Taxi-hailing platforms

Sun, Luoyi; Teunter, Ruud H.; Hua, Guowei; Wu, Tian

Published in:
Transportation Research. Part B: Methodological

DOI:
10.1016/j.trb.2020.10.001

Transportation Research Part B 142 (2020) 197–212

Contents lists available at ScienceDirect

Transportation Research Part B

journal homepage: www.elsevier.com/locate/trb

Taxi-hailing platforms: Inform or Assign drivers?

Luoyi Sun[a,1], Ruud H. Teunter[b,2], Guowei Hua[a,*], Tian Wu[c,3]

[a] Beijing Jiaotong University, Beijing, China
[b] University of Groningen, The Netherlands
[c] Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China

A R T I C L E I N F O

Article history:
Received 27 August 2019
Revised 4 September 2020
Accepted 4 October 2020
Available online 7 November 2020

Keywords:
Sharing economy
Two-sided market
Peer-to-peer market
Ride-sourcing platform

A B S T R A C T

Online platforms for matching supply and demand, as part of the sharing economy, are becoming increasingly important in practice and have seen a steep increase in academic interest. Especially in the taxi/travel industry, platforms such as Uber, Lyft, and Didi Chuxing have become major players. Some of these platforms, including Didi Chuxing, operate two matching systems: Inform, where multiple drivers receive ride details and the first to respond is selected; and Assign, where the platform assigns the driver nearest to the customer. The Inform system allows drivers to select their destinations, but the Assign system minimizes driver-customer distances. This research is the first to explore: (i) how a platform should allocate customer requests to the two systems and set the maximum matching radius (i.e., customer-driver distance), with the objective to minimize the overall average waiting times for customers; and (ii) how taxi drivers select a system, depending on their varying degrees of preference for certain destinations. Using approximate queuing analysis, we derive the optimal decisions for the platform and drivers. These are applied to real-world data from Didi Chuxing, revealing the following managerial insights. The optimal radius is 1-3 kilometers, and is lower during rush hour. For most considered settings, it is optimal to allocate relatively few rides to the Inform system. Most interestingly, if destination selection becomes more important to the average driver, then the platform should not always allocate more requests to the Inform system. Although this may seem counter-intuitive, allocating too many orders to that system would result in many drivers opting for it, leading to very high waiting times in the Assign system.

1. Introduction

Online platforms have been developed for many industries in recent years, and their popularity and use are rapidly increasing (Wang and Yang, 2019). For instance, the emergence of Uber, Lyft, and Didi Chuxing has provided passengers with a new, convenient travel mode. According to the *China Online Takeaway Market Monitoring Report for the First Half of 2018*, the market size of food delivery in China alone exceeded 26 billion euros in 2017 (iiMedia Research, 2018).

* Corresponding author at: No. 3 Shangyuancun, Haidian District, Beijing, China.
E-mail addresses: luoyisun@bjtu.edu.cn (L. Sun), r.h.teunter@rug.nl (R.H. Teunter), gwhua@bjtu.edu.cn (G. Hua), wutian@amss.ac.cn (T. Wu).
[1] No. 3 Shangyuancun, Haidian District, Beijing, China.
[2] University of Groningen, PO Box 72, 9700 AB Groningen, the Netherlands.
[3] No. 55 East Zhongguancun Road, Haidian District, Beijing, China.

https://doi.org/10.1016/j.trb.2020.10.001
0191-2615/

Different from the one-way value stream from suppliers to consumers for traditional enterprises, platforms' costs and benefits come from both ends. That is to say, platforms gain profit by building a communication bridge between suppliers and customers. Thus, the efficiency of matching supply and demand is the key factor that drives platforms' operational efficiency and optimizes resource allocation.

There are basically two main matching systems in use by platforms: the Inform (also referred to as Grab Single) system, where the platform acts as an information platform only and suppliers decide who to offer their services to; and the Assign system where the platform itself assigns a supplier to a customer (Gao, et al., 2016; Wang et al., 2016; Dai et al., 2017; Li and Wang, 2017). For online car hailing platforms, in the Inform system, the first driver to respond (in a certain area) is selected and matched to the customer. It is a fair system that allows drivers to control their destination area, by only responding to associated customer requests (Sina, 2017a, 2017b). A disadvantage of this system, however, is that the assigned driver might not be the one closest to the passenger, increasing the time until the customer is picked up (Dai et al., 2017). In the Assign system, this can be prevented by the platform, which typically selects the closest available driver. However, in this system, drivers lose their freedom of choice and can't refuse requests to transport customers to an area that they dislike.

A number of online car hailing platforms, including Didi Chuxing, Shouqi Limousine & Chauffeur and Yidao Yongche, allow drivers to opt between these two systems. Some other platforms, such as Uber, use one system that can be seen as a mixture of the two that we consider. Uber inquires nearby drivers in turn, from close to distant, until one agrees to accept the order. This has the advantage of limiting driving times, whilst still allowing drivers to reject rides. However, this system cannot reward more flexible drivers by assigning more rides to them. Also, the average matching time will increase, negatively affecting the average customer waiting time. The pros and cons of a mixed system compared to that with two separate systems is interesting to analyse. However, that is beyond the scope of this study, which compares the two different matching systems.

Obviously, how the platform allocates rides to the two systems affects how attractive those systems are to the drivers, and so affects the driver system selection. Ride allocation should ensure a good matching of supply and demand on both systems. Besides, to avoid large customer-driver distances and so long en route times, an important control lever is to set a maximum 'radius' (Castillo et al., 2017). Our main goal is to find out what allocation rates and radius minimize the overall expected waiting time. To the best of our knowledge, we are the first authors to do so.

Not all drivers have the same degree of preference for certain destinations over others. For instance, some drivers may want to avoid remote areas of a city more than others. Also, not all drivers are equally knowledgeable about different areas. Some are more experienced and/or have better access to real-time information, allowing them to differentiate better between good/preferred areas and bad/non-preferred areas. As a result, drivers are expected to be heterogeneous with respect to their preferences. We will take this heterogeneity into account when analysing the allocation decision for the platform.

To analyse the matching efficiency on the platform with the two described systems and with heterogeneous drivers, we adopt an *M/M/1* queuing approximation for both systems where drivers act as the server. We explore the strategic interactions between the platform and the drivers. In the presence of drivers' heterogeneity, drivers are offered a choice between the two operating systems and decide based on their regional preference. The platform sets the maximum customer-driver distance and decides what fractions of customer requests for good areas to assign to each system, which affects a driver's probability to reach to a good area and thereby the system selection decision. In order to avoid congestion, the platform minimizes the customers' expected waiting time (i.e. the sum of the matching time and the driver en route time to the customer), and we derive the optimal distribution rate of rides to good areas over the two systems as well as the optimal radius. We remark that although we refer to the case of ride-sourcing platforms in this study, our results also apply to other online platforms where platforms can either assign suppliers itself to a customer request or allow a selection of suppliers to service the customer.

A key finding is that, based on real life data collected for the Chengdu car hailing market, a radius of 1-3 kilometers is optimal. This provides the optimal balance between ensuring a sufficient number of available drivers to avoid long matching times and avoiding long en route times (of the driver to the customer). Moreover, the optimal radius is smaller during rush hour, because of the lower driving speed. Furthermore, for most considered settings, a relatively small fraction of rides (of about 10%) should be allocated to the Inform system, because the Assign system minimizes customer-driver distances. However, a much higher allocation rate (50%-60%) is optimal when either there is a small proportion rides to bad destinations under bad traffic conditions, or there is a small proportion of rides to good destinations under good traffic condition. Also, and most interestingly, if the degree of preference for good rides increases for the average driver and so the Inform system becomes more popular, then the platform should not always increase the allocation rate to that system. Otherwise, the Inform system may attract too many drivers, leading to very long waiting times in the Assign system.

The remainder of the paper is organized as follows: In Section 2, we review the related literature. In Section 3, we present the problem and formulate the model. Section 4 contains the analysis. In Section 5, we conduct an empirical study based on the ride-hailing market in Chengdu to derive managerial insights. Finally, Section 6 provides a conclusion.

2. Literature review

Online car hailing platforms are a typical example of two-sided markets in the growing peer-to-peer sharing economy (Djavadian and Chow, 2017; Nourinejad and Ramezani, 2019). There is a growing stream of literature that focuses on the effective matching of supply and demand for such markets, while keeping search frictions low. As our research also centers around matching, this is the main focus of our review. We next discuss key contributions on matching in general in Section 2.1, before moving on to matching under spatial references in Section 2.2. Thereafter, in Section 2.3, we point out our contribution.

2.1. Matching buyers and sellers

We first discuss studies that primarily focused on matching, as our study is. Yang et al. (2010) introduced an equilibrium model to explore the bilateral searching and meeting frictions between customers and taxis on road networks. Yang and Yang (2011) further explored properties of the market equilibrium and the meeting function between customers and taxis for the traditional taxi market. Their analysis is based on the observation that, in equilibrium, the expected number of times that a customer randomly encounters a driver and vice versa in some area during a certain period of time (e.g. 1 hour) must be equal. He and Shen (2015) made the first attempt on determining a matching function for online car hailing, assuming that customers are homogenous with respect to waiting time and that the ride length is constant. Others extended their work by considering platform profits, social welfare, subsidies strategies, and customer cancellation behaviour when determining equilibrium matching functions (Zha et al., 2016; Wang et al., 2016; Wang and Yang, 2019; Wang et al., 2020).

To analyse waiting times, many researchers modelled the matching process for the sharing economy as a queuing system (Banerjee et al., 2015; Feng et al., 2017; Hu and Zhou, 2017b; Bai et al., 2018; Benjaafar et al., 2018; Taylor, 2018; Nourinejad and Ramezani, 2019; Xu et al., 2020), where buyers arrive according to a Poisson process, and sellers function as servers. Benjaafar et al. (2018) considered car sharing and modelled the matching between car owners and renters by a multi-server loss queuing system. They further compared two types of platforms based on profit-maximizing and social-welfare-maximizing, respectively, in terms of general impacts, including ownership, usage, and social welfare. They found that the transition to a sharing economy can result in either lower or higher ownership and usage levels, with higher ownership and usage levels more likely when the cost of ownership is high. Sun et al. (2019) explored the matching of supply and demand for ride-sourcing platforms by providing per service price determination for any specific ride request. They derived the optimal pricing strategy by taking ride details and driver location into account, and found that the optimal pricing structure for successful matchings includes three parts: (a) a base fare based on the ride length, (b) a rush hour congestion fee, and (c) an emergency fee. Mo et al. (2020) modelled the matching of drivers and riders in a duopoly competition between two ride-sourcing platforms to study optimal subsidization of electric vehicles. Their numerical results revealed that when the returns to scale are significant, the platforms revenue, ride price, and consumer surplus may not be monotone functions of the subsidy. Using bilateral meeting functions, Wang et al. (2020) modelled the matching rates for ride-sourcing and taxi markets. They found that the order cancellation rate is negatively correlated with the customer waiting time, as customers may switch to a taxi whilst waiting to be picked up by a ride-sourcing car. Yang et al. (2020) considered the matching time interval and matching radius as two decision variables in order to optimize the matching process in a ride-sourcing platform, considering the matching rate, expected waiting time, and pick-up time. Theoretical and numerical findings both showed that the matching rate and expected pick-up time increase with the matching radius up to some threshold for that radius, and then become independent of the radius. Moreover, when the supply is considerably larger than the demand, an optimal matching time interval can maximize the system performance.

Many matching studies have considered heterogeneity at the demand side, for instance because of customer impatience under congestion (Bai et al., 2018; Ibrahim, 2019) and under time-related uncertainty of demand (Jiang and Tian, 2019; Hu and Zhou, 2017a). Some researchers have centered on the heterogeneity at the supply side (Hu and Zhou, 2017a; Ibrahim, 2019; Chen et al., 2020). For example, Ibrahim (2019) explored the problems of staffing and controlling queuing systems in sharing economies, taking the randomness in the number of servers into account, where a shortage of supply creates congestion in the system. Hu and Zhou (2017a) considered matching (for transportation, but also more generally) between two types of supply and demand during multiple time periods, where the rewards for each potential matched pair are given and costs are incurred at the end of each period for unmatched supply and demand. Although they allowed supply to be heterogeneous, they assumed that the number of suppliers of the different types as well as the demand are exogenous and observed.

2.2. Matching with spatial variability features

A key aspect of our work is that we include spatial preferences on the supply side. Only in the last few years, researchers have started to pay attention to spatial variability when matching supply and demand. Rayle et al. (2016) focused on the spatial distribution of passengers and drivers, capturing the spatial distribution of trip origins and destinations within San Francisco through a survey. In order to achieve a higher efficiency of matching, they tried to eliminate the differences between regions through spatial pricing, which provides drivers with an incentive to move to over-demand zones. Bimpikis et al. (2016) showed that differentiating the price based on customer location can indeed increase profits for drivers and the platform as well as the consumer surplus. Guda and Subramanian (2019) considered the strategic interaction amongst drivers in their decisions to move between two zones, and found that surge pricing can be profitable even in a zone where the supply of drivers exceeds demand. They focused on whether a driver will move to another area due to surge prices, where the remaining customers are lost in congestion. Zha et al. (2018) divided an urban region into different geographic zones and further investigated the optimal price in each zone under demand surges. They aimed at the same profitable level across different regions and determined the corresponding optimal pricing scheme for each region. Nair et al. (2020) developed a nonlinear-in-parameters multinomial logit (NPMNL) model to forecasting the destinations of deadheading trips on ride-sourcing, taking location specific characteristics (i.e., the built environment, employment opportunities, and socio-demographic characteristics) into account. Using publicly available data from Ride Austin, the showed increased accuracy compared to a multinomial logit (MNL) model.

2.3. Contribution

As is transparent from our literature review, so far studies about matching in sharing economy have mainly focused on the heterogeneity at the demand side. Studies that did consider heterogeneous supply, develop optimal work schedules in the presence of self-scheduling suppliers that have the flexibility to choose whether and when to work (Zha et al., 2018; Gurvich et al., 2019; Ibrahim, 2019).

We consider the flexibility of workers/drivers from a different and new perspective, by allowing them a choice between operating systems, motivated by this choice being available in a number of real life car hailing platforms, as discussed in Section 1. We model how destination preference affects the choice of a driver between the systems. We derive insights into how the platform should allocate rides to the two systems and how the platform should set the radius (i.e., the maximum driver-customer distance), in order to minimize the overall expected waiting time for customers across systems.

3. The model

The model is applicable to platforms that connect supply and demand using two systems, Inform (where the first driver to respond is assigned) and Assign (where the platform directly assigns a driver), which from now on we will also refer to as systems I and A, respectively, for notational ease. Furthermore, we classify destinations as good or bad (e.g. central or remote, respectively) and consider heterogeneity in driver preference between both types of destinations.

In what remains of this section, we first present some model basics (Section 3.1), then discuss platform (Section 3.2) and driver (Section 3.3) specifics and decisions, and finally describe our approximate queuing approach (Section 3.4).

3.1. Basic considerations

We consider a single period (of e.g. one hour), where the platform decides what fractions of customer requests to allocate to either system and the maximum driver-customer distance (radius), and drivers determine their system choice. We assume that drivers are fully informed on how the platform allocates rides, and that drivers make their system decisions based on the allocation rate announced by the platform at the start of the period. In real life, dependent on changing traffic conditions and customer request rates, the platform may alter the allocation fraction from one period to the next and drivers may adjust their system choice accordingly. Analysing such dynamic behaviour over time is of interest, but outside our scope. This exploratory paper considers a single period, i.e., a snapshot of the system. We also remark that the platform could set different radiuses for the two systems, but as the main purpose in both systems is to limit the driver-customer distance, it seems natural to use the same radius for both systems.

We consider an area with known driver density E, i.e., with E taxi drivers per square kilometer. Furthermore, taxi drivers are uniformly distributed over (parts of) the area. In our analysis in the next section, we will particularly focus on circular areas around a customer request, because drivers are only interested in and/or considered in a request if their distance to the customer is not too large. Letting R denote the radius of such an area, $\mathfrak{r}$ the distance to the customer, and Θ the angle between a driver and the positive $\mathfrak{r}$-axis, the uniform distribution implies that the polar coordinates of a driver, $(\mathfrak{r}, \Theta)$, follow the probability density functions:

$$f_r(\mathfrak{r}) = \frac{2\mathfrak{r}}{R^2},\ \ 0 \le \mathfrak{r} \le R, \tag{1}$$

$$f_\theta(\Theta) = \frac{1}{2\pi},\ \ 0 \le \Theta \le 2\pi,$$

and the corresponding probability distribution function of the distance between any driver in this area and the customer is

$$F_\mathfrak{r}(\mathfrak{r}) = \frac{\mathfrak{r}^2}{R^2},\ \ 0 \le \mathfrak{r} \le R.$$

Customer requests arrive according to a Poisson process at a rate of Λ per square kilometer. The corresponding locations are distributed uniformly over the considered area, i.e., equally likely to occur at any point of that area. Each request is for a good (destination) area with probability f_g and so for a bad area with probability $f_b = 1 - f_g$. Thus, customer requests for good and bad areas ($\mathcal{Z} = g, b$) arrive according to Poisson processes with rates $\Lambda_\mathcal{Z} = f_\mathcal{Z} \cdot \Lambda$.

All customers wait for the first taxi that becomes available - so we do not consider a maximum waiting time. In practice, customers may of course leave the system after a long wait, and probably show heterogeneity in their behaviour. However, modelling such behaviour is outside the scope of this paper.

3.2. Platform: systems, customer allocation and objective

The platform operates two systems. If a customer request is allocated to System I, then all drivers of that system within radius R are informed of the ride request, and the first driver to respond is assigned to the customer. In System A, the closest driver is selected within radius R. In Section 3.3, we will derive the maximum radius to be considered, denoted by R_M, ensuring a nonnegative profit for drivers.

The platform decides on what fraction φ ($0 \leq \varphi \leq 1$) of customer requests for good areas to allocate to System I (Inform), where the remainder $1 - \varphi$ is allocated to System A. All customer requests for bad areas must also be assigned to System A, because drivers opt for System I to avoid those rides. So, we get that the platform allocates requests allocated per square kilometre to System I with rate

$$\lambda_I = \varphi \Lambda_g,$$

and to System A with rate

$$\lambda_A = (1 - \varphi)\Lambda_g + \Lambda_b.$$

Note that, correspondingly, the probability that a ride allocated to System A is to a good area is

$$p_g = \frac{(1 - \varphi)\Lambda_g}{(1 - \varphi)\Lambda_g + \Lambda_b}.$$

The platform aims to minimize the expected waiting time over all customers, denoted by T. Denoting the expected waiting time for System s, $s = I, A$, by ϖ_s, the aim is to

$$\min_{\varphi,\ R} T = \frac{\varphi \Lambda_g}{\Lambda_g + \Lambda_b} \cdot \varpi_I + \frac{(1 - \varphi)\Lambda_g + \Lambda_b}{\Lambda_g + \Lambda_b} \cdot \varpi_A. \tag{2}$$

3.3. Drivers: profit, utility and system selection

The profit for a ride is calculated as the (expected) ride price Y minus the ride cost, where the ride cost equals the driving time (both en route to the customer, e, and the ride time from there to her destination, r) multiplied by the driving cost per time unit, c. Denoting the en route time by e and the ride time by r, the ride profit is given by $Y - c(e + r)$. For the profit to be positive, it must hold that $e \leq Y/c - r$. Denoting the average driving speed (during some part of the day) by v, this implies that the maximum distance from a potential driver to the customer, for the expected profit to be nonnegative, is $R_M = (Y/c\text{-}r)v$. Because drivers are not likely to accept loss-generating rides, we assume that the platform only considers values for the radius of at most R_M.

Besides profit, a driver derives utility for driving to a good destination. Drivers differ in their degree of preference, θ, towards a good area, which we model as being uniformly distributed between 0 and 1, i.e., $\theta \sim U[0, 1]$, where a higher value for θ indicates a stronger preference by a driver for a good area. Denoting by $\tau (\tau \geq 0)$ the maximum destination utility, the destination utility is given by $\tau\theta$. Obviously, by scaling τ, the relative importance of regional preference can be set. Therefore, combining the driver's profit and destination utility, the total utility per ride for a driver is $Y - c(e_s + r) + \tau\theta$, where e_s, $s = I, A$, denotes the expected en route time per ride for System s.

We assume that drivers are fully informed about the discussed preference distribution, the customer distribution over areas, and also on how the platform allocates rides. They are rational and opt for the system that brings the largest expected utility (for the considered period). By k_s, $s = I, A$, we denote the fractions of drivers that opt for System s ($k_I + k_A = 1$).

3.4. Queuing system approximation

For each system, customers waiting for a taxi to be assigned are modelled as a queue. Arrivals are according to a Poisson process (i.e. inter-arrival times are exponential) and the platform allocates them to either system with certain probabilities, implying that allocations to both systems also follow Poisson processes.

The drivers act as the server in our system. We will derive (in Section 4.1) their average service time, for each system, as the sum of the en route time to the customer and the ride time to the destinations. Both components are stochastic in real life and hence so is the service time. Deriving the exact service time distribution would be very difficult (if tractable at all) and also require the specification of the ride time distribution, which would be hard in real life. Instead of doing so and in line with the literature (Feng et al. 2017), we will derive the average service times and assume exponential service time distributions.

For System s, $s = I, A$, the number of drivers per request (so drivers within radius R) is easily obtained as $K_s = k_s \pi R^2 E$. Here and in our later derivation of the expected waiting times (and related service rates), we treat K_s as a continuous variable for tractability, despite its discrete nature in real life.

Summarizing, we approximate the matching process for System s, $s = I, A$, as an *M/M/1* queue. We illustrate the determination of the customer waiting time and the driver service time for our queuing model in Fig. 1, and summarize the notations in Table 1.

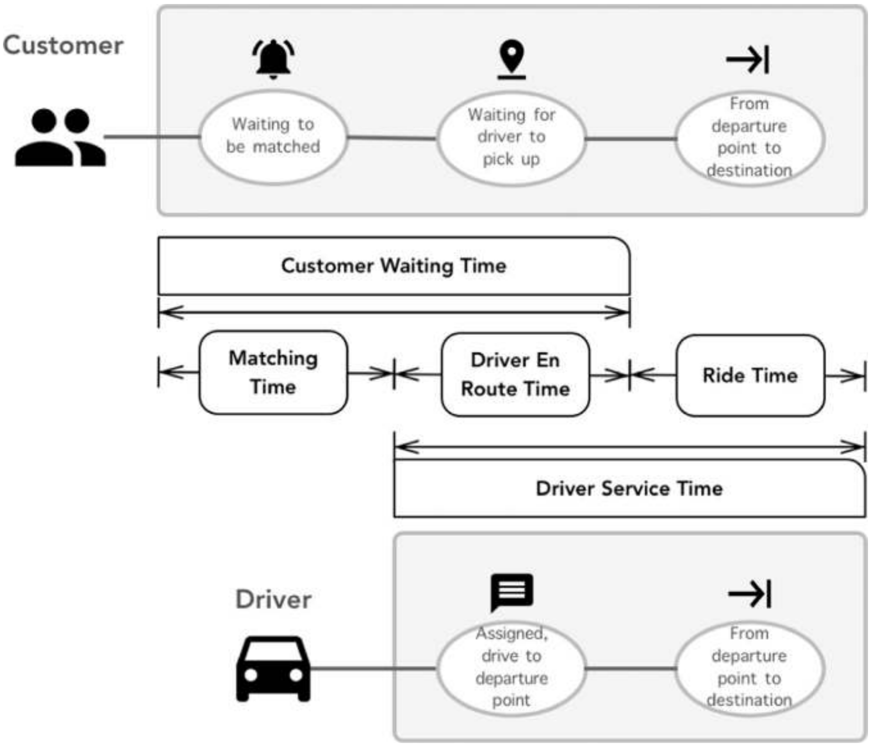


Fig. 1. The composition of time spent for a ride.

4. Analysis

We start in Section 4.1 by analysing what fraction of drivers opt for Systems I and A given a certain allocation policy of the platform. Then, in Section 4.2, we use the results to optimize the allocation policy.

4.1. Driver en route time and system decision

A driver opts for the system that maximizes his expected utility, which depends on the expected time to ride to a customer. In Section 4.1.1 and Section 4.1.2, we derive that expected time for both systems. We then use the results to derive the expected utility for both systems in Section 4.1.3, and also to derive drivers' system selection, i.e., the fractions of drivers in each system.

4.1.1. Driver en route time and driver service rate for System I

For System I, using (1), we find that the expected driver en route distance is

$$\int_0^R \mathfrak{r} f_r(\mathfrak{r}) d\mathfrak{r} = \frac{2R}{3}.$$

At speed v, the expected driver en route time is

$$e_I = \frac{1}{v} \int_0^R \mathfrak{r} f_r(\mathfrak{r}) d\mathfrak{r} = \frac{2R}{3v}. \tag{3}$$

Thus, the average time that a taxi drives to and with a matched customer is $e_I + r$, and so we have a System I service rate of

$$\mu_I = \frac{k_I E}{\frac{2R}{3v} + r}.$$

4.1.2. Driver en route time and driver service rate for System A

In System A, the nearest driver within radius R is matched to the customer. Intuitively, if there (usually) is at least one driver available in a certain radius, then selecting a larger radius has (almost) no effect on the expected driver-customer distance, as the nearest driver is selected anyway. It is the driver density rather than the radius that determines that average distance, as will be confirmed by our analysis.

Table 1
Notations.

Notification	
Decision variable	
φ	fraction of rides to good areas orders assigned to System I (Inform)
R	radius set by the platform for informing drivers
Parameters	
λ_s	customer arrival rate per square kilometre for System s ($s = I, A$)
μ_s	driver expected service rate in System s ($s = I, A$)
k_s	fraction of drivers that opt for System s ($s = I, A$)
K_s	number of informed drivers for System s ($s = I, A$)
e_s	expected driver en route time in System s ($s = I, A$)
m_s	expected matching time in System s ($s = I, A$)
$\tilde{\lambda}_s$	customer arrival rate allocated in the considered area to System s ($s = I, A$)
$\tilde{\mu}_s$	driver expected service rate in the considered area to System s ($s = I, A$)
ϖ_s	customer average waiting time for System s ($s = I, A$)
Λ	customer arrival rate per square kilometer
$\Lambda_{\mathcal{Z}}$	customer arrival rate to good/bad region $\mathcal{Z}$ ($\mathcal{Z} = g, b$)
$f_{\mathcal{Z}}$	probability of a request is for a good/bad area (destination) $\mathcal{Z}$ ($\mathcal{Z} = g, b$)
E	drivers density
R_M	maximum radius (driver-customer) distance with a nonnegative ride profit
v	driving speed
r	expected ride time from departure point to destination
p_g	probability that a request allocated to System A is for a good area (destination)
θ	driver degree of preferences towards the good area
c	driver' waiting cost per time unit
τ	driver utility related to regional preference
Y	ride price

Let us denote the number of drivers within radius R who have opted for System A by K_A (which will be linked to the driver density later in the analysis), so that $K_A = k_A \pi R^2 E$. Let $\mathfrak{r}_i,\ i = 1, 2, \ldots, K_A$, denote the distance of driver i (arbitrarily ordered) to the customer, and y_i, $1 \le i \le K_A$, denote the distance to the customer for closest driver i. Then the probability distribution function of the distance of the nearest driver to the customer is

$$F_Y(y) = P\left(\min[\mathfrak{r}_1, \mathfrak{r}_2, \ldots, \mathfrak{r}_{K_A}] \le y\right) = 1 - \left(1 - \frac{y^2}{R^2}\right)^{K_A}.$$

Differentiating $F_Y(y)$ with y, the probability density function is obtained as

$$f_Y(y) = \frac{2yK_A}{R^2}\left(1 - \frac{y^2}{R^2}\right)^{K_A - 1}. \tag{4}$$

Using (4), we get the expected en route distance as

$$E(y) - \int_0^R y f_Y(y) dy = \int_0^R y \frac{2yK_A}{R^2}\left(1 - \frac{y^2}{R^2}\right)^{K_A - 1} dy = \frac{K_A\sqrt{\pi} \cdot \Gamma(K_A) \cdot R}{2\Gamma\left(K_A + \frac{3}{2}\right)}. \tag{5}$$

Following Davis (1959), we use

$$\Gamma(K_A) = (K_A - 1)!$$

and

$$\Gamma\left(K_A + \frac{3}{2}\right) = \left(K_A + \frac{1}{2}\right)\Gamma\left(K_A + \frac{1}{2}\right) = \left(K_A + \frac{1}{2}\right)2^{-K_A}\sqrt{\pi}(2K_A - 1)!!,$$

where $x!!$ stands for the double factorial (Smarandache, 1991), to rewrite (5) as

$$E(y) = \frac{K_A\sqrt{\pi} \cdot (K_A - 1)! \cdot R}{2\left(K_A + \frac{1}{2}\right)2^{-K_A}\sqrt{\pi}(2K_A - 1)!!} = \frac{2^{K_A} \cdot K_A! \cdot R}{(2K_A + 1) \cdot (2K_A - 1)!!} = \frac{(2K_A)!!}{2(K_A + 1) \cdot (2K_A - 1)!!}R.$$

To approximate $E(y)$, we use the Wallis-product (Hazewinkel, 1994), which states that $\frac{(2K_A)!!}{(2K_A - 1)!!}$ rapidly converges to $\sqrt{\pi K_A}$ as K_A increases. Since the number of competing drivers is typically large (see also our empirical study in Section 5 based on real life data; in Appendix A, we further show numerically that the effect on the driver distribution is also negligible), we get

$$E(y) \approx \frac{\sqrt{\pi K_A}}{(2K_A + 1)}R, \tag{6}$$

which can be further approximated as

$$E(y) \approx \frac{\sqrt{\pi K_A}}{2K_A}R = \frac{\sqrt{\pi}}{2\sqrt{K_A}}R = \frac{1}{2\sqrt{k_A R^2 E}}R = \frac{1}{2\sqrt{k_A E}}. \tag{7}$$

At speed v, the expected driver en route time is

$$e_A = \frac{E(y)}{v} \approx \frac{1}{2v\sqrt{k_A E}}. \tag{8}$$

Given the number of drivers per square kilometre opting for System A is $k_A E$, the expected service rate for System A is approximately

$$\mu_A \approx \frac{k_A E}{1/\left(2v\sqrt{k_A E}\right) + r}.$$

4.1.3. Drivers' system decision

Using the results of Section 4.1.1, we find that the expected utility per ride is

$$Y - c(e_I + r) + \tau \cdot \theta$$

for System I and

$$Y - c(e_A + r) + \tau \cdot \theta \cdot p_g$$

for System A.

Besides the profit per ride, the total expected utility per driver also depends on the expected number of rides per driver per period for each system. The expected number of rides allocated per square kilometre to System I is $\lambda_I = \varphi\Lambda_g$ for System I and $\lambda_A = (1-\varphi)\Lambda_g + \Lambda_b$ for System A. Given the driver density E and letting k_I denote the fraction of drivers that opt for System I, we find that the expected number of rides per driver for Systems I and A are $\lambda_I/(k_I E)$ and $\lambda_A/(k_A E)$, respectively.

Therefore, the expected driver utilities are

$$U_I = \frac{\lambda_I}{k_I E}[Y - c(e_I + r) + \tau \cdot \theta] = \frac{\varphi\Lambda_g}{k_I E}[Y - c(e_I + r) + \tau \cdot \theta]$$

for System I and

$$U_A = \frac{\lambda_A}{k_A E}[Y - c(e_A + r) + \tau \cdot \theta \cdot p_g] = \frac{(1-\varphi)\Lambda_g + \Lambda_b}{k_A E}[Y - c(e_A + r) + \tau \cdot \theta \cdot p_g]$$

for System A.

Note that there are conditions under which either $U_I < U_A$ or $U_I > U_A$ for all values of θ, implying that all drivers opt for the same system (respectively Systems I and A). However, these cases are of limited practical and theoretical interest. Therefore, rather than also considering these extreme cases, we next provide the conditions for both systems to be selected by part of the drivers. Obviously, drivers without any preference should then opt for System A, implying that $U_I \leq U_A$ for $\theta = 0$. Also, drivers with the highest preference should opt for System I, implying that $U_I \geq U_A$ for $\theta = 1$. This situation is depicted in Fig. 2. In Appendix B, we show that these conditions can be rewritten as

$$\varphi \geq k_I \tag{9}$$

and

$$k_A(Y - c(e_I + r))p_g \leq k_I(Y - c(e_A + r)). \tag{10}$$

Drivers are indifferent between both systems if $U_I = U_A$, which gives

$$\theta^* = \frac{k_A(cr - Y + ce_I)\lambda_I - k_I(cr - Y + ce_A)\lambda_A}{\tau(k_A\lambda_I - k_I p_g \lambda_A)}. \tag{11}$$

Since driver preferences are uniformly distributed over [0,1], we must have $k_A = \theta^*$. Combining this with (11) and substituting k_I with $1 - k_A$, we can then obtain the driver distribution in each system, and the approximate number of drivers opting for System s, $s = I, A$, as $K_s = k_s \pi R^2 E$.

4.2. Calculation of the customer waiting time

Recall from Section 3.4 that we approximate the matching process for either system as an $M/M/1$ queue. The customer arrival rate in the considered area is $\tilde{\lambda}_s = \lambda_s \cdot \pi R^2$, where the arrival rate (λ_s) per square kilometre was derived in Sections 3.3. Similarly, the service rate is $\tilde{\mu}_s = \mu_s \cdot \pi R^2$, where μ_s was determined in Section 4.1. The associated system utilization is

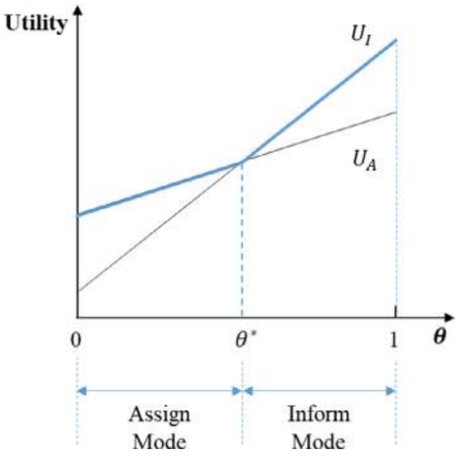


Fig. 2. Driver catchment areas of Inform system and Assign system.

Table 2
Parameter values.

Parameter	Value
Customer arriving rate per square kilometre (in hours)	$\Lambda = 119.10$
Drivers' waiting cost (in RMB/h)	$c = 17.24$
Driver's utility led by his regional preference (in RMB/h)	$\tau = 15$
Ride time, from departure point to destination (in hours)	$r = 0.15$
Drivers' expected profit (in RMB/h)	$Y = 31.93$
Driver speed (in km/h)	$v = 14.75$; 25.7
Drivers density (in km^2)	$E = 22.44$

$\rho_s = \tilde{\lambda}_s/\tilde{\mu}_s$, where the condition for the existence of a steady-state solution is $\rho_s < 1$ or $\tilde{\lambda}_s < \tilde{\mu}_s$. Under that condition, using standard queuing results, we obtain the (approximate) matching time for a System s customer as

$$m_s = \frac{\rho_s}{\tilde{\mu}_s - \tilde{\lambda}_s}. \tag{12}$$

Adding the expected en route time before the driver reaches the customer, we then get the total waiting time for the customer before the driver arrives as

$$\varpi_s = m_s + e_s. \tag{13}$$

By inserting (13) into (2), both the optimal distribution rate and radius can be determined. Due to the complexity, this does not lead to a closed-form results. However, we can get some analytical validation on the effects of the radius. Intuitively, a larger radius leads to longer en route times, but reduces expected matching times by pooling supplier (driver) resources over a larger area. This is formalized in Proposition 1 of which the proof is given in Appendix C.

Proposition 1. For a constant allocation rate, increasing the radius leads to a reduced expected matching time, but an increased expected en route time.

Therefore, optimizing the radius implies a trade-off between the expected matching and driver en route times. This trade-off will be illustrated empirically in Section 5.

5. Empirical analysis

We apply our model to the ride-hailing market in Chengdu, Didi's largest market in China, to get insight into how the Didi Chuxing should decide both the optimal distribution rate and maximum matching radius, and how the different model parameter affect the results. Table 2 lists all estimated model parameter values, where interested readers can refer to Appendix D for sources and explanations. Note that we consider two driving speeds related to bad (rush hour) and good traffic conditions, with driving speeds of 14.75 km/h and 25.7 km/h, respectively.

The maximum drive-customer distance, R_M, for which taxi drivers do not operate at a loss is 37.94 km during rush hour and 66.10 km outside the rush hour. Obviously, these maximum drive-customer distances are far beyond customer's expectations in real life. To ensure that we consider practically relevant settings, we focus on a smaller radius range of [0, 10] (kilometres). The results will show that the optimal radius is indeed never outside that range.

Fig. 3 shows how the allocation rates and considered radius affect the drivers' system decisions with bad (left) and regular (right) traffic conditions. Moreover, each picture shows the results for three different values of the fraction of rides to good destinations (25%, 50% or 75%).

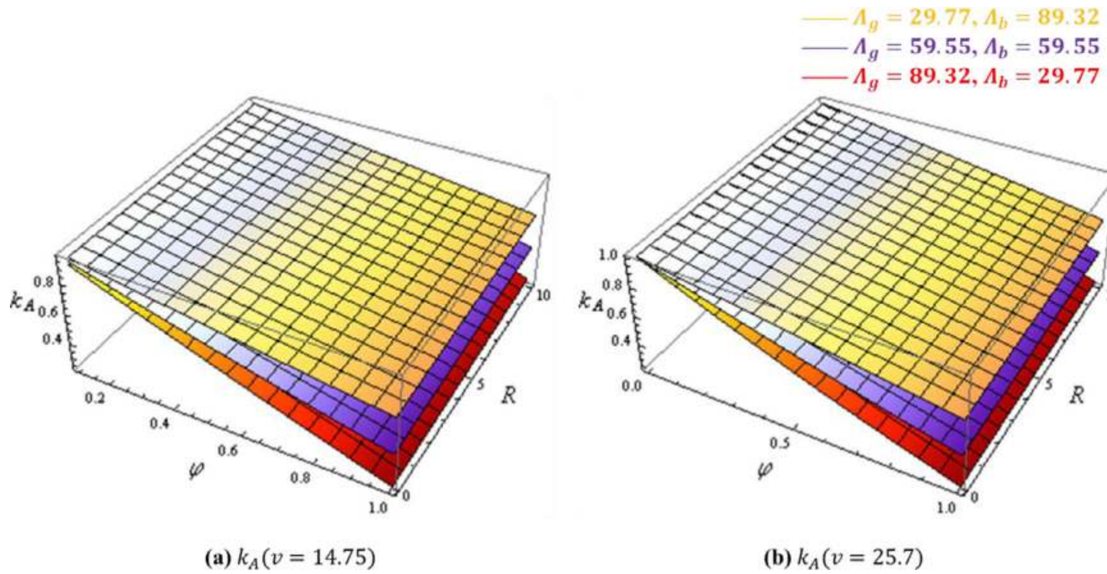


Fig. 3. Proportions of drivers opting for the System *A*.

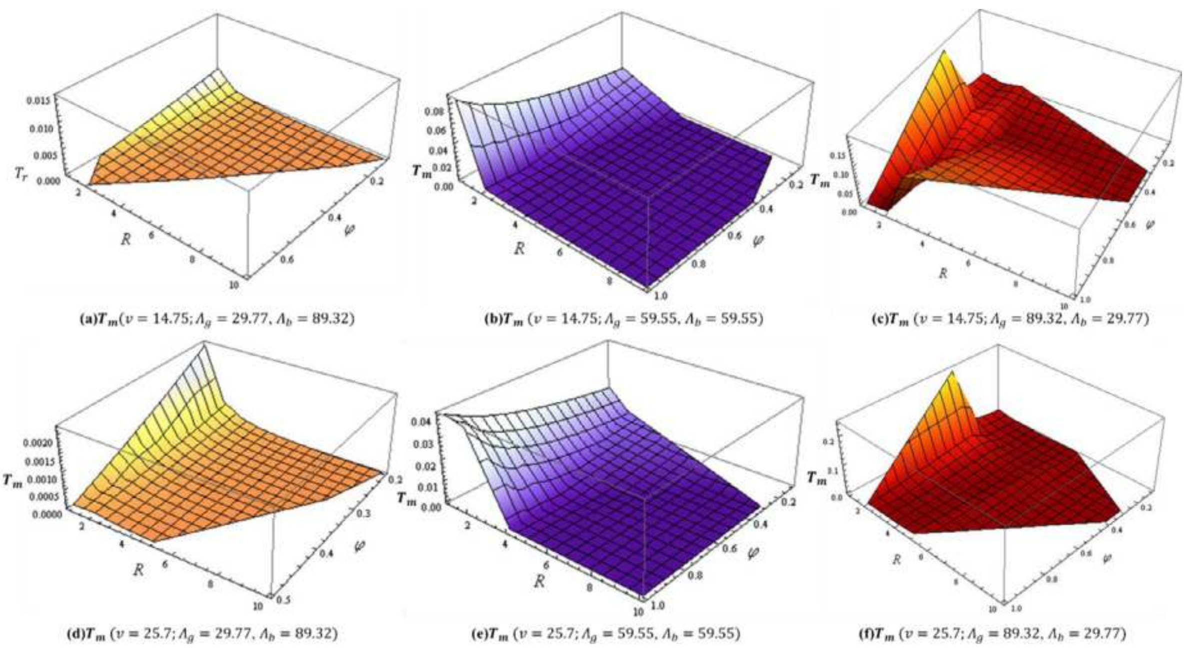
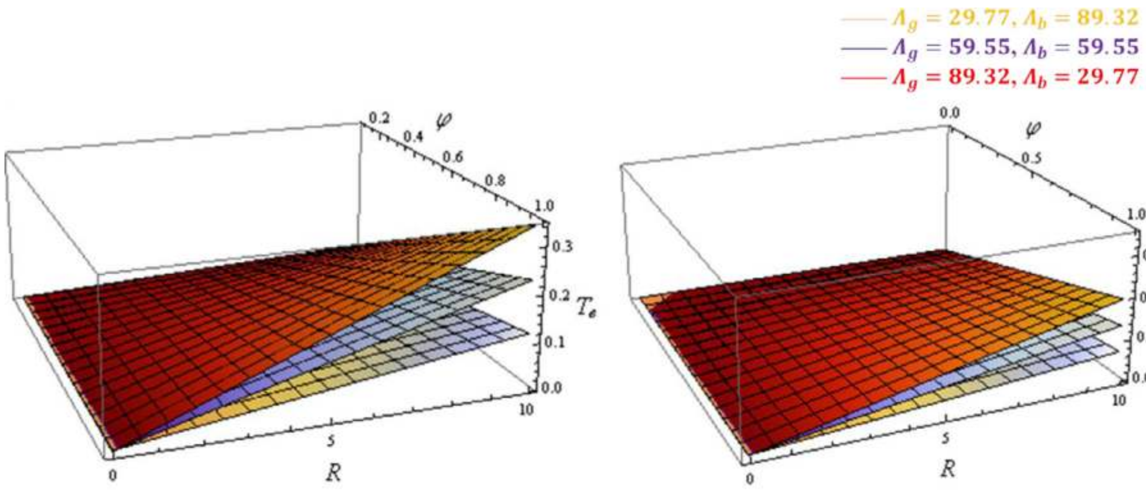


Fig. 4. Effects of radius (R) and allocation rate (φ) on the expected matching time.

It appears from Fig. 3 that, compared with the considered radius, drivers are much more sensitive to the allocation rate when deciding their system choice. Obviously, fewer drivers opt for System *A* if more rides are allocated to System *I*, and so that allocation rate should not be set too high in order to avoid long queues for System *A* customers.

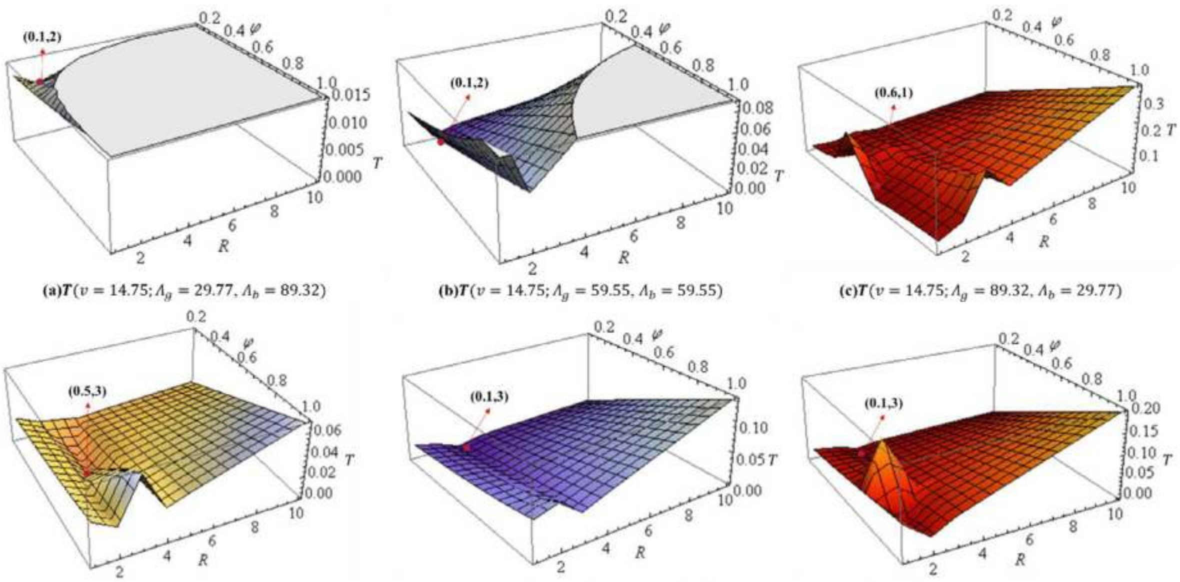
To analyse this further, Fig. 4 displays the effects of the radius *R* and allocation rate φ on the expected matching time. We observe that the matching time generally decreases with *R*, which is expected and in line with the analytical results of Section 4.3. Importantly, we observe that under both bad and good traffic conditions, setting a radius of at least 1-3 kilometers (depending on the traffic conditions and allocation rate) can greatly shorten the expected matching time.

At the same time, the range should not be set too large, as this will increase the expected en route time for the System *I* and thereby also the overall expected en route time. This is illustrated in Fig. 5. Note that the effect is more pronounced for a larger allocation rate, since more rides are then allocated to the System *I*.



(a) $T_e(v = 14.75)$ **(b)** $T_e(v = 25.7)$

Fig. 5. Effects of radius (R) set by the platform and allocation rate (φ) on the expected driver en route time.



(a)$T(v = 14.75; \Lambda_g = 29.77, \Lambda_b = 89.32)$ **(b)**$T(v = 14.75; \Lambda_g = 59.55, \Lambda_b = 59.55)$ **(c)**$T(v = 14.75; \Lambda_g = 89.32, \Lambda_b = 29.77)$

(d)$T(v = 25.7; \Lambda_g = 29.77, \Lambda_b = 89.32)$ **(e)**$T(v = 25.7; \Lambda_g = 59.55, \Lambda_b = 59.55)$ **(f)**$T(v = 25.7; \Lambda_g = 89.32, \Lambda_b = 29.77)$

Fig. 6. Effects of radius (R) and allocation rate (φ) on the expected customer waiting time.

So, especially for setting R, an optimal balance must be sought between the matching time and the en route time. Therefore, in Fig. 6, we consider the effects R and φ on the overall expected waiting time (matching plus en route) per customer. Recall that this is what the platform aims to minimize, and the optimal solution is indicated. This is done for bad (top) and good (bottom) traffic conditions, as well as for percentages of rides to good destinations of 25% (left), 50% (middle) and 75% (right).

For all considered settings, the optimal radius is between 1 and 3 kilometers. As discussed before, too small values result in high matching times, whereas too large values lead to much higher expected en route times. We also observe that the radius should be set lower under bad traffic conditions. This is explained by the lower speed under such conditions, making it more important to limit driver-distances and thereby en route times by lowering R.

For most considered settings, setting a low allocation rate (10%) is optimal. However, it is much higher for two considers settings: (i) few rides to bad destinations under bad traffic conditions, or (ii) few rides to good destinations under good traffic condition. Under these two settings, 50%-60% good rides should be allocated to System I. This is explained as follows. (i) If there are few rides to bad destinations, then a higher allocation rate to the System I still leaves a considerable number

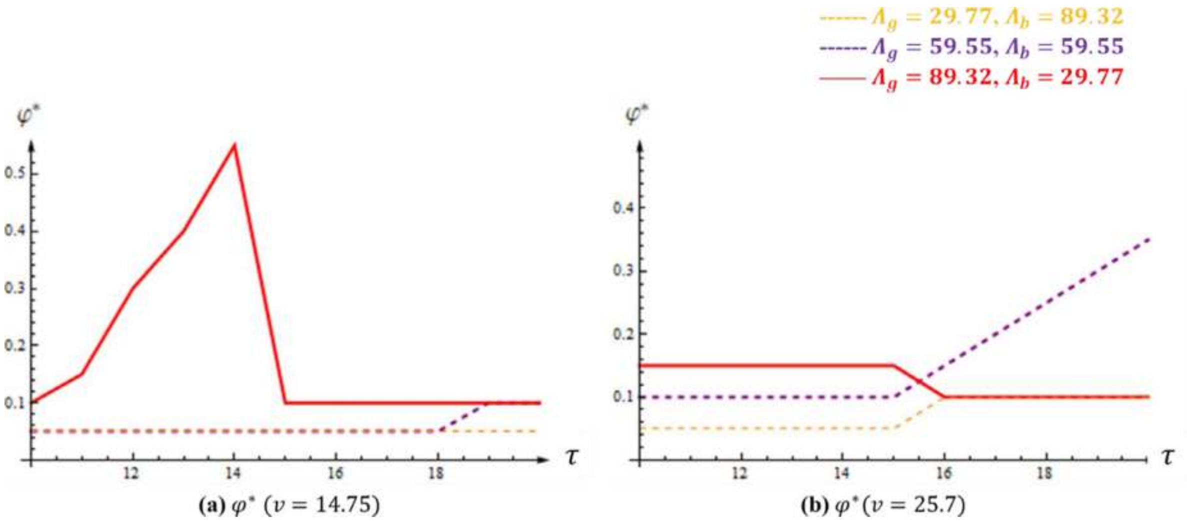


Fig. 7. Effects of drivers' utility led by their regional preference (τ) on the optimal distribution rate.

of good rides for the System *A*. Moreover, under bad traffic conditions, drivers have more benefit from reduced en route times in the System *A*. So, despite a higher allocation rate, enough drivers will still opt for the System *A*. (ii) If there are very few rides to good destinations, then competition for those rides in the System *I* will intensify, especially under good traffic conditions where Inform drivers suffer less from the increased en route time compared to Assign drivers. Hence, the platform allocated a larger fraction of good rides to the System *I*.

Finally, in Fig. 7, we consider the effects of the strength of the regional preference (τ). This is a key parameter for driver choice and so can obviously affect the optimal allocation rate in particular. Indeed, it has relatively little effect on the optimal radius. The results so far showed that the optimal radius is around 3 kilometres under regular traffic conditions and 2 kilometres under bad traffic conditions, and so use these values in Fig. 7.

Interestingly and somewhat unexpectedly, Fig. 7 shows that except for the case where the fraction of rides to good destinations is 50% under a regular traffic condition, the platform should not keep allocating more rides to System *I* as drivers get stronger regional preferences. In fact, for preferences higher than a certain value, the opposite happens. The explanation is that the combination of strong preferences and many rides allocated to System *I* would result in too many drivers opting for System *I*, in turn leading to high waiting times in System *A*. In other words, by not allocating too many good rides to System *I*, the platform ensures that enough drivers opt for System *A*. Those drivers will have a smaller expected utility per ride, but more rides to compensate for that.

6. Conclusion

We studied how taxi sharing platforms with matching power should allocate demand to different operating systems and set the radius (maximum customer-driver distance) in the presence of supply heterogeneity. Motivated by the way that e.g. Didi Chuxing operates and contrary to existing studies, we allowed drivers to choose between operating systems. The Inform system informs all drivers within a certain radius from the customer, and selects the first driver to respond. The Assign system assigns the driver nearest to the customer. Using approximate queuing analysis, we explored how destination preference affects a driver's system choice, and how the platform optimally allocated rides to both systems and sets the radius under the objective to minimize the overall expected customer waiting time across systems.

Key observations are as follows. Matching efficiency is sensitive to both ride allocation and radius, which need to be determined carefully. For the real life case of the Chengdu market that we considered, generally 10% of rides to good areas should be allocated to the Inform system; but a much higher allocation rate (50%-60%) should be set when either there are few rides to bad destinations under bad traffic conditions, or few rides to good destinations under good traffic conditions. Moreover, the platform should only inform drivers that are relatively close to a customer (1-3 kilometers). Interestingly, if destination selection becomes more important (to the average driver), then the platform allocate fewer rides to the Inform system for most considered settings. Although this may seem counter-intuitive, allocating too many orders to that system would result in too many drivers opting for it, leading to very high waiting times in the Assign system.

There are a number of limitations of our research and findings, linking to avenues for further research. First, some model elements can be altered/relaxed. We assumed that all customers are patient and do not leave the system during congestion, whereas they may cancel the order after a long wait, and probably show heterogeneity in their behaviour. We assumed drivers make their system decision before each time slot, whereas they may leave the system at any time in real life. We assumed that the goal of the platform was to minimize the expected customer waiting time, but other objectives can also be considered, e.g. maximizing platform profit. The platform could also set different radiuses for the two systems.

Secondly, our model is designed for the case that two operating systems are provided for suppliers, but other sharing platforms may have different systems of operation. For example, Uber has a different matching process using one system that can be seen as a mixture of the two we considered. They inform drivers one at a time, starting with the one closest to the customer, and assign the request to the first informed driver willing to take it. Our model and analysis could be adapted for such a system, allowing a comparative performance evaluation of different matching systems.

Thirdly, our analysis is static, e.g. we consider an exogenous number of drivers in each period, but in real life the number of drivers assigned to a platform may depend on the traffic condition, location, time of a day, and real-time rewards from other competing platforms. One could consider a dynamic analysis over multiple periods and also consider competition between platforms.

A fourth interesting direction for future research is to explore the usage of incentive schemes for customers and/or drivers to improve matching of supply and demand. Our study takes the driver density and request intensity as given, but customers could for instance be offered discounts for selecting less busy parts of a day and drivers could be rewarded for transferring to areas with a shortage of drivers.

Finally, we only considered the drivers' short-term system choice, and so did not consider dynamic behaviour over time, where drivers react to each other and to changing traffic conditions. Taking such a long-term perspective is certainly worthwhile.

CRediT authorship contribution statement

Luoyi Sun: Conceptualization, Methodology, Writing - original draft, Investigation. **Ruud H. Teunter:** Conceptualization, Methodology, Writing - review & editing, Supervision. **Guowei Hua:** Project administration, Funding acquisition, Writing - review & editing, Supervision. **Tian Wu:** Conceptualization.

Acknowledgements

We thank the editor and two anonymous referees for their helpful comments on an earlier version of our paper. This work was supported by the National Natural Science Foundation of China (Grant number 71831001 and 71804181). We would like to thank the Didi Chuxing GAIA Initiative for access to the data on which the empirical analysis of this paper is based.

Appendix A. Comparison between the approximated and exact results

In this appendix, to prove the validity of the adoption of approximation in (7), we compare the driver distribution over both systems for the approximated and exact results per square kilometre. Table 3(a)-(c) shows the results under different market situations (with 25%, 50% or 75% of rides to non-preferred areas) and regular driving speed, and with other parameters set to the base case values. For all considered settings, we find that the differences are negligible.

Table 3
Number of drivers that opt for either system, for different request rates for good and bad areas and other parameters set to the base case values, resulting from both the exact and approximate analysis.

Calculation $k_s \pi R^2 E$	φ	0	0.2	0.4	0.6	0.8	1
(a) $\Lambda_g = 89.32,\ \Lambda_b = 29.77$							
Exact results	System I	0	11.43	22.61	33.50	44.08	54.30
	System A	70.50	59.07	47.89	37.00	26.42	16.20
Approximate results	System I	0	11.29	22.37	33.23	43.82	54.13
	System A	70.50	59.21	48.13	37.27	26.68	16.37
(b) $\Lambda_g = 59.55,\ \Lambda_b = 59.55$							
Exact results	System I	0	8.36	16.43	24.22	31.71	38.91
	System A	70.50	62.14	54.07	46.28	38.79	31.59
Approximate results	System I	0	8.24	16.23	23.96	31.43	38.63
	System A	70.50	62.26	54.27	46.54	39.07	31.87
(c) $\Lambda_g = 29.77,\ \Lambda_b = 89.32$							
Exact results	System I	0	4.66	9.19	13.57	17.82	21.95
	System A	70.50	65.84	61.31	56.93	52.68	48.55
Approximate results	System I	0	4.59	9.06	13.39	17.61	21.70
	System A	70.50	65.91	61.44	57.11	52.89	48.80

Appendix B. Proof of conditions for both systems to be in operation

In this Appendix, we will provide the proofs of conditions (9) and (10). Given the conditions of (i) $U_I \le U_A$ for $\theta = 0$ and (ii) $U_I \ge U_A$ for $\theta = 1$, which are equivalent to the conditions of (i) the slope of U_I is no smaller than U_A and (ii) $U_I \le U_A$ for $\theta = 0$, we can rewrite the conditions as

$$\frac{\varphi\Lambda_g}{k_I E}\cdot\tau \ge \frac{(1-\varphi)\Lambda_g+\Lambda_b}{(1-k_I)E}\cdot\frac{(1-\varphi)\Lambda_g}{(1-\varphi)\Lambda_g+\Lambda_b}\tau, \tag{14}$$

and

$$\frac{\varphi\Lambda_g}{k_I E}[Y-c(e_I+r)] \le \frac{(1-\varphi)\Lambda_g+\Lambda_b}{(1-k_I)E}[Y-c(e_A+r)]. \tag{15}$$

which can easily be simplified to (9) and (10), respectively.

Appendix C. Proof of Proposition 1.

We first show that the expected en route time increases with the radius, and thereafter that the expected matching time decreases with the radius.

Using (3) and (8), we obtain the expected en route time across all customers as

$$T_e = \frac{\varphi\Lambda_g}{\Lambda_g+\Lambda_b}\cdot\frac{2R}{3v}+\frac{(1-\varphi)\Lambda_g+\Lambda_b}{\Lambda_g+\Lambda_b}\cdot\frac{1}{2v\sqrt{k_A E}}.$$

By taking the derivative of T_e with respect to R, we get

$$\frac{dT_e}{dR} = \frac{\varphi\Lambda_g}{\Lambda_g+\Lambda_b}\cdot\frac{2}{3v} > 0.$$

and so the driver en route time increases with the radius.

The expected matching time across all customers is given by

$$T_m = \frac{\varphi\Lambda_g}{\Lambda_g+\Lambda_b}\cdot m_I+\frac{(1-\varphi)\Lambda_g+\Lambda_b}{\Lambda_g+\Lambda_b}\cdot m_A. \tag{16}$$

By taking the derivative of T_m with respect to R, using (12), and letting $X = 3Evk_I$, $Q = 3rv$, and $Z = \frac{3E(1+2rv\sqrt{Ek_A})^2(\Lambda_b+(1-\varphi)\Lambda_g)^2}{(Ek_A)^{3/2}[2v(Ek_A)^{3/2}-(1+2rv\sqrt{Ek_A})(\Lambda_b-(-1+\varphi)\Lambda_g)]}$, we get

$$\frac{dT_m}{dR} = \frac{2R\varphi^2\Lambda_g^2}{k_I\left(\frac{X}{2R+Q}-\varphi\Lambda_g\right)}+\frac{2(2R+Q)\varphi^2\Lambda_g^2\left(-X(R+Q)+(2R+Q)^2\varphi\Lambda_g\right)}{k_I(-X+(2R+Q)\varphi\Lambda_g)^2}-Z. \tag{17}$$

From this, we get

$$\begin{aligned}\frac{dT_m}{dR} &= \frac{2R\varphi^2\Lambda_g^2}{k_I\left(\frac{X}{2R+Q}-\varphi\Lambda_g\right)}+\frac{2(2R+Q)\varphi^2\Lambda_g^2\left(-X(R+Q)+(2R+Q)^2\varphi\Lambda_g\right)}{k_I(-X+(2R+Q)\varphi\Lambda_g)^2}\\ &\quad-\frac{3E\left(1+2rv\sqrt{Ek_A}\right)^2(\Lambda_b-(-1+\varphi)\Lambda_g)^2}{(Ek_A)^{3/2}\left(2v(Ek_A)^{3/2}-\left(1+2rv\sqrt{Ek_A}\right)(\Lambda_b-(-1+\varphi)\Lambda_g)\right)}\\ &= \frac{2(2R+Q)\varphi^2\Lambda_g^2(-XQ+(R+Q)(2R+Q)\varphi\Lambda_g)}{k_I(X-(2R+Q)\varphi\Lambda_g)^2}\\ &\quad-\frac{3E\left(1+2rv\sqrt{Ek_A}\right)^2(\Lambda_b-(-1+\varphi)\Lambda_g)^2}{(Ek_A)^{3/2}\left(2v(Ek_A)^{3/2}-\left(1+2rv\sqrt{Ek_A}\right)(\Lambda_b-(-1+\varphi)\Lambda_g)\right)}.\end{aligned}$$

So, we have that $\frac{dT_m}{dR} < 0$ if $(R+Q)(2R+Q)\varphi\Lambda_g - XQ < 0$ and $2v(Ek_A)^{3/2}-(1+2rv\sqrt{Ek_A})[\Lambda_b+(1-\varphi)\Lambda_g] > 0$. Using the definitions of X and Q, these conditions can be rewritten as $\frac{Ek_I}{(R/3rv+1)}-(\frac{2R}{3v}+r)\varphi\Lambda_g > 0$, and $Ek_A-(\frac{1}{2v\sqrt{Ek_A}}+r)[\Lambda_b+(1-\varphi)\Lambda_g] > 0$.

Since $\frac{R}{3rv}+1 > 1$, the following conditions are also sufficient:

$$Ek_I-\left(\frac{2R}{3v}+r\right)\varphi\Lambda_g > 0$$

and

$$Ek_A - \left(\frac{1}{2v\sqrt{Ek_A}} + r\right)[\Lambda_b + (1-\varphi)\Lambda_g] > 0.$$

Now note that these conditions simply state that the number of drivers per square kilometre is larger than the number of drivers needed to complete all rates, for both systems. That is, there is no congestion for either system. However, this assumption underlies our queuing analysis, and the platform would never consider solutions for which these conditions are not satisfied.

Appendix D. Data (sources) for the Chengdu market

Customer arriving rate (in hours) and Drivers density: We used drivers' trajectory data from Didi Chuxing of the first week of October, 2016. It comes from Chengdu's Second Ring Area of 65-square-kilometer. According to the data, the number of orders per hour is around 119.10 (within a circular area of radius 1 km), and the drivers density is about 22.44 drivers per km^2.

Drivers' waiting cost: According to the Didi Chuxing's report of "2016 Chengdu Mobile Travel for Employment and Social Development Report" (http://www.sohu.com/a/112174427_399538), the average salary of Didi Chuxing' drivers in Chengdu is 17.24 Yuan per hour. We take this as the drivers' waiting cost per hour.

Drivers' expected profit: According to the website of the Chengdu Bureau of Statistics (http://www.cdstats.chengdu.gov.cn/htm/detail_51951.html), the average salary of urban employees in Chengdu is 31.93 Yuan per hour. We take this as the drivers' expected profit per hour.

Driver speed: According to Didi Chuxing's report of '2016 Big Data Report for Smart Travel' (https://sichuan.scol.com.cn/cddt/201701/55800890.html), the average driving speed in Chengdu is about 25.7 km/h; according to the statistical data from Chengdu Traffic Management Bureau (http://scnews.newssc.org/system/20151012/000608216_2.html), the driving speed in peak hour is about 14.75 km/h.

Ride time, from departure point to destination (in hours): Again, we used drivers' trajectory data from Didi Chuxing of the first week of October, 2016. According to the data, the average ride time is about 0.15 hours. Note that we use this average ride time for both regular and rush hour conditions, which implies a shorter ride average distance in peak demand periods. This seems reasonable, as slower road traffic and higher decrease the competitiveness of taxis on especially long-distance rides compared to public transport alternatives such as BRT, metro, etc.

References

Bai, J., So, K.C., Tang, C.S., Chen, X.M., Wang, H., 2018. Coordinating Supply and Demand on an On-Demand Service Platform with Impatient Customers. Manufacturing & Service Operations Management, pp. 1–52. doi:10.1287/msom.2018.0707.

Banerjee, S., Johari, R., Riquelme, C., 2015. Pricing in ride-sharing platforms: a queueing-theoretic approach. In: Proceedings of the Sixteenth ACM Conference on Economics and Computation, p. 639. doi:10.1145/2764468.2764527.

Benjaafar, S., Kong, G., Li, X., Courcoubetis, C., 2018. Peer-to-peer product sharing: Implications for ownership, usage, and social welfare in the sharing economy. Manage. Sci. 65 (2), 477–493. doi:10.1287/mnsc.2017.2970.

Bimpikis, K., Candogan, O., Saban, D., 2016. Spatial pricing in ride-sharing networks. Oper. Res.. forthcoming https://www.gsb.stanford.edu/faculty-research/working-papers/spatial-pricing-ride-sharing-networks .

Castillo, J.C., Knoepfle, D., Weyl, G., 2017. Surge pricing solves the wild goose chase. In: Proceedings of the 2017 ACM Conference on Economics and Computation, pp. 241–242. doi:10.1145/3033274.3085098.

Chen, X., Zheng, H., Ke, J., Yang, H., 2020. Dynamic optimization strategies for on-demand ride services platform: Surge pricing, commission rate, and incentives. Transp. Res. Part B 138, 23–45. doi:10.1016/j.trb.2020.05.005.

Dai, G., Huang, J., Wambura, S.M., Sun, H., 2017. A balanced assignment mechanism for online taxi recommendation. In: Proceedings - 18th IEEE International Conference on Mobile Data Management, MDM 102-11 doi:10.1109/MDM.2017.23.

Davis, P.J., 1959. Leonhard Euler's integral: a historical profile of the gamma function. Memoriam 66 (10), 849. doi:10.2307/2309786.

Djavadian, S., Chow, J.Y.J., 2017. An agent-based day-to-day adjustment process for modeling 'Mobility as a Service' with a two-sided flexible transport market. Transp Res. Part B 104, 36–57. doi:10.1016/j.trb.2017.06.015.

Feng, G., Kong, G., Wang, Z., 2017. We are on the way: analysis of on-demand ride-hailing systems. SSRN Electron. J. doi:10.2139/ssrn.2960991.

Gao, G., Xiao, M., Zhao, Z., 2016. Optimal multi-taxi dispatch for mobile taxi-hailing systems. In: Proceedings of the International Conference on Parallel Processing, pp. 294–303. doi:10.1109/ICPP.2016.41.

Guda, H., Subramanian, U., 2019. Your uber is arriving: managing on-demand workers through surge pricing, forecast communication, and worker incentives. Manag. Sci. Artic. Adv. 1–20. doi:10.1287/mnsc.2018.3050.

Gurvich, I., Lariviere, M., Moreno, A., 2019. Operations in the on-demand economy: staffing services with self-scheduling capacity. In: Sharing Economy. Springer, Cham, pp. 249–278.

Hazewinkel, M., 1994. Encyclopedia of Mathematics. Springer Science+Business Media B.V., Kluwer Academic Publishers.

He, F., Shen, Z., 2015. Modeling taxi services with smartphone-based e-hailing applications. Transp. Res. Part C 58, 93–106. doi:10.1016/j.trc.2015.06.023.

Hu, M., Zhou, Y., 2017a. Dynamic Type Matching. SSRN https://ssrn.com/abstract=2592622 or http://dx.doi.org/10.2139/ssrn.2592622.

Hu, M., Zhou, Y., 2017b. Price, Wage and Fixed Commission in On-Demand Matching. SSRN, pp. 1–42.

Ibrahim, R., 2019. On queues with a random capacity: some theory, and an application. Sharing Economy. Springer, Cham.

iiMedia Research, 2018. China Online Takeaway Market Monitoring Report for the First Half of 2018 https://www.iimedia.cn/c400/62229.html.

Jiang, B., Tian, L., 2019. The strategic and economic implications of consumer-to-consumer product sharing. Sharing Economy. Springer, Cham.

Li, J., Wang, Z., 2017. Online car-hailing dispatch: deep supply-demand gap forecast on spark. In: IEEE, International Conference on Big Data Analysis IEEE, pp. 811–815.

Mo, D., Yu, J., Chen, X.M., 2020. Modeling and managing heterogeneous ride-sourcing platforms with government subsidies on electric vehicles. Transp. Res. Part B 139, 447–472.

Nair, G.S., Bhat, C.R., Batur, I., Pendyala, R.M., Lam, W.H.K., 2020. A model of deadheading trips and pick-up locations for ride-hailing service vehicles. Transp. Res. Part A 135, 289–308. doi:10.1016/j.tra.2020.03.015.
Nourinejad, M., Ramezani, M., 2019. Ride-Sourcing modeling and pricing in non-equilibrium two-sided markets. Transp. Res. Part B 132, 340–357. doi:10.1016/j.trb.2019.05.019.
Rayle, L., Dai, D., Chan, N., Cervero, R., Shaheen, S., 2016. Just a better taxi? A survey-based comparison of taxis, transit, and ridesourcing services in San Francisco. Transp. Policy 45, 168–178.
Sina, 2017a. There are also some different voices about the Assign mode (in Chinese). http://www.sohu.com/a/201097335_442726.
Sina, 2017b. Didi is accused of monopoly - disputes about the Assign mode (in Chinese). https://top.sina.cn/zx/2017-10-28/tnews-ifynfvar4801397.d.html?cre=tianyi&mod=wpage&loc=12&r=32&doct=0&rfunc=0&tj=none&tr=32.
Smarandache, F., 1991. Only Problems, not Solutions! Xiquan. Publishing House, Chicago: Phoenix.
Sun, L., Teunter, R.H., Babai, M.Z., Hua, G., 2019. Optimal pricing for ride-sourcing platforms. Eur. J. Oper. Res. 278 (3), 783–795. doi:10.1016/j.ejor.2019.04.044.
Taylor, T., 2018. On-demand service platforms. Manufact. Serv. Oper. Manage. 20 (4), 704–720. doi:10.1287/msom.2017.0678.
Wang, X., Liu, W., Yang, H., Wang, D., Ye, J., 2020. Customer behavioural modelling of order cancellation in coupled ride-sourcing and taxi markets. Transp. Res. Part B 132, 358–378. doi:10.1016/j.trb.2019.05.016.
Wang, H., Yang, H., 2019. Ridesourcing systems: A framework and review. Transp. Res. Part B 129, 122–155. doi:10.1016/j.trb.2019.07.009.
Wang, X., He, F., Yang, H., Gao, H., 2016. Pricing strategies for a taxi-hailing platform. Transp. Res. Part E 93, 212–231. doi:10.1016/j.tre.2016.05.011.
Xu, Z., Yin, Y., Ye, J., 2020. On the supply curve of ride-hailing systems. Transp. Res. Part B 132, 29–43. doi:10.1016/j.trb.2019.02.011.
Yang, H., Yang, T., 2011. Equilibrium properties of taxi markets with search frictions. Transp. Res. Part B 45 (4), 696–713. doi:10.1016/j.trb.2011.01.002.
Yang, H., Leung, C.W.Y., Cong, S.C., Bell, M.G.H., 2010. Equilibria of bilateral taxi–customer searching and meeting on networks. Transp. Res. Part B 44 (8-9), 1067–1083. doi:10.1016/j.trb.2009.12.010.
Yang, H., Qin, X., Ke, J., Ye, J., 2020. Optimizing matching time interval and matching radius in on-demand ride-sourcing markets. Transp. Res. Part B 131, 84–105. doi:10.1016/j.trb.2019.11.005.
Zha, L., Yin, Y., Du, Y., 2018. Surge pricing and labor supply in the ride-sourcing market. Transp. Res. Part B 117, 708–722. doi:10.1016/j.trb.2017.09.010.
Zha, L., Yin, Y., Yang, H., 2016. Economic analysis of ride-sourcing markets. Transp. Res. Part C 71, 249–266. doi:10.1016/j.trc.2016.07.010.
Zha, L., Yin, Y., Xu, Z., 2018. Geometric matching and spatial pricing in ride-sourcing markets. Transp. Res. Part C 92, 58–75. doi:10.1016/j.trc.2018.04.015.