Transportation Research Part B 44 (2010) 1067–1083

Contents lists available at ScienceDirect

Transportation Research Part B

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

Equilibria of bilateral taxi–customer searching and meeting on networks

Hai Yang [a,*], Cowina W.Y. Leung [a], S.C. Wong [b], Michael G.H. Bell [c]

[a] Department of Civil and Environmental Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, PR China
[b] Department of Civil Engineering, The University of Hong Kong, Pokfulam Road, Hong Kong, PR China
[c] Department of Civil and Environmental Engineering, Imperial College, Imperial College Road, London, SW7 2BU, United Kingdom

ARTICLE INFO

Article history:
Received 9 July 2008
Received in revised form 30 November 2009
Accepted 2 December 2009

Keywords:
Competitive equilibria
Traffic network
Meeting function
Market friction
Taxi industry

ABSTRACT

This paper proposes an equilibrium model to characterize the bilateral searching and meeting between customers and taxis on road networks. A taxi driver searches or waits for a customer by considering both the expected searching or waiting time cost and ride revenue, and a customer seeks a taxi ride to minimize full trip price. We suppose that the bilateral taxi–customer searching and meeting occurs anywhere in residential and commercial zones or at prescribed taxi stands, such as an airport or a railway station. We propose a meeting function to spell out the search and meeting frictions that arise endogenously as a result of the distinct spatial feature of the area and the taxi–customer moving decisions. With the proposed meeting function and the assumptions underlying taxi–customer search behaviors, the stationary competitive equilibrium achieved at fixed fare prices is determined when the demand of the customers matches the supply of taxis or there is market clearing at the prevailing searching and waiting times in every meeting location. We establish the existence of such an equilibrium by virtue of Brouwer's fixed-point theorem and demonstrate its principal operational characteristics with a numerical example.

1. Introduction

With their 24-h-a-day availability and capacity to provide door-to-door service, taxis are an important complement to regularly scheduled services provided by other forms of public transportation. Taxis are regulated in two main ways in most cities: entry restrictions and price control. In an attempt to understand the manner in which the demand and supply are equilibrated in the presence of regulations, economists have examined the problem using highly aggregate models, without considering the spatial structures of the market. It is commonly realized that the expected customer-waiting time is an important service quality measure that affects whether or not a potential customer takes a taxi and thus plays a crucial role in the determination of the price level and the resulting market equilibrium as well as government decision-making with regard to taxi regulation or deregulation (Douglas, 1972; De vany, 1975; Hackner and Nyberg, 1995; Arnott, 1996; Cairns and Liston-Heyes, 1996; and, recently, Flores-Guri, 2003; Fernandez et al., 2006; Moore and Balaker, 2006). In view of the fact that demand and supply of taxi services varies substantially over time and space, Yang and his colleagues have embarked on modeling taxi services in a network context (Yang and Wong, 1998). A few extensions have been added to deal with demand elasticity, multi-class taxi services with service area regulation and congestion effects, and multi-period dynamic taxi services with endogenous service intensity (Wong et al., 2001, 2008; Yang et al., 2002, 2005a,b).

These aggregate and network models have focused almost exclusively on the effect of taxi availability on customer-waiting time and the resulting market equilibrium. Only very limited attention has been paid to the analysis of the

* Corresponding author. Tel.: +852 2358 7178; fax: +852 2358 1534.
E-mail address: cehyang@ust.hk (H. Yang).

0191-2615/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.trb.2009.12.010

prevailing bilateral searching and meeting between taxis and customers in real-world taxi markets. In fact, bilateral searching and matching between agents is a popular phenomenon in economics such as trading and labor markets, for which the matching function is widely used to explicitly account for the nature of the matching frictions and their implications for market outcomes (Mortensen and Pissarides, 1994; Petrongolo and Pissarides, 2001; Burdett et al., 2001). The matching function summarizes a trading technology between agents who possess limited information, so time and resources have to be spent seeking trading partners. In labor market applications, the matching function approach proceeds by assuming the existence of a well-behaved function that gives the number of jobs formed at any moment in time in terms of the number of workers looking for jobs, the number of firms looking for workers, and a small number of other variables. A matching function was first introduced by Schroeter (1983) in the analysis of taxi services in a regulated market where radio dispatch and airport taxi stands are the primary modes of operations. Lagos (2000) provided a theoretical microeconomic model of the meeting frictions and processes that can give rise to an aggregate meeting function. He further used the dynamic equilibrium model of searching and meeting frictions to quantify the impact of a recent change in taxicab regulations on the process that rules the meetings between customers and taxis in New York City (Lagos, 2003). Without explicit use of a meeting function, Wong et al. (2005) examined bilateral searching and meeting by taking into account the stochastic micro-searching behaviors of both taxis and customers in a network. The bilateral search was described using the absorbing Markov chain approach. Matsushima and Kobayashi (2006) modeled endogenous market formation with meeting externality in a single spot market (a single taxi stand), where waiting and meeting between taxis and customers was modeled with a double-queuing model.

In this paper, we propose an equilibrium model to characterize the bilateral searching and meeting between taxis and customers in road networks. In contrast to the absorbing Markov chain approach, which focuses on the micro-searching process of taxi drivers for customers (Wong et al., 2005), a searching-and-meeting model is developed and integrated with earlier network models, thus making the combined model more operational in practical applications. A taxi driver searches for a customer by considering both the expected searching or waiting time cost and ride revenue, and a customer seeks a taxi ride to minimize full trip price. Movements of vacant taxis in search of customers are formulated by the early simple network model proposed by Yang and Wong (1998) and Wong and Yang (1998). Suppose that the bilateral taxi–customer meetings occur anywhere in residential and commercial zones or at prescribed taxi stands such as an airport or a railway station or a hotel. We propose a meeting function to spell out the meeting frictions that arise endogenously as a result of the distinct spatial feature of the meetings and the taxi–customer moving decisions. With the assumption of the underlying taxi–customer search behaviors, the proposed meeting function and the early network model of taxi services, we determine that the stationary competitive equilibrium is achieved at fixed fare prices when the demand of customers matches the supply of taxis at prevailing searching and waiting times in every location. By regarding each meeting location as a market, our conceptualization is similar to that of the economy employed in competitive general equilibrium theory (Ellickson, 1993; Starr, 1997), except that waiting times rather than prices clear market. We establish the existence of such a bilateral searching and meeting equilibrium by virtue of Brouwer's fixed-point theorem and provide an iterative numerical algorithm to find an equilibrium solution.

In what follows, we set out the basic model of central interest including the meeting function in the next section. In Section 3, we briefly define the stationary bilateral searching and meeting equilibria, prove its existence and provide an iterative algorithm to find an equilibrium solution. Section 4 provides a simple example for a brief numerical assessment of the operational characteristics of the proposed model. General conclusions and a discussion of further researches are given in Section 5.

The following major symbols are used in the paper (additional symbols are introduced as necessary):

I	set of customer origin zones
J	set of customer destination zones
K	set of taxi-customer meeting locations
N	number of taxis in service
N_k^c	number of unserved customers at location $k \in K$
N_k^{vt}	number of unoccupied taxis at location $k \in K$
Q_{ij}^c	total customer demand (person/h) for trips (given exogenously) by taxis and other public transportation modes from origin zone $i \in I$ to destination zone $j \in J$
Q_{ikj}^c	customer demand (person/h) from origin zone $i \in I$ to destination zone $j \in J$ via a meeting location $k \in K$; $\mathbf{Q}^c = (Q_{ikj}^c, i \in I, j \in J, k \in K)$
Q_k^c	number of customer trips per unit time (person/h) with meeting location $k \in K$
Q_{kj}^c	number of customers per unit time (person/h) moving from meeting location $k \in K$ to destination $j \in J$
T_k^{vt}	arrival rate (veh/h) of vacant taxis at location $k \in K$
T_{kj}^{ot}	occupied taxi movements (veh/h) from meeting location $k \in K$ to destination node $j \in J$; $\mathbf{T}^{ot} = (T_{kj}^{ot}, k \in K, j \in J)$
T_{jk}^{vt}	vacant taxi movements (veh/h) from destination node $j \in J$ to meeting location $k \in K$; $\mathbf{T}^{vt} = (T_{jk}^{vt}, j \in J, k \in K)$
h_{ik}	customer walking time (h) from origin zone $i \in I$ to meeting location $k \in K$
h_{kj}	occupied taxi travel time (h) from meeting location $k \in K$ to customer destination $j \in J$
h_{jk}	vacant taxi travel time (h) from customer destination $j \in J$ to meeting location $k \in K$

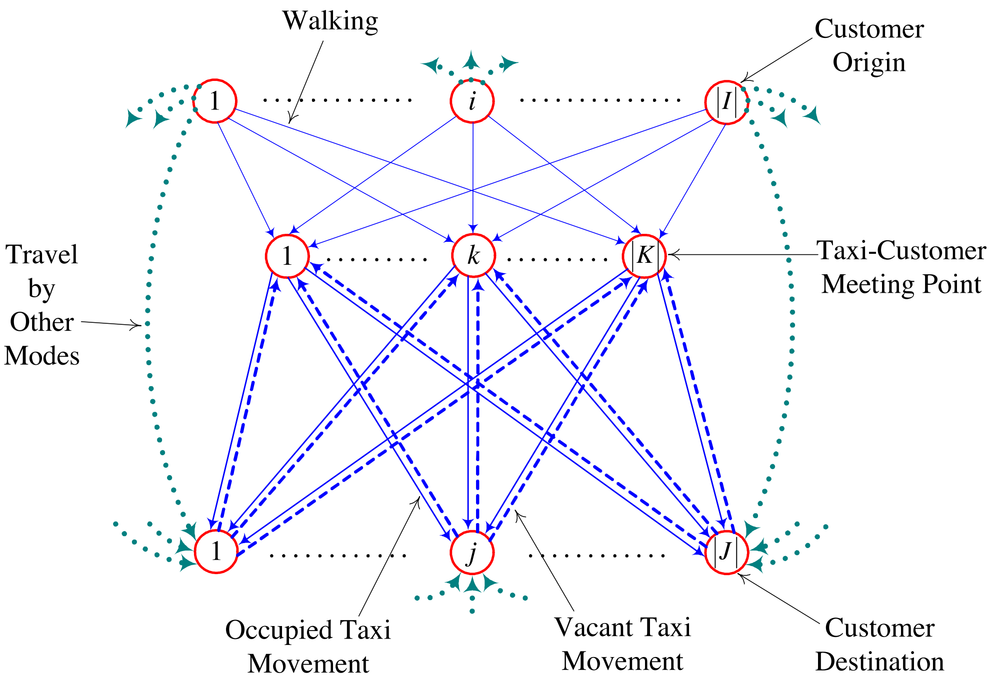

Fig. 1. Characterization of movements of customers and vacant and occupied taxis on networks.

w_k^c customer-waiting time (h) at meeting location $k \in K$
w_k^t taxi waiting time (h) at meeting location $k \in K$
F_0 taxi flag-drop charge (HK$)
τ taxi charge per unit ride time (HK$/h)
F_{kj} fare (HK$) of a taxi ride from meeting location $k \in K$ to destination node $j \in J$
$P_{k/ij}^c$ probability that a customer traveling from origin zone $i \in I$ to destination zone $j \in J$ patronizes meeting location $k \in K$
$P_{k/j}^t$ probability that a vacant taxis originating at node $j \in J$ meets a customer at location $k \in K$
m_k^{c-t} number of meetings per unit time or meeting rate (customer–taxi/h) between customers and taxis at meeting location $k \in K$

2. Characterization of taxi–customer searching and meeting

2.1. Basic definitions

Consider movements of customers and taxis in an urban area. Let I and J be the set of customer origins and destinations, respectively, and K be the set of the spatially distinct meeting locations where customers and taxis meet each other. Here, a meeting location can be anywhere in a residential or commercial zone or a prescribed taxi stand, such as at an airport or a railway station, and the corresponding taxi–customer meetings are referred to as area and point meetings respectively. To model the bilateral searching and meeting equilibrium, contacts between customers and taxis are assumed to occur only in the locations of the given set, K. A customer with an initial origin, $i \in I$, has to walk to a meeting location, $k \in K$, to find a taxi and then travel to his/her destination, $j \in J$, by a taxi. A taxi, once taking a customer at a meeting location, $k \in K$, will move to the customer's destination, $j \in J$, via the shortest path. Once a customer ride is completed, the taxi becomes vacant and is free to choose a meeting location to find the next customer. For expositive purpose, a schematic representation of the bilateral meeting and movements of customers and taxis is given in Fig. 1.[1] The occupied taxi movements take place along the solid lines connecting nodes $k \in K$ and $j \in J$; and the vacant taxi movements take place along the broken lines connecting nodes $j \in J$ and $k \in K$.

Note that the proposed schematic representation does not exclude the possibility that a customer's initial origin is identical with or sufficiently close to a taxi–customer meeting location and/or coincides with another customer's destination (a node is both an origin and a destination location for customers and also a meeting location). In this case dummy nodes can be

[1] It would be interesting to look at the taxi market model considered here from the perspective of competitive general equilibrium theory (Ellickson, 1993; Starr, 1997). On one hand, fare prices are fixed and searching and waiting times of taxis and customers adjust to clear the markets (meeting locations). On the other hand, excess customer demands exist due to searching and meeting frictions and are served via other mass transit modes.

added to separate such physically shared origin and destination and/or meeting locations for the proposed schematic network representation. For example, if origin node i^*, taxi stand k^* and destination node j^* are at the same location, then the corresponding movement time will be zero ($h_{i^*k^*} = 0, h_{k^*j^*} = 0, h_{j^*k^*} = 0$). In the absence of congestion (congestion effects are not taken into account for simplicity), customer walking time $h_{ik}, i \in I, k \in K$, occupied taxi travel time $h_{kj}, k \in K, j \in J$ and vacant taxi travel time $h_{jk}, j \in J, k \in K$ are all constants. Here, h_{ik}, h_{kj} and h_{jk} are the times via the shortest paths and are measured as a fraction of an hour. The proposed schematic representation will greatly simplify the subsequent exposition.

2.2. Customer movements

Consider taxi movements and customer demands in a stationary state in any given hour in the network. Let Q_{ikj}^c be the customer demand for taxi from $i \in I$ to $j \in J$ via a meeting location, $k \in K$, where 'c' stands for customer. Let $Q_k^c = \sum_{i \in I} \sum_{j \in J} Q_{ikj}^c$ be the number of customer trips with meeting location $k \in K$ and $Q_{kj}^c = \sum_{i \in I} Q_{ikj}^c$ be the number of customers moving from meeting location $k \in K$ to destination $j \in J$.

Given the customer-waiting time, w_k^c, at meeting location $k \in K$, a customer chooses a meeting location, $k \in K$, to minimize his/her (subjective) full trip price. The full price for traveling by taxi through taxi stand $k \in K$ is given as

$$g_{ikj} = F_{kj} + \beta_1 h_{ik} + \beta_2 h_{kj} + \beta_3 w_k^c, \quad \text{for any } i \in I, \ j \in J, \ k \in K, \tag{1}$$

where $F_{kj} = F_0 + \tau h_{kj}$ is the fare of a taxi ride from $k \in K$ to $j \in J$ and is assumed to be a given constant as congestion is not considered, F_0 is the taxi flag-drop charge and τ is the variable charge per unit ride time; β_1, β_2 and β_3 are the customers' monetary values of unit walking time, in-vehicle time and waiting time, respectively. In order to consider elasticity of customer demand for taxi services, we suppose that a *virtual* mode (e.g., other public transportation services such as trains and buses) is available from each $i \in I$ to each $j \in J$ with a constant generalized price $\bar{g}_{ij}$. Now, we make the following assumption regarding the customers' choice of modes and taxi-meeting locations.

A1: *The probability of a customer from origin zone $i \in I$ to destination zone $j \in J$ patronizing meeting location $k \in K$ is given by the logit model*:

$$P_{k/ij}^c = \frac{\exp\{-\theta^c g_{ikj}\}}{\exp\{-\theta^c \bar{g}_{ij}\} + \sum_{m \in K} \exp\{-\theta^c g_{imj}\}}, \quad \text{for any } k \in K, \ i \in I, \ j \in J, \tag{2}$$

where θ^c is a nonnegative parameter.

Note that the value of parameter θ^c reflects the degree of uncertainty in customer demand and taxi services in the whole market from the perspective of individual customers. With the virtual mode simply treated as an alternative parallel 'nontaxi service' with constant generalized price, one may use a more realistic two-level logit model, where a primary model split is carried out between the virtual and the taxi modes, based on the constant virtual mode travel cost and the minimum expected disutility of using the taxi mode; the second level is then simply a logit-based choice model of taxi-meeting locations. Such an addition does not alter the ongoing examination of the bilateral taxi–customer searching and meeting models, which is our central concern.

Suppose that the total number, Q_{ij}^c (person/h), of customers for trips in a given hour by taxis and the virtual mode (other public transportation modes) from origin zone i to destination zone j is given. With the logit-based customer choice model (2), the number of customers traveling from $i \in I$ to $j \in J$ by taxis through meeting location $k \in K$ is given by

$$Q_{ikj}^c = Q_{ij}^c P_{k/ij}^c, \quad i \in I, \ j \in J, \ k \in K, \tag{3}$$

and the total number of customers arriving at location $k \in K$ is given by

$$Q_k^c = \sum_{i \in I} \sum_{j \in J} Q_{ikj}^c = \sum_{i \in I} \sum_{j \in J} Q_{ij}^c P_{k/ij}^c, \ \ k \in K. \tag{4}$$

2.3. Vacant and occupied taxi movements

Suppose that there are N cruising taxis operating in the network and consider one *unit* period (1 h) operations of taxis in the network with a given customer demand pattern in a stationary state.[2] The total occupied time of all taxis is the taxi-hours given by $\sum_{k \in K} \sum_{j \in J} T_{kj}^{ot} h_{kj}$, where T_{kj}^{ot} is the occupied taxi movement (veh/h) from node (meeting location) $k \in K$ to destination $j \in J$ and 'ot' stands for occupied taxis. The total unoccupied taxi time consists of the moving times of vacant taxis from nodes of customer destinations to nodes of meeting locations and waiting (searching) times at each meeting

[2] While taxi supply at individual meeting locations is elastic (internal non-zero supply elasticity) the system-wide supply elasticity is zero (the total number of taxis in service is fixed). A supply elasticity could be introduced by making the number of taxis dependent on equilibrium taxi utility (to be introduced later), or by employing the multi-period dynamic model of taxi services with endogenous service intensity (Yang et al., 2005b).

location. This time is given by $\sum_{j\in J}\sum_{k\in K}T^{vt}_{jk}\{h_{jk}+w^{t}_{k}\}$, where 'vt' and 't' stand for vacant taxis and taxis; T^{vt}_{jk} is the vacant taxi movement (veh/h) from node $j\in J$ to node $k\in K$; $w^{t}_{k}, k\in K$ is the taxi waiting time at location $k\in K$ if k is a prescribed taxi stand and the taxi searching time if k denotes a meeting zone. The sum of total occupied taxi-hours and total vacant taxi-hours should be equal to the total taxi service hour. Therefore, the following taxi service time constraint must be satisfied in view of the 1-h period modeled here:

$$\sum_{k\in K}\sum_{j\in J}T^{ot}_{kj}h_{kj}+\sum_{j\in J}\sum_{k\in K}T^{vt}_{jk}\{h_{jk}+w^{t}_{k}\}=N. \tag{5}$$

As assumed before (Yang and Wong, 1998), all occupied taxies follow an 'all-or-nothing' routing behavior. A taxi, once occupied by a customer in location $k\in K$, will move to the customer's destination, $j\in J$, via the shortest path. Once a customer ride is completed, the taxi becomes vacant and is free to choose the next meeting location (possibly at the same location of the customer's destination) to seek the next customer. In doing this, each taxi driver is assumed to maximize expected utility, which depends positively on expected net revenue (revenue net of operating costs) and negatively on moving and waiting time (Wong et al., 2003). Let κ denote the cost per unit taxi operating time, and let $\kappa(h_{jk}+w^{t}_{k})$ be the searching cost (disutility) for a vacant taxi originating from node $j\in J$ to meet a customer in location $k\in K$, and $(\tilde{F}_k-\kappa\tilde{h}_k)$ be the expected net revenue (utility) of individual taxis upon meeting a customer in location $k\in K$. We have $\tilde{h}_k=\sum_{j\in J}T^{ot}_{kj}h_{kj}/\sum_{j\in J}T^{ot}_{kj}$ and $\tilde{F}_k=\sum_{j\in J}T^{ot}_{kj}F_{kj}/\sum_{j\in J}T^{ot}_{kj}$, respectively, which are the average trip length and the average fare revenue of the occupied taxi ride originating from location $k\in K$ to all destination nodes, $j\in J$.

Here, we should note that we do not distinguish the unit operating cost between occupied and vacant (cruising and waiting) taxis, taxi drivers are not allowed to differentiate customers at each meeting location and one taxi is occupied by one customer only.

With the trade-off between the expected net revenue and the searching cost, the net utility function for a taxi driver originating at node $j\in J$ and searching and meeting a customer at location $k\in K$ can be defined as

$$U_{k/j}=\frac{\lambda}{\kappa}\left(\tilde{F}_k-\kappa\tilde{h}_k\right)-\left(h_{jk}+w^{t}_{k}\right) \tag{6}$$

or rewritten to $U_{k/j}=\tilde{r}_k-\left(h_{jk}+w^{t}_{k}\right),\ k\in K,\ j\in J$, where $\tilde{r}_k$ is regarded as the perceived profitability of meeting location k

$$\tilde{r}_k=\frac{\lambda}{\kappa}(\tilde{F}_k-\kappa\tilde{h}_k),\quad k\in K. \tag{7}$$

Here, λ is a behavioral parameter (termed as the perceived profitability parameter hereinafter) that reflects the taxi drivers' perceptions of the importance of the net revenue consideration relative to that of the searching cost. This parameter measures the aggressiveness of taxi drivers in trying to maximize their expected net revenue of individual rides, which depends on the information on the profitability of different meeting locations available to taxi drivers. The larger the value of λ, the more information is available for taxi drivers to maximize the net revenue of individual trips. In the case of λ equal to zero, the expected net revenue of taxi drivers conditional on having contacted a customer is identical in any meeting location, and taxi drivers in this case maximize the expected net utility by minimizing the expected searching and waiting time for meeting the next customer. This is exactly the early taxi behavior model initially proposed by Yang and Wong (1998). Note that the above consideration of the expected profitability of a taxi ride from a specific meeting location is critical when some major meeting spots exist in the model. Taxi drivers are often willing to wait in airport queues in Hong Kong for a few hours for a single ride. This is a direct result of the driver's response to the much higher expected profitability of the airport taxi service.

The expected net revenue and searching and waiting cost in each meeting location are a random variable due to the variation in perceptions of taxi drivers and the random arrival of customers. This random variable can be assumed to be identically distributed with a Gumbel density function. With these considerations, we have the following behavior assumption on the searching and meeting of vacant taxis on the network.

A2: *The probability that a vacant taxi originating at node $j\in J$ chooses to meets its next customer in location $k\in K$ is given by the logit model*:

$$P^{t}_{k/j}=\frac{\exp\{-\theta^{t}[(h_{jk}+w^{t}_{k})-\tilde{r}_k]\}}{\sum\limits_{m\in K}\exp\{-\theta^{t}[(h_{jm}+w^{t}_{m})-\tilde{r}_m]\}},\quad \text{for any } k\in K,\ j\in J, \tag{8}$$

where θ^{t} is a nonnegative parameter.

Note that the value of parameter θ^{t} reflects the degree of uncertainty about customer demand and taxi services in the whole market from the perspective of individual taxi drivers, and it can be calibrated from observational data. Finally, for a given hour, the number of taxis to complete services at a customer destination zone must be identical to the number of vacant taxis available from that destination. We thus have

$$T^{vt}_{jk}=\sum_{l\in K}T^{ot}_{lj}P^{t}_{k/j},\ k\in K, j\in J. \tag{9}$$

2.4. Taxi–Customer meetings and waiting times

We now investigate the technical characteristics of the production of taxi services by using a meeting function to characterize the contacts between customers searching for taxis and taxis searching for customers. The meeting function is then used to derive the customer and taxi searching and waiting time functions in each location.

The rate of meetings in a specific location between customers and taxis depends on the size of two pools: the number of waiting or unserved customers and the number of vacant or unoccupied taxis at a given instant. Clearly, the number of waiting customers, denoted by N_k^{c}, at location $k \in K$ is given by $N_k^{\mathrm{c}} = w_k^{\mathrm{c}} Q_k^{\mathrm{c}}$, where Q_k^{c} and w_k^{c} are, respectively, as defined before, the customer arrival rate and the waiting time at location $k \in K$. The number of unoccupied taxis, denoted by N_k^{vt}, at location $k \in K$ is given by $N_k^{\mathrm{vt}} = w_k^{\mathrm{t}} T_k^{\mathrm{vt}}$, where T_k^{vt} is the arrival rate of vacant taxis at location $k \in K$ with $T_k^{\mathrm{vt}} = \sum_{j \in J} T_{jk}^{\mathrm{vt}}$ and w_k^{t} is, again, as defined before, the taxi searching and waiting time at location $k \in K$. Thus, the meeting rate, $m_k^{\mathrm{c-t}}$, between customers and taxis is given as a function of N_k^{c} and N_k^{vt} below:

$$m_k^{\text{c-t}} = M_k\left(N_k^{\mathrm{c}}, N_k^{\mathrm{vt}}\right) = M_k\left(w_k^{\mathrm{c}} Q_k^{\mathrm{c}}, w_k^{\mathrm{t}} T_k^{\mathrm{vt}}\right), \quad k \in K. \tag{10}$$

We now have the following assumption regarding the property of the meeting function.

A3: *For any meeting location (area or point) $k \in K$, the meeting function, $m_k^{\mathrm{c-t}}$, in (10) is nonnegative and continuously differentiable with $\partial m_k^{\mathrm{c-t}}/\partial N_k^{\mathrm{c}} > 0$ and $\partial m_k^{\mathrm{c-t}}/\partial N_k^{\mathrm{vt}} > 0$ in their domain $N_k^{\mathrm{c}} \geqslant 0$, $N_k^{\mathrm{vt}} \geqslant 0$. Furthermore, $m_k^{\mathrm{c-t}} \to 0$ as either $N_k^{\mathrm{c}} \to 0$ or $N_k^{\mathrm{vt}} \to 0$.*

Note that differentiability of the meeting assumption is not essential; strictly increasing meeting functions help to guarantee the existence of the bilateral meeting equilibrium. Moreover, meeting rates might be expected to increase faster (slower) than linearly with proportionate increases in N_k^{c} and N_k^{vt}, which reflects an increasing (decreasing) return to scale of the production of taxi–customer meetings. The exact nature of the return to scale may well depend on the characteristics of the meeting location (point or area meeting) and the availability of demand and supply information to taxi drivers and customers in the market.

Suppose that each taxi can carry at most one customer. The inter-meeting times, that is, the intervals between successive taxi–customer meetings at a given location are thus given by:

$$t_k^{\text{c-t}} = \frac{1}{M_k\left(N_k^{\mathrm{c}}, N_k^{\mathrm{vt}}\right)} = \frac{1}{M_k\left(w_k^{\mathrm{c}} Q_k^{\mathrm{c}}, w_k^{\mathrm{t}} T_k^{\mathrm{vt}}\right)}, \quad k \in K. \tag{11}$$

The expected waiting times of customers and vacant taxis in steady-state equilibrium are then respectively

$$w_k^{\mathrm{c}} = \frac{N_k^{\mathrm{c}}}{M_k\left(N_k^{\mathrm{c}}, N_k^{\mathrm{vt}}\right)} = \frac{w_k^{\mathrm{c}} Q_k^{\mathrm{c}}}{M_k\left(w_k^{\mathrm{c}} Q_k^{\mathrm{c}}, w_k^{\mathrm{t}} T_k^{\mathrm{vt}}\right)}, \quad k \in K, \tag{12}$$

$$w_k^{\mathrm{t}} = \frac{N_k^{\mathrm{vt}}}{M_k\left(N_k^{\mathrm{c}}, N_k^{\mathrm{vt}}\right)} = \frac{w_k^{\mathrm{t}} T_k^{\mathrm{vt}}}{M_k\left(w_k^{\mathrm{c}} Q_k^{\mathrm{c}}, w_k^{\mathrm{t}} T_k^{\mathrm{vt}}\right)}, \quad k \in K. \tag{13}$$

Now, we consider the following specific form of the Cobb–Douglas type production function (Varian, 1992) of taxi–customer meetings that satisfies Assumption A3:

$$M_k\left(N_k^{\mathrm{c}}, N_k^{\mathrm{vt}}\right) = A_k\left(w_k^{\mathrm{c}} Q_k^{\mathrm{c}}\right)^{\alpha_1}\left(w_k^{\mathrm{t}} T_k^{\mathrm{vt}}\right)^{\alpha_2}. \tag{14}$$

In this production function, α_1 and α_2 are the elasticities of the meeting rate with respect to the number of unserved customers and the number of vacant taxis, respectively. These values are constants determined by available meeting technology. The location-dependent constant A_k can be specified as

$$A_k = a_0(\Phi_k)^{\alpha_0} \tag{15}$$

where Φ_k is a characteristic variable of the meeting location, $k \in K$ and a_0, $a_0 > 0$, and α_0 are two parameters. For example, Φ_k could represent the size of the meeting location, k, to reflect the spatially distinct meeting locations (point versus zonal meeting). In this case it is expected that $\alpha_0 < 0$.

Note that the function parameters, a_0 and α_0 in (15) and α_1 and α_2 in (14), are common to all meeting locations, and, of course, they can be made location-specific to further capture the heterogeneous meeting characteristics of individual locations.

For the specific Cobb–Douglas meeting function (14), it is generally expected that $0 < \alpha_1 \leqslant 1$, and $0 < \alpha_2 \leqslant 1$. The bilateral meeting function is said to exhibit increasing (constant or decreasing) returns to scale if $\alpha_1 + \alpha_2 > 1$ (=1 or <1). The increasing returns to scale property, as prevailing in most taxi markets like a queuing system, is an economic justification for the necessity of subsidizing taxi services in a first-best environment (Douglas, 1972; Arnott, 1996).

With the above specification of the Cobb–Douglas meeting functions (14), customer and taxi waiting time functions (12) and (13) in a steady-state equilibrium then become:

$$w_k^c = \frac{w_k^c Q_k^c}{A_k (w_k^c Q_k^c)^{\alpha_1} (w_k^t T_k^{vt})^{\alpha_2}}, \quad k \in K, \tag{16}$$

$$w_k^t = \frac{w_k^t T_k^{vt}}{A_k (w_k^c Q_k^c)^{\alpha_1} (w_k^t T_k^{vt})^{\alpha_2}}, \quad k \in K. \tag{17}$$

Simplifying the two equations yields:

$$w_k^c = \left[A_k (Q_k^c)^{\alpha_1 - 1} (w_k^t T_k^{vt})^{\alpha_2}\right]^{-\frac{1}{\alpha_1}} = (A_k)^{-\frac{1}{\alpha_1}} (Q_k^c)^{\left(\frac{1-\alpha_1}{\alpha_1}\right)} (w_k^t T_k^{vt})^{\left(-\frac{\alpha_2}{\alpha_1}\right)}, \quad k \in K, \tag{18}$$

$$w_k^t = \left(A_k (T_k^{vt})^{\alpha_2 - 1} (w_k^c Q_k^c)^{\alpha_1}\right)^{-\frac{1}{\alpha_2}} = (A_k)^{-\frac{1}{\alpha_2}} (T_k^{vt})^{\left(\frac{1-\alpha_2}{\alpha_2}\right)} (w_k^c Q_k^c)^{\left(-\frac{\alpha_1}{\alpha_2}\right)}, \quad k \in K. \tag{19}$$

If we further assume that $\alpha_1 = \alpha_2 \equiv 1.0$ (corresponding to the increasing returns to scale in the meeting function, because $\alpha_1 + \alpha_2 = 2.0 > 1.0$), Eqs. (18) and (19) then take the following simple forms:

$$w_k^c = \frac{1}{A_k w_k^t T_k^{vt}}, \quad k \in K, \tag{20}$$

$$w_k^t = \frac{1}{A_k w_k^c Q_k^c}, \quad k \in K. \tag{21}$$

Eq. (20) tells us that the customer-waiting time in a specific location is inversely proportional to the vacant taxi-hours (equal to the average number of vacant taxis available, as a 1-h modeling period is considered). This assumption has been implicitly or explicitly used by most previous authors who have made waiting time a function solely of the vacant taxi-hours (for instance, Douglas, 1972; De vany, 1975; Beesley and Glaister, 1983; Cairns and Liston-Heyes, 1996; Yang et al., 2002, 2005a,b) and is central to the theoretical results obtained so far in the literature on the taxi industry.[3]

In parallel to Eqs. (20), (21) states that the taxi searching and waiting time in a specific location is inversely proportional to the average number of unserved customers.

Now, we proceed to look further at the property of the meeting functions and the expected customer and taxi waiting times. Multiplying both sides of Eqs. (18) and (19) gives rise to:

$$(w_k^c)^{\alpha_1} (w_k^t)^{\alpha_2} = \frac{1}{A_k} (Q_k^c)^{\frac{1}{2} - \alpha_1} (T_k^{vt})^{\frac{1}{2} - \alpha_2}, \quad k \in K. \tag{22}$$

Fig. 2 depicts the general shape of the curve of w_k^c versus w_k^t for given alternative values of Q_k^c and T_k^{vt} in Eq. (22) (k is omitted). One should be aware that in a steady-state equilibrium, $Q_k^c \equiv T_k^{vt}$, $k \in K$, from the flow conservation conditions. In this case the term on the right-hand side of Eq. (22) becomes a constant, $1/A_k$, for $\alpha_1 + \alpha_2 = 1.0$, the case of constant returns to scale in the meeting function. The equilibrating variables of waiting times, w_k^c versus w_k^t, can nevertheless take any values on the curve in Fig. 2, depending on the outcome of the equilibration.

To proceed, we again assume that $\alpha_1 = \alpha_2 \equiv 1.0$ and consider the case where $\Phi_k \to 0$ (point meeting). In this case, $A_k \to \infty$ in (15) due to $a_0 > 0$ and $\alpha_0 < 0$. In view of the fact that Q_k^c and T_k^{vt} are of limited values, we thus arrive at

$$w_k^c w_k^t = 0, \quad k \in K. \tag{23}$$

Eq. (23) represents point meeting or only holds at a taxi stand. It is readily observed from this equation that if $w_k^c > 0$, then $w_k^t = 0$, or alternatively if $w_k^t > 0$, then $w_k^c = 0$, which means that either customers are waiting for taxis or taxis are waiting for customers at point meeting location $k \in K$. In terms of the searching and matching literature (Petrongolo and Pissarides, 2001), equilibrium will be said to exhibit frictions if the corresponding allocation simultaneously exhibits vacant taxis and unserved customers. In the present case, an equilibrium exhibits frictions at location $k \in K$ if $w_k^c w_k^t > 0$, and it is frictionless if $w_k^c w_k^t = 0$. The former corresponds to the bilateral taxi–customer search and meeting in, for example, a residential area without fixed taxi stands, whereas the latter corresponds to the meetings between taxis and customers at a prescribed taxi stand (with unlimited queuing and boarding capacity), such as an airport where taxi drivers usually wait in queues for hours and the average customer-waiting time is essentially zero or negligible.

3. Existence and solution algorithm of stationary equilibria

3.1. Definition of stationary equilibrium of searching and meeting

Suppose the behaviors of individual customers and taxi drivers are guided solely by their personal self-interest under the given rules of market operations, and each customer and taxi is individually insignificant in the market with no market

[3] The specification with $\alpha_1 = \alpha_2 \equiv 1.0$ gives rise to the traditional customer-waiting time function (20) used in a number of previous studies. This specification is somewhat flawed, since it fails to take into account that waiting time depends not only on the number of vacant taxis but also on the number of people who are waiting for vacant taxis.

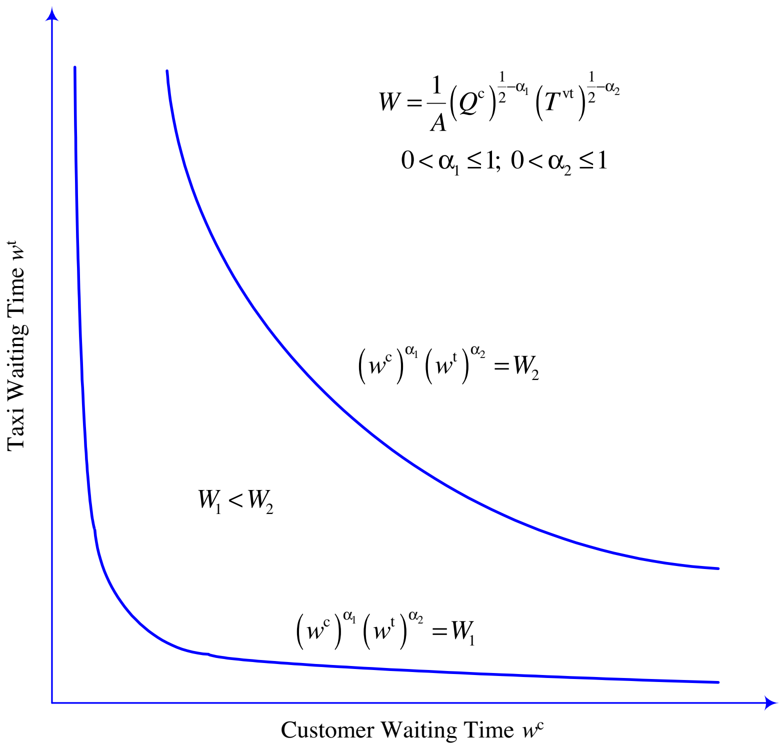

Fig. 2. The Cobb–Douglas meeting function and customer–taxi waiting time.

power to affect prevailing searching and waiting times. Then, under this perfectly competitive market system, a stationary equilibrium is achieved when the demands of customers match the supplies of (vacant) taxis at prevailing searching and waiting times in every meeting location simultaneously. In other words, the stationary equilibrium means a time-invariant distribution of taxis and customers across locations such that, given this distribution, taxi drivers choose where to locate themselves to maximize net utilities and customers choose meeting locations for taxi rides to minimize full trip prices. Specifically, the bilateral equilibrium of searching and meeting in the taxi market is given below.

Definition. With a Cobb–Douglas type production function (14) of taxi–customer meetings, the customer and taxi movements in a network are said to achieve a competitive stationary equilibrium if the following conditions, C1, C2 and C3, hold.

C1: *For all* $k \in K$, (Q^{c}_{ikj}, Q^{c}_{k}) *satisfy* Eqs. (1)–(4) given w^{c}_{k} (*choice of customers*).
C2: *For all* $k \in K$, $(T^{ot}_{kj}, T^{vt}_{jk})$ *satisfy* Eqs. (5)–(9) *given* w^{t}_{k} (*choice of taxi drivers*).
C3: *The customer and taxis match exactly or there is market clearing at each taxi-customer meeting location* $k \in K$, *namely* $T^{ot}_{kj} = Q^{c}_{kj}, k \in K, j \in J$, *and* (w^{c}_{k}, w^{t}_{k}) *satisfy the meeting function* (22) (*or equivalently*, (18) or (19)) *at any meeting location*, $k \in K$.[4,5]

3.2. Existence of equilibria

With the underlying market setting and modeling assumptions, we now prove the existence of a steady-state equilibrium that meets the above conditions by virtue of Brouwer's fixed-point theorem (Fuente, 2000) for the highly nonlinear system introduced so far. Let $\mathbf{Q}^{c} = (Q^{c}_{ikj}, i \in I, j \in J, k \in K)$ be the vector of customer demands for taxi services and Ω be the feasible set of $\mathbf{Q}^{c}$. Brouwer's fixed-point theorem states that: if $\Gamma : \Omega \rightarrow \Omega$ is a continuous function mapping a compact and convex

[4] Note that a perfect analogy between the Arrow–Debreu model of competitive general equilibrium (Arrow and Debreu, 1954) and the taxi "economy" can not be made. In the Arrow–Debreu model, aggregate demand and aggregate supply for each market are functions of the price vector, and prices adjust to equilibrate demand and supply in each market. Here the customer demand functions depend on the vector of customer-waiting times, the taxi supply functions depend on the vector of taxi waiting times, and the searching and waiting times of taxis and customers are related through the meeting function.

[5] Walras' Law plays a very important role in competitive general equilibrium theory (Ellickson, 1993; Starr, 1997). In its proper form, it states that, in or out of equilibrium, if all consumers are on their budget constraints, the value of excess demand is zero. It is often stated in terms of one of its implications: If all markets but one clears, that one must clear too, which is why general equilibrium determines only relative prices. Walras' Law does not apply here for the taxi market because of the existence of the virtual mode, mass transit.

set Ω into itself, there is some $\mathbf{Q}^c$ in Ω such that $\mathbf{Q}^c = \Gamma(\mathbf{Q}^c)$. Our proof proceeds as follows: First, according to the generalized travel cost (1) by taxi mode and the logit-based mode choice model (2), we show that Ω is a compact and convex set. Second, with resort to an artificial entropy-type convex minimization problem, we establish two lemmas to show that the continuity condition required for Brouwer's fixed-point theorem is satisfied. Namely, we take customer demands, $\mathbf{Q}^c$, as arguments and show that both taxi and customer-waiting times, w_k^t, $w_k^c, k \in K$, are their continuous functions. Finally, from the continuity result and the logit-based customer demand model (2) again, we establish a continuous mapping from Ω into itself, which, by Brouwer's Theorem, has a fixed point.

First, let $\bar{g}_{ikj} = F_{kj} + \beta_1 h_{ik} + \beta_2 h_{kj}$ in the generalized travel cost function (1) of taxi mode, we have $\bar{g}_{ikj} \leqslant g_{ikj} \leqslant \infty$ because $0 \leqslant w_k^c \leqslant \infty$. From the logit-based mode choice model (2), we have

$$0 \leqslant P_{k/ij}^c \leqslant \bar{P}_{k/ij}^c = \frac{1}{1 + \exp[-\theta^c(\bar{g}_{ij} - \bar{g}_{ikj})]}, \tag{24}$$

which implies that

$$0 \leqslant Q_{ikj}^c \leqslant \bar{Q}_{ikj}^c = Q_{ij}^c \bar{P}_{k/ij}^c, \quad i \in I,\ j \in J,\ k \in K \tag{25}$$

where $\bar{P}_{k/ij}^c$ and $\bar{Q}_{ikj}^c$ are the upper bounds of $P_{k/ij}^c$ and Q_{ikj}^c, respectively. Note that $P_{k/ij}^c \to 0$ as $w_k^c \to \infty$. We simply define $P_{k/ij}^c = 0$ if $w_k^c = \infty$. Thus, in what follows, we can limit our discussion to the following feasible set of customer demands for taxi services:

$$\Omega = \left\{ \mathbf{Q}^c = \left(Q_{ikj}^c, i \in I, j \in J, k \in K \right) \middle| 0 \leqslant Q_{ikj}^c \leqslant \bar{Q}_{ikj}^c \right\}. \tag{26}$$

Clearly, the feasible set (26) of customer demands is closed and bounded and hence compact.

Next, we consider taxi search and movement for given $Q_k^c = \sum_{i \in I} \sum_{j \in J} Q_{ikj}^c, k \in K$. Consider the following artificial entropy-type convex minimization program that is constructed so as to generate the logit-based taxi choice probability, $P_{k/j}^t, j \in J, k \in K$, given by Eq. (8).

$$\min_{\mathbf{T}^{vt}} \quad \sum_{k \in K} \sum_{j \in J} \left\{ T_{jk}^{vt} h_{jk} + \frac{1}{\theta^t} T_{jk}^{vt} \left(\ln T_{jk}^{vt} - 1 \right) - T_{jk}^{vt} \tilde{r}_k \right\} \tag{27}$$

subject to

$$\sum_j T_{jk}^{vt} \leqslant Q_k^c, \quad k \in K, \tag{28}$$

$$T_{jk}^{vt} \geqslant 0, \quad k \in K,\ j \in J, \tag{29}$$

where $\mathbf{T}^{vt} = (T_{jk}^{vt}, j \in J, k \in K)$ and inequality (instead of equality) constraint (28) is introduced to ensure feasibility of the problem for a given $\mathbf{Q}^c \in \Omega$ and the nonnegativity of the corresponding Lagrange multipliers, which can be interpreted as taxi waiting times at various meeting locations as explained later.

We form the following Lagrangian,

$$L = \sum_{k \in K} \sum_{j \in J} \left\{ T_{jk}^{vt} h_{jk} + \frac{1}{\theta^t} T_{jk}^{vt} \left(\ln T_{jk}^{vt} - 1 \right) - T_{jk}^{vt} \tilde{r}_k \right\} + \sum_{k \in I} \alpha_k \left(\sum_{j \in J} T_{jk}^{vt} - Q_k^c \right). \tag{30}$$

Note that $\lim_{T_{jk}^{vt} \to 0^+} T_{jk}^{vt} \ln T_{jk}^{vt} = 0$ and hence one can define $T_{jk}^{vt} \ln T_{jk}^{vt} = 0$ for $T_{jk}^{vt} = 0$. As long as $Q_k^c > 0, \quad k \in K,\ T_{jk}^{vt} > 0$ is always satisfied at the minimum due to the shape of the objective function (Erlander and Stewart, 1990). Consider the case $Q_k^c > 0, \quad k \in K$. The optimality condition for the minimization program can be written as (Bazaraa et al., 1993):

$$\alpha_k \geqslant 0, \quad \frac{\partial L}{\partial \alpha_k} \leqslant 0 \to \sum_j T_{jk}^{vt} \leqslant Q_k^c; \tag{31}$$

$$\alpha_k \frac{\partial L}{\partial \alpha_k} = 0 \to \alpha_k \left(\sum_j T_{jk}^{vt} - Q_k^c \right) = 0, \quad k \in K; \tag{32}$$

$$\frac{\partial L}{\partial T_{jk}^{vt}} = 0 \to h_{jk} + \frac{1}{\theta^t} \ln T_{jk}^{vt} - \tilde{r}_k + \alpha_k = 0, \quad j \in J, \quad k \in K. \tag{33}$$

Therefore, from (33)

$$T_{jk}^{vt} = \exp\left\{ -\theta^t \left[(h_{jk} + \alpha_k) - \tilde{r}_k \right] \right\}, \quad j \in J,\ k \in K. \tag{34}$$

This result leads to

$$P_{k/j}^t = \frac{T_{jk}^{vt}}{\sum_{m \in K} T_{jm}^{vt}} = \frac{\exp\{-\theta^t[(h_{jk} + \alpha_k) - \tilde{r}_k]\}}{\sum_{m \in K} \exp\{-\theta^t[(h_{jm} + \alpha_m) - \tilde{r}_m]\}}, \quad j \in J,\ k \in K. \tag{35}$$

By comparing Eqs. (35) and (8) of the original logit-based taxi choice model, the Lagrange multiplier, α_k, can now be interpreted as the taxi waiting time in meeting location k, i.e., $w_k^t = \alpha_k, k \in K$. With this in mind and with the underlying market setting and Assumptions A1–A3, we are now ready to establish two lemmas.

Lemma 1. *Under Assumption A1–A3, constraint (28) is binding for all $k \in K$.*

Proof. Consider a representative location, $k \in K$. If $Q_k^c = 0$, then we must have $T_{jk}^{vt} = 0$ for all $j \in J$ because of $T_{jk}^{vt} \geqslant 0$. Thus constraint (28) is binding. Note that, in this case, the objective function, as mentioned earlier, is still well defined with $T_{jk}^{vt} \ln T_{jk}^{vt} = 0$ for $T_{jk}^{vt} = 0$. Next consider the case when $Q_k^c > 0$. If constraint (28) is not binding, i.e. $\sum_{j \in J} T_{jk}^{vt} < Q_k^c, k \in K$, then we have the corresponding Lagrange multiplier, $\alpha_k = w_k^t = 0$, and the optimal solution of the convex program (27)–(29) $T_{jk}^{vt} > 0$. By Assumption A3 we have $w_k^c \to \infty$, and then, by Assumption A1, we have $P_{k/ij}^c = 0$ for all $i \in I, j \in J$. From Eq. (4) we have $Q_k^c = 0$, which contradicts our assumption of $Q_k^c > 0$. Lemma 1 thus holds. □

Lemma 2. *Under Assumptions A1–A3, the taxi waiting time, $w_k^t = \alpha_k$, at any location, $k \in K$, is a continuous function of $\mathbf{Q}^c = (Q_{ikj}^c, i \in I, j \in J, k \in K)$.*

Proof. First we note that the objective function (27) is strictly convex in $\mathbf{T}^{vt}$ and thus the solution T_{jk}^{vt} of the program (27)–(29) is unique, which means that $P_{k/j}^t$ is unique as well for any $j \in J$ and $k \in K$. From Eq. (35) alone, we also observe that the Lagrange multiplier, α_k, is not unique, because an arbitrary constant, η, can be always added to α_k to form a new set of Lagrange multipliers $(\alpha_k + \eta)$ such that the solution to the minimization problem in Eq. (35) remains unchanged. Therefore, the problem is subject to one degree of freedom for the solution of Lagrange multipliers and hence the Lagrange multipliers are additive to a constant (Wong and Yang, 1998). Nonetheless, the constant is fixed and hence the taxi waiting times are unique for a given $\mathbf{Q}^c$. To see this, we first note that at the optimum of the convex program (27)–(29), constraint (28) is binding, and we have $\sum_{j \in J} T_{jk}^{vt} = Q_k^c = \sum_{i \in I}\sum_{j \in J} Q_{ikj}^c = \sum_{j \in I} Q_{kj}^c$ for any $k \in K$. The occupied taxi movements are thus specified by $T_{kj}^{ot} = Q_{kj}^c, k \in K, j \in J$. From the time conservation Eq. (5) for both occupied and vacant taxi movements, taxi waiting times must satisfy the following time conservation equation:

$$\sum_{k \in K} \sum_{j \in J} Q_{kj}^c h_{kj} + \sum_{j \in J} \sum_{k \in K} T_{jk}^{vt} \{h_{jk} + \alpha_k + \eta\} = N, \tag{36}$$

and thus,

$$\eta = \frac{N - \sum\limits_{k \in K} \sum\limits_{j \in J} Q_{kj}^c h_{kj} - \sum\limits_{j \in J} \sum\limits_{k \in K} T_{jk}^{vt} \{h_{jk} + \alpha_k\}}{\sum\limits_{j \in J} \sum\limits_{k \in K} T_{jk}^{vt}}. \tag{37}$$

Therefore, for a particular set of multipliers, $\alpha_k, k \in K$, determined from the convex program (27)–(29), the adjustment constant, η, to ensure fulfillment of the taxi time conservation Eq. (5), is uniquely determined by (37), and the taxi waiting times can then be obtained as

$$w_k^t = \alpha_k + \eta, \quad k \in K. \tag{38}$$

The taxi waiting time in (38) is independent of the choice of the particular set of multipliers. One can readily show that for any alternative set of multipliers, $\alpha'_k, k \in K$, with the corresponding constant, η', satisfying the logit model (35) and the time conservation Eq. (5), we must have $w_k^t = \alpha_k + \eta = \alpha'_k + \eta', k \in K$. We thus confirm that the taxi waiting time, $w_k^t, k \in K$, is unique for given $\mathbf{Q}^c = (Q_{ikj}^c, i \in I, j \in J, k \in K)$.

We still need to show the continuity of the taxi waiting times, $w_k^t, k \in K$, as functions of $\mathbf{Q}^c$. This can be proved by contradiction using the above known uniqueness properties. One may simply consider a special case with one single origin, i, and one single destination, j, and assume that $w_k^t, k \in K$ is not a continuous function of $\mathbf{Q}^c$. From the logit-based choice model of taxi drivers (Assumption A2), we know that T_{jk}^{vt} is a continuous function of $w_k^t, k \in K$; and, furthermore, from Lemma 1, we know that constraint (28) is binding under Assumptions A1–A3, namely $T_{jk}^{vt} = Q_k^c, k \in K$ or $Q_k^c, k \in K$ is a continuous function of $w_k^t, k \in K$. This continuity result and the assumption of discontinuity of $w_k^t, k \in K$ in $\mathbf{Q}^c$ imply that $w_k^t, k \in K$ must be nonunique for given $\mathbf{Q}^c$. This contradicts the earlier result that $w_k^t, k \in K$ is unique for given $\mathbf{Q}^c$. □

After establishing the continuity of taxi waiting times given by Lemma 2, we are now in a position to state and prove the following existence theorem of the bilateral taxi–customer searching and meeting equilibrium.

Existence Theorem. Under Assumptions A1–A3, there exists an equilibrium customer demand vector, $\mathbf{Q}^c = (Q_{ikj}^c, i \in I, j \in J, k \in K)$, and vacant and occupied taxi movement vectors, $\mathbf{T}^{vt} = (T_{jk}^{vt}, j \in J, k \in K)$ and $\mathbf{T}^{ot} = (T_{kj}^{ot}, k \in K, j \in J)$, that meet the equilibrium conditions C1–C3 defined before.

Proof. As established in Lemma 2, taxi waiting time $w_k^t, k \in K$ is a continuous function of customer demand vector, $\mathbf{Q}^c$, for taxi services. Furthermore, from the assumption of the meeting function, we know that customer-waiting time, w_k^c, varies continuously with w_k^t at each meeting location, $k \in K$. We thus conclude that customer-waiting time $w_k^c, k \in K$ is a continuous function of $\mathbf{Q}^c$. Let $w_k^c = w_k^c(\mathbf{Q}^c), k \in K$ represent such a continuous mapping, where $\mathbf{Q}^c \in \Omega$ defined in (26) and $0 \leqslant w_k^c \leqslant +\infty$. From the logit-based customer demand models, we thus have the following self-map

$$\mathbf{Q}^c = \Gamma(\mathbf{Q}^c) \tag{39}$$

where

$$\Gamma_{ikj}(\mathbf{Q^c}) = \frac{\exp\left\{-\theta^c g_{ikj}[w_k^c(\mathbf{Q}^c)]\right\}}{\exp\{-\theta^c \bar{g}_{ij}\} + \sum\limits_{m\in K} \exp\left\{-\theta^c g_{imj}\left[w_m^c(\mathbf{Q}^c)\right]\right\}}, \quad \text{for any} \quad k \in K,\ i \in I,\ j \in J \tag{40}$$

and

$$g_{ikj}\left[w_k^c(\mathbf{Q}^c)\right] = F_{kj} + \beta_1 h_{ik} + \beta_2 h_{kj} + \beta_3 w_k^c(\mathbf{Q}^c), \quad \text{for any} \quad k \in K,\ i \in I,\ j \in J. \tag{41}$$

By Lemma 2, the mapping $\Gamma : \Omega \subset \Re^m \rightarrow \Re^m$ defined by Eq. (40) is a continuous mapping on the compact and convex set Ω in $\Re^m$, where $m = |I||K||J|$ with $|I|, |K|, |J|$ denoting the cardinalities of the sets I, J and K. Also, for any $\mathbf{Q}^c \in \Omega$, clearly $\Gamma(\mathbf{Q}^c) \in \Omega$ so that $\Gamma\Omega \subset \Omega$. We can thus apply Brouwer's fixed-point theorem to conclude that Γ has at least one fixed point in Ω. Consequently, the existence of bilateral searching and meeting equilibria is guaranteed. □

We conclude this section by pointing out that there might be multiple equilibria of the bilateral searching and meeting, because of the nonlinear mapping or the nonuniqueness of the solution to the simultaneous nonlinear equations. Indeed, the possibility of both stable and unstable equilibria can arise even in a single spot market, as demonstrated by Matsushima and Kobayashi (2006).

3.3. A solution algorithm

The general idea for numerically finding a steady-state equilibrium is to solve both the customer choice subproblem and the vacant taxi moving subproblem iteratively until a convergence criterion is met. The general procedure is set out below.

Equilibration algorithm of taxi–customer searching and meeting

Step 0. Initialization: set an initial set of customer-waiting times, $w_k^{c(0)}, k \in K$. Let $n := 1$.
Step 1. Customer demand updating: update $Q_{ikj}^{c(n)} = Q_{ij}^c P_{k/ij}^{c(n)}, i \in I, j \in J, k \in K$ where $P_{k/ij}^{c(n)}$ is given by Eq. (2).
Step 2. Vacant taxi movement: solve the taxi movement sub-model for given $Q_{ikj}^{c(n)}, i \in I, j \in J, k \in K$ by the iterative balancing method and thus obtain a set of taxi searching/waiting times, $w_k^{t(n)}, k \in K$.
Step 3. Updating customer-waiting time: update $w_k^{c(n)}, k \in K$ with $w_k^{t(n)}, k \in K$ according to the meeting function (18).
Step 4. Convergence criterion: if convergence is attained, then stop. Otherwise, let $n := n + 1$ and go to Step 1.

In Step 2, the iterative balancing method for solving the taxi movement sub-model (27)–(29), with inequality constraint (28) replaced by equality constraint and subject to the taxi time conservation Eq. (36), is fully described by Wong and Yang (1998). In Step 4, a plausible convergence criterion is based on the change in the weighted customer-waiting times in all meeting locations between two successive iterations. That is,

$$\sqrt{\sum_{k\in K} \frac{Q_k^{c(n)}}{\sum\limits_{m\in K} Q_m^{c(n)}} \left(w_k^{c(n)} - w_k^{c(n-1)}\right)^2} \leqslant \varepsilon, \tag{42}$$

where ε is a predetermined convergence tolerance.

4. A numerical example

We now present a simple numerical example to investigate the convergence of the algorithm and conduct sensitivity analysis of the equilibrium outcomes of searching and meeting. In particular, we look at the impact of the two major factors, the taxi fleet size and the perceived profitability parameter, on the average taxi and customer-waiting times in spatially distinct meeting locations.

4.1. Example setting and input data[6]

We consider a simple four node network with full connections between all node pairs. Each node represents an origin and a destination location for customers and also a taxi–customer meeting location. Nodes 1–4 represent in sequence an airport, a railway station, a residential area and a commercial area, respectively. We assume that $\alpha_1 = \alpha_2 \equiv 1.0$, which means that the

[6] Values of parameter selected in the example do not necessarily represent the Hong Kong situation but are simply for illustrative purpose.

Table 1
Zone (node) to zone (node) customer walking time.

h_{ik} (h)	1	2	3	4
1	0.05	4.00	2.00	4.00
2	4.00	0.03	4.00	1.50
3	2.00	4.00	0.30	4.00
4	4.00	1.50	4.00	0.20

Table 2
Zone (node) to zone (node) taxi movement time.

h_{kj} (h)	1	2	3	4
1	0.02	1.00	0.60	1.20
2	1.00	0.01	1.20	0.20
3	0.60	1.20	0.15	1.00
4	1.20	0.20	1.00	0.10

waiting time functions, (20) and (21), are used, respectively, for customers and taxis in each location, representing a situation of increasing returns to scale meetings. At the airport and railway station (nodes 1 and 2), the meetings between customers and taxis occur only at fixed taxi stands (point meetings), and the corresponding meeting functions are assumed to have a large value of the location-specific term: $A_1 = a_0(\Phi_1)^{\alpha_0} = A_2 = a_0(\Phi_2)^{\alpha_0} = 10^4$ (1/veh-h) ($\Phi_1 \to 0$ and $\Phi_2 \to 0$ with $\alpha_0 < 0$). For the residential and commercial zones (nodes 3 and 4), much smaller values of $A_3 = a_0(\Phi_3)^{\alpha_0} = 0.01$ $(1/\mathrm{veh}-\mathrm{h})$ and $A_4 = a_0(\Phi_4)^{\alpha_0} = 0.1$ $(1/\mathrm{veh}-\mathrm{h})$ are used to represent the 'zonal nature' of searching, where taxi–customer meetings can occur anywhere in the zone. The matrices of customer walking times (inter-zone and intra-zone), taxi movement times and total customer demands by taxis and public transport modes are provided in Tables 1–3, respectively. A constant generalized travel cost by public transport modes between each customer origin and destination is assumed and given by

$$\bar{g}_{ij} = \min_{k \in K} \left\{ \frac{1}{15} F_{kj} + 1.1\beta_1 h_{ik} + 2.0\beta_2 h_{kj} + \beta_3 \bar{w}_k \right\}, \quad i \in I, \ j \in J, \tag{43}$$

where $\bar{w}_k, k \in K$ is assumed to be 0.15 (h) and the coefficients, 1/15 and 2.0, reflect the difference in fares and in-vehicle times between taxis and other public transport modes. The rest of the input data used in the example are given below (1US$ $\approx$ 7.8HK$): θ^{t} = 10.0 (1/h), θ^{c} = 0.05 (1/h), F_0 = 0.0 (HK$), τ = 150 (HK$/h), κ = 50 (HK$/h), β_1 = 120 (HK$/h), β_2 = 35 (HK$/h), β_3 = 60 (HK$/h). In addition, $\lambda = 1.0$ and N = 1000 (taxis) for the base case, unless the two parameters are varied for sensitivity analysis.

It should be noted here that the airport (node 1) is located in a remote rural area, the railway station (node 2) is located in the urban centre within the commercial area (node or zone 4), and the residential district (node or zone 3) is in the suburban area. Appropriate intra-zonal and inter-zonal moving times in Table 2 are assumed to reflect this geographical setting. For example, the taxi moving time between the airport and the three other locations are set much longer than the rest. This setting aims to demonstrate the spatially distinct meeting locations that exhibit different search frictions and considerable variation in the expected trip lengths and average net revenue per taxi ride. For clarity, Table 4 summarizes the characteristics and the nature of taxi–customer meetings in each node or zone.

4.2. Result analysis

With an initial starting point of customer-waiting times being zero in all four locations, Fig. 3 shows the convergence of the solution algorithm in terms of the error measure given by Eq. (42) versus the number of iterations. A sharp convergence of the proposed algorithm is observed in the initial few iterations, and the larger the taxi fleet size is, the higher the rate of convergence.

Table 3
Matrix of customer demand by taxi and public transport modes.

Q^{c}_{ij} (person/h)	1	2	3	4	Total
1	0	1000	1500	2000	4500
2	2000	0	1000	5000	8000
3	1000	2500	2000	3000	8500
4	2000	4000	5000	5000	16,000
Total	5000	7500	9500	15,000	37,000

Table 4
The nature of taxi–customer meeting in each node/zone.

Node or zone no.	1	2	3	4
Zonal nature	Airport	Railway station	Residential area	Commercial area
Location type	Rural	Urban	Suburban	Urban
Meeting nature	Point meeting	Point meeting	Zonal meeting	Zonal meeting

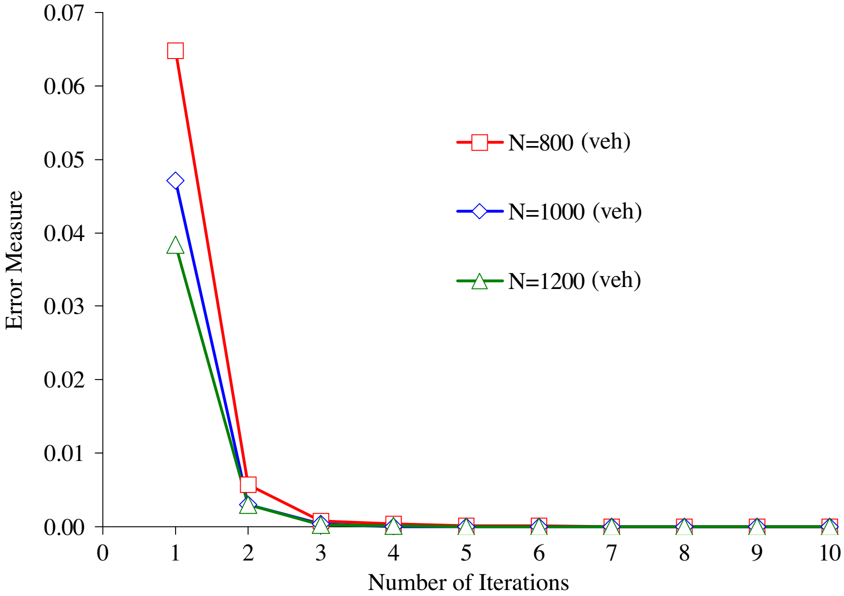

Fig. 3. The convergence error measure versus the number of iterations (Lambda = 1.0).

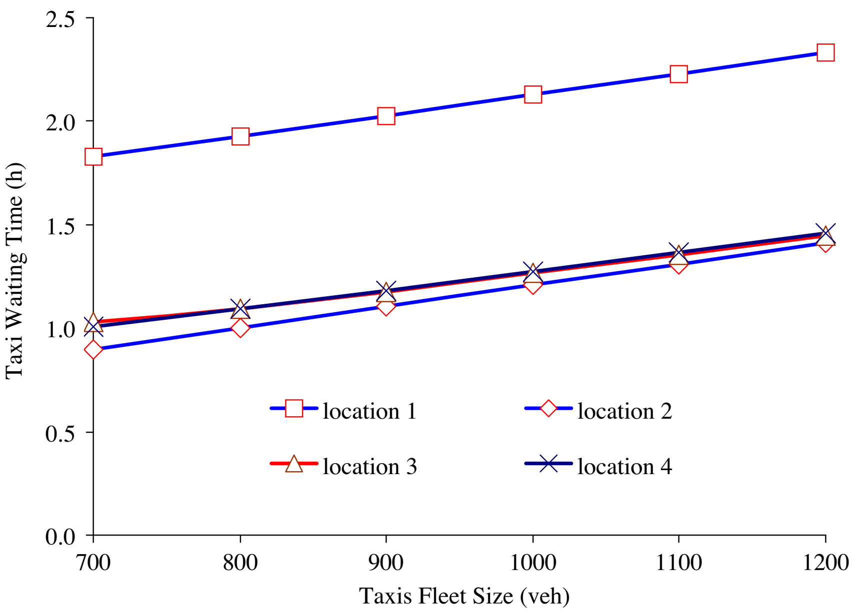

Fig. 4. Average taxi waiting times in different locations versus taxi fleet size.

Figs. 4 and 5 depict, respectively, the change in taxi and customer-waiting times in all four locations versus the total taxi fleet size with a given value of the perceived profitability parameter, $\lambda = 1.0$. From Fig. 4 it is observed that taxi waiting times in all four locations are strictly positive and increase nearly linearly with taxi fleet size, because the extra taxi service time

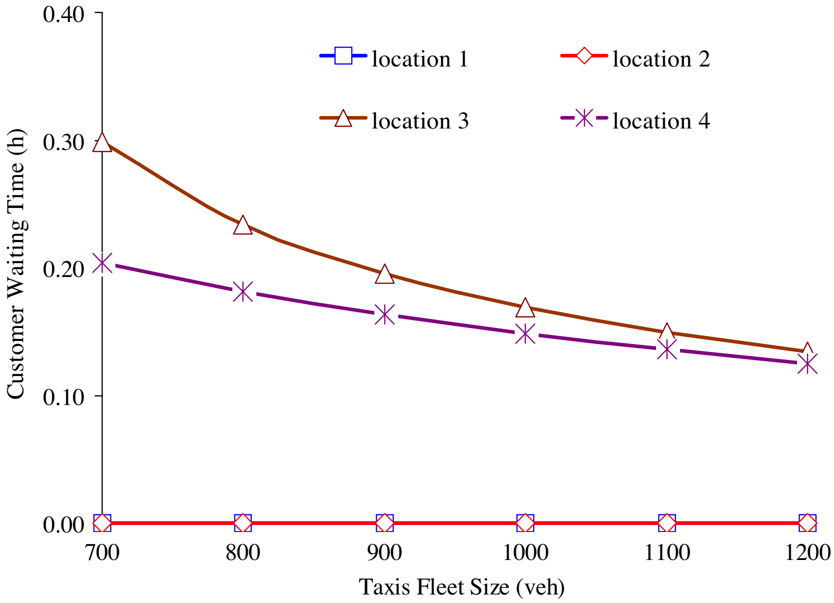

Fig. 5. Average customer-waiting times in different locations versus taxi fleet size.

from additional taxis is mostly spent in searching and waiting for customers. Nevertheless, the exact nature of the dependence of taxi waiting times on the number of taxis in service may well depend on the current market conditions (e.g., current taxi fleet size and customer demand elasticity at the current service level), because the new demand (and hence an increase in the occupied taxi time) brought about by the enhanced taxi availability due to an increase in the taxi fleet size depends on the sensitivity of the current demand with respect to service quality or customer-waiting time. From Fig. 5, it can be seen that the customer-waiting time in locations 3 and 4 decreases nonlinearly with taxi fleet size at a diminishing rate. This means that customer-waiting time is more sensitive to the number of taxis in service with smaller fleet size or the marginal effect of increasing the taxi fleet size in improving service quality is diminishing. Of particular note is that the customer-waiting time in locations 1 and 2 is almost zero, regardless of the taxi fleet size within the range considered. This is clearly a direct result of the nature of point meetings at locations 1 (airport) and 2 (railway station) (both locations are assumed to have unlimited queuing and boarding capacity), which attract a sufficient number of taxis waiting for customers because of high demand and high expected profitability at those spot locations.

Fig. 6 shows the change in average taxi waiting times in all four locations versus the value of the perceived profitability parameter, λ, with a fixed taxi fleet size of $N = 1000$ (veh). It can be observed that the average taxi waiting time at location 1 (airport) increases rapidly, but the average taxi waiting times in other zones including the railway station exhibit only very minor declines as the perceived profitability parameter increases, i.e., more taxis are attracted to the airport to wait for customers who want to go to the residential and commercial areas and the railway station. This can be explained by taking the airport and railway station as examples. Both the airport and railway station here are point meeting locations, but they exhibit quite different average taxi waiting times as a function of the perceived profitability parameter. This stems from the fact that the expected net revenue of a taxi ride from the airport and the railway station is quite different (the airport is located in a distant rural area but the railway station is located in the urban centre). Now, we consider an approximately deterministic choice of taxi drivers for meeting locations to maximize their expected net utility (this corresponds to a sufficiently large value of parameter θ^{t} in the logit choice model (8) and indeed the current value of $\theta^{t} = 10.0(1/\mathrm{h})$ is quite large). In this case, the expected net utility moving from any location, $j \in J$, to wait and meet a customer at the airport ($k = 1$) and railway station ($k = 2$) should be identical at market equilibrium. From Eq. (6), we thus have $\lambda(\tilde{F}_1 - \kappa\tilde{h}_1)/\kappa - (h_{j1} + w_1^{t}) = \lambda(\tilde{F}_2 - \kappa\tilde{h}_2)/\kappa - (h_{j2} + w_2^{t})$. As the inter-zonal taxi moving time, h, is constant, taxi drivers trade off the expected net revenue ($\tilde{F}$) of a taxi ride with the waiting time (w^{t}) that is required for such a ride, resulting in identical net utility received anywhere in the market. Because the expected net revenue of a taxi ride from the airport is much higher than that from the railway station, taxi drivers are willing to wait longer at the airport than at the railway station. In particular, when taxi drivers attach more importance (larger λ value) to the expected net revenue of a taxi ride, they are willing to wait even longer at the airport. Indeed, in Hong Kong, a considerable number of taxis are drawn to the airport to wait for customers for a number of hours (about 4 h). A similar phenomenon is observed at many other airports, such as Dallas and San Jose, with drivers waiting 2–3 h for their next customers (Schaller, 2007). The over-supply of taxis at the airport results in a substantial waste of valuable taxi service hours. To divert excess taxi supply from the airport to other areas and thus to enhance the overall taxi service quality, the Hong Kong SAR government has recently enacted, after consultation with the public, a new nonlinear taxi fare structure, in which long-haul fares are lowered but short-haul fares are raised (HKSAR Government, 2008).

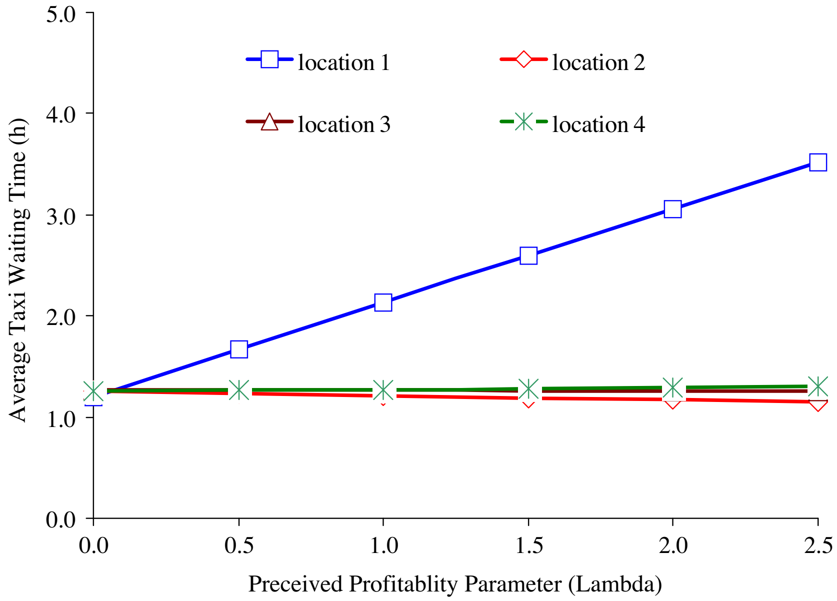

Fig. 6. Average taxi waiting times in different locations versus the perceived profitability parameter (Lambda) for a taxi fleet size of $N = 1000$ (veh).

Finally, Fig. 7 displays the proportion or probability of potential customers who actually patronize taxi services against taxi fleet size ($\lambda = 1.0$). At locations 1 and 2, the probability of potential customers taking taxis remains constant, regardless of the taxi fleet size. This is due to the fact that at these point meeting locations, the customer-waiting time is already nearly zero, and thus an increase in the taxi fleet size does not further reduce the customer-waiting time and hence there is no reduction in the cost of taking a taxi from the airport and the railway station. The proportion of customers using taxi services increases as the taxi fleet size increases at locations 3 and 4, because of the reduced customer-waiting times or improved service quality in the residential and commercial areas. The increase in demand is marginal largely because most potential customers are already exploited through provision of better services by a sufficiently large taxi fleet size. Indeed, in our example the taxi fleet size (700–1200 vehicles) is set a little large, resulting in somewhat unrealistically large taxi waiting times in the commercial and residential zones, as depicted in Fig. 4.

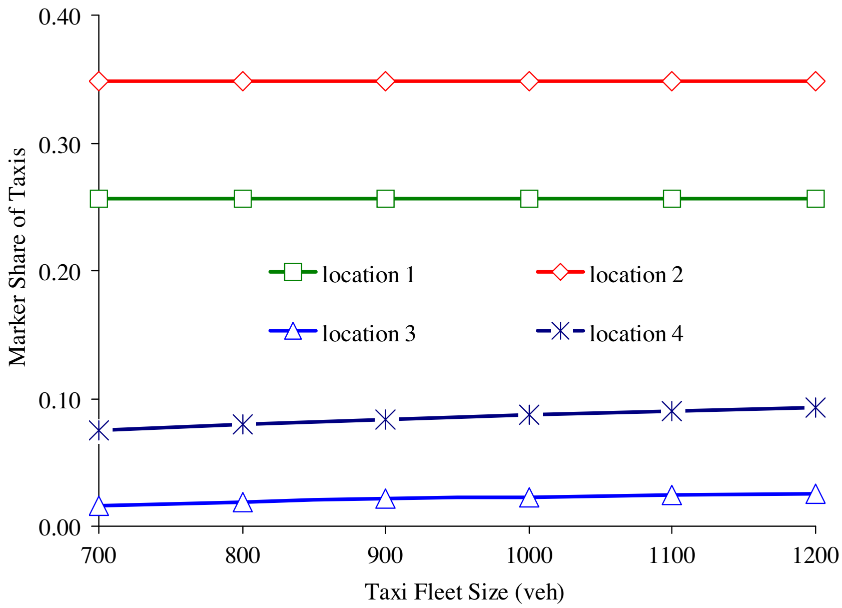

Fig. 7. Proportion of customers patronizing taxi services versus the taxi fleet size.

5. Conclusion and further research

We provided an equilibrium model for the problem of the bilateral searching and meeting between taxis and customers in a general network. A general meeting function is proposed that can capture the different natures of meetings in spatially distinct locations and includes the waiting time functions widely used in previous studies as a special case. An iterative numerical procedure is developed to obtain a stationary equilibrium solution at which a time-invariant distribution of taxis and customers across locations is established such that, given this distribution, taxi drivers maximize their net utility by optimally choosing where to locate themselves and customers minimize their full trip prices by optimally choosing where to ride a taxi. Our numerical results, although hypothetical and limited, show that the proposed equilibrium model does work well as expected to capture the bilateral search behaviors of taxi drivers and customers in a spatial context.

With the established bilateral searching and meeting functions, the network model could be usefully extended in various ways. First, our model is limited to cruising taxis only. It would be meaningful to consider dispatching taxis as well and look into their interactions such as through competition. In this case, one would expect endogenous determination of the customer demand and taxi supply for the two modes of operation. Second, the meeting function captures the distinct spatial feature of individual point (taxi stand) and zonal meeting locations by spelling out their searching and meeting frictions in the market. A practical application in policy analysis would be the determination of the number and locations of taxi stands to enhance overall market meeting efficiency, which can be formulated as a network design problem using a bi-level programming approach (Yang and Bell, 1998). Third, to have a clear and concise exposition of bilateral taxi and customer searching and meeting, we have simply assumed a constant speed of taxi movements in the networks. This assumption can be relaxed straightforwardly by incorporating the effect of traffic congestion into the network model of taxi movements, together with normal vehicular traffic (Wong et al., 2001). By doing so, we can investigate what regulatory reforms may benefit taxi drivers and customers simultaneously in the presence of congestion externality (Yang et al., 2005a). Fourth, with the bilateral searching and meeting function, we can explore the occurrence of Pareto-improvement in both service quality and market profitability (Manski and Wright, 1976; Schroeter, 1983; Yang et al., 2005b) theoretically and empirically, which is common to many queuing systems and was initially derived for scheduled transit systems (Mohring, 1972). Specifically, within some regions, simultaneous increases in the number of taxis and arrival rates of customers can result in both a reduction in expected customer-waiting times and an increase in taxi utilization rates and the resulting average net revenue per taxi. Clearly, when the customer arrival rate is endogenous and given by a demand function, occurrence of this win–win or Pareto-improving phenomenon is related to the current customer demand elasticity and the taxi fleet size in service, as well as the returns to scale in the meeting function.

Acknowledgements

The research described in this paper was supported by a Grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. RGC HKUST 6212/07E). The authors wish to express their thanks to an enthusiastic anonymous reviewer, whose constructive and insightful comments have significantly improved the exposition of the paper. Thanks also go to K.I. Wong for his earlier assistance with the computer code used in the computation.

References

Arnott, R., 1996. Taxi travel should be subsidized. Journal of Urban Economics 40 (3), 316–333.
Arrow, K.J., Debreu, G., 1954. Existence of an equilibrium for a competitive economy. Econometrica 22 (2), 265–290.
Bazaraa, M.S., Sherali, H.D., Shetty, C.M., 1993. Nonlinear Programming: Theory and Algorithms. John Wiley & Sons, Inc., New York.
Beesley, M.E., Glaister, S., 1983. Information for regulation: the case of taxis. The Economic Journal 93 (370), 594–615.
Burdett, K., Shi, S., Wright, R., 2001. Pricing and matching with frictions. Journal of Political Economy 109 (5), 1060–1085.
Cairns, R.D., Liston-Heyes, C., 1996. Competition and regulation in the taxi industry. Journal of Public Economics 59 (1), 1–15.
De vany, A.S., 1975. Capacity utilization under alternative regulatory constraints: an analysis of taxi markets. Journal of Political Economy 83 (1), 83–94.
Douglas, G.W., 1972. Price regulation and optimal service standards: the taxicab industry. Journal of Transport Economics and Policy 6 (2), 116–127.
Ellickson, B., 1993. Competitive Equilibrium: Theory and Applications. Cambridge University Press, Cambridge.
Erlander, S., Stewart, N.F., 1990. The Gravity Model in Transportation Analysis: Theory and Extensions. VSP, Utrecht, The Netherlands.
Fernandez, J.E., De Cea, J., Briones, J., 2006. A diagrammatic analysis of the market for cruising taxis. Transportation Research Part E 42 (6), 498–526.
Flores-Guri, D., 2003. An economic analysis of regulated taxicab markets. Review of Industrial Organization 23 (3–4), 255–266.
Fuente, Angel de La, 2000. Mathematical Methods and Models for Economists. Cambridge University Press, Cambridge.
Hackner, J., Nyberg, S., 1995. Deregulating taxi services: a word of caution. Journal of Transport Economics and Policy 29 (2), 195–207.
HKSAR Government, 2008. Report on Review of Taxi Operation. Transportation Advisory Committee, HKSAR Government, Hong Kong.
Lagos, R., 2000. An alternative approach to search frictions. Journal of Political Economy 108 (51), 851–873.
Lagos, R., 2003. An analysis of the market for taxicab rides in New York City. International Economic Review 44 (2), 423–434.
Manski, C.F., Wright, J.D., 1976. Nature of equilibrium in the market for taxi services. Transportation Research Record 619, 296–306.
Matsushima, K., Kobayashi, K., 2006. Endogenous market formation with matching externality: an implication for taxi spot markets. In: Kobayashi, K., Lakshmanan, T.R., Anderson, W.P. (Eds.), Structural Change in Transportation and Communications in the Knowledge Economy. Edward Elgar, pp. 313–336.
Mohring, H., 1972. Optimization and scale economies in urban bus transportation. American Economic Review 62 (4), 591–604.
Moore, A.T., Balaker, T., 2006. Do economists reach a conclusion on taxi deregulation? Econ Journal Watch 3 (1), 109–132.
Mortensen, D., Pissarides, C.A., 1994. Job creation and job destruction in the theory of unemployment. Review of Economic Studies 61 (3), 397–415.
Petrongolo, B., Pissarides, C.A., 2001. Looking into the black box: a survey of the matching function. Journal of Economic Literature 39 (2), 390–431.
Schaller, B., 2007. Entry controls in taxi regulation: implications of US and Canadian experience for taxi regulation and deregulation. Transport Policy 14 (6), 490–506.

Schroeter, J.R., 1983. A model of taxi service under fare structure and fleet size regulation. Bell Journal of Economics 14 (1), 81–96.
Starr, R.M., 1997. General Equilibrium Theory: An Introduction. Cambridge University Press, Cambridge.
Varian, H.R., 1992. Microeconomic Analysis, third ed. W.W. Norton & Company, Inc., New York.
Wong, S.C., Yang, H., 1998. A network equilibrium model of urban taxi services: an improved algorithm. Transportation Research Record 1623, 27–30.
Wong, K.I., Wong, S.C., Yang, H., 2001. Modeling urban taxi services in congested road networks with elastic demand. Transportation Research Part B 35 (9), 819–842.
Wong, K.I., Wong, S.C., Yang, H., Tong, S.C., 2003. The effect of perceived profitability on the level of taxi service in remote areas. Journal of the Eastern Asia Society for Transportation Studies 5, 79–94.
Wong, K.I., Wong, S.C., Bell, M.G.H., Yang, H., 2005. Modeling the bilateral micro-searching behavior for urban taxi services using the absorbing Markov chain approach. Journal of Advanced Transportation 39 (1), 81–104.
Wong, K.I., Wong, S.C., Yang, H., Wu, J.H., 2008. Modeling urban taxi services with multiple user classes and vehicle types. Transportation Research Part B 42 (10), 985–1007.
Yang, H., Bell, M.G.H., 1998. Models and algorithms for road network design: a review and some new developments. Transport Review 18 (3), 257–278.
Yang, H., Wong, S.C., 1998. A network model of urban taxi services. Transportation Research Part B 32 (4), 235–246.
Yang, H., Wong, S.C., Wong, K.I., 2002. Demand-supply equilibrium of taxi services in a network under competition and regulation. Transportation Research Part B 36 (9), 799–819.
Yang, H., Ye, M., Tang, W.H., Wong, S.C., 2005a. Regulating taxi services in the presence of congestion externalities. Transportation Research Part A 39 (1), 17–40.
Yang, H., Ye, M., Tang, W.H., Wong, S.C., 2005b. A multi-period dynamic model of taxi services with endogenous service intensity. Operations Research 53 (3), 501–515.