JOURNAL OF ADVANCED TRANSPORTATION
J. Adv. Transp. 2015; **49**:96–113
Published online 18 February 2014 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/atr.1267

Taxis in road pricing zone: should they pay the congestion charge?

Jincheng Zhu, Feng Xiao* and Xiaobo Liu

School of Transportation and Logistics, Southwest Jiaotong University, Chengdu, Sichuan, China

SUMMARY

This study aims at investigating the impact and feasibility of charging taxis with toll fee in the pricing zone when designing congestion pricing scheme. A bi-level programming model is developed to compare the maximum social welfares before and after the congestion charge is imposed on taxis. The lower level is a combined network equilibrium model formulated as a variational inequality program, which considers the logit-based mode split, route choice, elastic demand, and vacant taxi distributions. The upper level is to maximize the social welfare when toll rates vary. The bi-level problem can be solved by the genetic algorithm, whereas the lower level is solved by the block Gauss–Seidel decomposition approach together with the method of successive averages and diagonalization algorithm. An application with numerical examples is conducted to demonstrate the effectiveness of the proposed model and algorithm and to reveal some interesting findings. Copyright © 2014 John Wiley & Sons, Ltd.

KEY WORDS: taxis; road pricing; social optimum; bi-level programming

1. INTRODUCTION

Since Pigou first proposed the concept of congestion pricing in 1920 [1], abundant studies have been focused on different aspects of this subject, such as optimal toll rate, minimum toll revenue, equity effects, selection of charging locations, and public acceptance [2–8].

However, little attention in the literature has been paid to modeling the effects of charging taxi on social welfare within the context of the competition of multiple modes. It is a non-negligible issue when designing the road pricing policies, because taxis have accounted for higher and higher percentage of the overall traffic demand in urban area over years. And taxi always operates on the street searching for the next customer even when it is vacant; thus, its impact on the traffic congestion is consistent. Especially for metropolitan area such as Hong Kong, taxis contribute as high as 50–60% of the traffic [9, 10]. In recent years, taxis in Hong Kong have served more than one million passengers each day [11, 12]. In London, there are about 55,998 taxis cruising in the road network, which is equivalent to eight taxis per thousand inhabitants on average [13]. Because of the high demand of taxi in the urban area, whether charging taxis or not in road pricing zone could significantly change the composition of the traffic as well as the traffic flow distribution in a network, so that leads to different system performances.

Many studies have been focused on the taxi services and market regulations. Yang and Wong [14] made an initial attempt to model taxi services in a road network with a given customer OD demand pattern. This model simultaneously described occupied and vacant taxi movements in a steady state of equilibrium considering the effects of taxi fleet size. As an extension to the work of Yang and Wong [14], Wong *et al.* [15] incorporated congestion effects and demand elasticity into the model and

*Correspondence to: Feng Xiao, School of Transportation and Logistics, Southwest Jiaotong University, Chengdu, Sichuan, China. E-mail: xiaofeng@swjtu.edu.cn

reformulated the problem as a bi-level programming formulation. Wong *et al*. [16] further extended the work of Yang and Wong [14] and Wong *et al*. [15] to the case of multi-user, multiple taxi modes (normal, luxury, and restricted area taxis), and customer hierarchical logit-based modal choice for a given OD demand pattern. The distance-based and congestion-based taxi fare charging mechanisms are also introduced in addition to the existing taxi services.

Yang *et al*. [17] analyzed the nature of demand–supply equilibrium for taxi services in a regulated market including competitive and monopoly market. Yang *et al*. [10] then extended this model by taking account of congestion externalities due to both taxi and normal vehicle movements. Yang and Yang [12] proposed a bilateral searching and meeting function to describe the equilibrium properties of taxi market.

However, few of previous studies have concentrated on the effect of congestion charge on taxi. King and Peters [18] examined the impact of road pricing on the route choice and travel time of taxi trips between lower Manhattan and LaGuardia airport in Queens in the USA. It is concluded that, for passengers, usage of toll road for taxi represents a cost greater than the benefit, unless one has extremely high value of time (VOT) (about $170 per hour). It is worth noting that the work of King and Peters [18] considered the impacts of charging taxis on occupied taxis only. The impacts of the toll scheme on other modes in the system and the overall performance of the system were not investigated.

The equilibrium of multi-modal network has also been widely explored over the past 30 years [19–22]. The asymmetric interactions among different modes are usually taken into considerations, and the logit-based model is commonly adopted to describe the mode choice behavior. Considering the performance of the whole system with multiple traffic modes, there will be two effects when toll is imposed on taxis. The first effect is positive: Because an optimized toll scheme can effectively internalized the external cost of taxis, so that reduces the congestion cost caused by the selfish-routing behavior of taxis. The second effect is negative: Because increasing the cost of taking taxi will encourage the use of private car, a less efficient mode with higher operating cost compared with taxi or public transit. Therefore, the overall effect of charging taxi on the social welfare is not yet clear.

This paper compares the maximum social welfares obtained before and after toll is imposed on taxi by considering multi-modal and variable demands. A bi-level model is introduced where the lower level is a combined network equilibrium model (CNEM) that involves logit-based mode split, route choice, elastic demand, and vacant taxi distributions and the upper level aims at social welfare maximization. Because the interactions of network flows are asymmetric, the lower level is formulated as an equivalent variational inequality (VI) program. The numerical example indicates that given a certain level of bus service, whether to charge taxis mainly relies on the ratio of operating cost of private car to that of taxi. The threshold of the ratio is 9 (which can hardly be reached in the real world) in the numerical example based on realistic data settings. So far, as the ratio is lower than the threshold value, charging taxis always outweighs exempting them from the toll in terms of the maximum social welfare. And when the ratio is lower than the threshold, the ratio between the maximum social welfares of charging and not charging the taxi decreases with the taxi fare levels but does not affect much by the taxi fleet size, while the conclusion that charging taxis gains higher maximum social welfare still holds.

In the next section, the CNEM is developed. Section 3 constructs a VI formulation for the CNEM. Section 4 presents the upper-level social welfare maximization problem. Section 5 introduces the block Gauss–Seidel decomposition method together with the method of successive averages and diagonalization algorithm to solve the CNEM, while the bi-level programming model is solved by the genetic algorithm (GA). In Section 6, a numerical example is provided to illustrate the effectiveness of the methodology proposed, and some interesting findings are highlighted. Conclusions and future expansions are given in Section 7.

2. THE COMBINED NETWORK EQUILIBRIUM MODEL

2.1. Preliminaries

Consider a road network $G(V,A)$ where V is the set of nodes and A is the set of links. $\overline{A}$ is the set of toll links, $\overline{A} \subseteq A$. Let R and S be the sets of origin and destination nodes. In the following, the superscripts

DOI: 10.1002/atr

"p," "b," "o," and "v" indicate private car, bus, occupied taxi, and vacant taxi, respectively. The traffic flow is composed by all of them in the road network. Additionally, $m \in M = (p,b,o)$ represents the combination of private car, occupied taxi, and bus mode.

Let q_{rs} be the total demand between OD pair $r \in R$ and $s \in S$. We then have

$$q_{rs} = q_{rs}^p + q_{rs}^b + q_{rs}^o, \quad r \in R, s \in S \tag{1}$$

where q_{rs}^p, q_{rs}^b, and q_{rs}^o are the traffic demands of private car, bus, and occupied taxi from origin $r \in R$ to destination $s \in S$, respectively. Furthermore, for the taxi mode, we have the following trip end equations

$$O_r^o = \sum_{s \in S} q_{rs}^o, \quad r \in R \tag{2}$$

$$D_s^o = \sum_{r \in R} q_{rs}^o, \quad s \in S \tag{3}$$

where O_r^o and D_s^o are the demands for taxi mode from origin zone $r \in R$ and to destination zone $s \in S$, respectively.

2.2. Generalized costs

2.2.1. Generalized costs of the private car and taxi mode

Let c_a^p, c_a^o, and c_a^v be the generalized costs on link $a \in A$ for private car, occupied taxi, and vacant taxi, respectively. And all of them are assumed to be a linear function of link length d_a, link travel time t_a, and toll y_a (if toll is charged). Let b^p and b^v be the operating costs per unit distance for private car and taxi. Without loss of generality, we assume $b^p > b^v$. Additionally, b_0^o, b_1^o, and b_2^o represent the preliminary flag-fall charge per ride and the mileage-based and delay-based[1] taxi fares that are charged to customers who take taxi. Then, we have the following cost structures if taxis are tolled in the road pricing zone [10,16].

$$c_a^p = \lambda t_a(x_a) + b^p d_a + y_a, \quad a \in A \tag{4}$$

$$c_a^o = \lambda t_a(x_a) + b_1^o d_a + b_2^o \left(t_a(x_a) - t_a^0\right) + y_a, \quad a \in A \tag{5}$$

$$c_a^v = \lambda_v t_a(x_a) + b^v d_a + y_a, \quad a \in A \tag{6}$$

where λ is the VOT for users taking private car or taxi while λ_v is for taxi drivers. $t_a(x_a)$ is the travel time and is supposed to be an increasing function of total flow x_a on link $a \in A$. t_a^0 is the free flow travel time. Note that if taxis are exempt from the congestion charge, there are no tolls for them but private cars still have to pay.

The total generalized costs for private car and taxi on route $k \in K_{rs}$ between origin $r \in R$ and destination $s \in S$ are presented as follows:

$$C_{rs,k}^p = \sum_{a \in A} c_a^p \delta_{a,k}^{rs}, \quad r \in R, s \in S, k \in K_{rs} \tag{7}$$

$$C_{rs,k}^o = \sum_{a \in A} c_a^o \delta_{a,k}^{rs} + b_0^o + \lambda_{ow} W_r^o, \quad r \in R, s \in S, k \in K_{rs} \tag{8}$$

[1]Delay-based taxi fare is an additional fee that customers have to pay when taxis are subject to congestion.

DOI: 10.1002/atr

$$C^v_{sr,k} = \sum_{a\in A} c^v_a \delta^{sr}_{a,k}, \quad r\in R, s\in S, k\in K_{sr} \tag{9}$$

where W^o_r is an endogenous variable that denotes the customer waiting time for taxi at zone r. λ_{ow} is the value of customer waiting time. Note that in practice, the taxi cost is not strictly linearly proportional to the travel distance because of the fixed flag-fall charge for the first two or three kilometers. Yet as the travel distance increases, the linear approximation becomes more accurate. Thus, for simplicity, we directly used the linear assumption as made by Yang *et al.* [10] and Wong *et al.* [16].

According to Wong *et al.* [15], we can specify the expected customer waiting time as a function of the cruising vacant taxi hours and the area of the zone.

$$W^o_r = \eta \frac{Z_r}{N^v_r w^v_r}, \quad r\in R \tag{10}$$

where Z_r is the area of zone $r\in R$ and η is a model parameter that is common to all zones. w^v_r is the waiting/searching time of vacant taxi in zone r. N^v_r is the number of vacant taxis meeting customers in zone r per hour. It is noteworthy that at equilibrium, we have $N^v_r = O^o_r$. Thus, Equation (10) can be represented as follows:

$$W^o_r = \eta \frac{Z_r}{O^o_r w^v_r}, \quad r\in R \tag{11}$$

2.2.2. Generalized cost of the bus mode

It is generally known that dedicated bus lanes are now common in many metropolises. And this kind of lanes can only be used by busses and thus can facilitate faster movement of busses. In this paper, we suppose that dedicated bus lanes are available in the network, and hence, there is no interaction between bus and private car or taxi.

Also, for each OD pair, it is assumed that there is one bus line [23]. Therefore, the generalized costs of bus passengers between origin $r\in R$ and destination $s\in S$ (denoted as C^b_{rs}) can be described as follows:

$$C^b_{rs} = \lambda_b T_{rs} + \zeta G_{rs}\left(q^b_{rs}\right) + \lambda_{bw} W^b_{rs} + \tau, \quad r\in R, s\in S \tag{12}$$

where T_{rs} is the bus travel time. With the assumptions described earlier as well as the given bus schedule and frequency, T_{rs} is constant here. λ_b is the VOT for bus passengers. $G_{rs}\left(q^b_{rs}\right)$ is the crowding discomfort experienced by bus passengers, which is an increasing function of the number of travelers choosing the bus. ζ is the unit cost of discomfort. W^b_{rs} is the waiting time of bus passengers, and we specify it as $W^b_{rs} = \frac{\alpha}{F_{rs}}$, where F_{rs} is the bus frequency. Generally, the value of α is set to 0.5 when the passenger arrival is assumed to be a uniform random distribution and the bus headway is constant. λ_{bw} is the waiting time value of bus passenger. τ is the bus fare.

Furthermore, the total flow on each link $a\in A$ can be obtained through the following equation:

$$x_a = \sum_{r\in R, s\in S} \sum_{k\in K_{rs}} \left(f^p_{rs,k} + f^o_{rs,k}\right)\delta^{rs}_{ak} + \sum_{s\in S, r\in R} \sum_{k\in K_{sr}} f^v_{sr,k}\delta^{sr}_{ak}, \quad a\in A \tag{13}$$

where $f^p_{rs,k}$ and $f^o_{rs,k}$ are the flow on route $k\in K_{rs}$ for private car and occupied taxi, respectively. $f^v_{sr,k}$ is vacant taxi flow on route $k\in K_{sr}$, where K_{rs} and K_{sr} are the sets of paths between zone $r\in R$ and zone $s\in S$. δ^{rs}_{ak} and δ^{sr}_{ak} are link-route indicator variables which are 1 if route k between OD pair $r\in R$ and $s\in S$ uses link a, and 0 otherwise.

2.3. *Taxi service time constraint*

We assume that there are N cruising taxis that operate in the network, and in one unit period (1 h) operations of taxis, the total taxi service time consists of occupied time (denoted as TO) and empty time (denoted as TV). In a stationary state, the total taxi occupied time is equal to the taxi hours that complete all the customer demands q_{rs}^{o} and thus is given by

$$TO = \sum_{r \in R} \sum_{s \in S} q_{rs}^{o} h_{rs}, \quad r \in R, s \in S \tag{14}$$

where h_{rs} is the average travel time from origin $r \in R$ to destination $s \in S$, and can be represented as $h_{rs} = \frac{\sum_{k \in K_{rs}} \left(f_{rs,k}^{o} \sum_{a \in A} t_a \delta_{a,k}^{rs} \right)}{\sum_{k \in K_{rs}} f_{rs,k}^{o}}$ [16]. The total empty time of taxis is composed of moving times from zone $s \in S$ to zone $r \in R$ and waiting/searching times in the zones. Thus, it can be computed by

$$TV = \sum_{s \in S} \sum_{r \in R} q_{sr}^{v} \left(h_{sr} + w_r^{v} \right), \quad r \in R, s \in S \tag{15}$$

where q_{sr}^{v} is the number of vacant taxis traveling from zone $s \in S$ to zone $r \in R$ to search for customers. Therefore, the following constraint should be satisfied within 1-h period [14].

$$\sum_{r \in R} \sum_{s \in S} q_{rs}^{o} h_{rs} + \sum_{s \in S} \sum_{r \in R} q_{sr}^{v} \left(h_{sr} + w_r^{v} \right) = N \tag{16}$$

where N is the taxi fleet size.

2.4. *Traffic assignment*

It is assumed that the route choices of all travelers including private car, bus passenger (although for each OD pair there is one bus line), occupied taxi, and vacant taxi follow user equilibrium. And we further suppose that the routes of occupied taxi are determined by the customers taking the taxi.

At equilibrium, the following conditions have to be satisfied.

$$C_{rs,k}^{m} = u_{rs}^{m}, \text{if } f_{rs,k}^{m} > 0, \quad r \in R, s \in S, k \in K_{rs}, m \in M \tag{17}$$

$$C_{rs,k}^{m} \geq u_{rs}^{m}, \text{if } f_{rs,k}^{m} = 0, \quad r \in R, s \in S, k \in K_{rs}, m \in M \tag{18}$$

and

$$C_{sr,k}^{v} = u_{sr}^{v}, \text{if } f_{sr,k}^{v} > 0, \quad r \in R, s \in S, k \in K_{sr} \tag{19}$$

$$C_{sr,k}^{v} \geq u_{sr}^{v}, \text{if } f_{sr,k}^{v} = 0, \quad r \in R, s \in S, k \in K_{sr} \tag{20}$$

where u_{rs}^{m} and u_{sr}^{v} are the minimum generalized costs for mode $m \in M$ and vacant taxi between origin $r \in R$ and destination $s \in S$, respectively.

2.5. *Logit mode split*

We propose the following logit-based mode choice function, which is able to give the proportion of trips taken by the mode $m \in M$ between origin $r \in R$ and destination $s \in S$ at equilibrium.

DOI: 10.1002/atr

$$P_{rs}^{m} = \frac{\exp\left[-\beta\left(u_{rs}^{m} - \varphi_{rs}^{m}\right)\right]}{\sum\limits_{i \in M} \exp\left[-\beta\left(u_{rs}^{i} - \varphi_{rs}^{i}\right)\right]}, \quad r \in R, s \in S, m \in M \tag{21}$$

Then, we have the number of travelers who take mode m as follows:

$$q_{rs}^{m} = \frac{\exp\left[-\beta\left(u_{rs}^{m} - \varphi_{rs}^{m}\right)\right]}{\sum\limits_{i \in M} \exp\left[-\beta\left(u_{rs}^{i} - \varphi_{rs}^{i}\right)\right]} q_{rs}, \quad r \in R, s \in S, m \in M \tag{22}$$

Here, φ_{rs}^{m} represents the attraction of mode m for travelers between origin $r \in R$ and destination $s \in S$. β is the dispersion coefficient. From this function, we can see that the number of travelers for each mode is proportional to the attraction and inversely proportional to the minimum generalized cost.

2.6. Vacant taxi distributions

In addition to private car and occupied taxi, vacant taxi also contributes significantly to the traffic congestion, because there are always considerable amount of vacant taxis searching for customers on the limited road space. Vacant taxi drivers may change their initial destinations and paths if the toll is imposed on taxi. In this section, the following logit model is proposed to describe the vacant taxi behaviors on the road network [14]. As in the paper of Yang and Wong [14], here, we suppose that every taxi driver attempts to spend the minimal expected search time in meeting customer and the expected search time is a random variable that is identically distributed with a Gumbel density function.

$$P_{r/s} = \frac{\exp\left[-\sigma\left(u_{sr}^{v} + \lambda_{v} w_{r}^{v}\right)\right]}{\sum\limits_{i \in R} \exp\left[-\sigma\left(u_{si}^{v} + \lambda_{v} w_{i}^{v}\right)\right]}, \quad s \in S, r \in R \tag{23}$$

where $P_{r/s}$ is the probability that vacant taxi departs from zone $s \in S$ and meets the customer in zone $r \in R$. σ is a non-negative parameter reflecting the degree of uncertainty for taxi drivers on customer demand and taxi services of the whole market.

And, in a steady state of equilibrium, every customer is eventually able to take a taxi after waiting and searching, and all occupied taxis will become available when passengers arrive at destinations [15]. Thus, we have

$$\sum_{r \in R} q_{sr}^{v} = D_{s}^{o}, \quad s \in S \tag{24}$$

$$\sum_{s \in S} q_{sr}^{v} = \sum_{s \in S} D_{s}^{o} \bullet P_{r/s} = O_{r}^{o}, \quad r \in R \tag{25}$$

2.7. Elastic demand

A demand function is presented here to describe the elasticity of the OD demands.

$$q_{rs} = D_{rs}(u_{rs}), \quad r \in R, s \in R \tag{26}$$

where q_{rs} is the total demand between origin $r \in R$ and destination $s \in S$, which is supposed to be a continuously and strictly decreasing function of users' minimum perceived generalized costs u_{rs}.

DOI: 10.1002/atr

$$u_{rs} = -\frac{1}{\beta}\ln\left\{\sum_{m\in M}\exp\left[-\beta\left(u_{rs}^{m} - \varphi_{rs}^{m}\right)\right]\right\}, \quad r\in R, s\in R \tag{27}$$

Correspondingly, $u_{rs} = D_{rs}^{-1}(q_{rs})$ is the inverse demand function.

3. AN EQUIVALENT VARIATIONAL INEQUALITY PROGRAM

Note that the interactions of network flows are asymmetric because of the delay-based taxi charge. Namely,

$$\frac{\partial c_a^p}{\partial x_a^o} \neq \frac{\partial c_a^o}{\partial x_a^p} \tag{28}$$

where $\frac{\partial c_a^p}{\partial x_a^o}$ and $\frac{\partial c_a^o}{\partial x_a^p}$ are given by

$$\frac{\partial c_a^p}{\partial x_a^o} = \lambda\frac{\partial t_a(x_a)}{\partial x_a^o} \tag{29}$$

$$\frac{\partial c_a^o}{\partial x_a^p} = \lambda\frac{\partial t_a(x_a)}{\partial x_a^p} + b_2^o\frac{\partial t_a(x_a)}{\partial x_a^p} \tag{30}$$

Thus, a VI program is formulated in this section, which is equivalent to the aforementioned CNEM.

The feasible region Ω of our VI formulation is stated later, and related dual variables are also provided in the brackets:

$$\sum_{s\in S} q_{sr}^{v} = O_r^o, \quad r\in R, \quad (e_r) \tag{31}$$

$$\sum_{r\in R} q_{sr}^{v} = D_s^o, \quad s\in S, \quad (a_s) \tag{32}$$

$$q_{rs} = \sum_{m\in M} q_{rs}^{m}, \quad r\in R, s\in S, \quad (\gamma_{rs}) \tag{33}$$

$$\sum_{k\in K_{rs}} f_{rs,k}^{m} = q_{rs}^{m}, \quad r\in R, s\in S, m\in M, \quad \left(u_{rs}^{m}\right) \tag{34}$$

$$\sum_{k\in K_{sr}} f_{sr,k}^{v} = q_{sr}^{v}, \quad s\in S, r\in R, \quad \left(u_{sr}^{v}\right) \tag{35}$$

$$f_{rs,k}^{m} \geq 0, \quad r\in R, s\in S, m\in M, k\in K_{rs}, \quad \left(\chi_{rs,k}^{m}\right) \tag{36}$$

$$f_{sr,k}^{v} \geq 0, \quad r\in R, s\in S, k\in K_{sr}, \quad \left(\chi_{sr,k}^{v}\right) \tag{37}$$

$$q_{rs}^{m} \geq 0, \quad q_{sr}^{v} \geq 0, \quad r\in R, s\in S, m\in M \tag{38}$$

Constraints (31) and (32) are the conservation conditions of flow for vacant taxis. Equation (33) is the conservation equation for total demand. Equations (34) and (35) show that the sum of all path flows

DOI: 10.1002/atr

for private car, bus, occupied taxi, and vacant taxi should be equal to their demands, respectively. Equations (36), (37), and (38) are the non-negativity constraints on path flows and demands, respectively.

With the feasible region described earlier, the VI problem can be stated as follows. Find $\left(f^{m,*}_{rs,k}, f^{v,*}_{sr,k}, q^{*}_{rs}, q^{m,*}_{rs}\right) \in \Omega$, which satisfies

$$\sum_{r\in R,s\in S}\sum_{m\in M}\sum_{k\in K_{rs}} C^{m}_{rs,k}(f^{*})\left(f^{m}_{rs,k}-f^{m,*}_{rs,k}\right)+\sum_{s\in S,r\in R}\sum_{k\in K_{sr}} C^{v}_{sr,k}(f^{*})\left(f^{v}_{sr,k}-f^{v,*}_{sr,k}\right)-\sum_{r\in R,s\in S} D^{-1}_{rs}\left(q^{*}_{rs}\right)\left(q_{rs}-q^{*}_{rs}\right)$$
$$+\sum_{r\in R,s\in S}\sum_{m\in M}\left(\frac{1}{\beta}\ln\frac{q^{m,*}_{rs}}{q_{rs}}-\varphi^{m}_{rs}\right)\left(q^{m}_{rs}-q^{m,*}_{rs}\right)+\sum_{s\in S,r\in R}\frac{1}{\sigma}\ln q^{v,*}_{sr}\left(q^{v}_{sr}-q^{v,*}_{sr}\right)\geq 0 \tag{39}$$

where $C^{m}_{rs,k}$ and $C^{v}_{sr,k}$, $r\in R$, $s\in S$, $m\in M$, $k\in K_{rs}$ are defined by Equations (7)–(9).

Proposition. *The optimality conditions of the proposed VI program are equivalent to the CNEM in Section 2.*

Proof. The KKT (Karush-Kuhn-Tucker) conditions of the VI formulation (39) are given later.

$$f^{m}_{rs,k}:\quad C^{m}_{rs,k}-u^{m}_{rs}-\chi^{m}_{rs,k}=0,\quad r\in R, s\in S, m\in M, k\in K_{rs} \tag{40}$$

$$f^{v}_{sr,k}:\quad C^{v}_{sr,k}-u^{v}_{sr}-\chi^{v}_{sr,k}=0,\quad r\in R, s\in S, k\in K_{sr} \tag{41}$$

$$q^{m}_{rs}:\quad \frac{1}{\beta}\ln\frac{q^{m}_{rs}}{q_{rs}}-\varphi^{m}_{rs}+u^{m}_{rs}-\gamma_{rs}=0,\quad r\in R, s\in S \tag{42}$$

$$q^{v}_{sr}:\quad \frac{1}{\sigma}\ln q^{v}_{sr}+u^{v}_{sr}+e_{r}+a_{s}=0,\quad r\in R, s\in S \tag{43}$$

$$q_{rs}:\quad -D^{-1}_{rs}(q_{rs})+\gamma_{rs}=0,\quad r\in R, s\in S \tag{44}$$

The complementarity conditions are

$$f^{m}_{rs,k}\bullet\chi^{m}_{rs,k}=0,\quad r\in R, s\in S, m\in M, k\in K_{rs} \tag{45}$$

$$f^{v}_{sr,k}\bullet\chi^{v}_{sr,k}=0,\quad r\in R, s\in S, k\in K_{sr} \tag{46}$$

$$\chi^{m}_{rs,k}\geq 0,\quad r\in R, s\in S, m\in M, k\in K_{rs} \tag{47}$$

$$\chi^{v}_{sr,k}\geq 0,\quad r\in R, s\in S, k\in K_{sr} \tag{48}$$

From Equations (45) and (47), we have $\chi^{m}_{rs,k}=0$, if $f^{m}_{rs,k}>0$. Then, from Equation (40), we can obtain that $C^{m}_{rs,k}=u^{m}_{rs}$. And if $f^{m}_{rs,k}=0$, $\chi^{m}_{rs,k}\geq 0$, then $C^{m}_{rs,k}\geq u^{m}_{rs}$. Therefore, we have the following conditions:

DOI: 10.1002/atr

$$C_{rs,k}^{m} = u_{rs}^{m}, \text{if } f_{rs,k}^{m} > 0, \quad r \in R, s \in S, k \in K_{rs}, m \in M \tag{49}$$

$$C_{rs,k}^{m} \geq u_{rs}^{m}, \text{if } f_{rs,k}^{m} = 0, \quad r \in R, s \in S, k \in K_{rs}, m \in M \tag{50}$$

which implies that the route choices of private car, bus commuters, and occupied taxi follow user equilibrium. Similarly, we can demonstrate that the route choice of vacant taxi also satisfies the user equilibrium condition by utilizing Equations (46), (48), and (41).

From Equation (42), we have

$$\frac{q_{rs}^{m}}{q_{rs}} = \exp\left[\beta\left(\varphi_{rs}^{m} - u_{rs}^{m} + \gamma_{rs}\right)\right], \quad r \in R, s \in S, m \in M \tag{51}$$

Taking the sum of m in both sides gives rise to

$$\exp(\beta\gamma_{rs}) = \frac{1}{\sum_{m \in M} \exp\left[\beta\left(\varphi_{rs}^{m} - u_{rs}^{m}\right)\right]}, \quad r \in R, s \in S \tag{52}$$

Substituting Equation (52) into Equation (51) leads to

$$q_{rs}^{m} = \frac{\exp\left[-\beta\left(u_{rs}^{m} - \varphi_{rs}^{m}\right)\right]}{\sum_{i \in M} \exp\left[-\beta\left(u_{rs}^{i} - \varphi_{rs}^{i}\right)\right]} q_{rs}, \quad r \in R, s \in S \tag{53}$$

which is consistent with the logit-based mode split model, Equation (22).

From Equation (52), we have that

$$\gamma_{rs} = -\frac{1}{\beta} \ln\left\{\sum_{m \in M} \exp\left[-\beta\left(u_{rs}^{m} - \varphi_{rs}^{m}\right)\right]\right\}, \quad r \in R, s \in S \tag{54}$$

Substituting Equation (54) into Equation (44) gives rise to

$$D_{rs}^{-1}(q_{rs}) = -\frac{1}{\beta} \ln\left\{\sum_{m \in M} \exp\left[-\beta\left(u_{rs}^{m} - \varphi_{rs}^{m}\right)\right]\right\}, \quad r \in R, s \in S \tag{55}$$

which indicates that the elastic demand function defined in Section 2.7 is satisfied.

Equation (43) can be rewritten as follows:

$$q_{sr}^{v} = \exp\left[-\sigma\left(u_{sr}^{v} + a_{s} + e_{r}\right)\right], \quad r \in R, s \in S \tag{56}$$

DOI: 10.1002/atr

Substituting Equation (56) into Equation (32), we have

$$\exp(-\sigma a_s) = \frac{D_s^o}{\sum_{r\in R} \exp\left[-\sigma\left(u_{sr}^v + e_r\right)\right]}, \quad s\in S \tag{57}$$

Substituting Equation (57) into Equation (56) leads to

$$q_{sr}^v = \frac{\exp\left[-\sigma\left(u_{sr}^v + e_r\right)\right]}{\sum_{i\in R} \exp\left[-\sigma\left(u_{si}^v + e_i\right)\right]} D_s^o, \quad r\in R, s\in S \tag{58}$$

Comparing Equation (58) with the logit-based vacant taxi distribution model (23), we know that $\lambda_v w_r^v$ is associated with e_r. Similar to Wong *et al*. [15], we can calculate taxi waiting/searching time w_r^v through Equations (16), (31), (32), and (56) (For details, one can see Wong *et al*. [15]).

Because the constraints (31) to (38) are non-negative and linear and the continuous formulation (39) has the region of non-negative flows and OD demands, we can conclude that at least one solution to the VI program exists (Wong *et al*. [16]).

4. SOCIAL WELFARE MAXIMIZATION

The upper level aims to derive the optimal congestion tolls under which the social welfare is maximized. Note that social welfare is expressed as the total willingness to pay of the travelers minus the total social cost.

$$\begin{aligned} &\max SW(\mathbf{y}) = \\ &\sum_{r\in R, s\in S} \int_0^{q_{rs}(\mathbf{y})} D_{rs}^{-1}(w)dw - \sum_{a\in A} (\lambda t_a(x_a) + b^p d_a) x_a^p(\mathbf{y}) - \sum_{r\in R, s\in S} C_{rs}^{b\,\prime} q_{rs}^b(\mathbf{y}) - \sum_{a\in A} \lambda t_a(x_a) x_a^o(\mathbf{y}) \\ &- \sum_{r\in R, s\in S} \lambda_{ow} q_{rs}^o(\mathbf{y}) W_r^o(\mathbf{y}) - \sum_{a\in A} \left[\lambda_v t_a(x_a) x_a^v(\mathbf{y}) + b^v d_a\left(x_a^o(\mathbf{y}) + x_a^v(\mathbf{y})\right)\right] - \sum_{s\in S, r\in R} \lambda_v q_{sr}^v(\mathbf{y}) w_r^v(\mathbf{y}) \end{aligned} \tag{59}$$

s.t.

$$y_a^{\min} \le y_a \le y_a^{\max}, \quad a\in \overline{A} \tag{60}$$

where the revenue of tolls, fares that customers pay taxis, and bus fares are not involved in the social costs, because these costs imply only a transfer within the system. $C_{rs}^{b\,\prime}$ is expressed as $C_{rs}^{b\,\prime} = \lambda_b T_{rs} + \zeta G_{rs}\left(q_{rs}^b\right) + \lambda_{bw} W_{rs}^b$. $x_a^p(\mathbf{y}), x_a^o(\mathbf{y})$, and $x_a^v(\mathbf{y})$ denote the flows on link $a\in A$ for private car, occupied taxi, and vacant taxi, respectively. $y_a^{\min}$ and $y_a^{\max}$ are the upper bound and lower bound of the toll on link $a\in\overline{A}$. Note that travel time $t_a(x_a)$, link flows $x_a^p(\mathbf{y}), x_a^o(\mathbf{y})$, and $x_a^v(\mathbf{y})$, demand matrices $q_{rs}(\mathbf{y}), q_{rs}^b(\mathbf{y}), q_{rs}^o(\mathbf{y})$, and $q_{sr}^v(\mathbf{y})$, customer waiting time $W_r^o(\mathbf{y})$, and taxi waiting/searching time $w_r^v(\mathbf{y})$, $a\in A$, $r\in R$, $s\in S$ are derived from solving the lower-level problem.

5. SOLUTION ALGORITHM

The block Gauss–Seidel decomposition approach together with the method of successive averages and diagonalization algorithm is developed to solve the VI formulation (39). One can refer to Florian *et al*. [24] and Wong *et al*. [16] for details. The algorithm follows the procedure later.

Step 0. Initialization. Let the iteration number $l = 0$. Select an initial feasible solution $f_{rs,k}^{m(l)}, f_{sr,k}^{v(l)}$, $r\in R$, $s\in S$, $m\in M$, $k\in K_{rs}$.

DOI: 10.1002/atr

Step 1. Computation of generalized costs. Compute the generalized costs $C^{p(l)}_{rs,k}$, $C^{b(l)}_{rs,k}$, and $C^{v(l)}_{sr,k}$, and then obtain the minimum generalized costs $u^{p(l)}_{rs}$, $u^{b(l)}_{rs}$, and $u^{v(l)}_{sr}$, $r \in R$, $s \in S$, $k \in K_{rs}$. Compute the demands for taxi from origin zone $r \in R$ and to destination zone $s \in S$, that is, O^o_r and D^o_s through $f^{o(l)}_{rs,k}$ and Equations (2) and (3). Therefore, we can find the vacant taxi distribution $q^{v(l)}_{sr}$ and taxi waiting/searching time $w^{v(l)}_r$ by solving the doubly constrained gravity model Equation (56) subject to Equations (31), (32), and (38). Customer waiting time $W^{o(l)}_r$, generalized cost of occupied taxi $C^{o(l)}_{rs,k}$, and its minimization $u^{o(l)}_{rs}$ can then be computed through Equations (11), (5), and (8).

Step 2. Computation of demand matrices. Compute total demand $q^{(l)}_{rs}$ according to elastic demand function Equation (26) and use the logit-based mode choice model, that is, Equation (22) to find demand matrices for private car $q^{p(l)}_{rs}$, bus $q^{b(l)}_{rs}$, and occupied taxi $q^{o(l)}_{rs}$, $r \in R$, $s \in S$, based on the minimum generalized costs obtained in step 1.

Step 3. Network equilibrium assignment. After obtaining the demand matrices $q^{m(l)}_{rs}$ and $q^{v(l)}_{sr}$, $r \in R$, $s \in S, m \in M$, apply diagonalization algorithm to the following fixed demand network equilibrium problem to compute the auxiliary path flows $H^{m(l)}_{rs,k}$ and $H^{v(l)}_{sr,k}$, $r \in R$, $s \in S$, $k \in K_{rs}$, $m \in M$.

$$\sum_{r\in R,s\in S}\sum_{m\in M}\sum_{k\in K_{rs}} C^m_{rs,k}(f^*)\left(f^m_{rs,k}-f^{m,*}_{rs,k}\right)+\sum_{s\in S,r\in R}\sum_{k\in K_{sr}} C^v_{sr,k}(f^*)\left(f^v_{sr,k}-f^{v,*}_{sr,k}\right)\geq 0$$

subject to Equations (34–37).

Step 4. Method of successive averages. Utilize the following equations to find the path flow pattern of next iteration.

$$\begin{aligned} f^{m(l+1)}_{rs,k} &= f^{m(l)}_{rs,k}+\frac{1}{l+1}\left(H^{m(l)}_{rs,k}-f^{m(l)}_{rs,k}\right), \quad r\in R, s\in S, k\in K_{rs}, m\in M \\ f^{v(l+1)}_{sr,k} &= f^{v(l)}_{sr,k}+\frac{1}{l+1}\left(H^{v(l)}_{sr,k}-f^{v(l)}_{sr,k}\right), \quad\quad r\in R, s\in S, k\in K_{rs} \end{aligned}$$

Step 5. Convergence test. If $\dfrac{\sqrt{\sum\limits_{r\in R,s\in S}\sum\limits_{k\in K_{rs}}\sum\limits_{m\in M}\left(f^{m(l+1)}_{rs,k}-f^{m(l)}_{rs,k}\right)^2+\sum\limits_{s\in S,r\in R}\sum\limits_{k\in K_{sr}}\left(f^{v(l+1)}_{sr,k}-f^{v(l)}_{sr,k}\right)^2}}{\sum\limits_{r\in R,s\in S}\sum\limits_{k\in K_{rs}}\sum\limits_{m\in M} f^{m(l)}_{rs,k}+\sum\limits_{s\in S,r\in R}\sum\limits_{k\in K_{sr}} f^{v(l)}_{sr,k}} < \varepsilon$, then go to step 6; otherwise, set $l = l + 1$ and return to step1.

Step 6. Computation of final results. Compute the final generalized costs $C^{m(l+1)}_{rs,k}$ and $C^{v(l+1)}_{sr,k}$, and demand matrices $q^{m(l+1)}_{rs}$ and $q^{v(l+1)}_{sr}$ according to the final path flows $f^{m(l+1)}_{rs,k}$ and $f^{v(l+1)}_{sr,k}$, $m \in M$, $r \in R$, $s \in S$, $k \in K_{rs}$.

It is worth noting that this bi-level problem is non-convex; thus, heuristic algorithms have to be considered to obtain a relatively good solution. The sensitivity analysis-based method would be a highly efficient solution algorithm to solve the bi-level programming problems when the first-order derivatives are easy to be obtained from the lower-level problem (Yang and Bell [25]). In this paper, however, to solve the sensitivities from the VI program Equation (39) is very difficult in view of the complexity of the formulation. Compared with some other algorithms such as projection-based algorithm, descent algorithm, or penalty function approach, GA would still be a straightforward and effective method to solve the bi-level problem, which has been proven in many existing literatures (Liu *et al*. [26]). Thus, in this paper, the GA is adopted for its simplicity and effectiveness.

DOI: 10.1002/atr

6. NUMERICAL EXAMPLE

In this section, we provide a numerical example solved by the proposed algorithm. Consider a road network depicted in Figure 1 with four OD pairs, six nodes, and 14 links. The bus lines are 1-2-4-6, 1-3-4, 5-6, and 5-3-4, respectively, which are represented as dotted lines in Figure 1. We assume that the travel time function for each link follows the traditional BPR (Bureau of Public Roads) function

$$t_a(x_a) = t_a^0\left[1 + 0.15\left(\frac{x_a^p + x_a^o + x_a^v}{C_a}\right)^4\right], \quad a \in A \tag{61}$$

where the free flow travel time t_a^0 and link capacity C_a as well as link length d_a, $a \in A$, are given in Table 1.

The negative exponential demand function is

$$q_{rs} = \overline{q}_{rs}\exp(-\kappa_{rs}u_{rs}), \quad r \in R, s \in S \tag{62}$$

where $\overline{q}_{rs}$ is the potential demand for each OD pair and κ_{rs} is a elasticity parameter that represents the sensitivity of demand to generalized costs. Let $\overline{q}_{16} = 4000$ veh/h, $\overline{q}_{14} = 3000$ veh/h, $\overline{q}_{56} = 4000$ veh/h, and

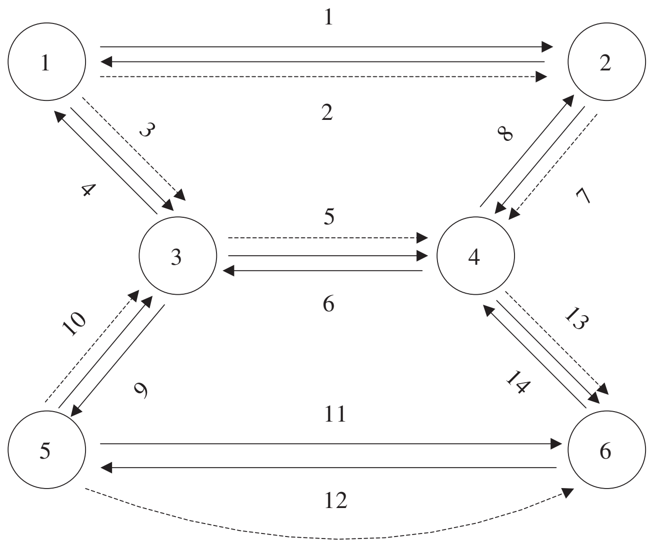

Figure 1. The network.

Table I. Link free flow, capacity and length.

Link	Star node	End node	t_a^0 (h)	C_a (veh/h)	d_a (km)
1	1	2	0.03	1200	2
2	2	1	0.03	1200	2
3	1	3	0.03	1200	2
4	3	1	0.03	1200	2
5	3	4	0.04	1200	3
6	4	3	0.04	1000	3
7	2	4	0.02	1800	1
8	4	2	0.02	1800	1
9	3	5	0.03	1200	2
10	5	3	0.03	1200	2
11	5	6	0.04	1200	3
12	6	5	0.04	1200	3
13	4	6	0.04	1200	3
14	6	4	0.04	1200	3

DOI: 10.1002/atr

$=3000$ veh/h, respectively. Also, let $\kappa_{16}=\kappa_{14}=\kappa_{56}=\kappa_{54}=0.03$. Similar to Huang [27], the following passenger crowding discomfort function is specified

$$G_{rs}\left(q_{rs}^{b}\right)=\theta_1\left(q_{rs}^{b}\right)^2+\theta_2 q_{rs}^{b},\quad r\in R, s\in S \tag{63}$$

where θ_1 and θ_2 are positive parameters. In this numerical example, let θ_1 and θ_2 be 0.001 and 0.0001, respectively. Furthermore, the bus frequency and bus travel time are shown in Table 2. Other parameters are set next:$b^p=3$\$/km, $b_0^o=10$\$, $b_1^o=2$\$/km, $b_2^o=30$\$/h, $b^v=1.5$\$/km, $\lambda=60$\$/h, $\lambda_v=40$\$/h, $\lambda_{ow}=120$\$/h, $\lambda_b=30$\$/h, $\lambda_{bw}=60$\$/h, $N=2000$, $\varphi_{16}^p=2$, $\varphi_{16}^b=1$, $\varphi_{16}^o=5$, $\varphi_{14}^p=2$, $\varphi_{14}^b=1$, $\varphi_{14}^o=5$, $\varphi_{56}^p=2$, $\varphi_{56}^b=1$, $\varphi_{56}^o=5$, $\varphi_{54}^p=2$, $\varphi_{54}^b=1$, $\varphi_{54}^o=5$, $\beta=0.06$, $\sigma=0.2$, $\eta Z_r=10$, $r\in R$, $\zeta=0.01$, $\tau=2$.

Figure 2 displays the maximum social welfare of charging and not charging taxi versus the number of iterations, where the execution time is 25 min. It can be observed in Figure 2 that there is an increase (8%) in the maximum social welfare as congestion toll is imposed on taxi, where the maximum social welfare is 1.60×10^5\$ for charging taxi and 1.48×10^5\$ for not charging it. This can be explained by the fact that without the toll, every taxi behaves selfishly when making the route choice, which causes efficiency loss.

Tables 3–5 as well as Figure 3 compares the performances between charging and not charging taxi for the congestion toll. We observe that the private car demand and bus demand increase by 445 and 176, respectively, after taxi is tolled, while the taxi demand decreases by 1467. This is because part of the demand shifts from taxi to private car and bus due to the congestion fee. The total demand decreases by 846 (7.9%), which is considerably affected by the charge on taxi.

Figure 4 portrays the change of the ratio of maximum social welfares (denoted as ρ) with the ratio of operating cost per unit distance for private car to that for taxi (b^p/b^v, denoted as ω) when preliminary flag-fall charge b_0^o is 10 and taxi fleet size N is 2000. As we can see from the figure, that ρ decreases with ω. The reason is that when toll is imposed on taxi, part of travelers will divert from taxi to private car, leading to a higher social cost because of the higher operating cost of private car. Thus, the greater the ω, the higher the social cost for charging taxi, and hence, the less the ratio of maximum social

Table II. Bus travel time and frequency.

Origin	Destination	Travel time T_{rs} (h)	Bus frequency F_{rs} (veh/h)
1	6	0.3	10
1	4	0.2	10
5	6	0.15	10
5	4	0.2	10

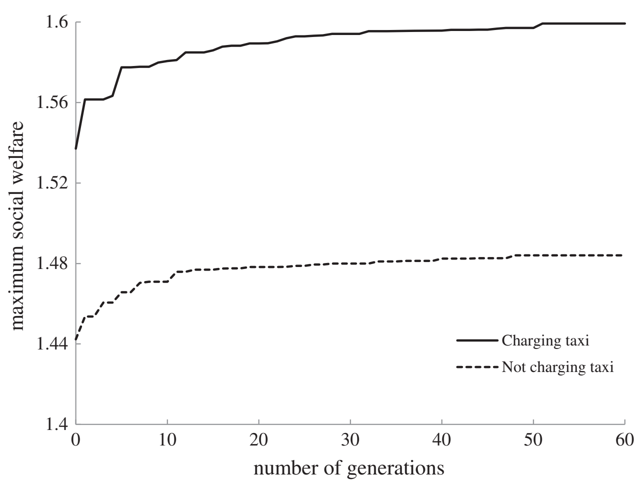

Figure 2. Maximum social welfare of charging and not charging taxi versus the number of iterations.

Table III. The OD matrix of private car when charging taxi (not charging taxi).

	6	4	$\sum_{s\in S} q_{rs}^p$
1	448 (318)	731 (590)	1179 (908)
5	1303 (1180)	570 (519)	1873 (1699)
$\sum_{r\in R} q_{rs}^p$	1751 (1498)	1301 (1109)	3052 (2607)

Table IV. The OD matrix of taxi when charging taxi (not charging taxi).

	6	4	$\sum_{s\in S} q_{rs}^o$
1	375 (833)	514 (864)	889 (1697)
5	828 (1078)	459 (868)	1287 (1946)
$\sum_{r\in R} q_{rs}^o$	1203 (1911)	973 (1732)	2176 (3643)

Table V. The OD matrix of bus when charging taxi (not charging taxi).

	6	4	$\sum_{s\in S} q_{rs}^b$
1	1288 (1233)	1070 (1030)	2358 (2263)
5	1201 (1185)	1107 (1042)	2308 (2227)
$\sum_{r\in R} q_{rs}^b$	2489 (2418)	2177 (2072)	4666 (4490)

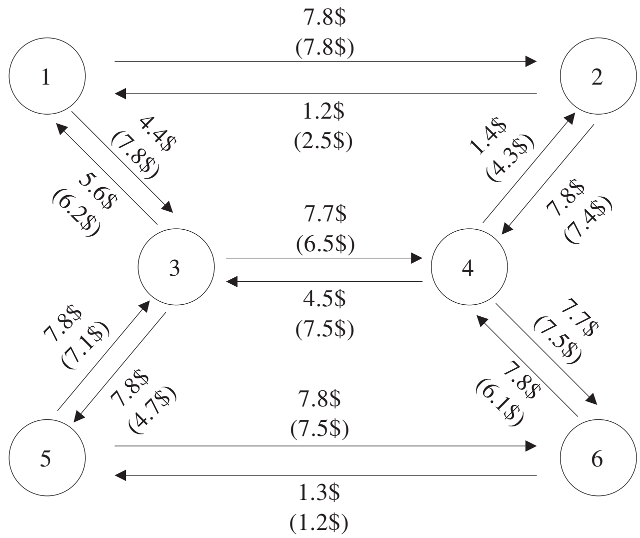

Figure 3. The optimal toll when charging taxi (not charging taxi).

welfares. It is also worth noting that as ω increases further beyond 9 (although it unlikely happens in real life), ρ will be less than 1, implying that the maximum social welfare for charging taxis will be less than that for not charging them.

Figure 5 depicts the change of ρ against preliminary flag-fall charge per ride, as ω equals 2 and N equals 2000. It can be seen in this figure that initially, ρ decreases dramatically with the preliminary flag-fall charge. This is because the demand for taxi decreases with the flag-fall charge, resulting in the decrease of the total external cost caused by taxis, so as ρ. Moreover, ρ tends to be steady and approach to 1 as the initial flag-fall charge exceeds 25$, because now the cost of taking taxi is too high and the taxi demand is too low to affect the system performance.

Figure 6 shows that ρ does not change much with respect to the taxi fleet size. This is because of the two opposite effects of raising the taxi fleet size on ρ. On the one hand, as taxi fleet size grows,

DOI: 10.1002/atr

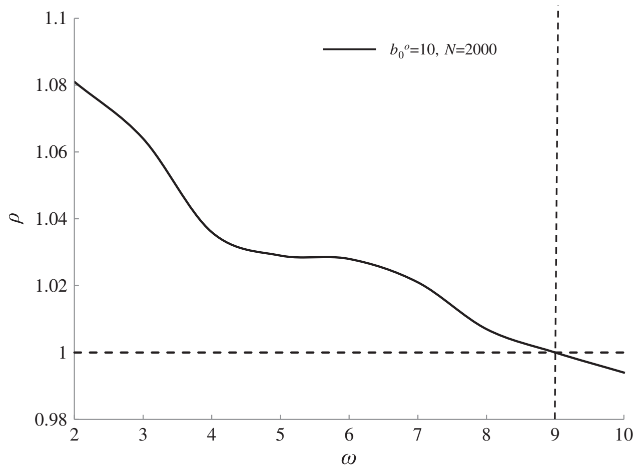

Figure 4. Ratio of maximum social welfare for charging taxi to that for not charging taxi versus ratio of operating cost per unit distance for private car to that for taxi.

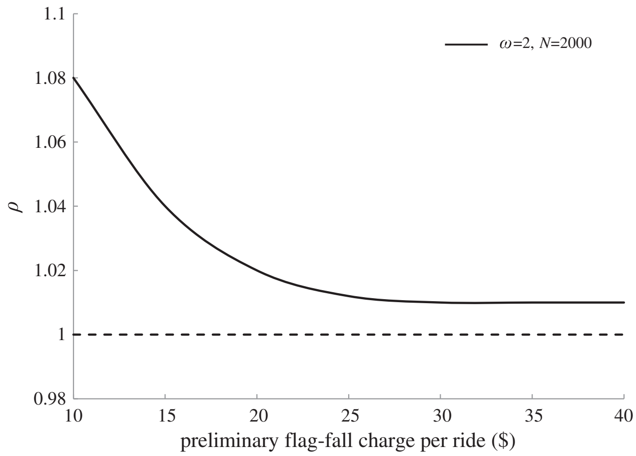

Figure 5. Ratio of maximum social welfare for charging taxi to that for not charging taxi against preliminary flag-fall charge per ride.

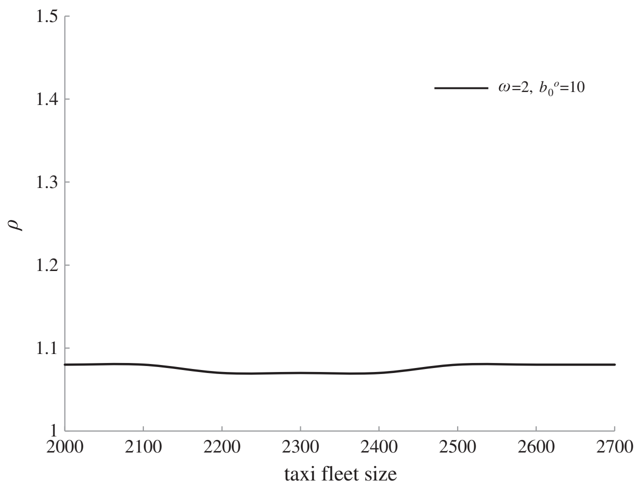

Figure 6. Ratio of maximum social welfare for charging taxi to that for not charging taxi against taxi fleet size.

DOI: 10.1002/atr

taxi service quality is improved and thus more travelers will be attracted to take taxi. Therefore, when charging taxi, the number of travelers shifting from taxi to private car will increase, resulting in a larger ρ. On the other hand, as taxi fleet size grows, the taxi waiting/searching time increases after taxi is tolled, which leads to the social welfare decrease. As a result, the value of ρ tends to be steady when taxi fleet size varies.

From all the previous discussions, we have the conclusion that, in this case, charging taxi is always a better choice as long as the operating cost of private car is not far beyond that of the taxi (not greater than nine times).

7. CONCLUSIONS

A mathematical model was presented to address the issue that whether or not taxis should pay the congestion charge in pricing zone. The model is developed as a bi-level programming problem where the lower level is formulated as a VI program, with the logit-based mode split, route choice, elastic demand, and vacant taxi distributions. The upper level aims to maximize the social welfare. The bi-level model is solved by GA, whereas the lower-level sub-problem is solved by the block Gauss–Seidel decomposition approach together with the method of successive averages and diagonalization algorithm.

The results of the numerical example indicated that whether taxis should be charged in pricing zone is mainly dependent on the ratio of operating costs for private car to that for taxi. Generally (when the operating cost of the private car is not extremely higher than that of the taxi), charging taxis for the congestion fee is always a better choice in terms of the maximum social welfare. This can be explained by the fact that as taxis are charged as well as the private cars, the external cost of the system is fully internalized, which outweighs the benefit from encouraging the usage of taxi mode by not charging them.

This study offers a useful methodology when designing congestion pricing schemes for multiple traffic modes for decision-makers. Real transportation networks and traffic data will be tested in the future instead of the synthetic one used in the current study.

8. NOTATION

The list of symbols used in the paper:

N	set of nodes
A	set of links
$\overline{A}$	set of toll links, $\overline{A} \subseteq A$
R	set of origin zones
S	set of destination zones
q_{rs}	total demand between zone $r \in R$ and zone $s \in S$
q_{rs}^{p}	demand of private car from origin $r \in R$ to destination $s \in S$
q_{rs}^{b}	demand of bus from origin $r \in R$ to destination $s \in S$
q_{rs}^{o}	demand of taxi from origin $r \in R$ to destination $s \in S$
O_r^o	demand for taxi mode from origin zone $r \in R$
D_s^o	demand for taxi mode to destination zone $s \in S$
q_{sr}^{v}	number of vacant taxis traveling from zone $s \in S$ to zone $r \in R$
x_a	flow on link $a \in A$
δ_{ak}^{rs}	link-route indicator variable, which is equal to 1 if route k between OD pair $r \in R$ and $s \in S$ uses link a, and 0 otherwise
K_{rs}	set of paths between zone $r \in R$ and zone $s \in S$
c_a^p	generalized cost on link $a \in A$ for private car
c_a^o	generalized cost on link $a \in A$ for occupied taxi
c_a^v	generalized cost on link $a \in A$ for vacant taxi
$f_{rs,k}^{m}$	flow on route $k \in K_{rs}$ for mode $m \in M$ between zone $r \in R$ and zone $s \in S$
$f_{sr,k}^{v}$	vacant taxi flow on route $k \in K_{rs}$ between zone $s \in S$ and zone $r \in R$

DOI: 10.1002/atr

$C^m_{rs,k}$	generalized cost for mode $m \in M$ on route $k \in K_{rs}$ from zone $r \in R$ to zone $s \in S$
$C^v_{sr,k}$	generalized cost for vacant taxi on route $k \in K_{rs}$ from zone $r \in R$ to zone $s \in S$
$t_a(x_a)$	travel time on link $a \in A$
t^0_a	free flow travel time on link $a \in A$
λ	VOT for users taking private car or taxi
λ_v	VOT for taxi drivers
λ_b	VOT for bus passengers
λ_{ow}	value of customer waiting time
λ_{bw}	value of bus passenger waiting time
b^p	operating cost per unit distance for private car
b^o_0	preliminary flag-fall charge per ride
b^o_1	mileage-based taxi charge
b^o_2	delay-based taxi charge
W^o_r	customer waiting time for taxi at zone r
b^v	operating cost per unit distance for taxi
d_a	length of link $a \in A$
w^v_r	waiting/searching time of vacant taxi in zone r
Z_r	the area of zone $r \in R$
η	a model parameter that is identical to all zones
T_{rs}	bus travel time from origin $r \in R$ to destination $s \in S$
W^b_{rs}	waiting time of bus passengers between zone $r \in R$ and zone $s \in S$
$G_{rs}\left(q^b_{rs}\right)$	crowding discomfort experienced by passengers
ζ	the unit cost of discomfort
F_{rs}	bus frequency between zone $r \in R$ and zone $s \in S$
τ	bus fare
N	taxi fleet size
P^m_{rs}	proportion of trips taken by mode $m \in M$ from zone $r \in R$ to zone $s \in S$
φ^m_{rs}	attraction of mode $m \in M$ between zone $r \in R$ and zone $s \in S$
β	dispersion coefficient
$P_{r/s}$	probability that vacant taxi departs from zone $s \in S$ and meets the customer in zone $r \in R$.
σ	non-negative parameter reflecting the degree of uncertainty for taxi drivers on customer demand and taxi services of the whole market
h_{rs}	average travel time from origin $r \in R$ to destination $s \in S$
a_s	Lagrange multiplier related to taxi demand constraint at destination zone $s \in S$
e_r	Lagrange multiplier related to taxi demand constraint at origin zone $r \in R$
γ_{rs}	Lagrange multiplier related to conservation equation of total demand from zone $r \in R$ to zone $s \in S$
$\chi^m_{rs,k}$	Lagrange multiplier related to the flow non-negativity constraint from zone $r \in R$ to zone $s \in S$ for mode $m \in M$ on route $k \in K_{rs}$
$\chi^v_{sr,k}$	Lagrange multiplier related to the flow non-negativity constraint from zone $s \in S$ to zone $r \in R$ for vacant taxi on route $k \in K_{sr}$
u_{rs}	users' minimum perceived generalized costs between zone $r \in R$ and zone $s \in S$
u^m_{rs}	minimal generalized costs for mode $m \in M$ between zone $r \in R$ and zone $s \in S$
u^v_{sr}	minimal generalized costs for vacant taxi between zone $r \in R$ and zone $s \in S$
y_a	toll charge on link $a \in \overline{A}$
$y^{\max}_a$	upper bound of toll rate for link $a \in \overline{A}$
$y^{\min}_a$	lower bound of toll rate for link $a \in \overline{A}$
ε	acceptable error in the block Gauss–Seidel decomposition method
C_a	capacity of link $a \in A$
$\overline{q}_{rs}$	potential total demand between zone $r \in R$ and zone $s \in S$
κ_{rs}	elasticity parameter between zone $r \in R$ and zone $s \in S$, representing the sensitivity of demand to the minimum generalized costs
$D_{rs}(u_{rs})$	demand function between zone $r \in R$ and zone $s \in S$
$D^{-1}_{rs}(q_{rs})$	inverse of demand function

DOI: 10.1002/atr

ACKNOWLEDGEMENTS

The research is supported by the National Natural Science Foundation of China (71201135) and the Specialized Research Fund for the Doctoral Program of Higher Education (20120184120017). The authors would like to express their thanks to the three anonymous reviewers for their constructive suggestions and helpful comments.

REFERENCES

1. Pigou AC. *Wealth and Welfare*. Macmillan: London, 1920.
2. Eliasson J, Mattsson LG. Equity effects of congestion pricing: quantitative methodology and a case study for Stockholm. *Transportation Research Part A* 2006; **40**(7):602–620.
3. Maruyama T, Sumalee A. Efficiency and equity comparison of cordon- and area-based road pricing schemes using a trip-chain equilibrium model. *Transportation Research Part A* 2007; **41**(7):655–671.
4. Guo X, Yang H. Pareto-improving congestion pricing and revenue refunding with multiple user classes. *Transportation Research Part B* 2010; **44**(8–9):972–982.
5. Mirabel F, Reymond M. Bottle congestion pricing and modal split: redistribution of toll revenue. *Transportation Research Part A* 2011; **45**(1):18–30.
6. Ekström J, Sumalee A, Lo HK. Optimizing toll locations and levels using a mixed integer linear approximation approach. *Transportation Research Part B* 2012; **46**(7):834–854.
7. Zheng N, Waraich RA, Axhausen KW, *et al*.. A dynamic cordon pricing scheme combining the macroscopic fundamental diagram and an agent-based traffic model. *Transportation Research Part A* 2012; **46**(8):1291–1303.
8. Meng Q, Liu Z, Wang S. Optimal distance tolls under congestion pricing and continuously distributed value of time. *Transportation Research Part E* 2012; **48**(5):937–957.
9. Transport Department. *The Level of Taxi Services*. TTSD Publication series. Hong Kong SAR Government: Hong Kong, 1986-2000.
10. Yang H, Ye M, Tang WH, Wong SC. Regulating taxi services in the presence of congestion externality. *Transportation Research Part A* 2005; **39**(1):17–40.
11. Transport Advisory Committee. *Report on review of taxi operation*. Hong Kong Government, Hong Kong SAR, 2008.
12. Yang H, Yang T. Equilibrium properties of taxi markets with search frictions. *Transportation Research Part B* 2011; **45**(4):696–713.
13. Salanova JM, Estrada M, Aifadopoulou G, Mitsakis E. A review of the modeling of taxi services. *Procedia social and behavioral sciences* 2011; **20**:150–161.
14. Yang H, Wong SC. A network model of urban taxi services. *Transportation Research Part B* 1998; **32**(4):235–246.
15. Wong KI, Wong SC, Yang H. Modeling urban taxi services in congested road networks with elastic demand. *Transportation Research Part B* 2001; **35**(9):819–842.
16. Wong KI, Wong SC, Yang H, Wu JH. Modeling urban taxi services with multiple user classes and vehicle modes. *Transportation Research Part B* 2008; **42**(10):985–1007.
17. Yang H, Wong SC, Wong KI. Demand-supply equilibrium of taxi services in a network under competition and regulation. *Transportation Research Part B* 2002; **36**(9):788–819.
18. King DA, Peters JR. Slow down, you move too fast: the use of tolls by taxicabs in New York City. *Proce., 91st Annual Meeting of the Transportation Research Board* (Compendium of Papers, CD-ROM) 2012.
19. Florian M. A traffic equilibrium model of travel by car and public transit modes. *Transportation Science* 1977; **11**(2):166–179.
20. Ferrari P. A model of urban transport management. *Transportation Research Part B* 1999; **33**(1):43–61.
21. Ahn K. Road pricing and bus service policies. *Journal of Transport Economics and Policy* 2009; **43**:25–53.
22. Li Z, Lam WHK, Wong SC. Modeling intermodal equilibrium for bimodal transportation system design problems in a linear monocentric city. *Transportation Research Part B* 2012; **46**(1):30–49.
23. Shi R, Li Z. Pricing of multimodal transportation networks under different market regimes. *Journal of transportation systems engineering and information technology* 2010; **10**(5):91–97.
24. Florian M, Wu JH, He S. A multi-class multi-mode variable demand network equilibrium model with hierarchical logit structures. In *Transportation and Network Analysis: Current Trends: Miscellanea in Honor of Michael Florian*, Marcotte P, Gendreau M(eds). Kluwer Academic: London, 2002; 119–133.
25. Yang H, Bell MGH. Models and algorithms for road network design: a review and some new developments. *Transport Review* 1998; **18**(3):257–278.
26. Liu H, Hu X, Yang S, *et al*. Application of complex network theory and genetic algorithm in airline route networks. *Transportation Research Record: Journal of the Transportation Research Board* 2011; **2214**:50–58.
27. Huang H. Pricing and logit-based mode choice models of a transit and highway system with elastic demand. *European Journal of Operational Research* 2002; **140**(3):562–570.

DOI: 10.1002/atr