A peak-period taxi scheme design problem: formulation and policy implications

Baicheng Li[a,b], W. Y. Szeto[c,*], Qin Luo[a,b]

[a] College of Urban Transportation and Logistics, Shenzhen Technology University, Shenzhen, PR China

[b] Guangdong Rail Transit Intelligent Operation and Maintenance Technology Development Center

[c] Department of Civil Engineering, The University of Hong Kong, Pokfulam, Hong Kong, PR China

*Email: ceszeto@hku.hk

Acknowledgments

This study is jointly supported by grants from the Research Grants Council of the Hong Kong Special Administrative Region of China (17200618), the Ordinary University Engineering Technology Development Center Project of Guangdong Province (No. 2019GCZX006), the Natural Science Foundation of Guangdong Province (No. 2018A030313119), and the Ordinary University Innovation Project of Guangdong Province (No. 2018KTSCX350). The authors are grateful to three anonymous reviewers for providing useful comments and suggestions to this study.

A peak-period taxi scheme design problem: formulation and policy implications

Abstract

Taxis are one of the most important urban transportation modes which provide prompt and comfortable service to customers. It is commonly known that demand for and supply of taxis fluctuate at different times of a day, leading to peak periods when customer waiting time for taxis is longer and the quality of taxi service is lower than that of the off-peak periods. There have been real-world practices to mitigate the demand-supply imbalance and improve the service quality of taxis during peak periods. For example, a peak-period surcharge is imposed on taxi passengers in Singapore; the city of Perth in Australia introduces a fleet of peak-period taxis (PTs) which are allowed to operate within specific hours as the additional supply to the market. However, there lacks theoretical evidence to tell which means (or both) should be implemented and it is also unclear which factor(s) is determinant to the optimal surcharge and the optimal fleet size and shift of PTs. Moreover, there is no methodology to design the optimal shifts (the permitted operating hours) and fleet size of PTs and the optimal peak-period surcharge. To tackle the above issues, this paper proposes a peak-period taxi scheme design problem (PTSDP) that aims to determine the optimal fleet size/shifts of PTs and a peak-period taxi surcharge. The problem is formulated as a bi-level optimization model in which the upper level is the regulator (government) problem and the lower level stands for the taxi driver problem. The model is solved by a brute force method combined with the Hooke-Jeeves pattern search and the Frank-Wolfe algorithm. Numerical examples are given to give policy implications and managerial insights into the regulation of taxi markets.

Keywords:
Taxi market regulation; taxi shift; peak-hour taxi; peak-hour surcharge; bi-level optimization

1. Introduction

Taxis offer round-the-clock and door-to-door services to customers with comforts and speed. As one of the most important urban transportation modes, taxis take up a large proportion of daily trips. For example, in Hong Kong, there are in total 18,163 taxis running in the city that takes nearly one million passengers daily (GovHK, 2020a). In New York City, over 130000 taxis serve around 1000000 trips every day (TLC, 2019).

It is well-known that the demand for and supply of taxi services vary during different hours of a day and that the demand-supply imbalance of taxi markets exists, resulting in peak periods in which customer waiting time for taxis is longer and the quality of taxi services is lower than that of the off-peak periods.

To tackle the demand-supply imbalance during peak-periods, two main ways can be found in several real-world practices, namely the implementation of peak-period taxi surcharge and the introduction of peak-period taxis (PTs). For example, Singapore implements a peak-period surcharge to taxi passengers with an additional 25% price on top of the meter fare of each ride. The peak-period surcharge is effective in two periods. The first one is from 6 a.m. to 9 a.m. from Monday to Friday and the second one refers to the time from 6 p.m. to 12 a.m. on any day. According to SG Observer (2019), the idea behind the surcharge is to adjust the prices of the taxi rides to match customer demand with driver supply. Raising the taxi fares during peak periods can ensure that there are sufficient cabs around for passengers during these times.

Another example can be found in the city of Perth in Australia, where the Department of Transport introduced a fleet of PTs to the city. There were 1493 conventional taxis and 293 PTs operating in the city in 2016 (GovWA, 2016). The PTs must operate on Friday and Saturday nights (5 p.m. to 6 p.m.) and may work in the other three optional time slots (see Table 1). According to the introduction from the Swan Taxi Limited (2019), the main purpose of PTs is to increase the taxi fleet size in Perth during peak periods to reduce the customer waiting time for taxis.

The above real-world practices, yet, bring some critical questions to our attention. First, there lacks theoretical evidence to tell which means (peak-period surcharge or PTs) or both should be implemented in resolving demand-supply imbalance during peak periods. On one hand, imposing a peak-period surcharge lowers the customer waiting time for taxis by pricing out a certain number of passengers, which implies that consumer surplus also decreases. On the other hand, although introducing PTs brings more taxis on streets so that customer waiting time for taxis falls, the existing taxis in the market may earn less as now more taxis are competing with them, which leads to a lower producer surplus. Therefore, it is still necessary to verify whether we should adopt a peak-period surcharge or PTs (or both) to reduce the customer waiting time for taxis while social welfare can be properly sustained.

Table 1 Optional operating time slots for PTs in Perth (GovWA, 2014)

Time slots	Effective day(s) of week	Permitted operating hours
1	Friday	From 4 p.m. onwards
2	Monday to Friday	4 a.m. to 9 a.m.
3	Sunday	6 p.m. to 12 a.m.

The second critical question is how to determine the optimal shifts of PTs, the optimal fleet size of PTs, and the optimal peak-period surcharge. Normally, a taxi shift refers to a set of consecutive hours during which taxi drivers are permitted to work. In many cities around the world (e.g., Beijing, Hong Kong, Sydney, and Barcelona), the same fleet of taxis is usually operated by more than one group of drivers, dividing a day into several shifts. Within a driver's shift, he can freely design his work schedule (when to start and end working). The same mechanism also applies to PTs, which are introduced to serve the passengers within confined periods. Although the introduction of PTs aims at addressing demand-supply imbalance during peak periods, it is worth investigating whether the PT shifts should cover off-peak hours to achieve better system performance (e.g., higher social welfare). It is important to have a methodology to answer this question. Moreover, it is essential to have a methodology to determine the optimal PT fleet size and peak-period surcharge to maximize social welfare.

The final question is what the determinant factors to the optimal peak-period surcharge and the optimal PT fleet size and shifts are. Understanding these factors can help the government to draw policy insights and select an appropriate strategy in different scenarios to regulate the taxi market. We believe that the above three questions are related to the regulation of taxi markets in terms of price, fleet size, shift design, and labor supply. Section 2 reviews the existing literature related to these topics and points out the contributions of this study.

2. Literature review, research scope, and contributions

This study is related to the temporally heterogeneous taxi fleet size and pricing regulations, in which the supply of taxi service (or the service intensity) is a result of the scheduling decisions by individual taxi drivers. There have been extensive studies on fleet size and pricing regulations of taxi markets, most of which followed the work by Douglas (1972). For example, Arnott (1996) proposed an aggregate taxi model that depicted demand-supply equilibrium without congestion effects in a simplified circle city and revealed that the first-best taxi fare per ride should only cover the operating cost of a taxi during ride time. Therefore, all taxis are operating at a loss equal to the vacant time and need to be subsidized. Yang et al. (2005a) incorporated congestion effects into the aggregate taxi model and investigated the corresponding optimal taxi fare and fleet size. They found that the first-best taxi fare per ride should cover, in addition to the operating cost of a taxi during ride time, a cost related to the congestion effects. Yang and Yang (2011) proposed a function with constant and nonconstant returns to scale to spell out the bilateral searching and meeting process between taxis and customers in a congestion-free market. They showed that the first-best taxi fare should be higher than, equal to, or lower than the operating cost of a taxi during ride time, depending on the returns to scale of the meeting function. Until most recently, more and more attention has been paid to the emerging e-hailing taxi platforms, which provide customers with more convenient taxi services compared with the traditional street-hailing mode. For example, He and Shen (2015) considered both the street-hailing and e-hailing modes in a taxi market. Customers could choose between the two hailing modes when traveling, while vacant taxi drivers could also choose to pick up a passenger through e-hailing or street-hailing. Wang et al. (2016) investigated the pricing strategies for a taxi e-hailing platform, which included the platform's charges on taxi drivers and passengers. They analytically showed the conditions under which a price perturbation (a small change in charges on taxi drivers and customers) can affect the system performance, including the platform profit, customer waiting time, and market equilibrium. He et al. (2018) studied the optimal pricing and penalty/compensation strategies for a taxi hailing-platform. The penalty/compensation strategy was designed to penalize customers who had already reserve a taxi through the taxi-hailing app but canceled the order by taking another taxi through street-hailing before being picked up by the reserved taxi. Two optimization models were formulated that maximized social welfare and platform's revenue, respectively.

The above studies assumed a one-hour modeling period to represent the market situation, which failed to capture the temporal dynamics of customer demand. Moreover, they assumed all drivers are mandated to work during the modeling period, which did not capture the supply variation of drivers in a day. Therefore, a series of studies have been conducted to model the temporal dynamics of taxi markets and to explore the scheduling behaviors of taxi drivers. Cairns and Liston-Heyes (1996) first modified the aggregate taxi model by assuming that taxi drivers can choose the number of hours to work each day, while the customer demand and the total number of taxis in service were still assumed to be uniformly distributed throughout the day. Camerer et al. (1997) investigated the relationship between the taxi drivers' work duration and the level of income with the data on trip sheets of New York City taxi drivers. They reported a negative wage elasticity among drivers (i.e., a higher wage leads to a shorter work duration) and argued that the drivers stopped working if they earned a targeted income. In contrast, Farber (2005) and Farber (2015) reported a positive wage elasticity among taxi drivers in New York based on taxi trip sheets and GPS-based taxi trip data. Yang et al. (2005b) proposed a multi-period dynamic model with service intensity as an endogenous variable. In their model, a day is divided into 24 hours in which customer demand varies. Taxi drivers were assumed to be homogeneous and could freely choose their working schedules so that the number of taxis in service also varies across the day. The scheduling behaviors of taxi drivers were formulated as an equilibrium problem in the time-expanded network. Recently, Qian and Ukkusuri (2017) proposed a time-of-day dynamic pricing scheme to increase the total taxi revenue in a day. The temporal dynamic of a taxi market was modeled as a semi-Markov process. By using the New York City taxi trip data, they found that the dynamic pricing scheme can increase the total taxi revenue by more than 10%.

In addition to the studies on the temporal dynamics of taxi markets, we notice that there is a series of similar studies on the ride-sourcing markets. Both the taxi drivers and the ride-sourcing drivers can design their own schedules, although ride-sourcing drivers may enjoy a higher degree of freedom because, in many cities, taxi drivers' schedules are usually confined to the shift they work in. As the shift ends, taxi drivers hand over the vehicles to those working in the next shift. Chen and Shelton (2015) reported a positive wage elasticity among ride-sourcing drivers using a randomly-chosen dataset from Uber. Zha et al. (2018) studied the surge pricing and labor supply with heterogeneous ride-sourcing drivers who have different preferences in the start/end time of work and the target income level. The time-expanded network proposed by Yang et al. (2005b) was adopted in their study to model the scheduling behaviors of the ride-sourcing drivers. Ke et al. (2019a) further used the time-expanded network to investigate the scheduling and recharging behaviors of ride-sourcing drivers considering both electric and gasoline vehicles. Sun et al. (2019a) simultaneously investigated the participating decisions and working-hour decisions of the ride-sourcing drivers. They empirically found a positive and significant elasticity in both participation and work hours. Sun et al. (2019b) modeled the drivers’ participating decisions and working hour decisions with an objective to maximize their utility from income and leisure time. They revealed that the participating and working hour decisions are dependent on drivers' heterogeneous characteristics such as their other income, idle time, and leisure time. Guda and Submaranian (2019) investigated how should the ride-sourcing platforms manage the drivers through surge pricing, forecast communication, and driver incentives. In their model, the time horizon was comprised of two successive periods, in which customer demand varies. Yang et al. (2020b) proposed a reward scheme integrated with surge pricing for ride-sourcing passengers to balance demand and supply. They considered two types of periods for trips, namely peak and off-peak periods with relatively low and high demand, respectively. Other studies that involved the temporal dynamics of ride-sourcing markets include demand forecasting (Ke et al., 2017, 2019b), optimal matching strategy between passengers and drivers (Yang et al., 2020a; Ke et al., 2020), etc.

The above studies on the temporal dynamics of taxi or ride-sourcing markets and the scheduling behaviors of drivers rarely considered the existence of shifts, the PTs, and peak-hour surcharges. As mentioned in Section 1, taxi shifts are commonly observed in taxi markets around the world and it is necessary to model taxi shifts, especially if we aim to design the optimal shifts for PTs. Unfortunately, few studies can be found on providing a methodology to design the optimal shift design for taxis. Salanova and Estrada (2019) investigated the optimal shifts and fares for the Barcelona taxi market. In Barcelona, the number of taxis is under strict regulation from the government, while the number of licensed drivers is unregulated. A taxi may be operated by more than one driver in a day so that the total working hour per taxi increases, which results in oversupply during some periods of a day. In their paper, regulating taxi shifts means to determine the maximum number of hours permitted for a taxi to run on streets to mitigate oversupply, which is different from the concept of our study defined in Section 1. Moreover, their study assumes the fleet size is known and does not consider another important regulation strategy, the peak-hour surcharge, to deal with the demand-supply imbalance during peak hours. In terms of surcharge, although surge pricing has been proposed in taxi or ride-sourcing markets (e.g., Qian and Ukkusuri, 2017; Zha et al., 2018), it is different from the concept of surcharge. Generally, surge pricing means to alter the trip fare in every short time interval, which is easy to implement based on the smartphone ride-sourcing apps. However, for traditional taxi industries in which trip fare is usually regulated by the government and taxi-hailing apps may not be very common, the applicability of such a highly time-dependent pricing scheme is still debatable. On the contrary, a surcharge that is only implemented on top of the regular taxi fare in several hours of a day with a fixed extra trip fare on passengers can be more applicable to traditional taxi markets.

Based on the above literature review and the introduction, we can summarize the following research gaps. First, there lacks theoretical evidence to justify which of the following means (or both) can address the demand-supply imbalance of taxi service during peak periods, namely implementing peak-period taxi surcharges and introducing PTs. Second, there is no methodology to design the optimal shifts and fleet size of PTs and optimal peak-period surcharge. Third, it is also unclear what the determinant factors are to the optimal surcharge, and the optimal PT fleet size and shifts.

To fill the research gaps, this study proposes a peak-period taxi scheme design problem (PTSDP) to simultaneously determine the peak-period taxi surcharge and the optimal fleet size and shifts of PTs in a regulated taxi market. We assume there is a fleet of normal taxis (NTs) in the market and the taxi fare per ride is given. A time-expanded network, which divides the span of a day into 24 periods with each equal to 1 hour, is adapted from the network of Zha et al. (2018) to depict the time-of-day dynamics of demand for and supply of taxis in the market. Customer demand for taxis in each period is assumed to be a monotonically decreasing function of taxi fare, in-vehicle travel time, and customer waiting time for taxis in that period. We assume that NT and PT drivers are mutually exclusive and all drivers are working on a shift basis. The NTs are driven by two groups of drivers with an equal number in two non-overlapping shifts, and each shift counts for 12 hours. Moreover, we assume that taxi drivers are mandated in terms of shift choice but can freely design their work schedules within the designated shift.

The PTSDP is formulated as a bi-level optimization program. The upper level is the regulator (government) problem and the lower level refers to the taxi driver problem. The upper-level objective is to maximize social welfare by determining the optimal fleet size and shifts of PTs and the optimal surcharge for taxi passengers. Meanwhile, we require that the level of taxi service, which is represented by the customer waiting time for taxis, must be higher than a pre-determined value throughout the day. The lower-level problem is adapted from the problem of Zha et al. (2018), which describes the equilibrium of taxi drivers' scheduling behaviors. The upper-level problem is solved by a brute force method with Hooke-Jeeves pattern search and the lower-level problem is solved by the famous Frank-Wolfe algorithm. Numerical experiments are conducted to give policy implications and managerial insights into the regulation of taxi markets.

The main contributions of this paper are summarized as follows.

1. We propose a new peak-period taxi scheme design problem that focuses on the management of taxi markets during peak-periods;
2. We propose a bi-level formulation of the problem and a solution method to solve the problem;
3. We justify the choice of implementing peak-period taxi surcharges or introducing PTs to resolve the demand-supply imbalance of taxi services during peak periods;
4. We identify determinant factors to the optimal fleet size and shifts of PTs, thereby providing insights into the regulation of taxi markets.

The remainder of this paper is organized as follows. Section 3 provides modeling preliminaries. Section 4 proposes the bi-level formulation of the PTSDP. Section 5 introduces the solution method to the proposed bi-level model. Section 6 presents the numerical examples and gives policy insights. Finally, Section 7 concludes the paper.

3. Preliminaries

This section presents the basic modeling preliminaries. Section 3.1 introduces our proposed time-expanded network adapted from the network of Zha et al. (2018). Section 3.2 gives the definitions and assumptions of taxi drivers in the market and their cost structure. Section 3.3 presents the aggregate taxi model, which spells out the demand-supply equilibrium of taxi service in each period defined in the time-expanded network. Appendix A summarizes the main notations used in this paper.

3.1. The time-expanded network

A time-expanded network to describe the temporal dynamics of a taxi market was first proposed by Yang et al. (2005b) and was modified later by Zha et al. (2018) and Ke et al. (2019a) to study the ride-sourcing market. We adapt the version of Zha et al. (2018) yet present the network in a different way to suit our application (see Figure 1). We denote the network as $G(T, A)$ in which T is the set of nodes and A is the set of links. A day is equally divided into 24 periods and each period accounts for one hour.

In the time-expanded network, all taxi drivers are assumed to travel from the origin node O to the destination node D by traversing nodes and links of the network. As shown in Figure 1, there are two sets of nodes indexed as $1, 2, ..., 24$ and $1', 2', ..., 24'$.

The nodes indexed with the same number represent the same time (e.g., both 1 and $1'$ represent 1 a.m., 13 and $13'$ represent 1 p.m., etc.).

The link set A can be further divided into five subsets, which are denoted as $A_0 - A_4$ and represent the sets of entry, work, rest, transition, and exit links. All links are directed links except for the transition links. Corresponding to the work and rest links, we call the nodes $1,2,...,24$, as the work nodes and $1',2',...,24'$ as the rest nodes. Time is consumed only on work and rest links, during which a driver chooses to work or rest. We use (e,u) as an alternative expression of links in the time-expanded network for a better presentation of our optimization model in Section 4, in which $e,u \in T$. A work (rest) link (e,u) stands for the period from e a.m. (or e-12 p.m.) to u a.m. (u-12 p.m.).

A work schedule of taxi drivers can be viewed as a path in the time-expanded network connecting O to D. For example, a path $O \to 1 \to 2 \to 3 \to 4 \to 5 \to 5' \to 6' \to 6 \to 7 \to 7' \to D$ means that a taxi driver starts to work at 1 a.m., takes a rest from 5 a.m. to 6 a.m., goes back to work until 7 a.m., and stops working.

In many cities, the same fleet of taxis is usually operated by more than one group of drivers, splitting a day into several shifts. For example, there are two shifts in Hong Kong for all taxi drivers, with each shift lasts for 12 hours. The day shift is from 4 a.m. to 4 p.m., while the night shift begins at 4 p.m. and ends at 4 a.m. Drivers working in a shift can freely design their work schedules, but the schedules cannot start earlier (end later) than the start (end) time of the shift. Generally, the start and end times of shifts are determined either by consensus among drivers (e.g., taxi shifts in Hong Kong) or by the government. We note that the existence of shifts creates a service time restriction to taxi drivers, which can be viewed as a restriction of traversing certain links/nodes in the time-expanded network.

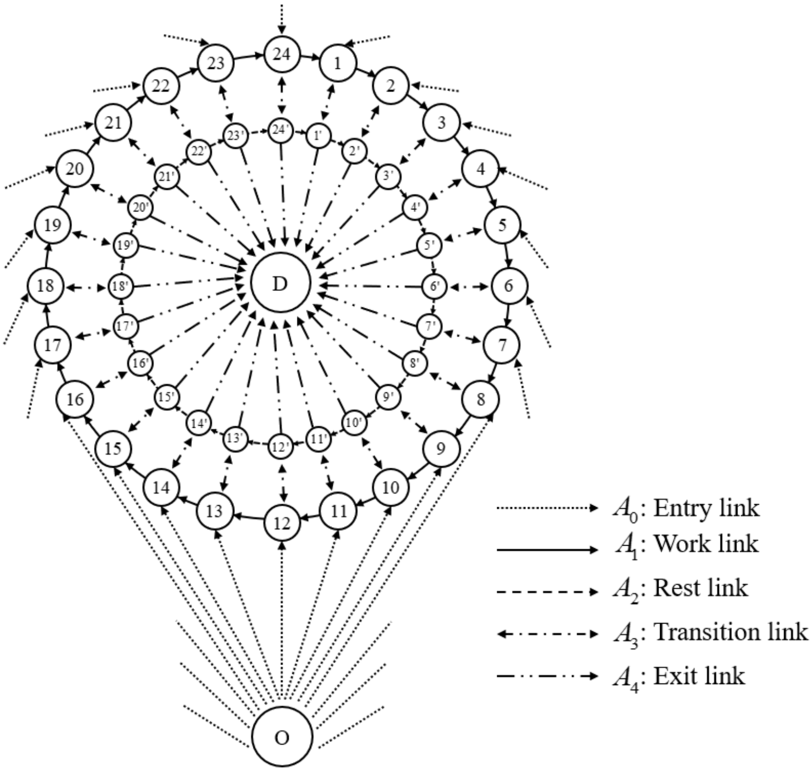


Figure 1 The time-expanded network

3.2. *Taxi drivers and their costs*

In this study, we assume two types of taxis in the market, namely NTs and PTs. NTs are those that already exist in the market, while PTs are to be introduced into the market by the government. We denote the fleet sizes of NTs and PTs as N^{n} (model parameter) and N^{p} (decision variable), respectively. Corresponding to NTs and PTs, there are NT and PT drivers that are mutually exclusive. We specify the following:

Assumption 1. NT drivers are full-time drivers and PT drivers are part-time drivers.

Full-time and part-time drivers are different in terms of their cost to work, which will be illustrated later in this section.

As introduced in Section 3.1, we consider different shifts of taxis, which means that NTs and PTs are operated by more than one group of NT and PT drivers, respectively. In reality, drivers may have the freedom to choose a shift to work in but for simplicity, we make the following assumption:

Assumption 2. Taxis drivers are mandated to work in a particular shift but are free to design their schedules within their designated shift. Each driver works for one shift in a day and the number of NT (PT) drivers working in each NT (PT) shift is equal to the fleet size of NTs (PTs).

Under Assumption 2, we further categorize taxi drivers based on the shift that they work in. We define $D=\{d \mid d=1,2,\ldots,|D|\}$ as the set of driver groups. Each $d \in D$ represents a group of taxi drivers working in a shift. The cardinality of D, i.e., $|D|$, indicates the number of driver groups in the market, including the NT and PT shifts. Figure 2 illustrates the relationships between taxis, taxi drivers, and driver groups when there are two shifts for NT and PT drivers. The NT (PT) fleet is assigned to two groups of NT (PT) drivers, with each group working in a shift.

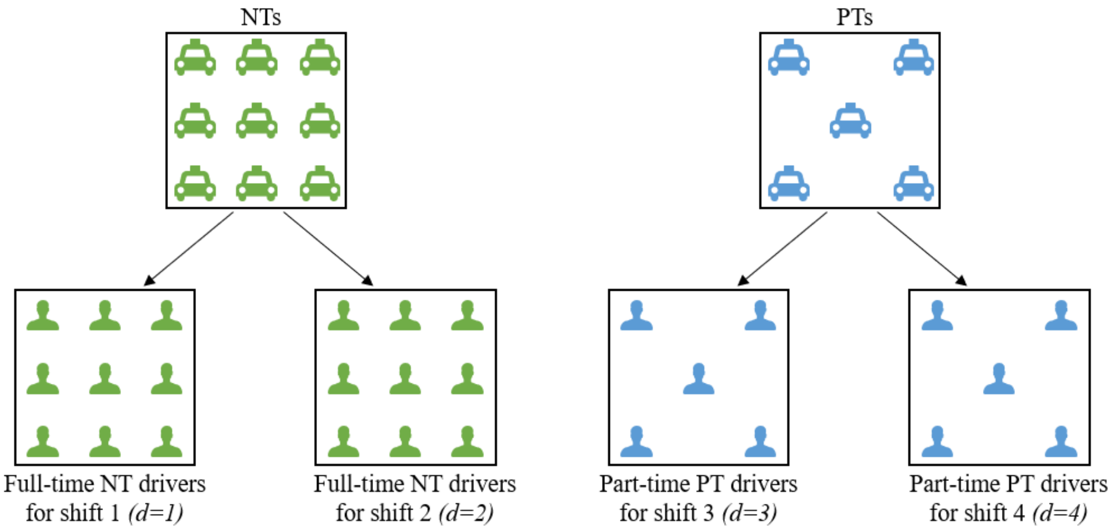


Figure 2 Relationship between taxis, taxi drivers, and driver groups

We also point out that the requirements of NT and PT shifts are different. First, for each driver group, their shifts are mutually non-overlapping. Second, NT shifts must cover the whole span of a day since NTs are expected to be available to customers at any time of the day. Such a requirement does not apply to PT shifts since PTs serve as the supplementary supply to customers. Therefore, PT shifts may only cover some hours of a day. A PT shift may start several hours after the last PT shift ends and terminates earlier than the start time of the next PT shift. Third, like the assumption about the fleet sizes of NTs, we assume that the NT shifts are given and fixed, while the government determines the start and end times of PT shifts.

On each link $(e,u) \in A$, we denote v_{eu} as the number of taxis (service intensity) traversing link (e,u), which can be calculated as

$$v_{eu} = \sum_{d \in D} v_{eu}^{d}, \forall (e,u) \in A, \qquad (1)$$

in which v_{eu}^{d} is the number of group d drivers traversing link (e,u). We let P be the path set that contains all paths connecting the O-D pair[1] in the time-expanded network and the number of group d drivers that choose $p \in P$ as their work schedule is denoted as f_{p}^{d}. In this regard, v_{eu}^{d} can be expressed as

$$v_{eu}^{d} = \sum_{p \in P} f_{p}^{d} \omega_{eu}^{p}, \forall (e,u) \in A. \tag{2}$$

ω_{eu}^{p} is the link-path incidence indicator which equals 1 if path p traverses link (e,u), and 0 otherwise. Substituting Eq. (2) into Eq. (1) gives

$$v_{eu} = \sum_{d \in D} \sum_{p \in P} f_{p}^{d} \omega_{eu}^{p}, \forall (e,u) \in A. \tag{3}$$

In terms of the cost of taxi drivers, we assume two types of costs incurred to each taxi driver in group $d \in D$. The first type of cost is called the link-specific cost $c_{eu}^{d}, (e,u) \in A$, meaning the cost incurred by traversing a particular link in the time-expanded network. The link-specific cost can be further divided into five types according to the link type, namely entry cost, work cost, transition cost, rest cost, and exit cost. Different types of link-specific costs have different meanings. For example, the entry cost can represent the fixed cost of a driver (e.g., the rental fee charged by taxi owners or the opportunity cost of being a taxi driver) and the work cost can stand for the hourly operating cost of a taxi.

The second type of cost, namely the path-specific (duration) cost $c_{p}^{d}, p \in P$, captures the cumulative effect of work duration on group d drivers and is expressed as

$$c_{p}^{d} = \sum_{l \in L} \alpha_{1}^{d} (h_{p}^{l})^{\alpha_{2}^{d}}, \forall p \in P, d \in D, \tag{4}$$

in which h_{p}^{l} is the length of sub-shift l in path p. For a path in the time-expanded network, a sub-shift $l \in L$ is defined as a consecutive period in which the drivers work. Figure 3 shows an example of sub-shifts in a path, in which a path consists of work nodes 13, 14, 15, 16, and 17. Therefore, there are two sub-shifts in this path, the first sub-shift is from node 13 to node 15 with a length as 2, and the second sub-shift is from node 16 to node 17, which lasts for 1 hour. α_{1}^{d} and α_{2}^{d} are positive parameters and

[1] As drivers' possible schedules (paths) are confined within their shifts, the path set that contains all possible paths within a shift is obviously a subset of *P*.

we assume $\alpha_2^d > 1$ to show that the duration cost incurred to drivers increases more than linearly as total work hour increases.

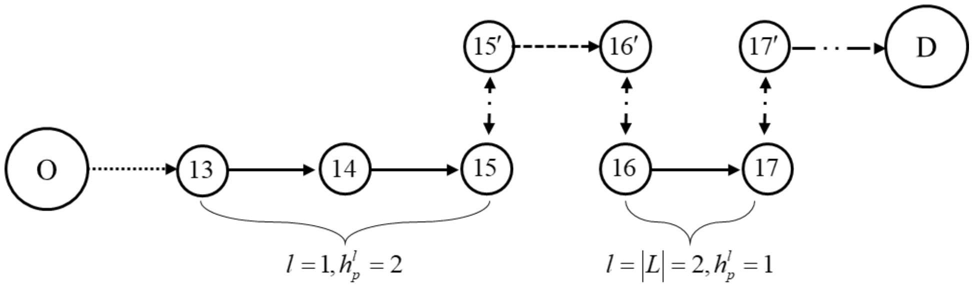


Figure 3 Example of sub-shifts in a path

3.3. Demand-supply equilibrium of the taxi market in each period

Customer demand varies across different periods (hours) of a day. We use an aggregate taxi model to describe the demand-supply equilibrium of the taxi market in each period. First, we assume that NTs and PTs have the same fare level and vehicle type, which means that they can be treated as a single traffic mode and customers have no preference between NTs and PTs. In any period $(e,u) \in A_1$, a Cobb-Douglas meeting function is used to quantify the meeting rate K_{eu} between taxis and customers as

$$K_{eu} = \Theta (N_{eu}^{\mathrm{v}})^{\beta_1} (N_{eu}^{\mathrm{c}})^{\beta_2}, \quad \forall (e,u) \in A_1, \tag{5}$$

in which $N_{eu}^{\mathrm{v}} = w_{eu}^{\mathrm{t}} V_{eu}$ and $N_{eu}^{\mathrm{c}} = w_{eu}^{\mathrm{c}} Q_{eu}$ are the numbers of vacant taxis and unserved customers in period (e,u), respectively. Note that N_{eu}^{v} and N_{eu}^{c} are period-specific and independent of those in other periods. The situation in which the unserved customers and vacant taxis in one period move into the next period is not considered in this study. In a stationary equilibrium, we have

$$K_{eu} = V_{eu} = Q_{eu}, \quad \forall (e,u) \in A_1, \tag{6}$$

which gives an expression of customer waiting time for taxis as

$$w_{eu}^{\mathrm{c}} = (\Theta)^{-\frac{1}{\beta_2}} (Q_{eu})^{\frac{1-\beta_1-\beta_2}{\beta_2}} (w_{eu}^{\mathrm{t}})^{-\frac{\beta_1}{\beta_2}}, \quad \forall (e,u) \in A_1. \tag{7}$$

The customer demand for taxis Q_{eu} is assumed to be a strictly decreasing function of the full price of taking taxis ρ_{eu} as

$$Q_{eu} = Q_{eu}(\rho_{eu}) = Q_{eu}(F_{eu}, l_{eu}, w_{eu}^{\mathrm{c}}), \quad \forall (e,u) \in A_1, \tag{8}$$

in which ρ_{eu} is a function of taxi fare per ride F_{eu}, the in-vehicle travel time l_{eu}, and the customer waiting time for taxis w_{eu}^{c} as

$$\rho_{eu} = F_{eu} + \delta l_{eu} + \kappa w_{eu}^{c}, \quad \forall (e,u) \in A_1. \tag{9}$$

In Eq. (9), δ and κ are values of customers' in-vehicle travel time and waiting time for taxis. The in-vehicle travel time l_{eu} is assumed to be given because congestion effects are not considered in this study for simplicity. The taxi fare per ride is expressed as

$$F_{eu} = F + \gamma \chi_{eu}, \quad \forall (e,u) \in A_1, \tag{10}$$

in which F is the flag fare and is assumed to be equal among different periods. γ is a decision variable to be determined by the government representing the taxi surcharge. χ_{eu} is a binary decision variable which equals 1 if taxi surcharge is implemented in period (e,u), and 0 otherwise. Therefore, the term $\gamma\chi_{eu}$ indicates that the taxi surcharge is implemented only within the PT shifts and is identical among periods that belong to the PT shifts.

For each period (e,u), the values of Q_{eu}, F_{eu}, and w_{eu}^{c} can be obtained by solving the system of equations (7)-(10), given the values of γ, χ_{eu}, and v_{eu}.

Once we obtain Q_{eu} and F_{eu}, the average taxi revenue in period (e,u) can be calculated as

$$R_{eu} = \frac{F_{eu} Q_{eu}}{v_{eu}}, \quad \forall (e,u) \in A_1, \tag{11}$$

in which $F_{eu} Q_{eu}$ stands for the total revenue collected from taxi trips in period (e,u).

It is necessary to note that the marginal average revenue with respect to v_{eu} is given as

$$\frac{\partial R_{eu}}{\partial v_{eu}} = \frac{F_{eu}}{v_{eu}} \frac{\partial Q_{eu}}{\partial v_{eu}} - \frac{F_{eu} Q_{eu}}{v_{eu}^2} = \frac{F_{eu} Q_{eu}}{v_{eu}^2} \left(\frac{\partial Q_{eu}}{\partial v_{eu}} \frac{v_{eu}}{Q_{eu}} - 1 \right), \quad \forall (e,u) \in A_1. \tag{12}$$

In Eq. (12), $\frac{\partial Q_{eu}}{\partial v_{eu}} \frac{v_{eu}}{Q_{eu}}$ is the elasticity of the customer demand with respect to the total taxi service hour in period (e,u). Clearly, the marginal average revenue of taxi drivers can be either positive or negative depending on the elasticity term $\frac{\partial Q_{eu}}{\partial v_{eu}} \frac{v_{eu}}{Q_{eu}}$. However, the situation in which $\frac{\partial R_{eu}}{\partial v_{eu}} > 0$ only occurs with an unrealistically small value of v_{eu}. It is more commonly observed that the increase in v_{eu} leads to a fall in R_{eu} (Yang et al., 2005a). Hence, we make the following assumption for the analysis hereinafter.

Assumption 3. For $\forall(e,u) \in A_1$, $\frac{\partial R_{eu}}{\partial v_{eu}} < 0$ is always satisfied.

4. A bi-level formulation of PTSDP

The PTSDP can be formulated as a bi-level optimization program, in which the upper level refers to the government problem and the lower level represents the problem of taxi drivers (both the NT and PT drivers). This section presents the formulation of the bi-level program.

4.1. Upper-level formulation

We assume the following market scenario. There is a fleet of NTs (the fleet size is N^{n}) operated by two groups of NT drivers in two shifts. The morning shift is from 4 a.m. to 4 p.m., while the evening shift is from 4 p.m. to 4 a.m. For simplicity, we assume two non-overlapping PT shifts to be determined by the regulator, with each shift at least covering a set of peak periods in a day (morning and afternoon peaks). The two sets of peak periods are the same as those reported by the Annual Traffic Census (GovHK, 2017). The morning peak lasts from 7 a.m. to 9 a.m., while the afternoon peak is from 4 p.m. to 7 p.m. Now, the government plans to introduce a fleet of PTs, design two non-overlapping PT shifts, and set a taxi surcharge during the PT shifts. We let $d=1$ and $d=2$ be the driver groups of morning and evening NT shifts, respectively, and let $d=3$ and $d=4$ be the driver groups of morning and afternoon PT shifts, respectively.

To facilitate the expression of our model, we denote $T_1 = \{t \mid t = 1,2,3,...,24\}$ as the set of work nodes of the time-expanded network. To design a shift is equivalent to selecting work nodes and links between each pair of adjacent work nodes to form a connected sub-network of the time-expanded network. In graph theory, a sub-network (or sub-graph) of a network is a network whose nodes (links) form a subset of the nodes (links) of the network. A connected sub-network must satisfy the contiguity condition, which requires that there exists at least one path connecting any two work nodes in the sub-network without traversing any work node that does not belong to the sub-network. Figure 4 shows the difference between connected and disconnected sub-networks, in which (a) is the work-node circle as shown in Figure 1. Both (b) and (c) are sub-networks of (a). (b) is a connected sub-network but (c) is a disconnected one (e.g., there is no path from node 1 to node 4).

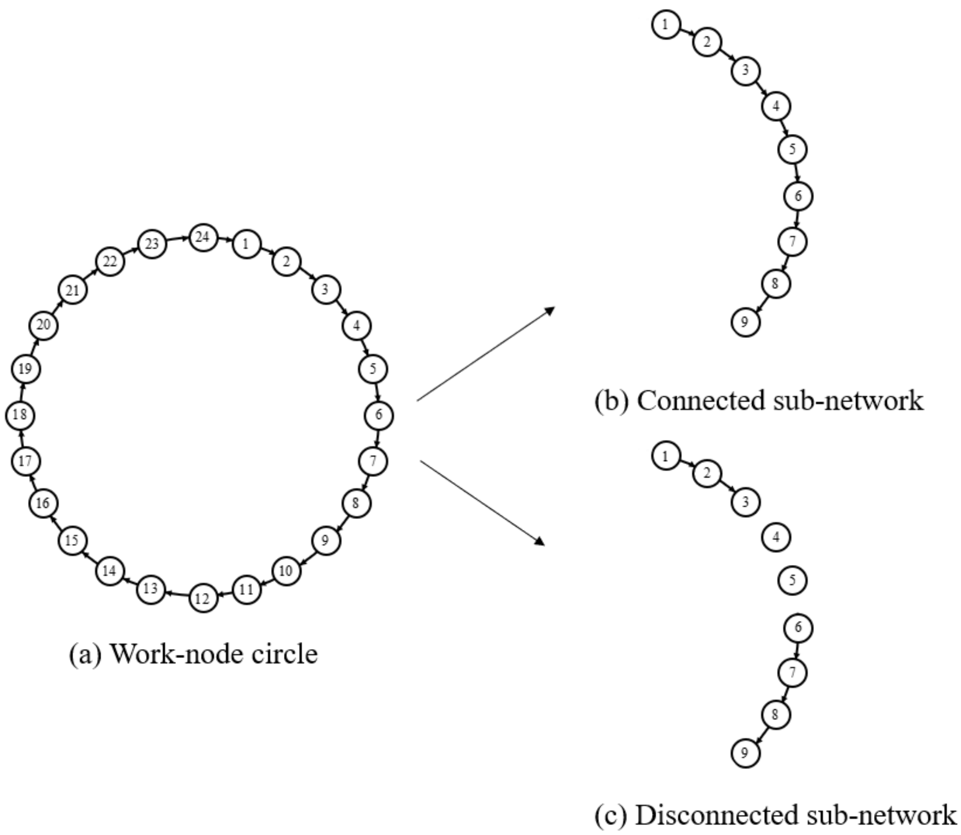


Figure 4 Illustrative examples of connected and dis-connected sub-networks

To model the contiguity condition, we adopted the formulation approach proposed by Shirabe (2005), which assumed the following mechanism. For any shift design, we arbitrarily choose one work node as the sink and every other work node provides at least one unit of supply (imaginary flow). Then for a shift to be contiguous (i.e., for a non-split shift), supply sent from every source node must ultimately arrive at the sink without passing through any work node or link that is not included in the sub-network.

The objective of the government is to maximize social welfare. We also require that customer waiting time for taxis in each period $w_{eu}^{c}, (e,u) \in A_1$ is not larger than a predetermined value to ensure the quality of taxi services throughout the day.

We give the upper-level mathematical program as follows.

$$\max_{\mathbf{X},\mathbf{y},\boldsymbol{\chi},\boldsymbol{\Phi},\mathbf{s},\gamma,N^{\mathrm{p}}} S = \sum_{(e,u)\in A_1} \int_{\rho_{eu}}^{+\infty} Q_{eu}(\varphi) d\varphi + \sum_{(e,u)\in A_1} F_{eu} Q_{eu} - \sum_{d\in D} \sum_{(e',u')\in A} v_{e'u'}^{d} c_{e'u'}^{d} - \mathbf{f}^{*}\mathbf{c}^{\mathrm{T}} \qquad (13)$$

s.t.

$$w_{eu}^{\mathrm{c}} \le \eta,\ \forall (e,u) \in A_1, \tag{14}$$

$$\sum_{t\in\{7,8,9\}} s_t^3 = 1, \tag{15}$$

$$\sum_{t\in\{16,17,18,19\}} s_t^4 = 1, \tag{16}$$

$$\sum_{\{u|(u,e)\in A_1\}} y_{ue}^d \le (K-1)X_e^d,\ \forall d \in \{3,4\},\ e \in T_1, \tag{17}$$

$$\sum_{\{u|(e,u)\in A_1\}} y_{eu}^d - \sum_{\{u|(u,e)\in A_1\}} y_{ue}^d \ge X_e^d - K s_e^d,\ \forall d \in \{3,4\},\ \ e \in T_1, \tag{18}$$

$$X_e^d + X_u^d \ge 2\phi_{eu}^d,\ \forall (e,u) \in A_1,\ d \in \{3,4\}, \tag{19}$$

$$\sum_{d\in\{3,4\}} \phi_{eu}^d \le 1,\ \forall (e,u) \in A_1, \tag{20}$$

$$\sum_{d\in\{3,4\}} \phi_{eu}^d = \chi_{eu},\ \forall (e,u) \in A_1, \tag{21}$$

$$X_t^d = \{0,1\},\ \forall d \in \{3,4\},\ t \in T_1, \tag{22}$$

$$s_t^d = \{0,1\},\ \forall d \in \{3,4\},\ t \in T_1, \tag{23}$$

$$\phi_{eu}^d = \{0,1\},\ \forall (e,u) \in A_1,\ d \in \{3,4\}, \tag{24}$$

$$\chi_{eu} = \{0,1\},\ \forall (e,u) \in A_1, \tag{25}$$

$$y_{eu}^d \ge 0,\ \forall (e,u) \in A_1,\ d \in \{3,4\}, \tag{26}$$

$$\gamma \ge 0\text{, and} \tag{27}$$

$$N^{\mathrm{p}} \ge 0. \tag{28}$$

In objective (13), the flow vector $\mathbf{f}^*$ is obtained by solving the lower-level problem defined in Section 4.2 for any fixed $\mathbf{X}, \mathbf{y}, \boldsymbol{\chi}, \boldsymbol{\Phi}, \mathbf{s}, \gamma,$ and N^{p}. $\mathbf{c}^{\mathrm{T}}$ is the transpose of the path cost vector $\mathbf{c} = \left[c_p^d\right], \forall p \in P, \forall d \in D$. According to our definition, the generalized cost ρ_{eu}, customer demand Q_{eu}, and link flow $v_{e'u'}^d$ are functions of $\mathbf{f}^*$. For variables and parameters, X_t^d is a binary decision variable which equals 1 if work node t is included in the sub-network for the shift of group d drivers, and 0 otherwise. ϕ_{eu}^d is a binary decision variable which equals 1 if work link (e,u) is selected into the sub-network for the shift of group d drivers and, 0 otherwise. s_t^d is a binary decision variable which equals 1 if work node t is chosen as the sink, and 0 otherwise. y_{eu}^d is a non-negative continuous decision variable that indicates the amount of imaginary flow of group d drivers from e to u. γ is the non-negative continuous decision variable that reflects the level of taxi surcharge. K is equal to the number of work nodes that can be selected, i.e., $K = 24$. Lastly, we have $\mathbf{X} = (X_t^d, \forall d \in \{3,4\}, t \in T_1)$, $\mathbf{y} = (y_{eu}^d, \forall d \in \{3,4\}, (e,u) \in A_1)$, $\mathbf{\Phi} = (\phi_{eu}^d, \forall (e,u) \in A_1, d \in \{3,4\})$, $\boldsymbol{\chi} = (\chi_{eu}, \forall (e,u) \in A_1)$, and $\mathbf{s} = (s_t^d, \forall d \in \{3,4\}, t \in T_1)$.

The objective (13) aims at maximizing social welfare, which is the sum of total consumer surplus $\sum_{(e,u)\in A_1} \int_{\rho_{eu}}^{+\infty} Q_{eu}(\varphi) d\varphi$ and total producer surplus $\sum_{(e,u)\in A_1} F_{eu} Q_{eu} - \sum_{d\in D} \sum_{(e',u')\in A} v_{e'u'}^d c_{e'u'}^d - \mathbf{f}^* \mathbf{c}^{\mathrm{T}}$. Total consumer surplus and total producer surplus are functions of $(\mathbf{X}, \mathbf{y}, \boldsymbol{\chi}, \mathbf{\Phi}, \mathbf{s}, \gamma, N^{\mathrm{p}})$ and $\mathbf{f}^*$ in which $\mathbf{f}^*$ solves the lower-level problem for a given $(\mathbf{X}, \mathbf{y}, \boldsymbol{\chi}, \mathbf{\Phi}, \mathbf{s}, \gamma, N^{\mathrm{p}})$ determined by the regulator. We use Figure 5 to elaborate on the calculation of consumer surplus and taxi revenue. Figure 5 shows a general demand curve Q_{eu} against the full-price of taking taxis ρ_{eu}. For a specific point on the curve $\left(\tilde{Q}_{eu}, \tilde{\rho}_{eu}\right)$, the figure shows that $\tilde{\rho}_{eu}$ is comprised of the trip fare $\tilde{F}_{eu}$, the travel time cost $\delta \tilde{l}_{eu}$, and the waiting time cost $\kappa \tilde{w}_{eu}^{\mathrm{c}}$. Clearly, the blue area stands for the taxi revenue $\tilde{F}_{eu} \tilde{Q}_{eu}$ and the yellow area represents the total waiting time cost $\kappa \tilde{w}_{eu}^{\mathrm{c}} \tilde{Q}_{eu}$. According to the economic theory, the consumer surplus is represented by the green area, which can be calculated by the integral $\int_{\tilde{\rho}_{eu}}^{+\infty} Q_{eu}(w) dw = \int_{\tilde{F}_{eu} + \delta \tilde{l}_{eu} + \kappa \tilde{w}_{eu}^{\mathrm{c}}}^{+\infty} Q_{eu}(w) dw$.

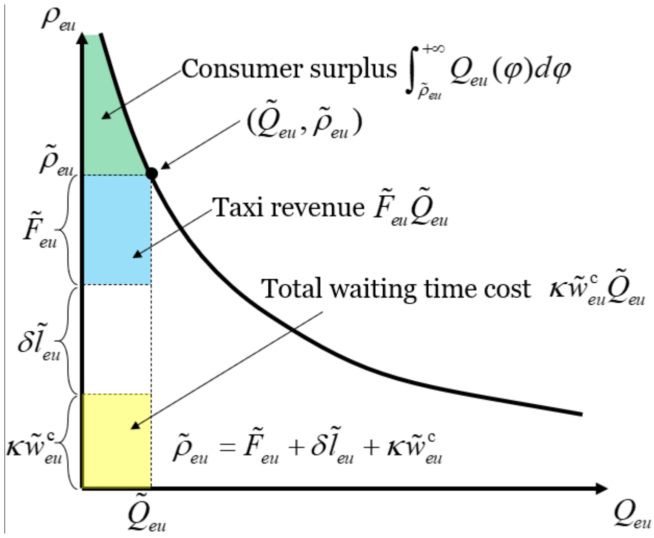


Figure 5 Diagrammatic representation of full-price and customer demand

Constraint (14) is the level-of-service (LOS) constraint for the taxi service, which requires that customer waiting time for taxis in each period must be no higher than a pre-determined level η. The parameter η can be viewed as the LOS index. The lower the value η, the higher the LOS for the taxi service. The LOS constraint is to assist the government in designing the optimal peak-period surcharge and the optimal PT shifts and fleet size. Yet, the LOS constraint and η do not necessarily affect customer demand. The customer demand Q_{eu} is affected by waiting time w^{c}_{eu} in a way described by the demand function Eq. (8). Constraints (15) and (16) indicate that the sink of each sub-network for a PT shift must be chosen within the corresponding peak period. Constraint (17) requires that the total imaginary inflow of the sub-network of each PT shift to any node e is non-positive if e is not included in the sub-network. It also requires that the total inflow cannot exceed the maximum total network supply $(K-1)$ if e is included in the sub-network. Moreover, the maximum total supply in a K-node network is equal to $K-1$ because we have one node selected as a sink that provides no supply and others provide at least one unit of supply. If all K nodes are chosen, then each non-sink node provides exactly one unit of supply and the total supply is $K-1$.

Constraint (18) represents the imaginary outflow of the sub-network of each PT shift from node e. It ensures that any selected non-sink node provides at least 1 unit of supply. Note that when $X_e^d = 1$ and $s_e^d = 1$ (node e is selected as a sink), Constraint (17) is still consistent with Constraint (18). Constraint (19) indicates that a work link is selected into the sub-network of group d drivers if and only if both its start and end nodes are included in that sub-network. Constraint (20) requires that a work link cannot be included in the sub-networks of both PT shifts. Constraint (21) means that the taxi surcharge is implemented on link (e,u) if and only if the corresponding work link of (e,u) belongs to the sub-network of a PT shift. Finally, Constraints (22) to (28) define the variable domains.

With the formulation of the upper-level problem, we have the following proposition regarding the lower bound of PT fleet size:

Proposition 1. Given any NT fleet size $N^{\mathrm{n}} \geq 0$, a lower bound of PT fleet size $\underline{N}^{\mathrm{p}}$ can be calculated as $\underline{N}^{\mathrm{p}} = \max\left(\arg\max\left\{v_{eu} \middle| w_{eu}^{\mathrm{c}}(v_{eu}, Q_{eu}) = \eta, \forall (e,u) \in A_1\right\} - N^{\mathrm{n}}, 0\right)$.

Proof. In view of the fact that customer demands in different periods are independent of each other and that customer waiting time in any period (e,u) monotonically decreases with service intensity (the number of taxis) v_{eu} (Yang et al., 2005b), there is a minimum required service intensity $\underline{v}_{eu}$ for each period (e,u) with which the LOS constraint is binding (i.e., $w_{eu}^{\mathrm{c}}(\underline{v}_{eu}, Q_{eu}) = \eta$). Clearly, the largest $\underline{v}_{eu}$ among all 24 periods is the minimum service intensity required so that all periods satisfy the LOS constraint (i.e., $\underline{v}_{eu} = \arg\max\left\{v_{eu} \middle| w_{eu}^{\mathrm{c}}(v_{eu}, Q_{eu}) = \eta, \forall (e,u) \in A_1\right\}$). Therefore, the total taxi fleet size $N^{\mathrm{n}} + N^{\mathrm{p}}$ should at least equal $\underline{v}_{eu}$, which implies $N^{\mathrm{p}} \geq \underline{v}_{eu} - N^{\mathrm{n}}$. Moreover, N^{p} is nonnegative. Therefore, a lower bound for the PT fleet size is $\underline{N}^{\mathrm{p}} = \max\left(\arg\max\left\{v_{eu} \middle| w_{eu}^{\mathrm{c}}(v_{eu}, Q_{eu}) = \eta, \forall (e,u) \in A_1\right\} - N^{\mathrm{n}}, 0\right)$.

Proposition 1 further gives rise to the following corollary:

Corollary 1. The fleet size lower bound $\underline{N}^{\mathrm{p}}$ increases as the LOS index η increases.

Corollary 1 is intuitive. As the expected taxi service level increases, more PTs are needed to satisfy the LOS constraint.

4.2. Lower-level formulation

The lower-level problem can be viewed as a multi-class network equilibrium problem that describes the scheduling behaviors of different groups of taxi drivers. Given the PT shifts, fleet size, and taxi surcharge from the upper-level problem (i.e., $\mathbf{X}, \mathbf{y}, \boldsymbol{\chi}, \boldsymbol{\Phi}, \mathbf{s}, \gamma$, and N^{p}), the scheduling equilibrium of taxi drivers can be defined as follows.

Definition 1. At the scheduling equilibrium of taxi drivers, for each driver group $d \in D$, all used paths (with positive flows) yield the same path profit (the difference between path revenue and path cost), which is no less than that of any unused path.

The formulation of the lower-level problem is given as follows.

$$\min_{\mathbf{f}} \; Z = -\sum_{(e,u)\in A_1} \int_0^{v_{eu}} R_{eu}(\psi)d\psi + \sum_{d\in D}\sum_{(e',u')\in A} v_{e'u'}^{d} c_{e'u'}^{d} + \mathbf{fc}^{\mathrm{T}} \tag{29}$$

s.t.

$$\sum_{p\in P} f_p^d = N^{\mathrm{n}}, \forall d \in \{1,2\}, \tag{30}$$

$$\sum_{p\in P} f_p^d = N^{\mathrm{p}}, \forall d \in \{3,4\}, \tag{31}$$

$$f_p^d \geq 0, \; \forall p \in P, d \in D, \quad \text{and} \tag{32}$$

$$M\phi_{eu}^{d} \geq \sum_{p\in P} f_p^d \omega_{eu}^{p}, \; \forall (e,u) \in A_1, \; d \in \{3,4\}, \tag{33}$$

where v_{eu} and v_{eu}^d are defined by Eqs. (1)-(3); $\mathbf{c} = \left[c_p^d\right]$ in which c_p^d is defined by Eq. (4); $R_{eu}(\psi)$ is defined by Eq. (11), which is in turn defined by Eqs. (7)-(10). M is a large constant. It should be noted that the impacts of surcharge rate and passenger demand on drivers' scheduling decisions are captured in this formulation. In the lower-level objective function (29), the first term on the left-hand side $-\sum_{(e,u)\in A_1} \int_0^{v_{eu}} R_{eu}(\psi)d\psi$ is related to the average taxi revenue $R_{eu} = \dfrac{F_{eu}Q_{eu}}{v_{eu}}$, in which the trip fare F_{eu} contains the surcharge rate γ and Q_{eu} is the customer demand.

The lower-level program treats the path flow vector $\mathbf{f} = \left[f_p^d\right], \forall p \in P, \forall d \in D$ as the decision variable. Constraints (30) and (31) require that the sum of all path flows of group d drivers must equal the corresponding taxi fleet size, which means that all drivers come out to work during their designated shift. Constraint (32) is the non-negativity constraint for path flows. Constraint (33) ensures that the path flow of group d drivers on work link (e,u) can be positive only if (e,u) is included in the sub-network for the shift of group d.

5. Solution method

5.1. A brute force method with Hooke-Jeeves pattern search for the upper-level problem

For the upper-level problem, which is formulated as a mixed-integer nonlinear program (MINLP), it can be solved by various well-known exact methods such as the branch-and-bound method. However, since the time-expanded network contains small numbers of work nodes and links, all feasible combinations of PT shifts can be easily enumerated. There are, in total, 3276 feasible shift combinations. For each feasible shift combination, we used a Hooke-Jeeves pattern search to determine the optimal PT fleet size N^{p} and the taxi surcharge γ. After determining the optimal PT fleet size and surcharge of each shift combination, the best solution among them was then selected as the final output of the PTSDP. Note that for each given pair of intermediate PT fleet and taxi surcharge found by the Hooke-Jeeves pattern search for a given PT shift, the famous Frank-Wolfe algorithm described in the next subsection was invoked to determine the objective value of the upper-level problem.

5.2. The Frank-Wolfe algorithm for the lower-level problem

For any given PT shifts, fleet size, and taxi surcharge from the upper-level program, the lower-level program is convex with linear equality constraints. The lower level program can, therefore, be solved to global optimality by the Frank-Wolfe algorithm. At each iteration i, a shortest-path problem for each driver group d, which finds the path with maximum profit within the corresponding taxi shift, is solved to expand the used path set (if needed). The decent direction of the current solution $\mathbf{f}^{(i)}$ is then obtained by performing the all-or-nothing assignment. The main steps of the Frank-Wolfe algorithm are given as follows.

Step 1. Set iteration count $i=0$. For each driver group d, select a path that traverses all work nodes and links belonging to the corresponding taxi shift to form the used path set $P_d^{(i)} \in P$. Load each group of drivers on the corresponding path to obtain the initial path flow vector as $\mathbf{f}^{(i)} = [f_p^d, \forall p \in P_d^{(i)}, d \in D]$;

Step 2. Update $R_{eu}, \forall (e,u) \in A_1$ based on $\mathbf{f}^{(i)}$;

Step 3. For each driver group d, solve the shortest-path problem to update the used path set $P_d^{(i)}$ and $\mathbf{f}^{(i)}$;

Step 4. Perform the all-or-nothing assignment to obtain the auxiliary flow pattern $\bar{\mathbf{f}}^{(i)}$;

Step 5. Calculate $\mathbf{f}^{(i+1)}$ by $\mathbf{f}^{(i+1)} = (1-\varphi)\mathbf{f}^{(i)} + \varphi\bar{\mathbf{f}}^{(i)}$, in which the step size φ is determined by solving the program $\min_{\varphi\in[0,1]} Z((1-\varphi)\mathbf{f}^{(i)} + \varphi\bar{\mathbf{f}}^{(i)})$;

Step 6. If $\nabla Z(\mathbf{f}^{i})(\mathbf{f}^{i+1} - \mathbf{f}^{i})^{\mathrm{T}} \geq \varsigma$ (ς is the convergence tolerance close to zero), output $\mathbf{f}^{(i)}$ and stop. Otherwise, set $i = i+1$ and return to Step 2.

6. Numerical examples

This section provides numerical examples with functions and values of parameters given in Section 6.1 unless specified otherwise. Section 6.2 illustrates three determinant factors to optimal taxi surcharges and PT fleet sizes and discusses which means (introducing taxi surcharge or PTs) is better to maximize social welfare. Sections 6.4 and 6.5 demonstrate how the LOS index and the duration cost of taxi drivers affect optimal PT fleet sizes/shifts and social welfare.

6.1. Function and parameter settings

6.1.1. Customer waiting time function

We first specify the customer waiting time function W_{eu}, which takes the following form:

$$w_{eu}^{\mathrm{c}} = \frac{1}{\Theta Q_{eu} w_{eu}^{\mathrm{t}}}, \tag{34}$$

in which $Q_{eu} w_{eu}^{\mathrm{t}}$ is the vacant taxi hours in period (e,u).

6.1.2. Customer demand function

We use a simple exponential function to describe the customer demand for taxis Q_{eu}:

$$Q_{eu} = \bar{Q}_{eu} \exp\left[\theta(F_{eu} + \delta l_{eu} + \kappa w_{eu}^{\mathrm{c}})\right]. \tag{35}$$

$\bar{Q}_{eu}$ is the total travel demand (trips/h) in each period, which is assumed to be a constant. θ (1/$) is the sensitivity of customer demand towards the full price of taking taxis.

6.1.3. Input parameters

We assume the total travel demand in each period as given in Table 2. The in-vehicle travel time is assumed to be 0.3 (h) for all non-peak periods and 0.4 (h) for peak periods. The LOS index η is assumed to be 0.1 (h), which is equal to the maximum acceptable waiting time for taxis adopted from the Traffic Characteristic Survey (GovHK, 2011).

The link-specific costs c_{eu}^{d} (HKD) are given in Table 3, in which the entry cost of NT drivers (*d*=1,2) is estimated based on the daily taxi rent in Hong Kong that a rentee-driver pays to taxi companies (Hong Kong Extras, 2020). The entry cost of PT drivers has rare empirical evidence but is expected to be no larger than that of the NT drivers. Therefore, we take the entry cost of PT drivers as $150. For the duration cost of drivers defined by Eq. (4), there is no existing data for calibration and hence we set the coefficients for NT drivers are $\alpha_1^d = 2$ and $\alpha_2^d = 1.3$ ($d \in \{1,2\}$). For PT drivers, we assume larger coefficients as $\alpha_1^d = 4$ and $\alpha_2^d = 2$ ($d \in \{3,4\}$) to depict that PT drivers have a higher duration cost than that of the NT drivers given the same work duration. This setting is valid in reality that part-time drivers spend their free time to do a part-time job so that they are more sensitive to work hours than full-time drivers. Similarly, the work, transition, rest, and exit costs are all estimated in this paper because there is no empirical data for our reference.

The values of other parameters are summarized in Table 4. Based on the Traffic Characteristic Survey (GovHK, 2011), we let the sensitivity of customer demand to the full price of taking taxis be $\rho_{eu} = 0.03$, the values of customers' in-vehicle travel time and waiting time are $\delta = 68$ and $\kappa = 50$, respectively. The fleet size of NTs is assumed to be identical to the fleet size of Hong Kong taxis as $N^{\mathrm{n}} = 18163$ (GovHK, 2020a). The taxi base fare is estimated based on the current fare structure of Hong Kong urban taxis as $F = 50$ (HKD) (GovHK, 2020b).

Table 2 Hourly total travel demand $\bar{Q}_{eu}$

Period	Total travel demand	Period	Total travel demand	Period	Total travel demand
1	169650	9	346478	17	450192
2	169650	10	296478	18	459845
3	169650	11	219180	19	248365
4	169650	12	219180	20	247065
5	200850	13	219180	21	235755
6	206895	14	224835	22	226200
7	441183	15	224835	23	226200
8	456112	16	447381	24	226200

Table 3 Link-specific cost c_{eu}^{d}

Driver group	$d = 1,2$	$d = 3,4$
Entry cost	400	150
Work cost	10	8
Transition cost	0	0
Rest cost	0	0
Exit cost	0	0

Table 4 Input parameters

Parameter	Value
Sensitivity of customer demand to the full price of taking taxis ρ_{eu}	0.03 (1/$)
Base fare F	50 (HKD)
Value of customers' in-vehicle travel time δ	68 (HKD/h)
Value of customers' waiting time for taxis κ	50 (HKD/h)
Parameter of the customer waiting time function Θ	0.0025 (veh·h)
The convergent tolerance for the Frank-Wolfe algorithm ς	0.01

6.2. Optimal solution to the PTSDP

We start by showcasing the optimal solution to the PTSDP using the functions and parameter setting in Section 6.1. The Frank-Wolfe algorithm introduced in Section 5 was coded and complied in MATLAB R2018a on a Desktop with Intel Core i7-7700 CPU 3.60GHz and 64 GB RAM. The Hooke-Jeeves pattern search was conducted with the MATLAB Optimization Toolbox. For comparative purposes, a benchmark (BM) scenario was also designed in which there were no PTs and taxi surcharge.

Table 5 displays the BM and PTSDP results and Figure 6 depicts the customer waiting times in different periods. We observe from Table 5 that the optimal PT fleet size is 3033, the optimal PT shifts are 7 a.m. – 10 a.m. (for $d = 3$) and 4 p.m. – 7 p.m. (for $d = 4$), and the optimal surcharge is zero. Moreover, we see that the introduction of PTs increases consumer surplus (customer waiting time is shorter with a higher service intensity) but decreases producer surplus (the individual profit of each driver in peak-periods is lower due to a higher service intensity), and the overall effect on social welfare is positive. This means that, in this example, introducing PTs helps improve social welfare. Furthermore, as observed in Figure 6, the BM customer waiting time violates the LOS constraint in periods from 8 a.m. – 10 a.m. and from 4 p.m. – 7 p.m. Hence, by introducing PTs, customer waiting times in the above periods are reduced and the LOS constraint is satisfied.

Table 5 Comparison of BM and PTSDP results

	Optimal PT fleet size	Optimal PT shifts	Optimal surcharge	Social welfare ($x10^7$)	Consumer surplus ($x10^7$)	Producer surplus ($x10^7$)
BM	N/A	N/A	N/A	3.86	2.37	1.49
PTSDP	3033	7 a.m. – 10 a.m. 4 p.m. – 7 p.m.	0	3.89	2.44	1.45

We then show the scheduling behaviors of drivers in BM and PTSDP solutions, which are shown in Table 6 and Table 7, respectively. It is interesting to see that in both the BM and PTSDP solutions, drivers in each group work for a full shift. The non-resting behaviors of drivers are probably the consequence of a low cumulative working cost compared with the cumulative revenue that a driver can earn by increasing his/her work duration. We also note that the optimal PT fleet size in this example is exactly equal to the lower bound as introduced in Proposition 1, i.e., $\underline{N}^{\mathrm{p}} = 3033$. With this lower bound of PT fleet size, the LOS constraint is satisfied in all periods and is binding in period $(18,19)$ with $w^{\mathrm{c}}_{(18,19)} = \eta = 0.1$.

Table 6 Results for BM driver scheduling equilibrium

Schedule	Driver group	Start time	End time	Duration	Path flow	Driver cost	Individual profit
1	1	4 a.m.	4 p.m.	12	18163	421.1	392.1
2	2	4 p.m.	4 a.m.	12	18163	397.1	400.0

Table 7 Results for PTSDP driver scheduling equilibrium

Schedule	Driver group	Start time	End time	Duration	Path flow	Driver cost	Individual profit
1	1	4 a.m.	4 p.m.	12	18163	571.6	394.1
2	2	4 p.m.	4 a.m.	12	18163	571.6	376.6
3	3	7 a.m.	10 a.m.	3	3033	211	87.8
4	4	4 p.m.	7 p.m.	3	3033	211	90.2

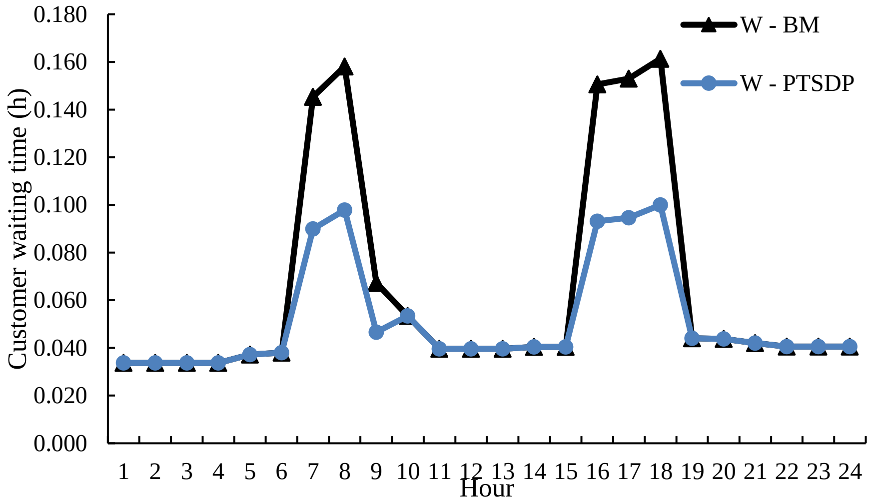


Figure 6 Customer waiting times for taxis in the BM scenario and PTSDP

6.3. Determinant factors to optimal taxi surcharges and PT fleet sizes under given PT shifts

We then examine the two means in real-world practice, namely implementing a peak-period surcharge only and introducing PTs only, to see which one is better in improving social welfare and what are the determinant factors to the optimal PT fleet size and surcharge. To capture the real-world situations and reduce the complexity of our analyses, we fixed the PT shifts to the peak periods only and omitted the LOS constraint (14). The morning PT shift is from 7 a.m. to 9 a.m. and the afternoon PT shift is from 4 p.m. to 7 p.m.

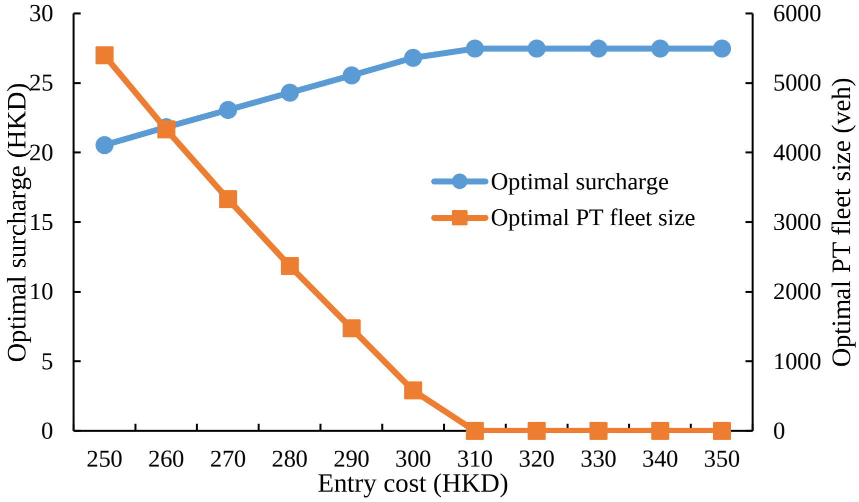
To clearly show the effects of the entry cost, the base fare, and the fleet size of NTs, we designed three scenarios (denoted as *I*, *II*, and *III*). The base fare F and the fleet size of NTs N^n differ among the three scenarios, which are listed in Table 8. In each scenario, we assume that the entry cost $c_{eu}^{d}, \forall (e,u) \in A_0$ (HKD) is the same for all driver groups and let it vary from 250 to 350 at an interval of 10. The corresponding optimal taxi surcharge and PT fleet size are shown in Figure 7.

Table 8 Base fare and fleet size of NTs

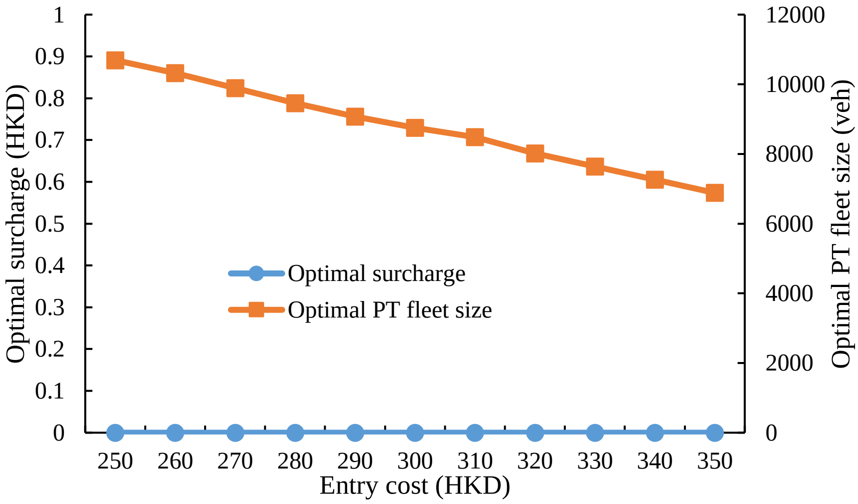
Scenario	Base fare F	Fleet size of NTs N^n
I	20 (HKD)	18163 (veh)
II	50 (HKD)	8000 (veh)
III	50 (HKD)	18163 (veh)

It is interesting to observe from Figure 7 that the values of the three parameters directly affect the optimal taxi surcharge and PT fleet size. Firstly, in Figure 7a (Scenario *I*) where F is relatively low, we observe that when $c_{eu}^{d} < 310$, the optimal surcharge increases with c_{eu}^{d} while the optimal PT fleet size decreases instead. Afterward, the optimal PT fleet size becomes zero and both the optimal surcharge and PT fleet size remain unchanged against c_{eu}^{d}. We note that Figure 7a shows two situations. The first one ($c_{eu}^{d} < 310$) is that both taxi surcharge and PTs are needed to maximize social welfare. The second one is that only a surcharge is needed ($c_{eu}^{d} \geq 310$). Secondly, Figure 7b (Scenario *II*) shows the opposite situation to Figure 7a where F is high but N^{n} is low. In this case, only PTs are needed to maximize social welfare and the surcharge is zero. Thirdly, Figure 7c (Scenario *III*) shows that when both F and N^{n} are high, neither surcharge nor PTs is necessary for the market.

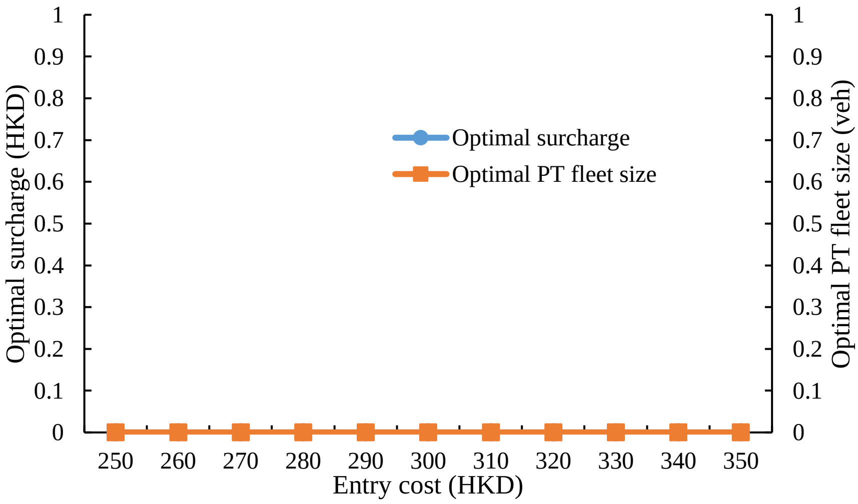
We thus conclude from Figure 7 that in terms of social welfare maximization, the peak-period surcharge and PTs are not always required, depending on the current levels of the entry cost of taxi drivers, the base fare, and the fleet size of NTs. Furthermore, there is no guarantee that which means is better than the other.



a. $F = 20, N^{n} = 18163$



b. $F = 50, N^{n} = 8000$



c. $F = 50, N^{n} = 18163$

Figure 7 Optimal surcharge and PT fleet size against the entry cost of taxi drivers

We use another example to further shed light on the implementation of taxi surcharge and PTs and their impacts on social welfare, consumer surplus, and producer surplus. We assume $c_{eu}^{d} = 250$, $F = 20$, and $N^{\mathrm{n}} = 18163$. Under these settings, we learn from Figure 7a that the optimal taxi surcharge and PT fleet size are around 20.5 (HKD) and 5397.5 (veh), respectively. Then, we obtain the corresponding social welfare, consumer surplus, and producer surplus that are shown in Table 9. In addition, Table 9 also gives the social welfare, consumer surplus, and producer surplus when the government 1) only implements a taxi surcharge, 2) only implements PTs, and 3) does nothing on the current market situation.

As shown in Table 9, implementing both the surcharge and PTs yields the largest social welfare among the four cases and the corresponding consumer and producer surpluses are both larger than those of the do-nothing case. The rise in consumer surplus can be seen as the consequence of the reduction in customer waiting time by pricing out some passengers through the taxi surcharge and introducing more taxis (PTs). Although implementing the surcharge also raises the trip fare so that consumer surplus decreases, the decrease cannot offset the increase by lowering customer waiting time. Besides, the increase in producer surplus compared with the do-nothing case is obviously due to the significant rise in taxi revenue by implementing the surcharge and providing more taxis. Although providing more taxis increases the total operating cost, the resultant increase in total taxi operating cost is insufficient to offset the increase in taxi revenue so that producer surplus increases. Moreover, it is interesting to observe that the two cases in which either the surcharge or PTs is introduced also lead to larger social welfare compared with that of the do-nothing case. However, the changes in consumer surplus and producer surplus are distinct. For the case with surcharge only, consumer surplus falls while producer surplus rises, which is respectively because of the increases in trip fare and taxi revenue resulted from the taxi surcharge. For the case of providing PTs only, consumer surplus is higher but producer surplus is lower than that of the do-nothing case. This is the consequence of a lower waiting time cost of customers and a higher taxi operating cost.

The above results clearly show the potential flaw in solely implementing a surcharge or PTs. Although either method can reduce customer waiting time and mitigate demand-supply imbalance, we see that either customers or taxis are made better off with the other party being made worse off compared with the do-nothing case. By contrast, the implementation of both the surcharge and PTs shows improvements in both the consumer surplus and producer surplus. Therefore, the situations in Singapore and Perth may be further improved if the government considers implementing both surcharge and PTs.

Table 9 System performance under different regulation regimes

Taxi surcharge	PTs	Social welfare (HKD)	Consumer surplus (HKD)	Producer surplus (HKD)
√	√	54142895.1	43006503.6	11136391.5
√	×	54039240.6	40018547.3	14020693.3
×	√	50608419.0	43299757.6	7308661.3
×	×	49643324.1	40598489.5	9044834.5

6.4. How the LOS index affects the optimal PT fleet size/shifts and social welfare

We then investigate how the LOS index η (h) affects the optimal PT fleet size/shifts to PTSDP and social welfare. The taxi surcharge is not considered in this section to better focus on how the LOS index changes the PT shifts and fleet size. As defined in Section 4.1, PTSDP must satisfy the LOS constraint (14). The lower the value of η is, the higher the taxi service quality is. We let η vary from 0.04 (2.4 minutes) to 0.16 (9.6 minutes) at an interval of 0.01. The corresponding optimal PT fleet size/shifts and social welfare were obtained. We also give social welfare in each hour of BM scenario, in which there is no PT nor taxi surcharge.

The optimal PT fleet size and social welfare of BM/PTSDP are depicted in Figure 8, whereas the optimal PT shifts are shown in Table 10. It can be observed from Figure 8 that the optimal PT fleet size decreases with the LOS index and reaches 2816 when $\eta \geq 0.11$. Moreover, the social welfare of PTSDP increases with η and is less than that of BM when $\eta < 0.06$. The lower social welfare of PTSDP compared with that of BM is clearly because of the presence of PTs, which leads to a lower customer waiting time so that both the consumer surplus and taxi revenue increase (taxi fare per ride is fixed). However, the increase in taxi operating cost resulted from the entry of PTs is higher than the increases in consumer surplus and taxi revenue, which results in a decrease in social welfare compared with BM. When $\eta \geq 0.06$, PTSDP yields larger social welfare than the BM. This can be explained by the fact that the BM fleet size can further increase to improve social welfare.

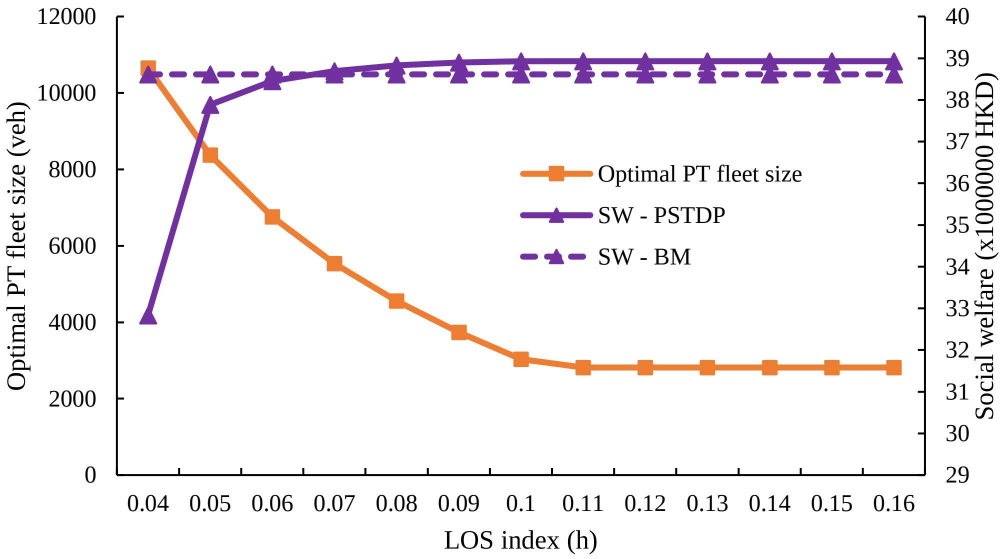


Figure 8 Optimal PT fleet size and social welfare of BM/PTSDP against the LOS index

In terms of optimal PT shifts, Table 10 shows that when $\eta = 0.04$, the two PT shifts are 7 a.m. to 12 p.m. and 2 p.m. to 12 a.m. As η grows, the optimal PT shifts shrink to 7 a.m. to 11 a.m. and 4 p.m. to 7 p.m. (when $\eta = 0.05$) and are further confined to 7 a.m. to 10 a.m. and 4 p.m. to 7 p.m. when $\eta \in [0.06, 0.16]$. Afterward, no PTs are needed and hence the PT shifts are unavailable.

We thus conclude from Figure 8 and Table 10 that the LOS index can influence the optimal PT shifts and fleet size. The presence of PTs improves the taxi service quality by reducing customer waiting time. Yet, the resulting social welfare decreases compared with that of BM. This reveals the possible trade-off that exists between welfare maximization and the level of taxi service. Therefore, whether or not to introduce PTs depends on how the government balance between the service level to taxi passengers and the benefit of the whole society.

Table 10 Optimal PT shifts against the LOS index

η	Morning PT shift	Afternoon PT shift
0.04	7 a.m. to 12 p.m.	2 p.m. to 12 a.m.
0.05	7 a.m. to 11 a.m.	4 p.m. to 7 p.m.
0.06 to 0.16	7 a.m. to 10 a.m.	4 p.m. to 7 p.m.

6.5. How the sensitivity of PT drivers towards work duration affects the optimal PT fleet size/shifts and social welfare

Lastly, we examine how the change of α_2^d ($d \in \{3,4\}$) defined in Eq. (4) alters the optimal PT fleet size/shifts and social welfare. Similar to Section 6.4, taxi surcharge is not considered in this section. The value of α_2^d ($d \in \{3,4\}$) reflects the sensitivity of PT drivers towards the work duration of their schedules. As mentioned in Section 6.1.3, a part-time driver is expected to be more sensitive to work hours compared with a full-time driver because he spends his spare time to do a part-time job. Therefore, by assuming that α_1^d is the same for all driver groups and letting α_2^d ($d \in \{3,4\}$) gradually increase, the situation when α_2^d ($d \in \{1,2\}$) is equal to α_2^d ($d \in \{3,4\}$) can be viewed as the scenario that PT drivers are full-time drivers. In this section, we set $\alpha_1^d = 2,\ \forall d \in D$ and let α_2^d ($d \in \{3,4\}$) vary from 1.3 to 2 at an interval of 0.1. The fleet size of NTs is assumed to be $N^n = 18163$ and the LOS index is set as $\eta = 0.1$. All unspecific parameters take the same values as given in Section 6.1.3.

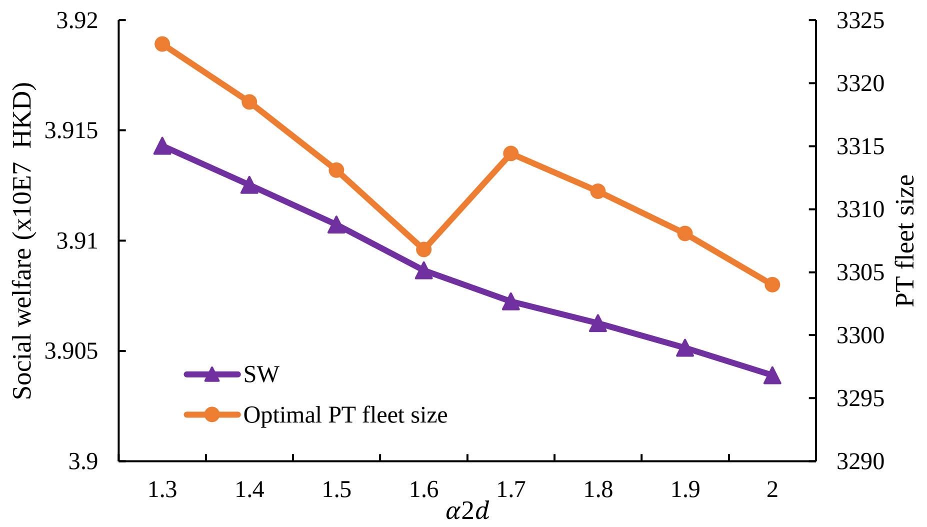


Figure 9 Optimal PT fleet size and social welfare against α_2^d

Figure 9 depicts the optimal PT fleet size and social welfare as α_2^d varies and Table 11 shows the optimal PT shifts. We observed from Figure 9 that, on one hand, the optimal PT fleet size decreases with α_2^d when $\alpha_2^d \in [1.3, 1.6]$ and $\alpha_2^d \in [1.7, 2]$. When α_2^d increases from 1.6 to 1.7, the optimal fleet size has a sudden rise from 3306 to 3314 because the optimal shifts also change. On the other hand, social welfare decreases with α_2^d, which is resulted from the rising path-specific cost (and therefore the falling producer surplus) as α_2^d increases.

Table 11 tells us that the optimal PT shifts shrink as α_2^d rises. This can be explained in a way that as the PT drivers become more sensitive to work duration, extending their shifts to off-peak hours results in a higher total taxi operating cost and the corresponding decrease in producer surplus cannot be off-set by the rise in consumer surplus of the off-peak hours resulted from a lower customer waiting time for taxis. In contrast, if the PT drivers are less sensitive to work duration, extending their shifts to off-peak hours does not contribute too much to the increase in their total operating costs. In this case, the rise in consumer surplus is greater than the fall in producer surplus, and therefore social welfare increases.

Table 11 Optimal PT shifts against α_2^d

α_2^d ($d \in \{3,4\}$)	Morning PT shift	Afternoon PT shift
1.3 to 1.6	7 a.m. to 11 a.m.	4 p.m. to 10 p.m.
1.7 to 2	7 a.m. to 9 a.m.	4 p.m. to 7 p.m.

The above results suggest that in some cases, it would be better to hire full-time drivers (or more generally, those who are less sensitive to work duration) as PT drivers bring higher social welfare and longer PT shifts that benefit the customers for longer periods in a day. However, full-time drivers driving PTs may earn less profit compared with those driving NTs because PT shifts are shorter than NT shifts. Therefore, the government should work out some methods (e.g., subsidizing the PT drivers) to balance the profits between PT and NT drivers. Otherwise, a full-time driver may not be willing to be a PT driver for the comparatively lower profit he can earn, causing a shortage of labor supply to PTs.

7. Conclusion

We propose a peak-period taxi scheme design problem to simultaneously determine the peak-period taxi surcharge and the optimal fleet size/shifts of PTs in a regulated taxi market. The problem is formulated as a bi-level program, in which the upper level is the government problem and the lower level refers to the taxi driver problem. The upper-level objective is to maximize social welfare and we require that customer waiting time for taxis in each period of a day must be lower than a predetermined value to guarantee the taxi service quality. The lower level is an equilibrium problem that describes the scheduling behaviors of taxi drivers working in different shifts. A time-expanded network is used to depict the time-of-day dynamics of demand for and supply of taxis. The bi-level program is solved by a brute force method combined with the Hooke-Jeeves pattern search and the Frank-Wolfe algorithm. Numerical experiments are conducted to give policy implications and managerial insights into the regulation of taxi markets. In summary, we have the following findings or insights.

1. In terms of social welfare maximization, the need for a peak-period taxi surcharge or PTs is highly dependent on the entry cost of taxi drivers, the current taxi fare, and the fleet size of NTs. Moreover, either implementing a surcharge or PTs can mitigate the demand-supply imbalance, but solely implementing a surcharge or PTs may yield a sub-optimal result to the market. Our experimental results show that implementing a surcharge and PTs simultaneously can reach social optimum;
2. The LOS index can directly affect the optimal fleet size and shifts of PTs. A smaller LOS index (i.e., a higher requirement of LOS) implies the need for more PTs but lower social welfare. Therefore, there is a trade-off between social welfare maximization and taxi service quality. The LOS index, which is based on the empirical evidence for taxi passengers' preference for waiting time, is critical to support the government in the decision-making process;
3. The optimal PT shifts are affected by the sensitivity of PT drivers towards the work duration. The optimal PT shifts become shorter and social welfare falls if PT drivers are more sensitive to work duration. Therefore, it is suggested that some full-time drivers should be hired as PT drivers (and working on a split-shift), but the government should work out some methods to address other issues such as balancing the profit levels between NT and PT drivers.

We believe that this study provides several directions for future studies. First, it would be meaningful to further analyze the mathematical properties of the proposed bi-level optimization model and to develop efficient solution methods. Second, this study only considers the drivers' decision on working hours, which is also known as the supply at the intensive margin, yet the drivers' participation decisions (also known as supply at the extensive margin) can still be considered (i.e., the decisions on whether to be taxi drivers or not) in future studies. Third, another group of essential stakeholders in the taxi market, i.e., the taxi companies, is not specified in this study. In many cities, taxi drivers do not own taxis but lease from the taxi companies with a rental fee. Therefore, one possible extension to this study is to investigate the interaction among the government (or regulator), taxi companies, and taxi drivers.

Appendix A

The following notations are used in this paper:

Sets

T	Set of nodes in the time-expanded network;
A	Set of links in the time-expanded network;
$A_0 - A_4$	Sets of entry, work, transition, rest, exit links in the time-expanded network;
P	Set of paths in the time-expanded network;
D	Set of taxi driver groups (shifts).

Indices

t	Index of node in the time-expanded network;
(e,u)	Index of link in the time-expanded network;
d	Index of driver group;
p	Index of path (work schedule) in the time-expanded network.

Decision variables

Upper-level decision variables

γ	Non-negative continuous decision variable which represents a taxi surcharge;
N^{p}	Non-negative continuous decision variable which represents the fleet size of PTs;
X_t^d	Binary decision variable which equals 1 if work node t is selected into the shift of group d drivers, and 0 otherwise;
y_{eu}^d	Non-negative continuous decision variable which indicates the amount of imaginary flow of group d drivers from e to u;

ϕ_{eu}^{d} Binary decision variable which equals 1 if work link (e,u) is selected into the shift of group d drivers, and 0 otherwise;

s_{t}^{d} Binary decision variable which equals 1 if work node t is chosen as the sink, and 0 otherwise;

χ_{eu} Binary decision variable which equals 1 if a taxi surcharge is implemented in period (e,u), and 0 otherwise;

$\mathbf{X}$ $[X_{t}^{d}]$;

$\mathbf{y}$ $[y_{eu}^{d}]$;

$\mathbf{\Phi}$ $[\phi_{eu}^{d}]$;

$\mathbf{s}$ $[s_{t}^{d}]$;

$\boldsymbol{\chi}$ $[\chi_{eu}]$.

Lower-level decision variables

f_{p}^{d} Non-negative continuous lower-level decision variable which indicates the number (flow) of group d drivers working in schedule (path) p;

$\mathbf{f}$ $[f_{p}^{d}]$.

Functions

Q_{eu} Customer demand for taxis in period (e,u) (trips/h);

F_{eu} Taxi fare per ride in period (e,u) (HKD);

v_{eu}^{d} Flow on link (e,u) with respect to group d drivers (veh/h);

v_{eu} Total flow on link (e,u) (veh/h);

w_{eu}^{c} Customer waiting time for taxis in period (e,u) (h);

R_{eu} Average revenue of all taxi drivers in period (e,u) (HKD);

c_{eu}^{d} Link-specific cost of group d drivers traversing link (e,u) (HKD);

c_{p}^{d} Path-specific (duration) cost of group d drivers on path p (HKD);

h_{p}^{l} Total working hour of sub-shift l in schedule p (h);

Parameters

$\bar{Q}_{eu}$	Total travel demand in period (e,u) (trips/h);
F	Flag fare (HKD);
θ	Parameter that reflects the sensitivity of customer demand towards the full price of taking taxis (1/HKD);
κ	Value of customer waiting time for taxis (HKD/h);
δ	Value of customer in-vehicle travel time (HKD/h);
l_{eu}	Average trip travel time in period (e,u) (h);
N^{n}	Fleet size of NTs (veh);
ω_{eu}^{p}	Link-path incidence which equals 1 if path p traverses link (e,u), and 0 otherwise;
η	LOS index of taxi service (h);
Θ, β_1, β_2	Parameters of the Cobb-Douglas meeting function.

References

Arnott, R., 1996. Taxi travel should be subsidized. Journal of Urban Economics, 40 (3), 316–333.

Cairns, R.D., Liston-Heyes, C., 1996. Competition and regulation in the taxi industry. Journal of Public Economics, 59 (1), 1–15.

Camerer, C., Babcock, L., Loewenstein, G., Thaler, R., 1997. Labor supply of New York City cabdrivers: one day at a time. The Quarterly Journal of Economics. 112 (2), 407–441.

Chen, M.K., Sheldon, M., 2016. Dynamic pricing in a labor market: Surge pricing and flexible work on the uber platform. In: EC, p. 455. <https://dl.acm.org/citation.cfm?id=2940798>.

Douglas, G.W., 1972. Price regulation and optimal service standards: the taxicab industry. Journal of Transport Economics and Policy, 6 (2), 116–127.

Farber, H.S., 2005. Is tomorrow another day? The labor supply of New York City cabdrivers. Journal of Political Economy, 113, 46–82.

Farber, H.S., 2015. Why you can't find a taxi in the rain and other labor supply lessons from cab drivers. The Quarterly Journal of Economics, 130, 1975–2026.

GovHK, 2011. Traffic Characteristic Survey. The Government of the Hong Kong Special Administrative Region.

GovHK, 2017. The Annual Traffic Census. The Government of the Hong Kong Special Administrative Region.

GovHK, 2020a. Transport in Hong Kong. The Government of the Hong Kong Special Administrative Region <https://www.td.gov.hk/en/transport_in_hong_kong/public_transport/taxi/index.html> (accessed on 20 July 2020).

GovHK, 2020b. Transport in Hong Kong. The Government of the Hong Kong Special Administrative Region <https://www.td.gov.hk/en/transport_in_hong_kong/public_transport/taxi/taxi_fare_of_hong_kong/index.html> (accessed on 20 July 2020).

GovWA, 2014. On-demand Transport Industry Status Report. The Government of Western Australia.

GovWA, 2016. On-demand Transport Industry Status Report. The Government of Western Australia.

Guda, H., Subramanian, U., 2019. Your Uber is arriving: Managing on-demand workers through surge pricing, forecast communication, and worker incentives. Management Science. (in press) 10.1287/mnsc.2018.3050.

He, F., Shen, Z.J.M., 2015. Modeling taxi services with smartphone-based e-hailing applications. Transportation Research Part C- Emerging Technologies, 58, 93–106.

He, F., Wang, X., Lin, X., Tang, X., 2018. Pricing and penalty/compensation strategies of a taxi-hailing platform. Transportation Research Part C- Emerging Technologies, 86, 263–279.

Hong Kong Extras, 2020. Taxis <https://www.hongkongextras.com/taxis.html> (accessed on 20 July 2020).

Ke, J., Zheng, H., Yang, H., Chen, X., 2017. Short-term forecasting of passenger demand under on-demand ride services: a spatio-temporal deep learning approach. Transportation Research Part C-Emerging Technologies, 85, 591–608.

Ke, J., Cen, X., Yang, H., Chen., X., Ye, J., 2019a. Modelling drivers' working and recharging schedules in a ride-sourcing market with electric vehicles and gasoline vehicles. Transportation Research Part E-Logistic and Transportation Review, 125, 160–180.

Ke, J., Yang, H., Zheng, H., Chen, X., Jia, Y., Gong, P., Ye, J., 2019b. Hexagon-based convolutional neural network for supply-demand forecasting of ride-sourcing services. IEEE Transactions on Intelligent Transportation Systems, 20 (11), 4160–4173.

Ke, J., Xiao, F., Yang, H. and Ye, J., 2020. Learning to delay in ride-sourcing systems: a multi-agent deep reinforcement learning framework. IEEE Transactions on Knowledge and Data Engineering. (in press) 10.1109/tkde.2020.3006084.

Salanova, J.M., Estrada, M., 2019. Social optimal shifts and fares for the Barcelona taxi sector. Transport Policy, 76, 111–122.

SG Observer, 2019. 9 Cab Fees and Taxi Surcharges in Singapore That You Can Easily Avoid <https://www.sgobserver.com/2019/03/9-taxi-surcharge-and-fees-in-singapore-that-you-need-to-avoid/> (accessed on 11 November 2019).

Shirabe, T., 2005. A model of contiguity for spatial unit allocation. Geographical Analysis, 37 (1), 2–16.

Swan Taxi Limited, 2019. Peak Period Taxis. <http://www.swantaxis.com.au/peak_period_taxis.php> (accessed on 11 November 2019).

Sun, H., Wang, H., Wan, Z., 2019a. Model and analysis of labor supply for ride-sharing platforms in the presence of sample self-selection and endogeneity. Transportation Research Part B-Methodological, 125, 76–93.

Sun, H., Wang, H., Wan, Z., 2019b. Flexible labor supply behavior on ride-sourcing platforms. Available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3357365.

TLC, 2019. About TLC. New York City Taxi & Limousine Commission <https://www1.nyc.gov/site/tlc/about/about-tlc.page> (accessed on 11 November 2019).

Qian, X., Ukkusuri, S., 2017. Time-of-day pricing in taxi markets. IEEE Transactions on Intelligent Transportation Systems, 18 (6), 1610–1622.

Wang, X., He, F., Yang, H., Gao, H.O., 2016. Pricing strategies for a taxi-hailing platform. Transportation Research Part E-Logistic and Transportation Review, 93, 212–231.

Yang, H., Qin, X., Ke, J., Ye, J., 2020a. Optimizing matching time interval and matching radius in on-demand ride-sourcing markets. Transportation Research Part B-Methodological, 131, 84–105.

Yang, H., Shao, C., Wang, H., Ye, J., 2020b. Integrated reward scheme and surge pricing in a ridesourcing market. Transportation Research Part B-Methodological, 134, 126–142.

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

Yang, H., Ye, M., Tang, W.H., Wong, S.C., 2005a. Regulating taxi services in the presence of congestion externality. Transportation Research Part A-Policy and Practice, 39 (1), 17–40.

Yang, H., Ye, M., Tang, W.H., Wong, S.C., 2005b. A multiperiod dynamic model of taxi services with endogenous service intensity. Operations Research, 53 (3), 501–515.

Zha, L., Yin, Y., Du, Y., 2018. Surge pricing and labor supply in the ride-sourcing market. Transportation Research Part B-Methodological, 117, 708–722.