, we divide the study area into zones and suppose that drivers are heterogeneous in their reservation earning rate, while customers are heterogeneous in valuation of service. We further assume that the actual earning rate of drivers is related to the wage set by the platform as well as the arrival rate of ride requests. The cost of customers are assumed to be related to the prices set by the platform as well as the waiting time, which is in turn dependent on the number of drivers and the arrival rate of ride requests. More importantly, the model endogenously determines the long-term demand characteristics, which change dynamically. Thus, the platform must take the interaction between drivers and customers and the changing demand characteristics into account to maximize its profit.

We conduct analysis of supply and request rates in two different cases. In the first case, we consider that the on-demand ride services platform is just entering a competitive market environment and is devoted to expanding its market penetration by maintaining its level of service to be identical across all of the zones. We first prove that in this case, our model is a biconcave problem, which can be solved by the alternating projection method. Our analytical results show that at a particular service zone, the platform should reduce the wage of local drivers as the number of available local drivers or service speed increases. This strategy will further help increase the optimal level of service as well as the participating number of drivers in the other zones and thus will help increase the optimal profit. We also show that at a given zone, when the local waiting time increases, the platform should offer a higher local wage rate to attract more drivers to join the platform. This decision will decrease the optimal level of service as well as the number of drivers at the other zones and thus will decrease the optimal profit. The numerical experiments with real-world data from DiDi coincide with the theoretical results.

In the second case, the on-demand ride services platform is on the stable stage of a monopoly market environment and aims at maximizing its profit by allowing spatial heterogeneity or flexible levels of service across all of the zones. In this case, our model becomes much more complicated and is no longer biconcave. Our further investigation shows that the optimal level of service should be less than 1/2. This is because a higher service level requires a higher wage rate, which results in a lower profit. We also show that once a zone's service level increases, the flow rate from the zone to other zones will increase. As a result, other zones' average speed tends to increase (decrease) if the speed between two zones is greater (less) than the original average speed. The same result also holds for the average travel distance. Thus, these changes (i.e., the changes in a particular zone's average speed and average travel distance) will further affect the the profit. It is worthy to mention that the profit of a particular zone increases as its average travel distance grows and average travel time drops. For instance, there might be an expressway connecting two regions, which are far from each other.

Finally, we numerically compared the platform profit under three different scenarios, that is, the flexible level of service and flexible payout ratio, identical level of service and flexible payout ratio, and identical level of service and identical payout ratio. The result shows that most flexible case is always the most profitable. We also show that the on-demand ride services platform has the potential to choose to expand its penetration by maintaining the same level of service across all geographies while ensuring flexible payment ratios, as it only loses a small portion of profits.

There are several limitations in this article. The cruising behavior of drivers is ignored, and correspondingly, the platform subsidy and more complex strategies are not examined. Besides, the OD-based pricing or path-based pricing is not taken into consideration. Therefore, there are two possible future directions. First, it would be interesting to investigate the short-term pricing strategy where the trip distribution equilibrium no longer exists and the number of customers and drivers in each time period will be affected by the strategies taken in the earlier period. Secondly, we also suggest theoretical and numerical comparisons between OD-based pricing and path-based pricing.

Appendix

A. Notation

Sets

(1) $\mathcal{G}(\mathcal{V},\mathcal{E})$: network with the set of nodes $\mathcal{V}$ and set of edges $\mathcal{E}$. Here, we assume $\mathcal{G}$ is a complete graph, which includes all the self-loops, that is, $\mathcal{E} = \mathcal{V} \times \mathcal{V}$

(2) S_i: the set of state of double queuing model at zone i

Variables

(1) β_i: the proportion of participating drivers to registered drivers
(2) $d_i(\mathbf{q})$: average amount of service units per service request at zone i
(3) k: the average driver number of all zones
(4) k_i: actual number of participating service providers at zone i
(5) λ_i: customer demand rate at zone i
(6) L_i: the average queuing length at zone i
(7) $\mu_i(\mathbf{q})$: average service speed of the service providers at zone $i \in \mathcal{V}$
(8) p: the average price of all zones
(9) p_i: price rate (price per service unit) charged from customers at zone i
(10) π_i: platform profit in zone i
(11) π: overall platform profit, which equals to the summation of π_i, $\forall i \in \mathcal{V}$
(12) q_{ij}: customer demand rate from origin zone i to destination zone j
(13) r_i: reservation (earning) rate of service providers at zone i
(14) ρ_i: the ratio of $\lambda_i d_i / k_i \mu_i$ at zone i
(15) s_i: target level of service at zone i
(16) v_i: value rate per service unit of a customer at zone i
(17) W_i: average queuing time for customers at zone i
(18) w_i: wage rate (wage per service unit) paid to service providers at zone i
(19) w: the average wage of all zones

Functions

(1) $N_i(t)$: the one-dimensional birth-and-death process representing the queuing process at zone i
(2) $U_i(\cdot)$: the surplus or utility function for customers in zone i

Parameters

(1) c_i: unit waiting cost of customers at zone i
(2) L_{ij}: average service unit from origin zone i to destination zone j
(3) $\overline{\lambda}_i$: customer demand rate who may opt to use the platform to request for service at zone i
(4) K_i: maximum number of potential service providers who may opt to participate at zone $i \in \mathcal{V}$
(5) V_{ij}: average service speed from origin zone i to destination zone j, $i, j \in \mathcal{V}$

B. Proofs of Propositions

Proof of Proposition 1. Note that model equations (13a)–(13d) are a strictly convex minimization problem subject to compact linear feasible region. Thus, there exists a unique optimal solution and strong duality holds Boyd et al. [38]. Let $\{\gamma_i\}_{i\in\mathcal{V}}, \{\beta_j\}_{j\in\mathcal{V}}$ be dual multipliers associated with constraints equations (13b) and (13c). Then, the Lagrangian function can be written as

$$\mathcal{L}(q,\gamma,\beta) = \sum_{i\in\mathcal{V}}\sum_{j\in\mathcal{V}}\left(q_{ij}\ln q_{ij} - q_{ij}\right) + \sum_{i\in\mathcal{V}}\gamma_i\left(\sum_{j\in\mathcal{V}} q_{ij} - s_i\overline{\lambda}_i\right) + \sum_{j\in\mathcal{V}}\beta_j\left(\sum_{i\in\mathcal{V}} q_{ij} - s_j\overline{\lambda}_j\right). \quad \text{(B.1)}$$

Thus, the first-order condition yields that

$$\frac{\partial L}{\partial q_{ij}}(q,\gamma,\beta) = \ln q_{ij} + \gamma_i + \beta_j = 0, \forall i, j \in \mathcal{V}. \quad \text{(B.2)}$$

Thus, the authors have

$$q_{ij}^* = e^{-\gamma_i-\beta_j}, \quad \text{(B.3)}$$

for each $i, j \in \mathcal{V}$.

By substituting equation (B.3) into equations (13b) and (13c), the authors have

$$s_i\overline{\lambda}_i = e^{-\gamma_i}\sum_{j\in\mathcal{V}} e^{-\beta_j}, \forall i \in \mathcal{V}, s_j\overline{\lambda}_j = e^{-\beta_j}\sum_{i\in\mathcal{V}} e^{-\gamma_i}, \forall j \in \mathcal{V}, \quad \text{(B.4)}$$

which implies that

$$\sum_{i\in\mathcal{V}} e^{-\gamma_i}\sum_{j\in\mathcal{V}} e^{-\beta_j} = \sum_{i\in\mathcal{V}} s_i\overline{\lambda}_i. \quad \text{(B.5)}$$

Thus, from equation (B.3), the authors arrive at

$$q_{ij}^* = e^{-\gamma_i-\beta_j} = \frac{s_i s_j\overline{\lambda}_i\overline{\lambda}_j}{\sum_{l\in\mathcal{V}} e^{-\gamma_l}\sum_{\tau\in\mathcal{V}} e^{-\beta_\tau}} = \frac{s_i s_j\overline{\lambda}_i\overline{\lambda}_j}{\sum_{\tau\in\mathcal{V}} s_\tau\overline{\lambda}_\tau}. \quad \text{(B.6)}$$

□

Proof of Proposition 2. The bi-level optimization model established in Section 3.4 is a constrained optimization problem. The existence of the optimal solution needs to be validated. The first-order partial derivative of objective function equation (12a) for k_i is

$$\frac{\partial \pi}{\partial k_i} = -c_i\lambda_i d_i\left(\frac{\lambda_i d_i}{\mu_i k_i}\right)^{k_i}\frac{\left(\log\left(\lambda_i d_i/\mu_i k_i\right)-1\right)\left(\mu_i k_i-\lambda_i d_i\right)-\mu_i}{\left(\mu_i k_i-\lambda_i d_i\right)^2}-\frac{2k_i}{K_i}. \tag{B.7}$$

$$\frac{\partial^2 \pi}{\partial k_i^2} = -c_i\lambda_i d_i\left(\frac{\lambda_i d_i}{\mu_i k_i}\right)^{k_i}\left[\frac{\left(\log\left(\lambda_i d_i/\mu_i k_i\right)-1\right)^2}{\mu_i k_i-\lambda_i d_i}+\frac{\left(\lambda_i d_i/k_i\right)-2\mu_i\log\left(\lambda_i d_i/\mu_i k_i\right)+\mu_i}{\left(\mu_i k_i-\lambda_i d_i\right)^2}+\frac{2\mu_i^2}{\left(\mu_i k_i-\lambda_i d_i\right)^3}\right]-\frac{2}{K_i}<0. \tag{B.8}$$

The authors now analyze the value of the first-order partial derivative when k_i is close to $\lambda_i d_i/\mu_i$. Let us denote $k_i\mu_i = \lambda_i d_i + \varepsilon$, where ε is a tiny positive number. Hence, the value of equation (B.7) is positive infinity. Since the second-order partial derivative is negative, the optimal k_i must be larger than $\lambda_i d_i/\mu_i$.

$$\lim_{k_i\longrightarrow\left(\lambda_i d_i/\mu_i\right)^+}\frac{\partial \pi}{\partial k_i} = \lim_{\varepsilon\longrightarrow 0}\frac{\partial \pi}{\partial k_i} = -c_i\lambda_i d_i\left(\frac{\lambda_i d_i}{\mu_i k_i}\right)^{k_i}\left(\frac{\log\left(\lambda_i d_i/\mu_i k_i\right)-1}{\varepsilon}-\frac{\mu_i}{\varepsilon^2}\right)-\frac{2\lambda_i d_i}{\mu_i K_i} = +\infty. \tag{B.9}$$

$$\lim_{k_i\longrightarrow K_i^-}\frac{\partial \pi}{\partial k_i} = -c_i\lambda_i d_i\left(\frac{\lambda_i d_i}{\mu_i K_i}\right)^{K_i}\left(\frac{\log\left(\lambda_i d_i/\mu_i K_i\right)-1}{\mu_i K_i-\lambda_i d_i}-\frac{\mu_i}{\left(\mu_i K_i-\lambda_i d_i\right)^2}\right)-2. \tag{B.10}$$

When k_i approaches the upper bound K_i, $\partial\pi/\partial k_i$ could be either negative (e.g., given $\lambda_i = 100\,\mathrm{h}^{-1}$, $d_i = 6\,\mathrm{km}$, $\mu_i = 20\,\mathrm{km/h}$, $c_i = 500\,\mathrm{h}^{-1}$, $K_i = 200$) or positive (e.g., when K_i is close to $\lambda_i d_i/\mu_i$). Hence, the optimal value of k_i is either the solution of $\partial\pi/\partial k_i = 0$ or K_i.

Based on the above analysis, the existence of k_i^* has been proven. When the arriving rate of requests is large and the drivers supply is not adequate for the demand, all available drivers should be attended, that is,, $k_i^* = K_i$. But k_i^* is smaller than K_i under the general condition. Overall, k_i^* is either K_i or the solution to $\partial\pi/\partial k_i = 0$.

Under the scenario that all s_i are equal, the first- and second-order partial derivatives of the objective function are

$$\frac{\partial \pi}{\partial s} = \sum_{i\in\mathscr{V}}\left[\overline{\lambda}_i d_i(1-2s)-c_i\overline{\lambda}_i d_i k_i\left(\frac{\lambda_i d_i}{\mu_i k_i}\right)^{k_i}\frac{\mu_i k_i+\mu_i-\lambda_i d_i}{\left(\mu_i k_i-\lambda_i d_i\right)^2}\right]. \tag{B.11}$$

$$\frac{\partial^2 \pi}{\partial s^2} = \sum_{i\in\mathscr{V}}\left[-2\overline{\lambda}_i d_i-c_i\overline{\lambda}_i d_i k_i\left(\frac{\lambda_i d_i}{\mu_i k_i}\right)^{k_i}\left(\frac{k_i\mu_i+\mu_i-\lambda_i}{\lambda_i d_i\left(\mu_i k_i-\lambda_i d_i\right)^2}+\frac{1}{\left(\mu_i k_i-\lambda_i d_i\right)^2}+\frac{2\mu_i}{\left(\mu_i k_i-\lambda_i d_i\right)^3}\right)\right]<0. \tag{B.12}$$

When $s \longrightarrow 0$, the first-order partial derivative is positive, because

$$\lim_{s\longrightarrow 0}\frac{\partial \pi}{\partial s} = \sum_{i\in\mathscr{V}}\overline{\lambda}_i d_i > 0. \tag{B.13}$$

Hence, s^* must be larger than 0. When $s \longrightarrow 1$, the authors have

$$\lim_{s\longrightarrow 1}\frac{\partial \pi}{\partial s} = -\sum_{i\in\mathscr{V}}\left[\overline{\lambda}_i d_i + c_i\overline{\lambda}_i k_i\left(\frac{\overline{\lambda}_i d_i}{\mu_i k_i}\right)^{k_i}\frac{\mu_i k_i+\mu_i-\overline{\lambda}_i d_i}{\left(\mu_i k_i-\lambda_i d_i\right)^2}\right]<0. \tag{B.14}$$

Hence, when $s^* = 1$, the first-order partial derivative must be negative. The unconstrained s^* might be smaller than 1, which implies that s^* exists between 0 and 1.

Since $\partial^2\pi/\partial k_i^2 < 0$, $\partial^2\pi/\partial k_i^2 < 0$, it is obvious that the objective function is concave in k_i and s, which indicates that the optimal $\mathbf{k}^*$ can be easily determined once s^* is fixed and the optimal s^* can be easily determined once $\mathbf{k}^*$ is fixed. □

Proof of Proposition 3. (1) Suppose that the optimal k_i^* is an optimal interior solution of the model, that is, $0 < k_i^* < K_i$. For simplicity, let us denote k^0 as the optimal value k_i^* when $K_i = K^0$. According to equation (B.7), the authors have

$$\frac{\partial \pi\left(k^0\right)}{\partial k_i} = -c_i\lambda_i\frac{\partial W_i}{\partial k_i}\Big|_{k_i=k^0} - \frac{2k^0}{K^0} = 0. \tag{B.15}$$

Clearly, $\partial\pi/\partial k_i$ increases with K_i. Let us denote $K^1 = K^0 + \epsilon$, in which ϵ is positive and arbitrarily small, and then, the authors have

$$\frac{\partial \pi\left(k^0\right)}{\partial k_i} = -c_i\lambda_i\frac{\partial W_i}{\partial k_i}\Big|_{k_i=k^0} - \frac{2k^0}{K^1} > 0. \tag{B.16}$$

According to equation (B.8), $\partial^2\pi/\partial k_i^2 < 0$, and thus, the optimal k_i must be greater than k^0 when $K_i = K^1$ to keep $\partial\pi/\partial k_i$ zero, which indicates that k_i^* increases in K_i.

The authors then analyze the optimal value of s^*, which satisfies $\partial\pi/\partial s = 0$, that is,

$$\frac{\partial \pi}{\partial s} = \sum_{i\in\mathcal{V}}\left[\overline{\lambda}_i d_i(1-2s) - c_i\overline{\lambda}_i d_i k_i\left(\frac{\lambda_i d_i}{\mu_i k_i}\right)^{k_i}\frac{\mu_i k_i + \mu_i - \lambda_i d_i}{\left(\mu_i k_i - \lambda_i d_i\right)^2}\right] = 0. \tag{B.17}$$

Clearly, the right side of equation (B.17) increases in k_i. Thus, if k_i^* increases and s is constant, $\partial\pi/\partial s$ becomes larger than 0. With the condition described in equation (B.12), $\partial^2\pi/\partial k_i^2 < 0$, the authors can conclude that s^* must increase with k_i to keep $\partial\pi/\partial s$ zero. Since k_i^* increases with K_i, s^* increases with K_i too. According to equation (12a), profit π increases with K_i even if k_i and s stay constant, and hence, the optimal profit π^* must increase with K_i.

The authors can derive that the waiting time W_i has the following properties:

$$\frac{\partial^2 W_i}{\partial k_i^2} = d_i\left(\frac{\lambda_i d_i}{\mu_i k_i}\right)^{k_i}\left[\frac{\left(\log\left(\lambda_i d_i/\mu_i k_i\right)-1\right)^2}{\mu_i k_i - \lambda_i d_i} + \frac{\left(\lambda_i d_i/k_i\right) - 2\mu_i\log\left(\lambda_i d_i/\mu_i k_i\right) + \mu_i}{\left(\mu_i k_i - \lambda_i d_i\right)^2} + \frac{2\mu_i^2}{\left(\mu_i k_i - \lambda_i d_i\right)^3}\right] > 0. \tag{B.18}$$

$$\frac{\partial^2 W_i}{\partial k_i\partial u_i} = d_i\left(\frac{\lambda_i d_i}{\mu_i k_i}\right)^{k_i}\left(-\frac{\log\left(\lambda_i d_i/\mu_i k_i\right)-1}{\mu_i k_i - \lambda_i d_i}\frac{k_i}{\mu_i} + \frac{-\mu_i k_i\log\left(\lambda_i d_i/\mu_i k_i\right) - 1/\mu_i k_i - \lambda_i d_i + \lambda_i d_i}{\mu_i\left(\mu_i k_i - \lambda_i d_i\right)^2} + \frac{\mu_i k_i + \lambda_i d_i}{\left(\mu_i k_i - \lambda_i d_i\right)^3}\right) > 0. \tag{B.19}$$

$$\frac{\partial^2 W_i}{\partial k_i\partial d_i} = \left(\frac{\lambda_i d_i}{\mu_i k_i}\right)^{k_i}\left[k_i\frac{\log\left(\lambda_i d_i/\mu_i k_i\right)-1}{\mu_i k_i - \lambda_i d_i} + \frac{\lambda_i d_i\log\left(\lambda_i d_i/\mu_i k_i\right) - 2\lambda_i d_i}{\left(\mu_i k_i - \lambda_i d_i\right)^2} - \frac{2\lambda_i d_i\mu_i}{\left(\mu_i k_i - \lambda_i d_i\right)^3}\right] < 0. \tag{B.20}$$

$$\frac{\partial^2 W_i}{\partial k_i\partial \lambda_i} = d_i^2\left(\frac{\lambda_i d_i}{\mu_i k_i}\right)^{k_i}\left[\frac{\left(k_i+1\right)\log\left(\lambda_i d_i/\mu_i k_i\right) - k_i}{\mu_i k_i - \lambda_i d_i} + \frac{\lambda_i d_i\left(\log\left(\lambda_i d_i/\mu_i k_i\right)-1\right) - \mu_i\left(k_i+1\right)}{\left(\mu_i k_i - \lambda_i d_i\right)^2} - \frac{2\lambda_i d_i\mu_i}{\left(\mu_i k_i - \lambda_i d_i\right)^3}\right] < 0. \tag{B.21}$$

Since $\partial^2 W_i/\partial k_i^2 > 0$, $\partial W_i/\partial k_i$ increases as k_i grows. Together with $\partial W_i/\partial k_i < 0$ (since the waiting time decrease with driver number), the first term in the following equation decreases with k_i. With the condition $\partial\pi(k^*)/\partial k_i = 0$, the authors have that the k_i^*/K_i decreases in K_i:

$$\frac{\partial \pi}{\partial k_i} = -c_i\lambda_i\frac{\partial W_i}{\partial k_i} - \frac{2k_i}{K_i}. \tag{B.22}$$

In order to explore the interaction between zones, the authors have analyzed the partial derivative $\partial\pi/\partial k_n$, $n \neq i$, as follows:

$$\begin{aligned}\frac{\partial \pi}{\partial k_n} &= -c_n\lambda_n d_n\left(\frac{\lambda_n d_n}{\mu_n k_n}\right)^{k_n}\frac{\left(\log\left(\lambda_i d_i/\mu_i k_i\right)-1\right)\left(\mu_n k_n - \lambda_n d_n\right) - \mu_n}{\left(\mu_n k_n - \lambda_n d_n\right)^2} - \frac{2k_n}{K_n}\\ &= c_n\lambda_n d_n\left(\rho_n\right)^{k_n}\left(\frac{\log\left(1/\rho_n\right)+1}{\mu_n k_n - \lambda_n d_n} + \frac{\mu_n}{\left(\mu_n k_n - \lambda_n d_n\right)^2}\right) - \frac{2k_n}{K_n}.\end{aligned} \tag{B.23}$$

Clearly, both items $c_n\lambda_n d_n(\rho_n)^{k_n}$ and $\mu_n/(\mu_n k_n - \lambda_n d_n)^2$ increase with s, and $2k_n/K_n$ is independent with s. Hence, if $\log(1/\rho_n) + 1/\mu_n k_n - \lambda_n d_n$ increases or is independent of s, the authors can infer that $\partial\pi/\partial k_n$ increases with s. By taking the partial derivative, the authors have

$$\frac{\partial \log\left(1/\rho_n\right) + 1/\mu_n k_n - \lambda_n d_n}{\partial s} = \overline{\lambda}_n d_n\frac{-1/\rho_n + 2 + \log\left(1/\rho_n\right)}{\left(\mu_n k_n - \lambda_n d_n\right)^2}. \tag{B.24}$$

If the numerator is not less than 0, that is, if $-1/\rho_n + 2 + \log(1/\rho_n) \geq 0$, then $\partial\pi/\partial k_n$ increases with s. Since the authors have deduced that s^* increases with K_i, combined with equation (B.8), the authors can indicate that as K_i increases, $k_n^*, \forall n \in \mathscr{V}, n \neq i$, must increase. Thus, as K_i increases, $k_n^*, \forall n \in \mathscr{V}$, must increase.

$$\frac{k_n}{\lambda_n d_n}\frac{\partial\pi}{\partial k_n} = c_n k_n (\rho_n)^{k_n}\left(\frac{\log(1/\rho_n)+1}{\mu_n k_n - \lambda_n d_n} + \frac{\mu_n}{(\mu_n k_n - \lambda_n d_n)^2}\right) - 2w_n. \tag{B.25}$$

To explore the sensitivity of wage in other regions, the authors formulated equation (B.25). If $-1/\rho_n + 2 + \log(1/\rho_n) \geq 0$, according to equation (B.24), $\log(1/\rho_n) + 1/\mu_n k_n - \lambda_n d_n$ increases in s. It is obvious that $\mu_n/(\mu_n k_n - \lambda_n d_n)^2$ increases with s too. With the first term in equation (B.25) increasing in s, the optimal wage w_n^* must increase in s to keep equation (B.25) zero. Thus, the optimal wage $w_n^*, \forall n \in \mathscr{V}, n \neq i$, increases in K_i, if $-1/\rho_n + 2 + \log(1/\rho_n) \geq 0$.

Suppose that the optimal k_i^* is an optimal boundary solution of the model. The authors have $k_i^* = K_i$; hence, k_i^* increases when K_i increases. Similarly, with equation (B.17), the authors can infer that the s^* increases in K_i and the $\mu_i k_i/\lambda_i d_i$ must increase in K_i too. According to equation (10) and $k_i^* = K_i$, the optimal wage w_i^* equals to

$$w_i = \frac{k_i^2}{K_i\lambda_i d_i(q^*)} = \frac{K_i}{\lambda_i d_i(q^*)}. \tag{B.26}$$

Since $\mu_i k_i/\lambda_i d_i$ increases in K_i, the wage w_i^* must increase in K_i.

According to equations (B.23) and (B.24), if $-1/\rho_n + 2 + \log(1/\rho_n) \geq 0$, then $\partial\pi/\partial k_n$ increases with s. Since the authors have deduced that s^* increases with K_i, combined with equation (B.8), the authors can indicate that as K_i increases, $k_n^*, \forall n \in \mathscr{V}, n \neq i$, must increase. Similarly, with equation (B.25), the optimal wage $w_n^*, \forall n \in \mathscr{V}, n \neq i$, increases in K_i, if $-1/\rho_n + 2 + \log(1/\rho_n) \geq 0$

(2) If V_{ij} increases, then μ_i increases. Denote k^0 as the optimal solution k_i^* when $\mu_i = \mu^0$. As proved above, $\partial^2 W_i/\partial k_i \partial u_i > 0$. Therefore, for any $\mu^1 = \mu^0 + \epsilon$ (ϵ is positive), the authors have

$$\frac{\partial\pi(k^0)}{\partial k_i} = -c_i\lambda_i \frac{\partial W_i}{\partial k_i}\Big|_{k_i=k^0,\mu_i=\mu^1} - \frac{2k_i}{K_i} < 0. \tag{B.27}$$

Since π is concave in k_i, the optimal k_i must be less than k^0, which indicates that k_i^* decreases in μ_i.

According to formulation equation (B.17), $\partial\pi/\partial s$ decreases in μ_i. Hence, the following equation holds:

$$\begin{aligned}\frac{\partial\pi(s^0)}{\partial s} &= \sum_{i\in\mathscr{V}}\Bigg[\overline{\lambda}_i d_i(1-2s^0) \\ &\quad - c_i\overline{\lambda}_i d_i k_i \left(\frac{\lambda_i^0 d_i}{\mu_i^1 k_i}\right)^{k_i} \frac{\mu_i^1 k_i + \mu_i^1 - \lambda_i^0 d_i}{(\mu_i^1 k_i - \lambda_i^0 d_i)^2}\Bigg] > 0.\end{aligned} \tag{B.28}$$

Since $\partial^2\pi/\partial s^2 < 0$, the authors know s^* must increase in μ_i. The first term of equation (B.17) decreases in s, and the second term increases in $\mu_i k_i/\lambda_i d_i$; thus, the optimal $\mu_i k_i/\lambda_i d_i$ must increase, while waiting time W_i decreases (since W_i decreases with $\mu_i k_i/\lambda_i d_i$).

The wage $w_i = k_i^2/K\lambda_i d_i$ increases in k_i and decreases in s; hence, the optimal wage w_i^* must decrease in V_{ij}. Clearly, the objective function increases in V_{ij}, while profit π^* increases.

According to equations (B.23) and (B.24), if $-1/\rho_n + 2 + \log(1/\rho_n) \geq 0$, $\partial\pi/\partial k_n$ increases with s. Since the authors have deduced that s^* increases with μ_i, combined with equation (B.8), the authors can indicate that as V_{ij} increases, $k_n^*, \forall n \in \mathscr{V}, n \neq i$, must increase. Similarly, with equation (B.25), the optimal wage $w_n^*, \forall n \in \mathscr{V}, n \neq i$, increases in μ_i, if $-1/\rho_n + 2 + \log(1/\rho_n) \geq 0$.

(3) Let us denote the optimal k_i^* as k_i^0, when $c_i = c^0$; and as k_i^1, when $c_i = c^0 + \epsilon$. The authors have

$$\frac{\partial\pi(k_i^0)}{\partial k_i} = -c^0\lambda_i \frac{\partial W_i}{\partial k_i}\Big|_{k_i=k^0} - \frac{2k_i^0}{K_i} > 0. \tag{B.29}$$

Hence, k_i^1 must be greater than k_i^0 in order to make the equation zero, indicating k_i^* increases in c_i.

According to equation (B.17), $\partial\pi/\partial s$ decreases in c_i. Hence, the following equation holds:

$$\begin{aligned}\frac{\partial\pi(s^0)}{\partial s} &= \sum_{i\in\mathscr{V}}\Bigg[\overline{\lambda}_i d_i(1-2s^0) \\ &\quad - c_i^1\overline{\lambda}_i d_i k_i \left(\frac{\lambda_i^0 d_i}{\mu_i k_i}\right)^{k_i} \frac{\mu_i k_i + \mu_i - \lambda_i^0 d_i}{(\mu_i k_i - \lambda_i^0 d_i)^2}\Bigg] < 0.\end{aligned} \tag{B.30}$$

Since $\partial^2\pi/\partial s^2 < 0$, s^* must decrease in c_i. With the increase of k_i^* and decrease of s^*, the authors can further infer that the optimal wage $w^* = k_i^* 2/K_i\overline{\lambda}_i s^* d_i$ must increase in c_i. And waiting time W_i^*, which decreases with k_i^* and increases with s^*, must decrease in c_i. From equation (12a), the optimal profit decreases with c_i.

According to equations (B.23) and (B.24), if $-1/\rho_n + 2 + \log(1/\rho_n) \geq 0$, $\partial\pi/\partial k_n$ increases with s. Since the authors have deduced that s^* decreases with c_i, combined with equation (B.8), the authors can indicate that as c_i increases, $k_n^*, \forall n \in \mathscr{V}, n \neq i$, must decrease. Similarly, with equation (B.25), the optimal wage $w_n^*, \forall n \in \mathscr{V}, n \neq i$, decrease in c_i, if $-1/\rho_n + 2 + \log(1/\rho_n) \geq 0$.

(4) If L_{ij} increases, then d_i increases. Suppose that k^0 denotes the optimal value of k_i, when $d_i = d^0$. When $d^1 = d^0 + \epsilon$, from equation (B.20), the authors have

$$\frac{\partial \pi(k^0)}{\partial k_i} = -c_i \lambda_i \frac{\partial W_i}{\partial k_i}\Big|_{k_i=k^0, d_i=d^1} - \frac{2k^0}{K_i} > 0. \tag{B.31}$$

Since π is concave in k_i, the optimal k_i must be greater than k^0, which indicates that k_i^* increases in L_{ij}.

Let us denote s^0 as the optimal s^*, when $d_i = d^0$; as s^1, when $d_i = d^1$. According to equation (B.17), $\partial\pi/\partial s$ decreases in μ_i. Hence, the following equation holds:

$$\frac{\partial \pi(s^0)}{\partial s} = \sum_{i\in\mathcal{V}} \left[\overline{\lambda}_i d_i^1 (1 - 2s^0) - c_i \overline{\lambda}_i d_i^1 k_i \left(\frac{\lambda_i^0 d_i^1}{\mu_i k_i} \right)^{k_i} \frac{\mu_i k_i + \mu_i - \lambda_i^0 d_i^1}{(\mu_i k_i - \lambda_i^0 d_i^1)^2} \right] < 0. \tag{B.32}$$

Since $\partial^2\pi/\partial s^2 < 0$, s^* must decrease in L_{ij}. From equation (12a), the optimal profit increases with L_{ij}.

According to equations (B.23) and (B.24), if $-1/\rho_n + 2 + \log(1/\rho_n) \geq 0$, $\partial\pi/\partial k_n$ increases with s. Since the authors have deduced that s^* decreases with L_{ij}, combined with equation (B.8), the authors can indicate that as L_{ij} increases, $k_n^*, \forall n \in \mathcal{V}, n \neq i$, must decrease. Similarly, with equation (B.25), the optimal wage $w_n^*, \forall n \in \mathcal{V}, n \neq i$, decreases in L_{ij}, if $-1/\rho_n + 2 + \log(1/\rho_n) \geq 0$. □

Proof of Proposition 4. Since the second-order partial derivative of π with respect to k_i remains negative, the objective function under Assumption 2 is concave in k_i with any fixed **s**. The first-order partial derivative of π with respect to k_i is unchanged; hence, the properties about k_i in Proposition 3 still hold.

$$\frac{\partial^2 \pi}{\partial k_i^2} = -c_i \lambda_i d_i \left(\frac{\lambda_i d_i}{\mu_i k_i} \right)^{k_i} \left[\frac{(\log(\lambda_i d_i/\mu_i k_i) - 1)^2}{\mu_i k_i - \lambda_i d_i} + \frac{(\lambda_i d_i/k_i) - 2\mu_i \log(\lambda_i d_i/\mu_i k_i) + \mu_i}{(\mu_i k_i - \lambda_i d_i)^2} + \frac{2\mu_i^2}{(\mu_i k_i - \lambda_i d_i)^3} \right] - \frac{2}{K_i} < 0. \tag{B.33}$$

$$\frac{\partial \pi}{\partial k_i} = -c_i \lambda_i d_i \left(\frac{\lambda_i d_i}{\mu_i k_i} \right)^{k_i} \frac{(\log(\lambda_i d_i/\mu_i k_i) - 1)(\mu_i k_i - \lambda_i d_i) - \mu_i}{(\mu_i k_i - \lambda_i d_i)^2} - \frac{2k_i}{K_i} = 0. \tag{B.34}$$

□

Proof of Proposition 5. For the conciseness of deduction, the profit is divided into the zone-based profit. The authors analyze the impact of s_i on zone i and other zones separately. Since distance d_i and travel speed μ_i are related to s_i, they are considered parameters, too. Let us denote the profit of zone i as π_i, $\mu_i k_i/\lambda_i d_i$ as ρ_i.

$$\frac{\partial \pi_i}{\partial s_i} = (1 - 2s_i)\overline{\lambda}_i d_i + \frac{\lambda_i (1 - s_i)}{\sum_{\tau\in\mathcal{V}} \lambda_\tau (L_{ii} - d_i) - c_i \frac{\rho_i^{k_i}}{\mu_i k_i (1 - \rho_i)} \left(k_i + 1 + \frac{\rho_i}{1 - \rho_i} \right) \left(\sum_{\tau\neq i} V_{i\tau} \lambda_\tau \frac{\sum L_{i\tau}\lambda_\tau}{(\sum V_{i\tau}\lambda_\tau)^2} + \frac{L_{ii}\lambda_i}{\sum V_{i\tau}\lambda_\tau} \right)}. \tag{B.35}$$

$$\frac{\partial \pi_j}{\partial s_i} = \overline{\lambda}_i (\lambda_j - s_j \lambda_j) \frac{L_{ji} - d_j}{\sum_{i\in\mathcal{V}} \lambda_i - c_j \overline{\lambda}_i \frac{\rho_j^{k_j}}{\mu_j k_j (1 - \rho_j)} \left(\frac{L_{ji} - V_{ji} d_j/\mu_j}{\sum V_{i\tau}\lambda_\tau} \right)}. \tag{B.36}$$

Since $0 < s_i < 1$, the third term of $\partial\pi_j/\partial s_i$ is negative. The authors already know the travel distance inside zone i is certainly less than the distance to other zones, that is, $L_{ii} - d_i < 0$, and then, the second term is negative, too. If $1 - 2s_i < 0$, i.e., $s_i > 1/2$, $\partial\pi_j/\partial s_i < 0$, the profit of zone i is decreasing with s_i. Hence, s_i must be greater than 1/2, if the valuation per service unit v is evenly distributed between $[0, 1]$.

The sign of $\partial\pi_j/\partial s_i$ depends on L_{ji}, that is, travel distance from zone j to i; V_{ji}, that is, travel speed from zone j to i; d_j, that is, average travel distance of zone j; and μ_j, that is, average travel speed of zone j. Thus, from equation (B.36), the authors can infer that if $L_{ji} > d_j$ and $L_{ji} > V_{ji} d_j/\mu_j$, that is, $L_{ji}/V_{ji} > d_j/\mu_j$, then $\partial\pi_j/\partial s_i < 0$. The authors can conclude that the value of profit in zone j increases in $\overline{\lambda_i}$, if $L_{ji} > d_j$ and $L_{ji}/V_{ji} > d_j/\mu_j$.

Correspondingly, if $L_{ji} < d_j$ and $L_{ji} > V_{ji} d_j / \mu_j$, that is, $L_{ji} / V_{ji} > d_j / \mu_j$, then $\partial \pi_j / \partial s_i < 0$. The authors can conclude that the value of profit in zone j decreases in s_i, if $L_{ji} < d_j$ and $L_{ji} / V_{ji} > d_j / \mu_j$. □

Data Availability

The ride-sourcing data used to support the findings of this study have not been made available because the data are not allowed to be public by the sponsor.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was financially supported by the National Natural Science Foundation of China (71922019, 72171210, and 71961137005) and CCF-DiDi GAIA Collaborative Research Funds for Young Scholars.

References

[1] J. Wirtz and C. Tang, "Uber: competing as market leader in the US versus being a distant second in China," in *Services Marketing: People Technology Strategy*, pp. 626–632, 2016.

[2] T. A. Taylor, "On-demand service platforms," *Manufacturing & Service Operations Management*, vol. 20, no. 4, pp. 704–720, 2018.

[3] K. Bimpikis, O. Candogan, and D. Saban, "Spatial pricing in ride-sharing networks," *Operations Research*, vol. 67, no. 3, pp. 744–769, 2019.

[4] J. Bai, K. C. So, C. S. Tang, X. Chen, and H. Wang, "Coordinating supply and demand on an on-demand service platform with impatient customers," *Manufacturing & Service Operations Management*, vol. 21, no. 3, pp. 556–570, 2019.

[5] Y. Wang, S. Peng, X. Zhou, M. Mahmoudi, and L. Zhen, "Green logistics location-routing problem with eco-packages," *Transportation Research Part E: Logistics and Transportation Review*, vol. 143, Article ID 102118, 2020a.

[6] Y. Wang, Y. Yuan, X. Guan et al., "Collaborative two-echelon multicenter vehicle routing optimization based on state-space-time network representation," *Journal of Cleaner Production*, vol. 258, Article ID 120590, 2020b.

[7] M. Behrend and F. Meisel, "The integration of item-sharing and crowdshipping: can collaborative consumption be pushed by delivering through the crowd?" *Transportation Research Part B: Methodological*, vol. 111, pp. 227–243, 2018.

[8] D. Çelebi, A. Yörüsün, and H. Işık, "Bicycle sharing system design with capacity allocations," *Transportation Research Part B: Methodological*, vol. 114, pp. 86–98, 2018.

[9] M. Lu, Z. Chen, and S. Shen, "Optimizing the profitability and quality of service in carshare systems under demand uncertainty," *Manufacturing & Service Operations Management*, vol. 20, no. 2, pp. 162–180, 2018.

[10] Y. Hara and E. Hato, "A car sharing auction with temporal-spatial OD connection conditions," *Transportation Research Part B: Methodological*, vol. 117, pp. 723–739, 2018.

[11] H. Xiao, M. Xu, and Z. Gao, "Shared parking problem: a novel truthful double auction mechanism approach," *Transportation Research Part B: Methodological*, vol. 109, pp. 40–69, 2018.

[12] S. X. Xu, M. Cheng, X. T. R. Kong, H. Yang, G. Q. Huang, and G. Q. Huang, "Private parking slot sharing," *Transportation Research Part B: Methodological*, vol. 93, pp. 596–617, 2016.

[13] T. Dong, Z. Xu, Q. Luo, Y. Yin, J. Wang, and J. Ye, "Optimal contract design for ride-sourcing services under dual sourcing," *Transportation Research Part B: Methodological*, vol. 146, pp. 289–313, 2021.

[14] J. Ke, X. Li, H. Yang, and Y. Yin, "Pareto-efficient solutions and regulations of congested ride-sourcing markets with heterogeneous demand and supply," *Transportation Research Part E: Logistics and Transportation Review*, vol. 154, Article ID 102483, 2021.

[15] J. J. Yu, C. S. Tang, Z.-J. Max Shen, and X. M. Chen, "A balancing act of regulating on-demand ride services," *Management Science*, vol. 66, no. 7, pp. 2975–2992, 2020.

[16] R. Ma and H. M. Zhang, "The morning commute problem with ridesharing and dynamic parking charges," *Transportation Research Part B: Methodological*, vol. 106, pp. 345–374, 2017.

[17] N. Masoud and R. Jayakrishnan, "A real-time algorithm to solve the peer-to-peer ride-matching problem in a flexible ridesharing system," *Transportation Research Part B: Methodological*, vol. 106, pp. 218–236, 2017.

[18] X. Qian, W. Zhang, S. V. Ukkusuri, and C. Yang, "Optimal assignment and incentive design in the taxi group ride problem," *Transportation Research Part B: Methodological*, vol. 103, pp. 208–226, 2017.

[19] M. Stiglic, N. Agatz, M. Savelsbergh, and M. Gradisar, "The benefits of meeting points in ride-sharing systems," *Transportation Research Part B: Methodological*, vol. 82, pp. 36–53, 2015.

[20] M. K. Chen and M. Sheldon, "Dynamic pricing in a labor market: surge pricing and flexible work on the uber platform," *Econews*, vol. 16, p. 455, 2016.

[21] L. Chen, A. Mislove, and C. Wilson, "Peeking beneath the hood of uber," in *Proceedings of the 2015 Internet Measurement Conference*, pp. 495–508, ACM, Tokyo Japan, October 2015.

[22] M. Hu and Y. Zhou, *Price, Wage and Fixed Commission in On-Demand Matching*, Social Science Electronic Publishing, London, UK, 2017.

[23] C. Riquelme, S. Banerjee, and R. Johari, "Pricing in ride-sharing platforms: a queueing-theoretic approach," in *Proceedings of the 16th ACM Conference on Economics and Computation*, p. 639, Portland Oregon USA, June 2015.

[24] G. P. Cachon, K. M. Daniels, and R. Lobel, "The role of surge pricing on a service platform with self-scheduling capacity," *Manufacturing & Service Operations Management*, vol. 19, no. 3, pp. 368–384, 2017.

[25] Y. Chen and H. Wang, "Pricing for a last-mile transportation system," *Transportation Research Part B: Methodological*, vol. 107, pp. 57–69, 2018.

[26] C. Lei, Z. Jiang, and Y. Ouyang, "Path-based dynamic pricing for vehicle allocation in ridesharing systems with fully compliant drivers," *Transportation Research Part B: Methodological*, vol. 132, pp. 60–75, 2020.

[27] Z. Xu, Y. Yin, and J. Ye, "On the supply curve of ride-hailing systems," *Transportation Research Part B: Methodological*, vol. 132, pp. 29–43, 2020.

[28] G. S. Chao and T. L. Friesz, "Spatial price equilibrium sensitivity analysis," *Transportation Research Part B: Methodological*, vol. 18, no. 6, pp. 423–440, 1984.

[29] C. S. Fisk, "Spatial price equilibrium on congested networks," *Transportation Research Part B: Methodological*, vol. 21, no. 3, pp. 175–182, 1987.

[30] C. Chen, F. Yao, D. Mo, J. Zhu, and X. Chen, "Spatial-temporal pricing for ride-sourcing platform with reinforcement learning," *Transportation Research Part C: Emerging Technologies*, vol. 130, Article ID 103272, 2021.

[31] L. Zha, Y. Yin, and Z. Xu, *Geometric Matching and Spatial Pricing in Ride-Sourcing Markets*, Social Science Electronic Publishing, London, UK, 2017.

[32] A. G. Wilson, "The use of the concept of entropy in system modelling," *Journal of the Operational Research Society*, vol. 21, no. 2, pp. 247–265, 1970.

[33] X. Chen, C. Chen, and W. Xie, *Optimal Spatial Pricing for an On-Demand Ride-Sourcing Service Platform*, SSRN, 2019, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3464228.

[34] K. Matsushima and K. Kobayashi, "Endogenous market formation with matching externality: an implication for taxi spot markets," *Structural Change in Transportation and Communications in the Knowledge Economy*, pp. 313–336, 2006, https://trid.trb.org/view/809446.

[35] Y. Shi and Z. Lian, "Optimization and strategic behavior in a passenger-taxi service system," *European Journal of Operational Research*, vol. 249, no. 3, pp. 1024–1032, 2016a.

[36] Y. Shi and Z. Lian, "Equilibrium strategies and optimal control for a double-ended queue," *Asia Pacific Journal of Operational Research*, vol. 33, no. 3, Article ID 1650022, 2016b.

[37] J. D. C. Little, "A proof for the queuing formula:L= λW," *Operations Research*, vol. 9, no. 3, pp. 383–387, 1961.

[38] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, UK, 2004.