Practical Taxi Sharing Schemes at Large Transport Terminals

Zhiyuan Liu [a], Shuaian Wang [b,c], Kai Huang [a,d], Jun Chen [a], Yingna Fu [a]*

[a]Jiangsu Key Laboratory of Urban ITS, Jiangsu Province Collaborative Innovation Center of Modern Urban Traffic Technologies, School of Transportation, Southeast University, China

[b]The Hong Kong Polytechnic University Shenzhen Research Institute, Nanshan District, Shenzhen, China

[c]Department of Logistics & Maritime Studies, The Hong Kong Polytechnic University, Kowloon, Hong Kong

[d]Institute of Transport Studies, Department of Civil Engineering, Monash University, Melbourne, Australia

* *Corresponding author: chenjun@seu.edu.cn*

Abstract

This paper addresses the taxi sharing at large transport terminals, where a large number of passengers are queuing for taxis and the demand for taxi sharing is high. A basic requirement of each passenger on the taxi sharing problem is: to gain from the sharing rather than to lose. Hence, this paper aims to develop an allocation method for the optimal taxi sharing plan, which is concise and efficient for practical use. Two key questions are addressed of the optimal sharing scheme: who should share a taxi and how much each of them should pay. Regarding the second question, for any two given passengers, a fair fare allocation method is first developed. Then, based on this method, the optimal sharing plan of two different charging types are dealt with: deterministic fare (only charged on trip distance) and stochastic fare (trip time is also charged). A system optimal static model and an efficient dynamic model are developed for the case of deterministic fare, both of which give better-off solutions. Regarding the stochastic fare, an averagely optimal solution based on the expectation value of fares is first provided. Consequently, to guarantee the better-off allocation and also to avoid the psychology concept "loss aversion", a chance constrained stochastic program is further developed for the optimal sharing plan. All the four proposed models are efficient for practical use. Some implementations issues are discussed to improve the practicality. Finally, a numerical study shows the significant gain of taxi passengers and drivers from the sharing scheme.

Key Words: transportation; taxi sharing; pricing and pairing; transport terminals

1 Introduction

Traffic congestion has become a severe hurdle in the development of a sustainable urban transport system, which has also enlarged the risks of traffic accidents (Kuang et al., 2015). As a consensus of transport planners, prompting the public transport usage is a key solution to mitigate the traffic congestion (Liu et al., 2013&2016; Huang et al., 2016). Hence, transport demand management schemes (e.g. congestion pricing, parking charge) have also been widely implemented in many cities to limit the private car use and shift the demand to the public transport system (Yang and Huang, 2005; Lawphongpanich et al., 2006; Meng et al., 2012; Liu et al., 2014&2017). However, for commuters living/working in an area with undeveloped/inconvenient public transport facilities, to drive or to ride then becomes a tough question. In such case, car sharing is a good alternative for them to complete their trip and fulfill travel needs (Furuhata et al., 2013; Hosni et al., 2014; Xu et al., 2015). Hence, car sharing tools like Uber have been very prevailing recently.

The existing taxi sharing services in the platform of Uber or Didi are for common users widely distributed in different locations. However, in a particular context, the demand for taxi sharing is very high: at big transport terminals (airport, train station, coach station, etc.), many arrival passengers prefer taking a taxi for comfort and door-to-door services. However, due to the limited number of taxis, it may take a long time that a queuing passenger can get a taxi, which is very exhausting and disturbing for an arrival passenger after a long trip. Therefore, at large transport terminals, the demand for taxi sharing is high. However, there is no convenient tool for such sort of taxi sharing. With a big fleet of queuing passengers, there are two key questions for taxi sharing: who should share one taxi; how much each of them should pay? With a reasonable solution to these two questions, a rationale taxi sharing scheme could be efficiently obtained. To this end, this paper aims to provide an optimal solution to these two questions.

1.1 Literature Review

The advantage of carpooling is to provide a more convenient, flexible, eco-friendly and cost-efficient travel mode, as a good supplement of the other travel modes. Minett and Pearce

(2011) asserted that a casual carpooling to additional locations would result in an increase in the average vehicle occupancy. Jacobson and King (2009) also suggested that the ridesharing can reduce the fuel consumption.

Many attempts on developing a carpooling system take ride-matching as an optimization problem. Agatz et al. (2012) claimed that the objectives should be maximizing the number of served passengers, minimizing the operating cost and minimizing passenger's inconvenience. Existing studies mainly focus on two objectives: minimizing total travel time of passengers and maximizing system profit (Jung et al., 2016; Ota et al., 2016; Jung et al., 2017). A breakthrough is to provide a more flexible carpooling in order to satisfy on-demand requests. Therefore, real-time and dynamic carpooling systems (Chan and Shaheen, 2012; Li et al., 2014) based on the platform of mobile and Global Positioning System (GPS) were developed with the help of communication technologies. According to ride-matching rules, ridesharing patterns can be classified into four groups, including identical carpooling, inclusive ridesharing, partial carpooling and detour carpooling.

The pricing issue among sharing passengers is also a key problem for the taxi sharing scheme. However, the pricing problem has drawn less attention, compared to the ride-matching optimization problem. Stoke exchange and ad exchange (Muthukrishnan, 2009) shed some lights on the ride-matching but cannot be incorporated into the pricing problem. There are three typical pricing rules including catalog price, rule-based pricing and negotiation-based pricing (Furuhata et al, 2013). Agatz et al. (2011) proposed a way to allocate the costs of the joint trip that is proportional to the distances of the separate trips. This rule does not consider the complex factors on the fairness and it should be used in the early stage of carpooling in order to improve the reputation of carpooling systems. Kleiner et al. (2011) proposed an auction-based mechanism to determine the driver's compensation. However, current pricing methods have some limitations such as multiple passengers in the same vehicle and detour ridesharing. Many carpooling participants are neighbors or close friends and they tend to use price negotiation. Naor (2005) treated the carpool problem as a coalitional game, solved the problem and found out four basic elements of the fair share: full coverage, symmetry, dummy and concatenation.

The studies on taxi modelling mainly focus on the network flow equilibrium and route choice of taxis, see, e.g., Yang et al. (2005), Wong et al. (2008), Yang et al. (2010), Yang and Yang (2011), Hosni et al. (2014), Long et al. (2014). Besides, some papers address the customer-search of vacant taxi: where to go after dropping off the preceding customers and how to search for the waiting customers (Wong et al., 2014a; 2014b; 2015a; 2015b). However, the studies on taxi sharing pricing are scarce. Some studies analyzed the macrocosm profit from sharing taxis. Fellows and Pitfield (2000) evaluated the effect of car-sharing scheme in some key aspects, including the reduction of trips, fuel use, vehicle hours, vehicle kilometers and average speed as well as the reduction of transit times and the number of vehicles on each link. Their study indicates that, with only a fraction of costs, the net benefits of car-sharing would be comparable to major road schemes and its economic value should be paid more attention by governments. Nevertheless, this research mainly provides macro-economic benefit but not micro pricing strategy. Based on the non-myopic sharing policies raised by Hyytiä et al. (2012), Sayarshad and Chow (2015) delivered a strategy of non-myopic dynamic pricing. The price for the initial route is set as the cost of the accepted service minus expected user benefit from accepted service plus the opportunity cost due to expected benefit from rejected service. Their research focuses on the improvements of social welfare by conducting taxi sharing. It can be seen from the literature review that the optimal pricing of taxi sharing problem at big transport terminals is still an open question, which is addressed in this paper.

1.2 Objectives and contributions

Taxi sharing is a personal behavior and the objective of such sharing is to reduce the passenger's own travel expenses. Hence, in this paper, a basic assumption is that each passenger involved should gain from each sharing.

Different cities have different taxi charging method; apart from the different charging rate, there is another major difference: some cities only charge on the travel distance, while some other cities also charge on the trip time (either waiting time or overall trip time). Hence, the distance-based charge is assumed to be predictable and deterministic, while the time-based charge is case by case and thus stochastic. For these two different types of charges, this paper aims to provide a methodology for the two key questions of taxi sharing.

Firstly, regarding “how much each sharing passenger should pay”, a fair fare allocation model/method is provided for any given two sharing passengers, which is determined based on the amount of taxi fare they should pay without sharing. To guarantee that each passenger involved could gain, the model takes the benefit of each passenger as a constraint.

Secondly, based on the method above, the optimal taxi sharing plan (who should share one taxi) is then solved. Two models are provided for each charging type. For the deterministic fare, a static model targeting at system optimum and a dynamic model efficient for practical use are developed. For the stochastic fare, a model based on expectation value is first proposed, which gives an averagely optimal solution and cannot guarantee the gain of each passenger; then, a chance constrained stochastic programming model is further proposed, and it also avoids a psychology concept “loss aversion”. Model properties and solution techniques of the four models are elaborated, which are all concise and efficient for practical implementations.

It should be further pointed out that this paper only addresses the taxi sharing problem at big transport terminals, where all the passengers gather at one origin. Other cases of passengers with multiple origins (e.g., en-route boarding) are not considered in this paper. The addressed problem has several unique features: (i) unlike matching agencies that require a large volume of participants of drivers and passengers, taxi sharing at a transport terminal simply works in a “business as usual” manner if the number of passengers is small; (ii) unlike service operators that use their own vehicles and drivers which usually require advanced booking, demand for taxi-sharing at a transport terminal can be satisfied in real time; (iii) the mechanism in our system is robust because there is virtually no cancelation, no change-of-schedule, and no “no show”; (iv) security concern is addressed in our system because there is always a licensed taxi driver in the vehicle rather than just a passenger and a stranger driver.

The remainder of the paper is organized as follows. Section 2 clearly defines the concept of taxi sharing and provides a method for the fair fare split between two sharing passengers. Section 3 presents two taxi sharing models for the deterministic taxi fares. The stochastic taxi fares are discussed in Section 4, where two optimization models are developed. Section 5 further discusses some practical issues of the implementations. A numerical example is conducted in Section 6. Section 7 finally concludes this paper.

2 Feasible taxi sharing and fair fare allocation

To have some contexts for the problem description, we consider the case at an airport, but the discussions and methods also hold for other types of transport terminals. When an airplane arrives, a number of arrival passengers will go to the taxi stand to take a taxi to their final destinations. As discussed in the Introduction, the taxi sharing problem should address two key questions: pairing (who should share a taxi) and pricing (how much each passenger should eventually pay).

We do not consider the passengers who do not want to share a taxi with other people, for example, a family or a group of friends. In this section and the next section, we assume that the taxi fare is only dependent on the travel distance. As a result, the taxi fare is deterministic for a passenger. We will consider in Section 4 more general cases in which the taxi fare also depends on waiting time (e.g., at red lights), which makes the taxi fare stochastic. For the ease of presentation, the list of notation is provided as follows:

Sets:

S_i	The set of suitable/potential sharing passengers of passenger i.
Ω_t	The set of passengers who are waiting to share at time t.

Parameters:

d_i	The distance from the station (origin) to passenger i's destination.
d_{ij}	The distance from i's destination to j's destination.
f_i	The deterministic fare that i should pay without sharing.
$f_i^{'}$	Passenger i's minimum requirement on the fare saving.
$\tilde{f}_i$	The stochastic fare if i travels alone without sharing.
F_{ij}	The deterministic fare (only based on distance) of a shared trip of passengers i and j (go to i's destination first).

$\tilde{F}_{ij}$	The stochastic fare (both trip distance and time are charged) of a shared trip of passengers i and j.
t_i	The travel time of the taxi to take passenger i without sharing.
t_{ij}	Travel time from i's destination to j's destination.
$t_i^{'}$	The upper-bound of passenger i's increased travel time caused by sharing.
v	The vth historical data records, $v = 1, 2 \cdots V$.
V	The number of historical data records.
β_{ij}	The percentage of the fare shown on the meter at i's destination to be paid by passenger i.
Δ_{ij}	The overall benefit gained by i and j after sharing.
π_j	The clock time when passenger j made the taxi-sharing request.
χ_{ij}	A binary indicator such that equals 1 if $j \in S_i$ and equals 0 otherwise.

Decision Variables

x_i	The fare paid by passenger i after sharing.
x_j	The fare paid by passenger j after sharing.
z_{ij}	A binary variable which equals 1 if i and j are assigned to share a taxi, and 0 otherwise.

2.1 Definitions and assumptions for feasible taxi sharing

The key reason of introducing taxi sharing at terminals is to consolidate the demand, reduce the queues and improve the efficiency of the taxi system. In this regard, the proposed scheme aims to encourage more passengers to share (especially when the demand is high), such that the efficiency of taxi system could be improved and also the waits of the passengers and taxis could be reduced. Therefore, providing the passengers more incentives to share is a good way to promote taxi pooling Hence, any two passengers i and j who are willing to share a taxi with someone should not have a considerable detour in the shared trip. That is to say, without consideration of some external factors (accidents, the driver loses his way), each passenger should gain from each sharing. Hence, we have Assumption 1:

Assumption 1. Both the sharing passengers and the driver require economic gain from the sharing, compared with the non-sharing case.

Note that the sharing behavior itself will give more benefit to the driver, since he/she travels a longer distance and hence collects more fare. We thus make a second assumption:

Assumption 2. The total fare paid by the two passengers is equal to the fare shown on the meter of the taxi when it arrives at the destination of the second passenger.

Therefore, we only need to provide a reasonable fare allocation method between i and j. Let F_{ij} denote the fare shown on the meter at the end of the trip if a taxi first takes i to her destination and then takes j to her destination; F_{ji} is defined similarly. It is reasonable to require the taxi to take the route with the lowest fare, i.e., if $F_{ij} \le F_{ji}$ then the taxi goes to the destination of i first, and vice versa. Without loss of generality, in the sequel, we assume that $F_{ij} \le F_{ji}$, $i < j$. Let x_i and x_j be the fare paid by passenger i and j after sharing. Assumption 2 implies that

$$x_i + x_j = F_{ij} \tag{1}$$

To ensure that both passengers are happy about the sharing, both of them must have remarkable fare savings and neither of them travels much longer than not sharing. To this end, the system needs to define the following parameters. After sharing, any passenger i should at least save f_i' dollars, and have a travel time increase of at most t_i'. Note that (i) f_i' and t_i' are only related to the destination of passenger i rather than any other characteristics of passenger i. (ii) f_i' and t_i' can be constant, for instance, 3 dollars and 5 minutes, or may depend on the destination of passenger i, for instance, 20% of the original fare and 10% of the travel time. (iii) As the real travel time t_i' depends on the traffic conditions, the system may set different values of t_i' during different time periods.

Define f_i and f_j as the fare of passenger i and j *without sharing*, respectively. Then, the overall benefit gained by i and j after sharing is defined as Δ_{ij}:

$$\Delta_{ij} = f_i + f_j - x_i - x_j = f_i + f_j - F_{ij} \tag{2}$$

Let t_i denote the travel time of the taxi to take passenger i if she travels alone, t_j the travel time of the taxi to take passenger j if she travels alone, and t_{ij} the time of the taxi to travel from the destination of passenger i to j's destination. Here t_i, t_j, and t_{ij} are estimated values based on the time of day. It is thus feasible for passenger i and passenger j to share a taxi if and only if:

$$\begin{aligned} & x_i \le f_i - f_i', x_j \le f_j - f_j', t_i \le t_i + t_i', t_i + t_{ij} \le t_j + t_j', x_i + x_j = F_{ij} \\ \Leftrightarrow\ & x_i \le f_i - f_i', x_j \le f_j - f_j', t_i + t_{ij} \le t_j + t_j', x_i + x_j = F_{ij} \\ \Leftrightarrow\ & F_{ij} \le f_i + f_j - f_i' - f_j', t_i + t_{ij} - t_j \le t_j' \\ \Leftrightarrow\ & \Delta_{ij} \ge f_i' + f_j', t_i + t_{ij} - t_j \le t_j' \end{aligned} \tag{3}$$

2.2 Fair fare allocation

If it is feasible for passenger i and passenger j to share a taxi, then the next question is how to allocate the taxi fare F_{ij} to the two passengers. It is fair to allocate F_{ij} such that the two passengers should pay the ratio of fare if they do not share, i.e.,

$$\frac{x_i}{x_j} = \frac{f_i}{f_j} \tag{4}$$

Together with $x_i + x_j = F_{ij}$, we can calculate the fare allocation:

$$\begin{cases} x_i^* = \dfrac{f_i F_{ij}}{f_i + f_j} \\ x_j^* = \dfrac{f_j F_{ij}}{f_i + f_j} \end{cases} \tag{5}$$

Note that the above allocation may not satisfy the requirements $x_i \le f_i - f_i', x_j \le f_j - f_j'$. If this occurs, we adjust the allocation so that $| f_i / f_j - x_i / x_j |$ is minimized:

$$\min_{x_i, x_j} \left| \frac{f_i}{f_j} - \frac{x_i}{x_j} \right| \tag{6}$$

subject to:

$$0 < x_i \le f_i - f_i' \tag{7}$$

$$0 < x_j \le f_j - f_j' \tag{8}$$

$$x_i + x_j = F_{ij} \tag{9}$$

Proposition 1: Define x_i^* and x_j^* as the ones shown in Eq. (5). Then the optimal solution to model (6) is:

$$(x_i, x_j) = \begin{cases} (x_i^*, x_j^*), & \text{if } x_i^* \le f_i - f_i', x_j^* \le f_j - f_j' \\ (f_i - f_i', F_{ij} - f_i + f_i'), & \text{if } x_i^* > f_i - f_i' \\ (F_{ij} - f_j + f_j', f_j - f_j'), & \text{if } x_j^* > f_j - f_j' \end{cases} \tag{10}$$

The proof is in Appendix 1.

3 Taxi sharing models

We have provided a commercially viable plan above for *two* passengers to share a taxi and to allocate the taxi fare. In this section, we further examine when there are more than two queuing passengers willing to share, how we can rationally allocate them in pairs. Here, a methodology for an optimal pairing plan is developed, which can maximize the total social welfare of all the sharing passengers. Note that once the system matches two passengers i and j, the fares x_i and x_j are determined in a way elaborated in Section 2.2.

3.1 A basic static model

To simplify, we assume that there are n passengers who are willing to share. Based on the location of their destinations, each passenger $i = 1, 2 \cdots n$ has a set of suitable sharing passengers, denoted by set S_i. To eliminate symmetrical solutions, we only consider the subsequent passengers $i+1, i+2 \cdots n$ in set S_i. For example, if there are three passengers and the first and the third passengers may share a taxi, we have $S_1 = \{3\}$, $S_2 = \varnothing$, and $S_3 = \varnothing$. Based on Section 2.1, set S_i is defined as

$$S_i := \left\{ j = i+1, i+2 \cdots n \mid \Delta_{ij} \geq f_i' + f_j', t_i + t_{ij} - t_j \leq t_j' \right\}, i = 1, 2 \cdots n-1 \tag{11}$$

We define a binary indicator χ_{ij} such that $\chi_{ij} = 1$ if $j \in S_i$ and $\chi_{ij} = 0$ otherwise. Note that the cardinality of the set S_i may be greater than 1 (a passenger may have more than one other passenger to share with). Therefore, the system should optimally match the passengers $1, 2 \cdots n$ to maximize the social welfare. In reality, it could be very difficult to estimate the social welfare increase caused by taxi sharing. Hence, in the sequel, we use the fare reduction Δ_{ij} in Eq. (2) as a surrogate.

Define a binary variable z_{ij} that equals 1 if $j \in S_i$ and the system assigns i and j to share a taxi, and 0 otherwise. The model is:

[P1]

$$\max \sum_{i=1}^{n-1} \sum_{j \in S_i} \Delta_{ij} z_{ij} \tag{12}$$

subject to:

$$\sum_{j \in S_k} z_{kj} + \sum_{i=1, k \in S_i}^{k-1} z_{ik} \leq 1, k \in \{1, 2 \cdots n\} \tag{13}$$

$$z_{ij} \in \{0,1\}, i \in \{1, 2 \cdots n-1\}, j \in S_i \tag{14}$$

Function (12) maximizes the total amount of increased social welfare. Constraints (13) guarantee that each passenger k will share with at most one passenger.

Model [P1] is an integer program and it has the following property, whose proof is in Appendix 2.

Proposition 2: The integrality constraints on z_{ij} in the above model cannot be removed. In other words, the linear programming relaxation of the above model may not have integer solutions.

However, after carefully examining the model, we find that model [P1] has the same structure as the maximum weight matching problem. Edmonds (1965) proved that the maximum weight matching problem can be solved in polynomial time. Hence, we have

Proposition 3: Model [P1] can be solved in polynomial time with regard to the number of passengers n.

We give a sketch of the idea of how [P1] is solved in polynomial time based on Schrijver (2003). The maximum weight matching problem is equivalent to the maximum weight perfect matching problem. The convex hull of all of the feasible solutions to the maximum weight perfect matching problem can be explicitly defined by a set of constraints, the number of which is exponential. However, there is a strongly polynomial time algorithm that can check whether a point satisfies all of the constraints, and if the answer is no, output a violated constraint. Thus, the separation oracle is polynomially solvable. The equivalence between optimization and separation via the ellipsoid method leads to a polynomial time algorithm for the maximum weight matching problem.

In practice, the passengers do not arrive in groups or regularly but randomly, e.g., following a Poisson distribution. The static model would give different results when different sets of passengers are dealt with. It is reasonable and convenient to solve the static model and update sharing results periodically based on a certain time interval denoted by T, say, $T=3$ minutes. Once the static model is calculated and the result is updated, the following coming passengers should wait in the queue for the next round of updates in time T.

If T is too large, then it may cause an unbearable waiting time for the early arriving passengers; on the other hand, if T is too small, there may be no new arriving passengers. Hence, for the static model, it is important to determine a suitable T.

3.2 Dynamic taxi sharing method

The taxi sharing scheme can also work in a real-time manner, termed as dynamic taxi sharing. Dynamic taxi sharing can be considered as the limit of static taxi sharing when the update time interval T approaches 0. The dynamic taxi sharing is easier to address algorithmically: when a passenger j requests at time t, the system just needs to check all of the waiting passengers, denoted by set Ω_t. If $\max_{i\in\Omega_t} \chi_{ij}\Delta_{ij} = 0$, i.e., $\chi_{ij} = 0$ for all $i \in \Omega_t$, then passenger j waits and the system sets $\Omega_t \leftarrow \Omega_t \cup \{j\}$; otherwise passenger j and the passenger identified below will be recommended to share a taxi:

[P2]
$$\arg\max_{i\in\Omega_t} \chi_{ij}\Delta_{ij} \tag{15}$$

The dynamic method is clearly more concise and a passenger can immediately know whether someone is assigned to share with her or she has to wait. However, compared with the static model, the dynamic model cannot give a system optimum. When there are many arrivals at the same time, the static model is more suitable to be used for a systematically optimal sharing plan. In such case, the sharing platform will receive several to many requests at the same time, and after a short stint of time (e.g., 5 seconds) the system can call the static method.

4 Taxi sharing with stochastic fare

If the taxi fare does not only depend on the distance but also the waiting time (and/or travel time), then due to fluctuations of traffic conditions, when the taxi arrives at j's destination, the fare shown on the meter is also case by case, which could be regarded as a stochastic value. As a result, the fare allocation in Section 2.2 is no longer applicable. In this section, we examine this stochastic case.

4.1 Expectation-based fare allocation

We define $\tilde{f}_i$ as the fare shown at i's destination, $\tilde{f}_j$ as the fare if j travels alone, and $\tilde{F}_{ij}$ the fare shown on the meter at j's destination if the two passengers share. $\tilde{f}_i$, $\tilde{f}_j$ and $\tilde{F}_{ij}$ are all stochastic values, which are actually unknown at the transport terminals before the trip. With a little abuse of notation, we define f_i, f_j, F_{ij} as the expected value of $\tilde{f}_i$,

$\tilde{f}_j$ and $\tilde{F}_{ij}$, respectively. Note that the system may give different values of f_i, f_j, F_{ij} in different days or in different periods of a day to account for traffic conditions.

Then, we propose a fare sharing scheme for the stochastic fares based on the expected fares:

Expectation-based fare allocation: When the taxi arrives at i's destination and the fare shown on the meter is $\tilde{f}_i$, passenger i pays $\tilde{x}_i(\tilde{f}_i) = x_i^* \tilde{f}_i / f_i$. When the taxi arrives at j's destination and the fare shown on the meter is $\tilde{F}_{ij}$, passenger j pays the rest fare, i.e., $\tilde{x}_j(\tilde{f}_i, \tilde{F}_{ij}) = \tilde{F}_{ij} - x_i^* \tilde{f}_i / f_i$.

Proposition 4: The expected value of $\tilde{x}_i(\tilde{f}_i)$ is equal to x_i^* and the expected value of $\tilde{x}_j(\tilde{f}_i, \tilde{F}_{ij})$ is equal to x_j^*.

Proof: $\mathbb{E}\left[\tilde{x}_i(\tilde{f}_i)\right] = \mathbb{E}\left(x_i^* \tilde{f}_i / f_i\right) = x_i^* \mathbb{E}\left(\tilde{f}_i / f_i\right) = x_i^* f_i / f_i = x_i^*$ and $\mathbb{E}\left[\tilde{x}_j(\tilde{f}_i, \tilde{F}_{ij})\right] = \mathbb{E}\left(\tilde{F}_{ij} - x_i^* \tilde{f}_i / f_{ii}\right) = \mathbb{E}\left(\tilde{F}_{ij}\right) - x_i^* \mathbb{E}\left(\tilde{f}_i / f_i\right) = F_{ij} - x_i^* f_i / f_i = F_{ij} - x_i^* = x_j^*$. □

Proposition 4 shows that the fare allocation is "stochastically" fair.

4.2 Using chance constraints to determine which passengers may share

Note that as the taxi fares $\tilde{f}_i$ and $\tilde{F}_{ij}$ are random values and unpredictable, there is no method to guarantee the gain of each passenger involved. Hence, the Expectation-based fare allocation gives a probabilistically better sharing plan; for instance, it is 60% possible that a passenger i would gain from the sharing. However, according to the psychology concept "loss aversion", this passenger would still be disappointed with the sharing plan, although she has a higher possibility to gain than to lose. Here, the loss aversion refers to people's tendency to strongly prefer avoiding losses to acquiring gains (Kahneman and Tversky, 1984); for example, the happiness from gaining $100 is smaller than the pain from losing $100.

To cope with this problem reflected by "loss aversion" and also to get an approximately better-off (each passenger could gain) solution, a new method is proposed. The new method imposes a service level, say $\alpha = 99\%$, meaning that both passengers benefit from taxi sharing with a probability of at least α. This method is termed chance constrained fare allocation, provided as follows.

Suppose that the system has V historical data records represented by $(f_i^{(v)}, f_j^{(v)}, F_{ij}^{(v)} t_i^{(v)}, t_j^{(v)}, t_{ij}^{(v)})$, $v = 1, 2 \cdots V$, where the superscript v indicates one particular data record. This method only needs one parameter β_{ij}, which is the percentage of the fare shown on the meter at i's destination to be paid by passenger i. That is, the fares paid by the two sharing passengers are:

$$x_i = \beta_{ij} \tilde{f}_i \tag{16}$$

$$x_j = \tilde{F}_{ij} - x_i \tag{17}$$

where $\tilde{f}_i$ and $\tilde{F}_{ij}$ can be observed from the metering in the trip.

Define function $\mathbf{1}(\cdot)$ to be equal to 1 if the logic in the parentheses is true and 0 otherwise. To achieve the service level α, we therefore solve the following optimization problem:

$$\begin{aligned} &\max_{0<\beta_{ij}<1} \frac{1}{V} \sum_{v=1}^{V} \mathbf{1}(\beta_{ij} f_i^{(v)} < f_i^{(v)}) \mathbf{1}(F_{ij}^{(v)} - \beta_{ij} f_i^{(v)} < f_j^{(v)}) \\ &= \max_{0<\beta_{ij}<1} \frac{1}{V} \sum_{v=1}^{V} \mathbf{1}(F_{ij}^{(v)} - \beta_{ij} f_i^{(v)} < f_j^{(v)}) \end{aligned} \tag{18}$$

subject to:

$$\frac{1}{V} \sum_{v=1}^{V} (f_i^{(v)} - \beta_{ij} f_i^{(v)}) \geq f_i' \tag{19}$$

$$\frac{1}{V} \sum_{v=1}^{V} [f_j^{(v)} - (F_{ij}^{(v)} - \beta_{ij} f_i^{(v)})] \geq f_j' \tag{20}$$

$$\frac{1}{V} \sum_{v=1}^{V} \left(t_i^{(v)} + t_{ij}^{(v)} - t_j^{(v)}\right) \leq t_j' \tag{21}$$

As $0 < \beta_{ij} < 1$, i will always pay a lower fare than not sharing. We thus compute the total number of records such that passenger j also pays a lower fare. If the probability is at least α, then it is commercially viable that the two passengers share. Constraint (19) enforces that passenger i on average saves at least f_i'. Constraint (20) guarantees that passenger j on average saves at least f_j'. Constraint (21) dictates that the travel time of passenger j increases averagely by at most t_j'.

Proposition 5: Let $\beta_{ij}^{(1)}$ be the largest value of β_{ij} such that Constraint (19) is binding, i.e.,

$$\beta_{ij}^{(1)} = 1 - \frac{Vf'}{\sum_{v=1}^{V} f_i^{(v)}} \tag{22}$$

Let $\beta_{ij}^{(2)}$ be the smallest value of β_{ij} such that Constraint (20) is binding, i.e.,

$$\beta_{ij}^{(2)} = \frac{Vf_j' + \sum_{v=1}^{V} (F_{ij}^{(v)} - f_j^{(v)})}{\sum_{v=1}^{V} f_i^{(v)}} \tag{23}$$

Let $\beta_{ij}^{(0)}$ be the smallest value of β_{ij} such that the objective value is at least α, i.e.,

$$\beta_{ij}^{(0)} = \arg\min_{\beta_{ij}} \left\{ \frac{1}{V} \sum_{v=1}^{V} \mathbf{1}(F_{ij}^{(v)} - \beta_{ij} f_i^{(v)} < f_j^{(v)}) \geq \alpha \right\} \tag{24}$$

Then (i) if Constraint (21) does not hold or $\beta_{ij}^{(0)} \geq 1$, the model has no feasible solution and the two passengers cannot share; (ii) else if $\beta_{ij}^{(2)} > \beta_{ij}^{(1)}$, the model has no feasible solution and the two passengers cannot share; (iii) else if $\beta_{ij}^{(0)} > \beta_{ij}^{(1)}$ and $\beta_{ij}^{(2)} \leq \beta_{ij}^{(1)}$, the model has feasible solutions but all of them correspond to service levels lower than α, and the two passengers cannot share; (iv) else, the optimal objective of the model is at least α, and it is commercially viable for the two passengers to share; moreover, any β_{ij} in the interval of $[\max(\beta_{ij}^{(0)}, \beta_{ij}^{(2)}), \beta_{ij}^{(1)}]$ leads to an objective value of at least α.

The proof is in Appendix 3.

If it is commercially viable for the two passengers to share, to determine β_{ij}, we maximize the fairness subject to the chance constraint (level-of-service constraint).

$$\min_{\beta_{ij}\in[\max(\beta_{ij}^{(0)},\beta_{ij}^{(2)}),\beta_{ij}^{(1)}]}\left|\frac{\sum_{v=1}^{V} f_i^{(v)}}{\sum_{v=1}^{V} f_j^{(v)}}-\frac{\sum_{v=1}^{V}\beta_{ij} f_i^{(v)}}{\sum_{v=1}^{V}(F_{ij}^{(v)}-\beta_{ij} f_i^{(v)})}\right| \tag{25}$$

With a little abuse of notation, we define

$$f_i := \frac{1}{V}\sum_{v=1}^{V} f_i^{(v)},\ f_j := \frac{1}{V}\sum_{v=1}^{V} f_j^{(v)}, F_{ij} := \frac{1}{V}\sum_{v=1}^{V} F_{ij}^{(v)} \tag{26}$$

Eq. (25) becomes

$$\min_{\beta_{ij}\in[\max(\beta_{ij}^{(0)},\beta_{ij}^{(2)}),\beta_{ij}^{(1)}]}\left|\frac{f_i}{f_j}-\frac{\beta_{ij} f_i}{F_{ij}-\beta_{ij} f_i}\right| \tag{27}$$

Proposition 6: Let $\beta_{ij}^{(3)}$ be the value of β_{ij} such that the objective function is 0, i.e.,

$$\beta_{ij}^{(3)} = \frac{F_{ij}}{f_i + f_j} \tag{28}$$

Then (i) if $\beta_{ij}^{(3)} \in [\max(\beta_{ij}^{(0)}, \beta_{ij}^{(2)}), \beta_{ij}^{(1)}]$, the optimal β_{ij} is $\beta_{ij}^{(3)}$; (ii) else if $\beta_{ij}^{(3)} > \beta_{ij}^{(1)}$, the optimal β_{ij} is $\beta_{ij}^{(1)}$; (iii) else, $\beta_{ij}^{(3)} < \max(\beta_{ij}^{(0)}, \beta_{ij}^{(2)})$ and the optimal β_{ij} is $\max(\beta_{ij}^{(0)}, \beta_{ij}^{(2)})$.

The proof to Proposition 6 follows a similar manner to the proof to Proposition 1 and hence is omitted.

5 Considerations in Practical Implementations

It should be pointed out that, using any methods/models proposed above, it is still common that some queuing passengers are not allocated to share with anyone. For example, the destination of passenger i is distant from all the other passengers. Mathematically, it means the optimal solution of Model (12)-(14) z_{ij}^* may give

$$\sum_{k=1}^{i-1} z_{ki}^* + \sum_{j=i+1}^{n} z_{ij}^* = 0 \tag{29}$$

for any particular passenger i. In such case, passenger i is suggested to further wait for updates or give up sharing. This is one phenomenon that may occur in practice. In this section, we further discuss some practical issues for better implementations of the taxi sharing platform.

5.1 Fairness considering passengers' waiting time

As discussed in the paragraph above, some passengers may be left over in several updates of the taxi sharing models/methods, which gives rise to a longer waiting time. It is very frustrating for such sort of passengers and would give them a very bad impression/experience on the sharing system. Consequently, if they refuse to use the sharing system next time, it is a loss to other passengers and the whole system.

To ensure fairness and also give them some incentive to further use the system, we should prioritize the passengers who request early and have waited longer. Let π_k be the time of request to share by passenger k. When the system solves the static sharing model [P1] at time t, passenger k has waited for a duration of $t-\pi_k$. We can add a weight to the pair of passengers i and j that is monotonically increasing with the waiting time $t-\pi_i$ and $t-\pi_j$. For instance, we can change the objective function (12) of [P1] as follows:

[P1']
$$\max \sum_{i=1}^{n-1} \sum_{j \in S_i} (t-\pi_i)(t-\pi_j)\Delta_{ij} z_{ij} \tag{30}$$

Similarly, in the dynamic model [P2], when passenger i arrives and requests at time t, the following passenger will be recommended to share a taxi with i:

[P2']
$$\arg\max_{j \in \Omega_t} (t-\pi_j)\chi_{ij}\Delta_{ij} \tag{31}$$

We also found another issue of the fairness, which is on the second passenger j: j's trip time is increased because of sharing, but passenger i's time is not affected. Hence, passenger i gets more benefit, especially on the deterministic case where the charge is only based on distance. To cope with this problem, an empirical parameter σ ($\sigma > 1$) can also be added to rationally increase i's share on the fare, namely, Eq. (4) can be modified as:

$$\frac{x_i}{x_j} = \frac{\sigma f_i}{f_j} \quad . \tag{32}$$

5.2 Fairness considering passengers' traveling time

We also found another issue of the fairness, which is on the second passenger j: j's trip time is increased because of sharing, but passenger i's time is not affected. Hence, passenger i gets more benefit, especially on the deterministic case where the charge is only based on distance. To cope with this problem, we need to incorporate a new parameter, θ, the value of time of passengers. Note that it is impossible to obtain the true value of θ for passenger j and hence θ is the estimated average value of time of all passengers.

Then, in the deterministic case of Section 2.2, the extra travel time for passenger j, which is equal to $t_i + t_{ij} - t_j$, implies an extra cost of $\theta(t_i + t_{ij} - t_j)$ for passenger j. Consequently, the fare allocation scheme in Eq. (4) can be modified as:

$$\frac{x_i}{x_j + \theta(t_i + t_{ij} - t_j)} = \frac{f_i}{f_j} \tag{33}$$

Together with $x_i + x_j = F_{ij}$, we can calculate the fare allocation:

$$\begin{cases} x_i^* = \dfrac{f_i F_{ij} + f_i \theta(t_i + t_{ij} - t_j)}{f_i + f_j} \\ x_j^* = \dfrac{f_j F_{ij} - f_i \theta(t_i + t_{ij} - t_j)}{f_i + f_j} \end{cases} \tag{34}$$

Then, in the expectation-based fare allocation of Section 4.1, when the taxi arrives at i's destination and the fare shown on the meter is $\tilde{f}_i$, passenger i pays $\tilde{x}_i(\tilde{f}_i) = x_i^* \tilde{f}_i / f_i$, where x_i^* is defined in Eq. (34) and is the expected fare that passenger i should pay considering the extra travel time of passenger j. When the taxi arrives at j's destination and the fare shown on the meter is $\tilde{F}_{ij}$, passenger j pays the rest fare, i.e., $\tilde{x}_j(\tilde{f}_i, \tilde{F}_{ij}) = \tilde{F}_{ij} - x_i^* \tilde{f}_i / f_i$.

Similarly, Eqs. (35) and (36) in the chance constrained model in Section 4.2 should be revised to the following two equations, respectively:

$$\begin{aligned}&\max_{0<\beta_{ij}<1}\frac{1}{V}\sum_{v=1}^{V}\mathbf{1}(\beta_{ij}f_i^{(v)}<f_i^{(v)})\mathbf{1}(F_{ij}^{(v)}-\beta_{ij}f_i^{(v)}+\theta(t_i^{(v)}+t_{ij}^{(v)}-t_j^{(v)})<f_j^{(v)})\\&=\max_{0<\beta_{ij}<1}\frac{1}{V}\sum_{v=1}^{V}\mathbf{1}(F_{ij}^{(v)}-\beta_{ij}f_i^{(v)}+\theta(t_i^{(v)}+t_{ij}^{(v)}-t_j^{(v)})<f_j^{(v)})\end{aligned} \tag{35}$$

$$\frac{1}{V}\sum_{v=1}^{V}[f_j^{(v)}-(F_{ij}^{(v)}-\beta_{ij}f_i^{(v)}+\theta(t_i^{(v)}+t_{ij}^{(v)}-t_j^{(v)}))]\geq f_j' \tag{36}$$

5.3 Sharing among passengers in team

The taxi sharing scheme has made the taxi to be a public transport mode, a mini-bus. Hence, similar to the bus fares, when there are more than two passengers sharing, each of them should still be charged based on her trip length. The above models/methods can be easily adjusted to handle this extended case. We take the case that three mutually unknown passengers i, j and k as a representative example. In this case, there are three destinations. Without loss of generality, we assume that these three passengers should be served in the sequence i, j, k. Then, Model (6) is revised to

$$\min_{x_i,x_j,x_k}\max\left\{\left|\frac{f_i}{f_j}-\frac{x_i}{x_j}\right|,\left|\frac{f_i}{f_k}-\frac{x_i}{x_k}\right|,\left|\frac{f_j}{f_k}-\frac{x_j}{x_k}\right|\right\} \tag{37}$$

subject to:

$$0<x_i\leq f_i-f_i' \tag{38}$$

$$0<x_j\leq f_j-f_j' \tag{39}$$

$$0<x_k\leq f_k-f_k' \tag{40}$$

$$x_i+x_j+x_k=F_{ijk} \tag{41}$$

where F_{ijk} is the final transit fare when the taxi arrives at the final passenger k's destination.

5.4 Payment methods and queuing

First, regarding the payment, as mentioned in Section 3.2, a mobile app is the most suitable platform for the taxi sharing scheme, which is similar to the other car sharing platform, e.g., Uber. It is also convenient and informative to handle the transactions through the mobile app;

for example, it gives passenger i the information about the final fare F_{ij} ($\tilde{F}_{ij}$) as well as total trip time to j's destination. Meanwhile, compared with paying by cash or card, it saves the passengers and taxi drivers from additional transaction time. The app should have a passenger version and driver version. When first using the app, the passengers and taxi drivers should be required to provide their phone number and then link the app to a bank account.

Note that when using such an app, not all passengers will request at the same time. The procedure should be as follows: (i) a passenger AA requests; (ii) the system either lets the passenger wait to find a suitable person to share; or finds a suitable person BB and informs the two passengers (tour, estimated time, estimated prices); (iii) the two passengers contact each other and take a taxi. In the above procedure, it is possible that a passenger, say BB, requests and then just forgets her request; it is also possible that, say BB, is unhappy with the tour/price of the sharing. In this case, either AA and/or BB can request the system to reassign a passenger (or wait if there is no suitable passenger) or just cancel the request. Of course, if AA is unhappy e.g. 5 times, then she has to log in again (or cannot log in in 5 minutes) to avoid malicious passengers.

Second, we further discuss the queuing. For all sorts of taxi calling and car sharing services, the drivers often complain about one bad behavior of passengers: the passengers sometimes request for service far before they are ready to board the vehicle, e.g., a passenger may submit a request when they are at home and quite far away from the road. It thus causes unbearable waiting time and loss to the drivers. This phenomenon may also occur for the taxi sharing scheme addressed in this paper: if a passenger requests for sharing before they join the queue (e.g., just get off the plane), it affects not only the drivers but also other sharing passengers. Hence, to solve this problem, a queuing machine can be equipped at the queue entrance, and each waiting passenger should get a queuing number/ID to make a request. Such a queuing ID also protects the passengers' privacy, which is very necessary considering that they are sharing with unknown passengers. Therefore, when two or three passengers achieve an agreed sharing plan, they can address each other by the queuing ID.

Another concern for the queuing issue is that the destination of each passenger is needed, and how to efficiently organize all the matched passengers is important for practical implementations. In this regard, the implemented taxi pooling scheme in Paddington and Euston stations at London (ATOC, 2017) has provided good examples, which is a marshalled service. Therefore, passengers travelling to the same destination zone could be clustered together. In addition, with the aids of the online platform, it is very efficient for the passengers to provide their exact destinations, and they can even pre-upload their destinations before joining the queue. Comparing with the long queuing time for taxis, it is not a time-consuming issue even if they provide the destination on site. In practice, on-site administrators are needed to maintain order and ensure the operations running smoothly.

6 Case Study

6.1 The Static Model

A case study based on the real survey data is used in this section to validate the static model [P1]. As shown in Figure 1, the survey is conducted in early November 2015 at the Wuxi Train Station in Wuxi, China. The demand for taxi is high and we see a long queue at the taxi stand, thus the waiting time for taxi is also quite long. Hence, most of the passengers are willing to share. The queuing passengers are asked about some relevant questions, the answers of which are given in Table 1. Totally 25 samples are collected, yet the destinations of 5 passengers are not tractable (this problem can be avoided when a mobile app is used), and thus these 5 samples are invalid.

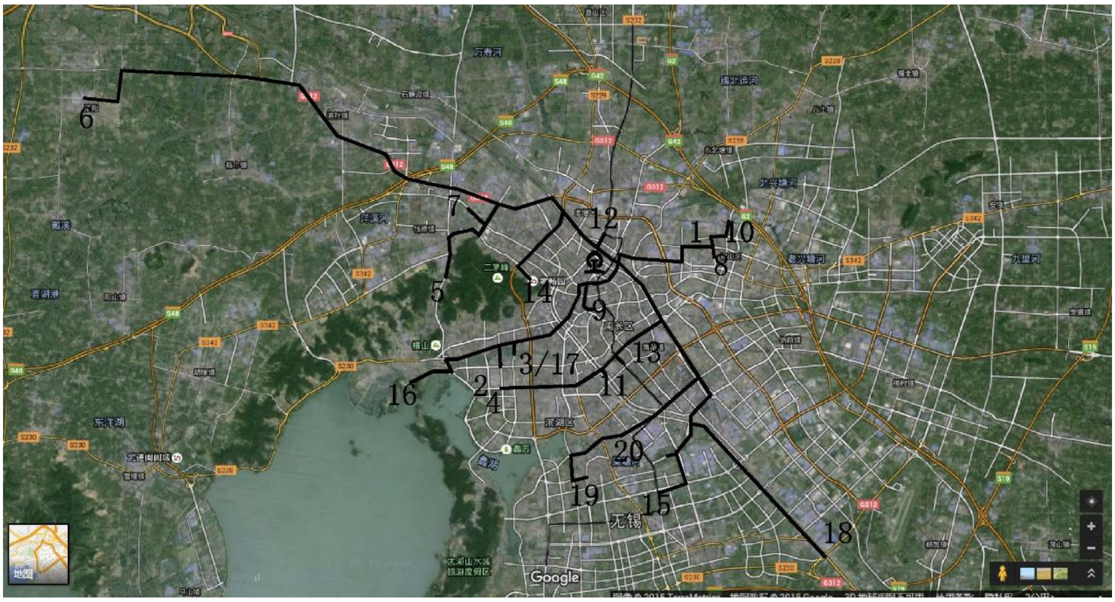

Figure 1 The destinations of taxi passengers at Wuxi Train Station on Google Map

Figure 1 shows the shortest-path tree (bold black lines) from the Wuxi Train Station to the destinations of the 20 passengers, where each destination is marked by the ID of the passenger. These destinations are widely distributed in the city of Wuxi. As inputs for the model calculation, the distance of shortest path between destinations of passenger i and j is also needed.

Table 1 Taxi passenger survey in Wuxi Train Station

No.	**Destination**	**Lat(N)**	**Lng(E)**
1	Sangdayuan Community	31°35'37.3"	120°20'55.7"
2	Xiyuanli Community	31°33'06.2"	120°15'24.4"
3	Wanda Plaza	31°33'38.7"	120°15'52.1"
4	Crowne Plaza	31°32'27.1"	120°15'27.3"
5	Zoo	31°35'00.6"	120°14'09.8"
6	Luoshe Building	31°38'50.2"	120°11'22.9"
7	Baijin Hotel	31°36'26.2"	120°14'46.6"
8	Columbus Plaza	31°35'34.5"	120°20'07.0"
9	Merchants Tower	31°34'56.9"	120°17'54.0"
10	Sevilla Flat	31°35'49.7"	120°21'24.1"

11	No.18, Yonghe Road	31°33'06.6"	120°18'00.9"
12	Qianqiao Market	31°36'07.5"	120°17'34.6"
13	Hongdou Flat	31°34'53.4"	120°17'55.6"
14	Huishan Town	31°35'06.0"	120°16'34.4"
15	Nanhu Community	31°30'27.5"	120°19'29.5"
16	Junlai Hotel	31°31'07.5"	120°15'51.6"
17	Wanda Plaza	31°33'38.7"	120°15'52.1"
18	Hongli Hotel	31°29'04.2"	120°23'32.7"
19	Zuibaxian Hotel	31°30'26.3"	120°17'56.8"
20	Donglin Plaza	31°34'42.4"	120°18'32.1"

In Table 1, the location of each destination is indicated by its Latitude (Lat) and Longitude (Lng). The Lat and Lng of the Wuxi Train station is 31°35'09.7" and 120°18'16.0", respectively. The taxi charge in Wuxi city is 1.9 RMB/km, which is used in this numerical example. In addition, the traffic conditions are not considered, thus the average travel speed of each taxi is assumed to be 60km/h. We assume that all of the passengers are ready at the taxi stand and we use the static model [P1].

In our first analysis, we assume that two passengers could share a taxi as long as overall they save money, i.e., we assume $f_i' = f_j' = 0$ and $t_j' = +\infty$. Hence, the set $S_i := \left\{ j = i+1, i+2 \cdots n \mid \Delta_{ij} \le 0 \right\}, i = 1, 2 \cdots n-1$. The results are below in Table 2.

Table 2. Optimal results of deterministic taxi sharing plan

Passenger ID	Pay if not sharing (RMB)	Share with passenger ID*	Pay in the sharing scheme (RMB)
1	9.77	10	5.34
2	13.48	4	7.61
3	10.75	17	5.38
4	15.29	2	8.63
5	14.79	14	10.47
6	29.24	7	20.13
7	13.67	6	9.41
8	6.88	**0**	**6.88**

9	1.60	13	0.91
10	11.63	1	6.35
11	8.72	20	7.34
12	4.77	**0**	**4.77**
13	1.68	9	0.96
14	6.10	5	4.32
15	20.36	18	14.04
16	19.14	19	13.30
17	10.75	3	5.38
18	32.00	15	22.06
19	20.00	16	13.90
20	2.15	11	1.81
Sum	**252.76**	—	**168.99**

*****"0" means not sharing

It can be seen from Table 2 that, among the 20 passengers, only two passengers (No. 8 and 12) are not allocated for pairing, and all the others are in pairs. Therefore, these 20 passengers are served totally by 11 taxis. The total amount of charge for the "not sharing" case is 252.76 RMB, which is 49.6% higher for the "sharing" case, indicating a big saving on the passengers' travel expenses. As to the taxi drivers, the average revenue in the "not sharing" case is 252.76/20=12.64 RMB, while the average revenue in the "sharing" case is 168.99/11=15.36 RMB, which is 22% higher. Hence, the drivers also have an evident gain from the sharing scheme.

Among all the sharing passengers, there are two having the biggest gain, which are passengers 3 and 17. We can see from Table 1 that passengers 3 and 17 are heading to the same destination "Wanda Plaza", hence giving rise to a 50% saving of their trip expenses.

In our second analysis, we assume $f_i' = \max(2, 0.1 f_i)$, $f_j' = \max(2, 0.1 f_j)$ and $t_j' = 0.5 t_j$. In other words, a passenger should save at least 2 RMB by sharing and at least 10% of her original fare; moreover, the second passenger's travel time cannot exceed the original travel time by over 50%. The results are provided in Table 3.

Table 3. Optimal results of deterministic taxi sharing plan with constraints

Passenger ID	**Pay if not sharing (RMB)**	**Share with passenger ID***	**Pay in the sharing scheme (RMB)**

1	9.77	10	5.34
2	13.48	4	7.61
3	10.75	17	5.38
4	15.29	2	8.63
5	14.79	14	10.69
6	29.24	7	20.13
7	13.67	6	9.41
8	6.88	**0**	**6.88**
9	1.60	**0**	**1.60**
10	11.63	1	6.35
11	8.72	**0**	**8.72**
12	4.77	**0**	**4.77**
13	1.68	**0**	**1.68**
14	6.10	5	4.10
15	20.36	18	14.04
16	19.14	19	13.30
17	10.75	3	5.38
18	32.00	15	22.06
19	20.00	16	13.90
20	2.15	**0**	**2.15**
Sum	**252.76**	—	**172.12**

*****"0" means not sharing

Compared with the results in Table 2, data in Table 3 have more "0"s in the third column, meaning that there are more passengers not allocated to any sharing plan. It is not surprising, as the passengers have more constraints on the sharing and they require a basic gain from the sharing. However, for others that can inherently benefit from the sharing, they are still allocated to the same partners. In particular, the only difference between the two results in Table 2 and 3 is that: the shared pairs "9-13" and "11-20" in Table 2 have not come to an agreement for sharing in Table 3. However, this does not result in any big change on the system costs, considering that the total amount of pay only increases from 168.99 to 172.12 RMB.

The total cost of the "not sharing" case is higher than that of the constrained sharing case shown in Table 3 by 46.9% (higher by 49.6% in Table 2). For the 20 passengers, totally 13 vehicles are needed in the sharing plan, and the passenger's average expense is reduced from 12.64 to 8.61. The average income of taxi drivers has increased from 252.76/20=12.64 to 172.12/13=13.24 RMB. Clearly, no one loses from the sharing scheme and some of them have an evident gain (up to 50%); hence, it is a better-off allocation that considers the benefit of everyone in the system.

Herein, we further test the model proposed in Section 5.2, where the increased travel time of the second passenger j is considered, for better fairness. In this example, the value of time θ is taken as 0.5 RMB/hour. The model proposed in Section 5.2 is then utilized for the optimal sharing plan. Without affecting results of interest, the circumstances for the case shown in Table 2 is followed, and the results of this new test are tabulated in Table 4.

Table 4. Optimal results of the taxi sharing plan considering the fairness for passenger j

Passenger ID	**Pay if not sharing (RMB)**	**time cost if not sharing (RMB)**	**Share with passenger ID***	**Pay in the sharing scheme (RMB)**	**time cost if sharing (RMB)**	**general cost saving if sharing**
1	9.77	2.57	10	5.34	2.57	4.43
2	13.48	3.55	4	7.61	3.55	5.75
3	10.75	2.83	17	5.38	2.83	5.38
4	15.29	4.02	2	8.63	4.27	6.52
5	14.79	3.89	14	10.47	3.89	4.32
6	29.24	7.70	7	20.13	7.78	9.05
7	13.67	3.60	6	9.41	3.60	4.23
8	6.88	**1.81**	**0**	**6.88**	**1.81**	**0**
9	1.60	0.42	13	0.91	0.42	0.66
10	11.63	3.06	1	6.35	3.08	5.27
11	8.72	2.29	20	7.34	2.41	1.28
12	4.77	**1.26**	**0**	**4.77**	**1.26**	**0**
13	1.68	0.44	9	0.96	0.49	0.70
14	6.10	1.61	5	4.32	1.61	1.78
15	20.36	5.36	18	14.04	5.36	5.90
16	19.14	5.04	19	13.30	5.04	4.91
17	10.75	2.83	3	5.38	2.83	5.38
18	32.00	8.42	15	22.06	9.50	9.28
19	20.00	5.26	16	13.90	7.16	5.13
20	2.15	0.57	11	1.81	0.57	0.32
Sum	**252.76**	**66.52**	—	**168.99**	**70.00**	**80.29**

*“0” means not sharing

By constrasting Table 2 and 4, it can be seen that the optimal sharing plan is not changed by the consideration of fairness for passenger j; while the fare paid by the sharing passenger

i and j are changed. We further point out that the overall fare (gain of the driver) for each sharing case is also not changed. The change on the split of fare between passenger i and j is because the increased loss/general-cost of j is larger when the trip time is considered.

6.2 Comparison of Static and Dynamic Methods

To provide a more in-depth test, we conduct randomly generated numerical experiments to compare different models proposed in the study. We consider a square city with the size of 20×20 km. This example assumes that the location of its railway station is at a corner of the city, and passengers' destinations are randomly generated in the city. We assume that the travel distance between two locations is 1.2 times the line segment connecting the two locations and that the travel time is proportional to the distance. Passengers arrive in a Poisson manner. We assume that the taxi charge is 1.9 RMB/km, and the parameters of feasible sharing are $f_i' = \max(2, 0.1 f_i)$, $f_j' = \max(2, 0.1 f_j)$ and $t_j' = 0.5 t_j$. We simulate a period of 5 hours, but discard the first 15 minutes which is used to warm up and the last 15 minutes which has the end-of-period effect. Hence, we actually consider 4 hours and a half. We assume that if a passenger does not find someone to share 10 minutes after arriving at the taxi stand, she will just take a taxi alone.

Three different models are considered: (i) the static system that is updated every 3 minutes, termed as Static-3; (ii) the static system that is updated every 1 minute, termed as Static-1; and (iii) the dynamic system that is updated whenever a customer arrives, termed as Dynamic. Moreover, we also considered weighted static and dynamic models, namely, models [P1'] and [P2'].

First, based on the empirical data, we assume that the arrival rate is 700 passengers/hour, reflecting peak periods, but only 500 of them would like to share a taxi. In the 4.5 hours of simulation, there are a total of 2284 customers, as shown in Figure 2, in which (0,0) is the railway station. Clearly, all the destinations are evenly distributed in the city area (study area).

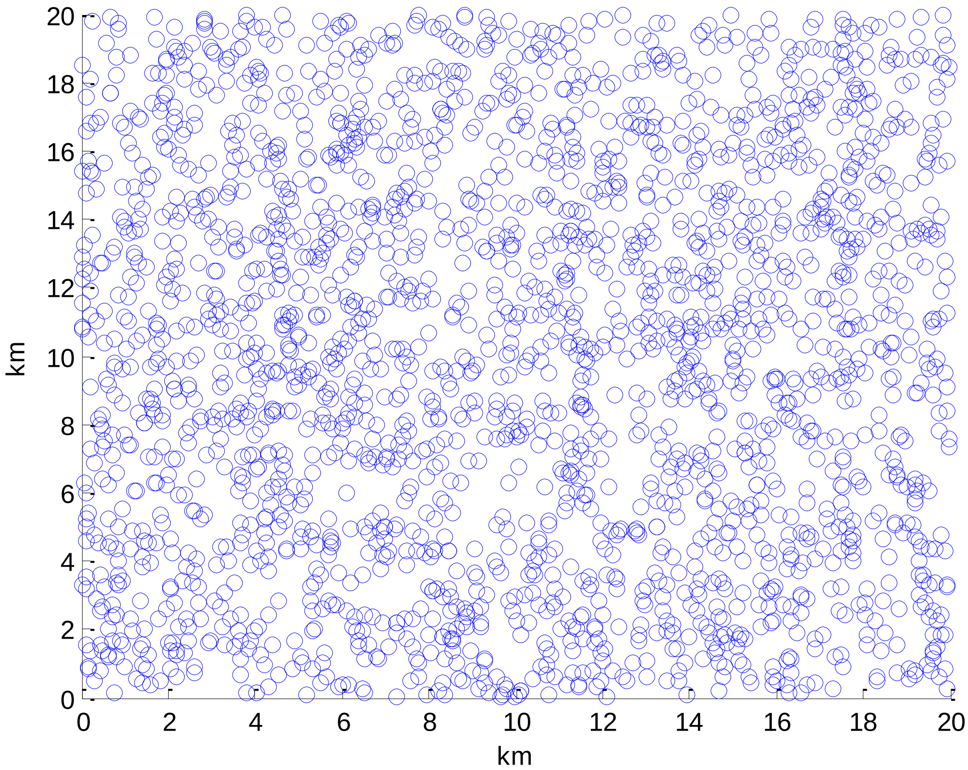

Figure 2. The Distribution of Passengers' Destinations (Peak Hour)

The results of different models are shown in Table 5. First, we observe that all of the taxi-sharing models significantly reduce the taxi fare a passenger pays by about 30% (or 12 RMB). Meantime, the travel time is only slightly increased by about 10% (2 minutes). These results are very encouraging, showing the considerable potential social impact by taxi sharing. Second, if the taxi sharing system is updated more frequently from every three minutes to every one minute to instantaneously, passengers wait for a shorter period to find someone to share; however, the partners to share are less ideal, as reflected by the higher fare and longer travel time. Hence, the system needs to balance the trade-off and find a suitable frequency of updating. From the results of weighted models, we can find that the average fare after sharing increase 6.5%, 5.0% and 0.5%, respectively. And, the average travel time after sharing increase 1.1%, 1.1% and 0.5%. Simultaneously, the number of passengers not matched has a clear decline: 14.3%, 14.3% and 0. The weighted models, albeit more fair as the waiting times of passengers are less diverse and fewer passengers could not find someone to share, are slightly inferior to their unweighted counterparts as passengers pay higher fare and travel longer.

Table 5. The Results of three models in peak hour (Arrival: 500 passengers/hour)

	Not weighted			Weighted Models		
	Static-3	Static-1	Dynamic	Static-3	Static-1	Dynamic
Average fare without sharing (RMB)	34.69	34.69	34.69	34.69	34.69	34.69
Average travel time without sharing (min)	18.3	18.3	18.3	18.3	18.3	18.3
Average fare with sharing	19.31	20.71	23.72	20.56	21.75	23.83
Average travel time after sharing (min)	18.6	18.9	19.7	18.8	19.1	19.8
Average waiting time to share (s)	98	39	11	99	39	11
Number of passengers not matched	21	21	18	18	18	18

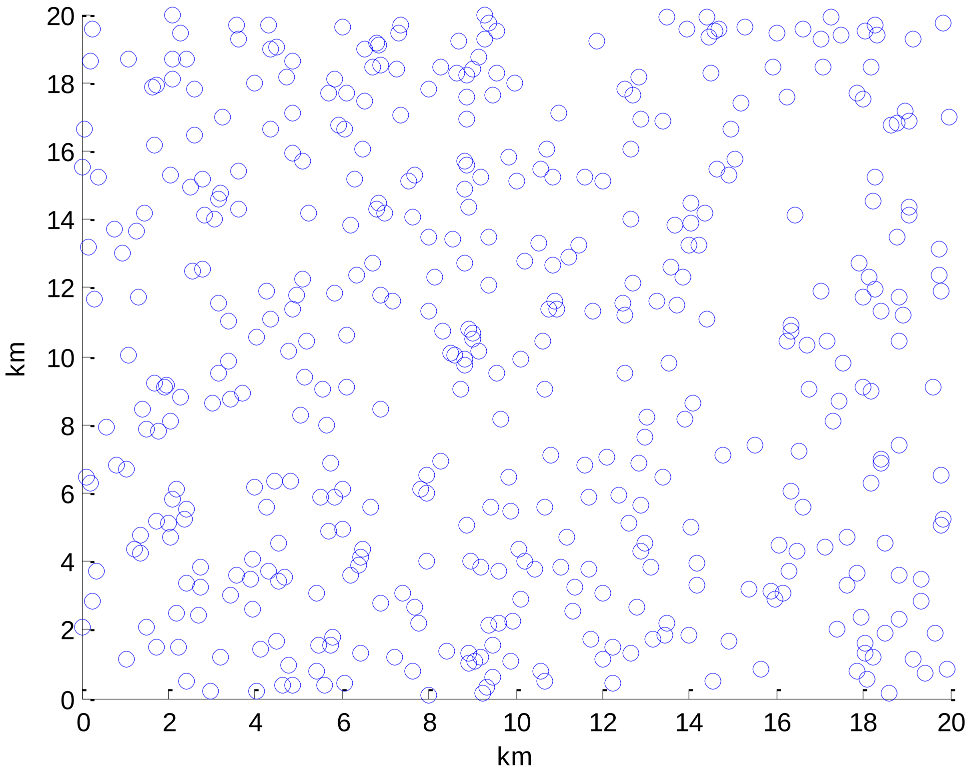

Figure 3. The Distribution of Passengers' Destinations (Off-Peak)

Then, we further consider the off-peak periods, and assume that the arrival rate is 100 passengers/hour who would like to share. In the 4.5 hours of simulation, there are a total of 443 customers, as shown in Figure 3, in which (0,0) is the railway station. Clearly, the density of destinations in Figure 3 is smaller than that in Figure 2. The results of different models are shown in Table 6. We can see the general conclusions are consistent with the results with an arrival rate of 500 potentially sharing passengers/hour. There is only one point that is noteworthy: when fewer passengers request to share, the average waiting time for a passenger to find someone to share is significantly longer, as shown by the row "average waiting time to share" of the two tables.

Furthermore, the results in Table 5 and Table 6 indicate that customers can benefit more with the increase of travel demand. In the case of 500 potentially sharing passengers/hour, especially in the Static-3, the average fare with sharing and average travel time after sharing are significantly less than the results in other models. Hence, we can achieve a better solution when there are more waiting passengers. It also implies that the static model can get a system optimal easier than the dynamic model.

Table 6. The Results of three models in off-peak hour (Arrival: 100 passengers/hour)

	Not weighted			Weighted Models		
	Static-3	Static-1	Dynamic	Static-3	Static-1	Dynamic
Average fare without sharing (RMB)	34.23	34.23	34.23	34.23	34.23	34.23
Average travel time without sharing (min)	18.0	18.0	18.0	18.0	18.0	18.0
Average fare with sharing	21.28	22.33	23.31	22.06	22.92	23.50
Average travel time after sharing (min)	18.7	18.9	19.3	18.9	19.1	19.4
Average waiting time to share (s)	120	60	30	121	61	30
Number of passengers not matched	6	1	1	1	1	1

7 Conclusion and Future Research

This paper addressed the better-off allocation of taxi sharing scheme at big transport terminals. The passengers waiting at the taxi stand can share a taxi and split the taxi fare to save their trip expenses. For a group of queuing passengers waiting to share, there are two key questions: how to assign them to share one taxi; how much each of them should pay. This paper first provided a fair fare allocation method for any two grouped passengers, based on which two models are developed to determine the optimal taxi sharing plan: the first one is a static model that maximizes the total social benefit of all the passengers requesting for sharing; the second one is a dynamic model that updates the sharing plan whenever a new arrival requested for sharing. This static model is more suitable for a crowded case when a big number of passengers come and request together, while the dynamic model is useful for an uncongested case.

The taxi fare addressed in the above two models are only based on distance, thus it is predictable and deterministic. However, the taxi fare in many cities also contains the time-based charge, and hence the taxi fare on a fixed itinerary is not predictable, which is case-by-case and stochastic. To cope with the stochastic fare, an optimization model based on the expectation value was then developed, which gives a profitable plan to the passengers on average. However, the averagely profitable plan implies that the passengers may still suffer a loss in some cases. As per the psychology concept "loss aversion", even a low chance of loss would also give a very negative impression to the participators of taxi sharing scheme. Consequently, a chance constrained stochastic program was proposed to guarantee a better-off allocation.

This paper provided an initial step on the optimal determination of taxi sharing scheme, where some practical issues (fairness, sharing in team, queuing, payment method, etc.) for real implementations were also discussed. However, the practical situations are more complicated than the cases discussed in this paper. Hence, future efforts are needed to further address other practical issues and incorporate them into the model; for instance, the impacts of shared taxis/vehicles on the fundamental characteristics of traffic flows (Qu et al. 2015&2017); furthermore, the taxi sharing scheme has reduced each passenger's trip expenses, thus it would change the modal split and increase the mode share of taxi users. Hence, for the sake of modelling, the demand elasticity and multimodal equilibrium should be further investigated.

Appendix 1: Proof of Proposition 1

Proof: If $x_i^* \le f_i - f_i', x_j^* \le f_j - f_j'$, then $(x_i, x_j) = (x_i^*, x_j^*)$ is a feasible solution with objective value of 0. Since the objective value must be nonnegative due to the absolute value operator, $(x_i, x_j) = (x_i^*, x_j^*)$ is optimal.

If $x_i^* > f_i - f_i'$, then we must have $x_i < x_i^*$ to make the model feasible and we can rewrite the objective function (6) as $\left|\frac{f_i}{f_j} - \frac{x_i}{x_j}\right| = \left|\frac{f_i}{f_j} - \frac{x_i}{F_{ij} - x_i}\right|$. By definition, $\frac{f_i}{f_j} = \frac{x_i^*}{x_j^*} = \frac{x_i^*}{F_{ij} - x_i^*}$; moreover, $\frac{x_i}{F_{ij} - x_i}$ is an increasing function of x_i. Hence, we have $\frac{x_i}{F_{ij} - x_i} < \frac{f_i}{f_j}$ for $x_i < x_i^*$. The objective function (6) is thus equal to $\frac{f_i}{f_j} - \frac{x_i}{F_{ij} - x_i}$. Hence, at the optimal solution x_i should take its largest possible value $f_i - f_i'$. The case when $x_j^* > f_j - f_j'$ can be proved similarly. □

Appendix 2: Proof of Proposition 2

Proof: We prove the proposition using the example in Figure 4. The red numbers are distance (cost). We have $\Delta_{12}' = (3+3) - (3+1) = 2$, $\Delta_{13}' = 2$, $\Delta_{23}' = 2$. The optimal solution to the linear programming relaxation is $z_{12} = z_{13} = z_{23} = 0.5$ with the objective value 3. However, an optimal solution to the original integer program is $z_{12} = 1, z_{13} = z_{23} = 0$ with the objective value 2.

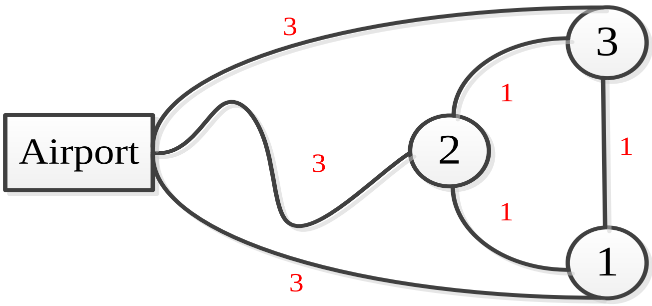

Figure 4. An illustrative example for taxi sharing

Appendix 3: Proof of Proposition 5

Proof: If Constraint (21) does not hold, then Eq. (3) is violated as passenger j experiences a very long travel time. The objective function (18) is an increasing function of β_{ij}; hence by definition if $\beta_{ij}^{(0)} \geq 1$, then there is no $\beta_{ij} \in (0,1)$ such that the objective function (18) is not smaller than α. This proves point (i) of the proposition.

The requirement that Constraint (19) is feasible is equivalent to $\beta_{ij} \leq \beta_{ij}^{(1)}$; the requirement that Constraint (20) is feasible is equivalent to $\beta_{ij} \geq \beta_{ij}^{(2)}$; the requirement that the objective function (18) is not smaller than α is equivalent to $\beta_{ij} \geq \beta_{ij}^{(0)}$. These facts imply points (ii), (iii) and (iv) of the proposition. □

Acknowledgment

This study is supported by the Youth Program (No. 71501038), General Project (No. 71771050), Key Projects (No. 51638004) of the National Natural Science Foundation of China, and the Natural Science Foundation of Jiangsu Province in China (BK20150603).

References

Agatz, N., Erera, A., Savelsbergh, M., & Wang, X. (2011). Dynamic ride-sharing: A simulation study in metro Atlanta. *Transportation Research Part B, 45*(9), 1450-1464.

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

ATOC (2017). Association of Train Operating Companies, Taxi share guide. Available at: https://www.raildeliverygroup.com/what-we-do/publications.html?task=file.download&id=469762528 [accessed 14 Sep 2017]

Chan, N., & Shaheen, S. (2012). Ridesharing in North America: Past, Present, and Future. *Transport Reviews, 32*(1), 93-112.

Edmonds, J. (1965). Maximum matching and a polyhedron with (0, l) - vertices. *Journal of Research National Bureau of Standards Section B, 69*, 125-130.

Fellows, N., & Pitfield, D. (2000). An economic and operational evaluation of urban car-sharing. *Transportation Research Part D, 5*(1), 1-10.

Furuhata, M., Dessouky, M., Ordóñez, F., Brunet, M. E., Wang, X., & Koenig, S. (2013). Ridesharing: The state-of-the-art and future directions. *Transportation Research Part B, 57*, 28-46.

Hosni, H., Naoum-Sawaya, J., & Artail, H. (2014). The shared-taxi problem: Formulation and solution methods. *Transportation Research Part B, 70*, 303-318.

Huang, D., Liu, Z., Liu, P, & Chen, J., (2016). Optimal transit fare and service frequency of a nonlinear origin destination based fare structure, *Transportation Research Part E*, 96, 1-19.

Hyytiä, E., Penttinen, A., & Sulonen, R. (2012). Non-myopic vehicle and route selection in dynamic DARP with travel time and workload objectives. *Computers & Operations Research, 39*(12), 3021-3030.

Jacobson, S., & King, D. (2009). Fuel saving and ridesharing in the US: Motivations, Limitations, and opportunities. *Transportation Research Part D, 14*(1), 14-21.

Jung, J., Jayakrishnan, R., & Choi, K. (2017). Dually sustainable urban mobility option: Shared-taxi operations with electric vehicles. *International Journal of Sustainable Transportation, 11*(8), 567-581.

Jung, J., Jayakrishnan, R., & Park, J. Y. (2016). Dynamic Shared‐Taxi Dispatch Algorithm with Hybrid‐Simulated Annealing. *Computer‐Aided Civil and Infrastructure Engineering, 31*(4), 275-291.

Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist, 39*(4), 341.

Kleiner, A., Nebel, B., & Ziparo, V. (2011). A mechanism for dynamic ride sharing based on parallel auctions. *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence*, 119-127.

Lawphongpanich S., Hearn, D., & Smith, M. (2006). *Mathematical and Computational Models for Congestion Charging*. Springer.

Li, B., Krushinsky, D., Reijers, H. A., & Van Woensel, T. (2014). The share-a-ride problem: People and parcels sharing taxis. *European Journal of Operational Research, 238*(1), 31-40.

Liu, Z., Wang, S., Chen, W. and Zheng, Y., (2016). Willingness to Board: A Novel Concept for Modeling Queuing Up Passengers, *Transportation Research Part B*, 90, 70-82.

Liu, Z., Wang, S., Meng, Q., (2014). Optimal Joint Distance and Time Toll for Cordon-based Congestion Pricing, *Transportation Research Part B*, 69, 81-97.

Liu, Z., Wang, S., Zhou, B., and Cheng, Q., (2017). Robust Optimization of Distance-based Tolls in a Network Considering Stochastic Day to Day Dynamics, *Transportation Research Part C*, 79, 58-72.

Liu, Z., Yan, Y., Qu, X., and Zhang, Y. (2013). Bus stop-skipping scheme with random travel time, *Transportation Research Part C*, 35, 46-56.

Long, J., Szeto, W. Y., & Huang, H. J. (2014). A bi-objective turning restriction design problem in urban road networks. *European Journal of Operational Research*, 237(2), 426-439.

Meng, Q., Liu, Z. and Wang, S., (2012). Optimal Distance-based Toll Design for Cordon-based Congestion Pricing Scheme with Continuously Distributed Value-of-time, *Transportation Research Part E*, 48(5), 937-957.

Minett, P., & Pearce, J. (2011). Estimating the energy consumption impact of casual carpooling.

Energies, *4*, 126-139.

Muthukrishnan, S. (2009). Ad exchanges: Research Issues. *Lecture Notes in Computer Science, 5929*, 1-12.

Naor, M. (2005). On fairness in the carpool problem. *Journal of Algorithms*, *55*(1), 93-98.

Ota, M., Vo, H., Silva, C., & Freire, J. (2016). Stars: Simulating taxi ride sharing at scale. *IEEE Transactions on Big Data*.

Qu, X., Wang, S., & Zhang, J. (2015). On the fundamental diagram for freeway traffic: A novel calibration approach for single-regime models. *Transportation Research Part B*, 73, 91-102.

Qu, X., Zhang, J., & Wang, S. (2017). On the stochastic fundamental diagram for freeway traffic: Model development, analytical properties, validation, and extensive applications. *Transportation Research Part B*, 104, 256-271.

Kuang, Y., Qu, X., & Wang, S. (2015). A tree-structured crash surrogate measure for freeways. *Accident Analysis & Prevention*, 77, 137-148.

Sayarshad, H., & Chow, J. (2015). A scalable non-myopic dynamic dial-a-ride and pricing problem. *Transportation Research Part B*, 81, 539-554.

Schrijver, A. (2003). Combinatorial Optimization: Polyhedra and Efficiency. Springer.

Wong, K.I., Wong, S.C., Yang, H., & Wu, J.H. (2008). Modeling urban taxi services with multiple user classes and vehicle types. *Transportation Research Part B*, 42(10), 985-1007.

Wong, R.C.P., Szeto, W.Y., & Wong, S.C. (2014a). Bi-level decisions of vacant taxi drivers traveling towards taxi stands in customer-search: Modeling methodology and policy implications. *Transport Policy, 33*, 73-81.

Wong, R.C.P., Szeto, W.Y., & Wong, S.C. (2014b). A cell-based logit-opportunity taxi customer-search model. *Transportation Research Part C, 48,* 84-96.

Wong, R.C.P., Szeto, W.Y., & Wong, S.C. (2015a). Behavior of taxi customers in hailing vacant taxis: a nested logit model for policy analysis. *Journal of Advanced Transportation*, *49*(8), 867-883.

Wong, R.C.P., Szeto, W.Y., & Wong, S.C. (2015b). A two-stage approach to modeling vacant taxi movements. *Transportation Research Part C, 59*, 147-163.

Xu, H., Pang, J., Ordóñez, F., & Dessouky, M. (2015). Complementarity models for traffic equilibrium with ridesharing. *Transportation Research Part B,* 81, 161-182.

Yang, H., Ye, M., Tang, W.H., & Wong, S.C. (2005). A multi-period dynamic model of taxi services with endogenous service intensity. *Operations Research*, 53(3), 501-515.

Yang, H., Leung, C., Wong, S, & Bell, M. (2010). Equilibria of bilateral taxi-customer searching and meeting on networks. *Transportation Research Part B, 44*(8-9), 1067-1083.

Yang, H., & Huang, H. (2005). Mathematical and Economic Theory of Road Pricing. Elsevier.

Yang, H., & Yang, T. (2011). Equilibrium properties of taxi markets with search frictions. *Transportation Research Part B*, 45(4), 696-713.