21st International Symposium on Transportation and Traffic Theory

A two-stage approach to modeling vacant taxi movements

R.C.P. Wong, W.Y. Szeto*, S.C. Wong

Department of Civil Engineering, The University of Hong Kong, Hong Kong

Abstract

In this paper, a two-stage modeling approach is proposed to predict vacant taxi movements in searching for customers. The taxi movement problem is formulated into a two-stage model that consists of two sub-models, namely the first and second stage sub-models. The first stage sub-model estimates the zone choice of vacant taxi drivers for customer-search and the second stage sub-model determines the circulation time and distance of vacant taxi drivers in each zone by capturing their local customer-search decisions in a cell-based network within the zone chosen in the first stage sub-model. These two sub-models are designed to influence each other, and hence an iterative solution procedure is introduced to solve for a convergent solution. The modeling concept, advantages, and applications are illustrated by the global positioning system data of 460 Hong Kong urban taxis. The results demonstrate that the proposed model formulation offers a great improvement in terms of root mean square error as compared with the existing taxi customer-search models, and show the model capabilities of predicting the changes in vacant taxi trip distributions with respect to the variations in the fleet size and fare. Potential taxi policies are investigated and discussed according to the findings to provide insights in managing the Hong Kong taxi market.

Keywords: Two-stage approach; Taxi customer-search behavior; Enhanced sequential logit model; Logit-opportunity model; Global positioning system data

1. Introduction

In many cities, taxis always circulate in search of customers, but this circulation consumes much road space. Especially, when many vacant taxis are concentrating at central business districts (CBDs) in customer-search, the local traffic congestion and air pollution problems are deteriorated. For better controlling the vacant taxi movements, numerous studies have been focused on examining the consequences on implementation of taxi regulatory policies in forms of entry restrictions and price controls (e.g., Douglas, 1972; De Vany, 1975; Manski and Wright, 1976; Beesley and Glaister, 1983; Schroeter, 1983; Arnott, 1996; Cairns and Liston-Heyes, 1996; Xu et al., 1999; Yang et al., 2000, 2002, 2005a, 2010a; Flores-Guri, 2003; Fernández et al., 2006; Moore and Balaker, 2006; and Loo et al.,

* Corresponding author. Tel.: +852-2857-8552; fax: +852-2559-5337.
E-mail address: ceszeto@hku.hk

2007). However, these studies have been developed based on an idealized market of conventional economic analysis in which the spatial structure of the market was ignored.

In an attempt to capture the spatial structure of the market, Yang and Wong (1998) developed a model to determine the taxi movements on a given road network. In this model, the customer-search behavior of taxi services is based on the assumption that each vacant taxi driver attempts to minimize his/her expected search time to find a customer. Later, the model was further improved by Wong and Yang (1998), Wong et al. (2001, 2002, 2005, 2008), Yang et al. (2001, 2005b, 2008, 2012), Yang and Yang (2011), Kim et al. (2005), and Hu et al. (2012) to capture congestion effects, multiple user classes, multiple taxi modes, customer hierarchical modal choice, taxi search behavior of taxi customers, day-to-day learning processes, stochastic travel time, and search frictions between vacant taxis and taxi customers. Yet, the assumption has ignored the case that vacant drivers travel to remote areas to pick up customers because of the high profit return.

Taken into account the profitability consideration, Wong et al. (2003) developed a taxi network model to explicit examined the effects of perceived profitability on the customer-search behavior of vacant taxis and the expected profit that taxi drivers could earn by picking up customers in particular zones. Yang et al. (2010b) further extended this concept to "profit per unit time" by incorporating the operational cost and time consumption involved in taxi trips to maximize profits from customer-search. However, the search behavior models in the preceding taxi network models are in logit form and have not been calibrated and validated.

Until recently, Sirisoma et al. (2010) provided an empirical evidence to validate the search behavior models. They determined the significant factors that affect vacant taxi drivers' customer-search decisions based on the data collected from a stated preference survey. Wong et al. (2014a) calibrated and validated multinomial logit (MNL) models based on their global positioning system (GPS) data obtained from urban taxis to predict the drivers' strategic zone choices for searching for customers in both peak and off-peak periods. Szeto et al. (2013) further extended the consideration to every hour in a day. Wong et al. (2015a) considered the sequential customer-search decisions of vacant taxi drivers on finding customers at intermediate zones while heading to their designated zones, and proposed an enhanced sequential logit (ESL) model to additionally predict search paths as compared with the model of Wong et al. (2014a). However, the preceding behavior models often predicted the strategic zonal decisions of vacant taxi drivers for customer-search and ignores the local (within zone) search behavior that as vacant taxi drivers search for customers, the probability of successfully meeting a taxi customer along the way increases. In addition, the probability of a vacant taxi driver meeting a taxi customer decreases if there are many vacant taxis nearby. This important fact has not yet been considered. Therefore, these models cannot be used for developing simulation-based models and simulation-based optimization models for depicting and managing taxi flows on the local streets.

For modeling the local search behavior of vacant taxi drivers, Hu et al. (2012) recently proposed a probabilistic dynamic programming vacant taxi routing model to depict the routing decisions of vacant taxi drivers at intersections according to the passenger arrival rate. Wong et al. (2014b) formulated a cell-based network and modeled the local search behavior of vacant taxi drivers based on the probability of successfully meeting the next taxi customers. Nevertheless, these studies assumed that vacant taxi drivers seek to minimize their expected time spent searching for their next customer nearby or to maximize the probability of successfully meeting a taxi customer. This assumption is realistic when modeling the search behavior in urban areas, where the length of customer trips may not significantly vary from zone to zone, and where taxi drivers may not have much information about the profitability of going to particular zones to search for customers. However, this assumption may not reflect the real situation in more remote areas with considerable potential demand (e.g., an airport or satellite town). Even though the search time to remote areas is longer than that at urban areas, numerous vacant taxis still go there to wait for customers because the drivers expect to earn more by doing so.

A two-stage taxi customer-search model is therefore proposed in this study to predict vacant taxi movements to address the above issues by integrating and formulating the zonal and local search problems of vacant taxi drivers. In fact, the integrated transportation modeling concept has been widely adopted for formulating multi-stage transportation problems to ensure that the obtained outcomes from different stages are consistent with each other. In urban taxi modeling, Wong et al. (2001) attempted to calibrate an integrated model for urban taxi services in a congested road network with elastic passenger demand. Later, Wong et al. (2008) further extended the model for multiple user classes, multiple taxi modes, and customer hierarchical modal choice. Their efforts of integrating different stages of transportation model into a single and simultaneous mathematical model gave inspirations to this study.

In our proposed two-stage model, the first stage model is formulated as an ESL model to determine the adjacent zone choice of vacant taxi drivers in customer-search. The choice is modeled as functions of the relative passenger demand at all candidate zones, cross-zonal travel distance between all the current and the candidate zones, circulation distance within each candidate zone, and the rate of return between all the current and the candidate zones. The second stage model is used to depict the local customer-search movements of vacant taxi drivers within the chosen zone in which a zone is divided into many equally sized square cells. The concept of logit-opportunity model is employed. The decision variable of the second stage model is the probability of successfully being hired in each cell, which is influenced by the outputs of the first stage model (because more trips distributed to a zone enhances the vacant taxi supply there, and consequently reduces the probability of success to pick up a customer). The outputs of the second stage model are the estimated circulation distance and time for finding customers in each zone, which are part of the attributes of the first stage model. Hence, these two sub-models influence each other. To solve for convergent solutions, an iterative solution procedure is introduced. This procedure introduces a feedback mechanism into the solution procedure that repeatedly updates the inputs of the upper stage (which are the outputs of the lower stage) by a certain number of iterations as in the literature (e.g. Florian et al., 1975; Evan, 1976).

To calibrate our proposed model, the GPS data of 460 Hong Kong urban taxis are adopted. A root mean square error (RMSE) evaluation is conducted in this study to measure the goodness of fit of the proposed model with respect to observed data, and to compare with that of the existing taxi customer-search models. Sensitivity analyses are also accomplished to illustrate the changes in the proportion of vacant taxi trips ended in the CBDs in customer-search with respect to the variations in fleet size and fare. Potential taxi policies are investigated and discussed according to the findings.

To sum up, the contributions of this paper include the following:

- Proposing a two-stage approach to modeling vacant taxi movements. This approach is capable of capturing the consequential effects of the zone choices of taxi drivers on the explanatory variables in the ESL model. The model structure is more suitable to describe the actual search options to taxi drivers to search for their customers;
- Developing an iterative solution procedure to solve the two-stage model;
- Demonstrating the model capabilities of predicting the consequences of various taxi policies, and explaining the effects of fleet size and fare on the proportion of vacant taxi trips ended in the CBDs for customer-search; and
- Allowing extensions on developing simulation-based models and simulation-based optimization models for depicting and managing taxi flow on local streets.

The remainder of this paper proceeds as follows. Section 2 describes the proposed two-stage model and iterative solution procedure. Section 3 depicts the data and network for model calibration, sensitivity analyses, and case studies. Section 4 discusses the results and illustrates the performance of the proposed solution procedure. Lastly, Section 5 concludes the paper and suggests a future study direction.

2. Methodology

The proposed two-stage model comprises into two sub-models: (1) the first stage zone choice model; and (2) the second stage local customer-search model.

2.1. First stage zone choice model

The first stage model is developed through an ESL modeling approach. The study area was divided into several zones. Each zone was connected to its adjacent zones which include both (1) the zones that can be reached directly from it via for example expressways or highways without passing the urban roads of other zones and (2) neighbor zones that are next to the current zone.

In the first stage model, a vacant taxi driver in each zone decides whether to search for customers (1) in the current zone or (2) in one of the adjacent zones, and this decision is modeled by the MNL model. If he/she cannot meet a customer in the zone from his/her first zonal search decision, he/she may make another decision in this zone based on the available search options there. The model assumes that all of the zonal customer-search decisions are made independently without depending on previous and subsequent decisions, and that the decisions are made sequentially. This process is repeated until he/she meets a customer. It allows us to trace the zones to be passed through before a customer is reached, or the search paths (in terms of a sequence of zones visited) taken to reach the customer. The zonal decisions were found to be influenced by the following four factors (Wong et al., 2015a).

Factor 1 – Relative Passenger Demand: The relative passenger demand (E_j) in zone $j \in I$ is defined as the number of customers (O_j) being picked up in that zone compared with the total number of customers being picked up in the entire study area, where I represents the set of zones in the study area, and can be mathematically represented by

$$E_j = \frac{O_j}{\sum_{m \in I} O_m}, \forall j \in I \text{ .} \tag{1}$$

This definition assumes that the pickup rate is the same for each zone. In that case, even though the pickup rate is not 100%, the relative passenger demand equals the relative *actual* passenger demand, because the *actual* passenger demand multiplied by the pickup rate gives the number of customers picked by taxis, and the pickup rate in each zone is cancelled out when we calculate the relative *actual* passenger demand.

Factor 2 – Intra-zonal Circulation Distance: It represents how far a vacant taxi driver travels in a chosen zone to reach the next customer on average. It is indirectly related to passenger demand. The intra-zonal circulation distance (D_j^{c}) in zone $j \in I$ is defined as the mean travel distance of all vacant trips for searching for a customer within the boundary of the zone.

Factor 3 – Cross-zonal Travel Distance: D_{ij}^{t} denotes the travel distance from zone $i \in I$ to zone $j \in I$ for searching for customers. When $i \neq j$, D_{ij}^{t} equals the average travel distance of all vacant trips from the starting point in zone $i \in I$ until each of the vacant taxis reaches the boundary of their chosen zone $j \in I$. When the vacant trips do not involve any cross-zonal travel (i.e., $i = j$), the travel distance equals zero.

Factor 4 – Rate of Return: The rate of return (R_{ij}), also known as profit per unit time, is calculated based on a search cycle that consists of the vacant taxi trip from the destination zone $i \in I$ of the previous occupied trip to zone $j \in I$ for picking up a customer, and the subsequent occupied trip from zone $j \in I$ for transporting the customer to his/her destination. The rate of return equals the mean profit (i.e., the difference between the expected in-pocket profit from an occupied trip and the total operational cost) obtained from the search cycle over the expected time taken to obtain that profit. Mathematically, the rate of return for a vacant taxi traveling from zone $i \in I$ to zone $j \in I$ to search for the next customer can be expressed by

$$R_{ij} = \frac{-C_{ij}^{\mathrm{t}} - C_j^{\mathrm{c}} + \overline{P}_j}{T_{ij}^{\mathrm{t}} + T_j^{\mathrm{c}} + \overline{T}_j^{\mathrm{t}}}, \ \forall i, j \in I \text{ .} \tag{2}$$

T_{ij}^{t} and T_j^{c} in Equation (2) are the associated time spent in cross-zonal travel from zone $i \in I$ to zone $j \in I$ and intra-zonal circulation within zone $j \in I$, respectively. The distances and time are, correspondingly, multiplied by the fuel cost per unit of distance (C^{f}) and the rental cost per minute (C^{r}), and their resultant products are added up to obtain the operational costs of cross-zonal travel (C_{ij}^{t}) from zone $i \in I$ to zone $j \in I$ and intra-zonal circulation within zone *j* (C_j^{c}):

$$C_{ij}^{\mathrm{t}} = C^{\mathrm{f}} D_{ij}^{\mathrm{t}} + C^{\mathrm{r}} T_{ij}^{\mathrm{t}} + \tau_{ij} \text{ and } C_j^{\mathrm{c}} = C^{\mathrm{f}} D_j^{\mathrm{c}} + C^{\mathrm{r}} T_j^{\mathrm{c}}, \ \forall i, j \in I \text{ ,} \tag{3}$$

where τ_{ij} denotes the toll charge (if any) associated with the trip from zone $i \in I$ to zone $j \in I$.

$\overline{P}_j$ in Equation (2) is the expected in-pocket profit from an occupied trip starting in zone $j \in I$ and equals the average taxi fare paid by passengers originated from that zone minus the corresponding average operational cost of the occupied trips. The expected profit can be formulated as

$$\bar{P}_j = \bar{F}_j - C^{\mathrm{f}}\bar{D}_j^{\mathrm{t}} - C^{\mathrm{r}}\bar{T}_j^{\mathrm{t}} - \bar{\tau}_j, \ \forall i, j \in I, \tag{4}$$

where $\bar{D}_j^{\mathrm{t}}$, $\bar{T}_j^{\mathrm{t}}$, and $\bar{\tau}_j$ denote the average travel distance and time and the toll charge associated with all occupied trips starting from zone $j \in I$, respectively. $\bar{F}_j$ denotes the average taxi fare paid by the taxi customers of all occupied trips starting in zone $j \in I$, which can be estimated by multiplying the occupied travel distance and waiting time (i.e., duration of travel speed less than 5 km/h) by the corresponding unit charge rate in a taxi fare structure.

Formulation – First Stage Zone Choice Model: It takes the following MNL form (McFadden, 1974):

$$P_q(j \mid i) = \begin{cases} 0 & , \text{if } j \notin H_i \cup \{i\}; \\ \dfrac{\exp\left(\beta^{\mathrm{E}} E_j + \beta^{\mathrm{t}} D_{ij}^{\mathrm{t}} + \beta^{\mathrm{c}} D_j^{\mathrm{c}} + \beta^{\mathrm{R}} R_{ij}\right)}{\sum_{n \in H_i \cup \{i\}} \exp\left(\beta^{\mathrm{E}} E_n + \beta^{\mathrm{t}} D_{in}^{\mathrm{t}} + \beta^{\mathrm{c}} D_n^{\mathrm{c}} + \beta^{\mathrm{R}} R_{in}\right)} & , \text{if } j \in H_i \cup \{i\}, \end{cases} \ \forall i, j \in I, q \in Q', \tag{5}$$

where $P_q(j \mid i)$ is the probability of an individual vacant taxi driver making a single zonal search decision on trip $q \in Q'$ in zone $i \in I$ and selecting zone $j \in I$ to search for the next customer and Q' is the set of trips considered. H_i denotes the set of zones adjacent to zone $i \in I$. β^{E} is the coefficient associated with the relative passenger demand. β^{t} is the coefficient associated with the cross-zonal travel distance. β^{c} is the coefficient associated with the intra-zonal circulation distance of the customer-search and β^{R} is the coefficient associated with the rate of return for the search cycle. To ensure that the selected zone is either adjacent to the starting zone or the starting zone itself, the recursive Equation (5) sets a criterion that if the selected zone is not the starting zone and not adjacent to the starting zone, then the probability that the trip will travel to the selected zone from the starting zone is zero. Otherwise, the probability is determined based upon the MNL model with the alternatives comprising the starting zone and its adjacent zones.

It is important to clarify that the above attributes are indeed the perceived values of a vacant taxi driver before he/she makes the decision to conduct a customer-search in a particular zone. We assume that every taxi driver has the same perception concerning the attributes. Therefore, for the sake of simplicity, the subscript *q* is omitted in each attribute.

According to the ESL model structure, vacant taxi drivers may make more than one zonal search decision to travel across zones for customer-search. Therefore, to predict the how trips generated from a zone is distributed to other zones (referred to as trip distribution) as well as the number of vacant taxi trips ended in each zone for customer-search, it is necessary to calculate the *cumulative* probability of a vacant taxi driver in zone $i \in I$ eventually selecting zone $j \in I$ (not necessarily to be an adjacent zone of zone $i \in I$ if $M > 1$) after making zonal search decisions *M* times, which can be expressed as

$$\hat{P}^M(j \mid i) = \begin{cases} P_q(j \mid i) & , \text{if } M = 1; \\ \sum_{k \in H_i \cup \{i\}} \hat{P}^{M-1}(k \mid i) P_q(j \mid k) & , \text{if } M > 1, \end{cases} \ \forall i, j \in I. \tag{6}$$

When the number of zonal search decisions (M) equals 1, we have $\hat{P}^1(j \mid i) = P_q(j \mid i), \forall i, j \in I$. This implies that the vacant taxi trip distribution is obtained simply using an MNL model.

When $M = 2$, we have $\hat{P}^1(j|i) = P_q(j|i), \forall i, j \in I$ and $\hat{P}^2(j|i) = \sum_{k \in H_j \cup \{i\}} \hat{P}^1(k|i) P_q(j|k), \forall i, j \in I$.

That is, when the number of zonal search decisions (M) equals 2, $\hat{P}^1(j|i)$ is obtained as in the case of M = 1, and the probability $\hat{P}^2(j|i)$ equals the probability of vacant taxi drivers moving from zone $i \in I$ to an intermediate zone $k \in H_i \cup \{i\}$ times the probability of the subsequent decision to travel to zone $j \in H_k \cup \{k\}$. (Note that zone k must be either adjacent to zone $i \in I$, or zone $j \in I$, or both of them.) Because there can be more than one possible search path linking zone $i \in I$ and zone $j \in I$, it is necessary to aggregate the probability associated with each search path.

Based on the cumulative probabilities defined by Equation (6), the total number of predicted vacant trips ending in zone $j \in I$ after making M decisions Q_j^M can be calculated as follows:

$$Q_j^M = \sum_{m \in I} \left[G_m \hat{P}^M(j|m) \right], \forall j \in I, \tag{7}$$

where G_m is the total number of vacant trips started (or equivalently, the total number of occupied taxis dropping off their customers) in zone $m \in I$. According to Equation (7), the total number of predicted vacant trips ending in zone $j \in I$ equals the sum of the product of the total number of trips generated from each zone times the corresponding probability that the trips end in zone $j \in I$.

2.2. Second stage zone choice model

The second stage model is a cell-based logit-opportunity model proposed by Wong et al. (2014b) and aims to 1) estimate the required (average) circulation distance and time for finding reaching the next customer for the first stage model/problem taking into account the number of vacant taxis in that zone and 2) improve the goodness of fit of the first stage model by predicting local customer-search movements. The model combines both the modeling principles of the logit-based search model and intervening opportunity model. The logit-based modeling concept is used to handle multi-directional, discrete search choices for vacant taxi drivers whereas the intervening opportunity model is reformulated and incorporated into the logit-based search model to capture the cumulative probability of successfully meeting a customer in each direction of customer-search. The second stage model assumes that vacant taxi drivers do not have clear destinations in customer-search after they drop off their preceding customers, but they travel towards an area within the zone concerned with a high probability of meeting a customer. It also assumes that the drivers continue to search for the customer in that zone until meeting the next one.

Each zone $j \in I$ was divided into many equally sized square cells to form a cell-based network. Each cell was connected to surrounding cells in at most four cardinal directions. When an urban taxi driver circulates in a zone locally for searching for new customers, he/she has two categories of choices: (1) searching for a customer in the current cell within the zone; and (2) traveling towards one of the adjacent cells in at most four cardinal directions.

Similar to the modeling concept of the first stage ESL model, a vacant taxi driver may not meet a customer in the cell selected once he/she reaches that cell. If so, he/she makes another search decision in this cell based on the available search options at that cell. The search decision made in a cell is assumed to be independent of the decision made previously. This process is repeated until he/she meets a customer. This process implies that the driver circulates in the cell-based network cell by cell to search for customers and the driver is assumed to make independent decisions sequentially to select the subsequent cells during his/her cruising within the cell-based network, with one decision made in each cell.

Experienced taxi drivers indeed would know an approximately average distance and time required to travel to meet their next customers. For this reason, they would not simply make their local customer-search decisions depending on the attractiveness (i.e., the probability of successfully find a customer) of the adjacent cells, but would also consider the subsequent possible choice of cells that could be reached later. The *cumulative* probability of success is hence introduced to represent the accumulated attractiveness to vacant taxi drivers searching on the way, and it was found to be the factor that affects the local movements of vacant taxi drivers in search of customers (Wong et al., 2014b).

Factor – Cumulative Probability of Success: The cumulative probability of success (S_y^l) represents the accumulated probability of a vacant taxi driver successfully picking up a customer, if the driver initially selects cell y and prepares to make local search decisions *l* times to meet a customer. The cumulative probability equals the probability of successfully picking up a customer in the first cell visited plus the additional probability gained from the subsequent cells. This can be expressed as

$$S_y^l = \begin{cases} s_y & , \text{if } l = 1; \\ s_y + \left(1 - s_y\right) \sum\limits_{m \in Z_y \cup \{y\}} \left[p^{l-1}(m \mid y) S_m^{l-1} \right] & , \text{if } 1 < l \le L, \end{cases} \quad \forall y \in X_j,\ j \in I, \tag{8}$$

where *L* is the number of local search decisions anticipated by a vacant taxi driver and its value should be calibrated. X_j is the set of cells within zone $j \in I$. Z_y denotes the set of adjacent cells that can be selected by vacant taxi drivers in cell $y \in X_j$. s_y is the probability of a vacant taxi driver to successfully pick up a customer in cell $y \in X_j$, and is referred to the probability of success of the driver in cell $y \in X_j$. Note that the superscript *j* is omitted for all the variables in the second state model for simplicity.

The probability $p^{l-1}(m \mid y)$ in Equation (8) is the probability that an individual vacant taxi driver in cell $y \in X_j$ selects subsequent cell $m \in Z_y \cup \{y\}$ in the ($L - (l-1) + 1$)-th *planned decision* to search for the next customer:

$$p^{l-1}(m \mid y) = \frac{\exp\left(\alpha^L S_m^{l-1}\right)}{\sum\limits_{n \in Z_y \cup \{y\}} \exp\left(\alpha^L S_n^{l-1}\right)}, \ \forall m \in Z_y \cup \{y\}, y \in X_j,\ j \in I, \text{ and } \forall l \in \{2, \dots, L\}, \tag{9}$$

where α^L denotes the non-negative parameter and its value depends on the number of local search decisions anticipated *L* made by vacant taxi drivers.

In fact, the second and subsequent decisions are only the initial (or conceptual) plan of the taxi driver made at his/her current position. He/she may not necessarily follow the plan and confine his/her remaining search decisions to the subsequent cells in the plan. The driver re-evaluates his/her plan after reaching the next cell and his/her decision made in each cell is independent of the decision made in other cells. *In this sense, the local search decisions made in all cells are independent and are made sequentially.*

For the case of $L = 1$, not more than five cells (including the current cell $x \in X_j$ and the cells next to it in at most four cardinal directions) are considered by the driver in cell $x \in X_j$ in his/her local customer-search plan. The associated probability of success of the driver successfully picking up a customer in cell $y \in Z_x \cup \{x\} \subset X_j$ is simply equal to the probability of being hired at that cell according to Equation (8).

For the case of $L > 1$, the number of local search decisions anticipated is more than one. Vacant taxi drivers make their local search decisions based on both the probability of success in each of the cells considered in the first decision and the additional probabilities gained from the subsequent cells after leaving the cells considered in the first decision. When the driver is currently in cell $x \in X_j$, the first term s_y at the lower right side of recursive Equation (8) can be interpreted as the probability of a vacant taxi driver successfully picking up a customer in cell $y \in Z_x \cup \{x\}$. The probability $\left(1 - s_y\right)$ equals the probability of the vacant taxi driver failing to meet a customer in cell $y \in Z_x \cup \{x\}$ and leave the cell. The function $\sum_{m \in Z_y \cup \{y\}} \left[p^{l-1}(m \mid y) S_m^{l-1} \right]$ represents the cumulative probability of success of the driver after leaving cell $y \in Z_x \cup \{x\}$. Therefore, the second term $\left(1 - s_y\right) \sum_{m \in Z_y \cup \{y\}} \left[p^{l-1}(m \mid y) S_m^{l-1} \right]$ represents the additional probability of success gained from the subsequent cells.

Formulation – Second Stage Local Customer-search Model: It takes the following MNL form (McFadden, 1974):

$$V_w^L(y \mid x) = \frac{\exp\left(\alpha^L S_y^L\right)}{\sum_{m \in Z_x \cup \{x\}} \exp\left(\alpha^L S_m^L\right)}, \forall y \in Z_x \cup \{x\}, x \in X_j, j \in I, w \in W', \tag{10}$$

where $V_w^L(y \mid x)$ denotes the probability of an individual vacant taxi driver, who is currently in cell $x \in X_j$ and makes a local movement $w \in W'$ from cell $x \in X_j$ to cell $y \in Z_x \cup \{x\}$ to search for the next customer. W' denotes the set of all local movements. Similar to Equation (5), every taxi driver was assumed to have the same perception concerning the cumulative probabilities of success for the corresponding connected cells. Therefore, for the sake of simplicity, the subscript *w* is omitted in the attribute. The model form is identical to that of the selection probability of a vacant taxi driver any subsequent cell as indicated in Equation (9) and the model has the same calibrated parameter (α^L) for each cell and zone to maintain a consistent search decision everywhere in the network.

Based on Equation (10), the predicted number of local movements eventually choosing cell $y \in X_j$ can be estimated by

$$W_y = \sum_{m \in Z'_y \cup \{y\}} U_m V_w^L(y \mid m), \forall y \in X_j, j \in I, \tag{11}$$

where U_m denotes the number of local movements started from cell $m \in Z'_m \cup \{y\}$, and Z'_y denotes the set of adjacent cells in which the drivers in those cells can select cell $y \in X_j$ as the next cell to visit. The total number of local movements in zone $j \in I$ can then be expressed as

$$W_j^{\mathrm{T}} = \sum_{y \in X_j} W_y, \forall j \in I. \tag{12}$$

The estimated circulation distance and time in zone $j \in I$ for the next customers (D_j^{c} and T_j^{c}) can be found by multiplying the mean values ($\bar{D}_j^{\mathrm{c}}$ and $\bar{T}_j^{\mathrm{c}}$) in one cell by the expected (i.e., average) number of cells traveled in zone $j \in I$, ($\bar{C}_j$) to meet the next customer:

$$D_j^{\mathrm{c}} = \bar{D}_j^{\mathrm{c}} \bar{C}_j \text{ and } T_j^{\mathrm{c}} = \bar{T}_j^{\mathrm{c}} \bar{C}_j, \forall j \in I. \tag{13}$$

The latter term $\bar{C}_j$ is estimated by the sum of the product of the proportion of trips selecting cell $y \in X_j$ as the next cell, $\frac{W_y}{W_j^{\mathrm{T}}}$, and the corresponding expected number of cells traveled to reach a customer after reaching that cell, $\left[s_y + \sum_{l=1}^{\infty} (l+1)\left(\hat{S}_y^{l+1} - \hat{S}_y^{l}\right)\right]$:

$$\bar{C}_j = \sum_{y \in X_j} \frac{W_y}{W_j^{\mathrm{T}}} \left[s_y + \sum_{l=1}^{\infty} (l+1)\left(\hat{S}_y^{l+1} - \hat{S}_y^{l}\right)\right], \ \forall j \in I \ . \tag{14}$$

The average number of cells traveled in Equation (14) should be based on the *predicted decisions* made by taxi drivers rather than the plans in their minds, because the drivers may change their mind once reaching the next cell due to having more accurate information for the next cell or those cells nearby. That is why $p^{l-1}(m \mid y)$ is no longer applicable for this case. Therefore, $p^{l-1}(m \mid y)$ in Equation (8) should be replaced by $V_w^L(m \mid y)$ to compute the cumulative probability of success ($\hat{S}_y^l$) based on the predicted search decisions.

2.3. Two-stage modeling concept

The modeling concept of the proposed two-stage model can be summarized in the flowchart in Figure 1. The inputs of the first stage model are (1) zonal passenger demand; (2) the time and distance of cross-zonal travel and intra-zonal circulation for searching for customers; (3) average occupied travel time and distance for transporting passengers; (4) unit operational costs; (5) tolls; (6) fare structure; and (7) the observed origins and destinations of zonal vacant taxi movements, where those with an asterisk are the inputs that can be obtained from the GPS data. Based on these inputs, we calculate the four model attributes for the first stage model as explained in Section 2.1. The output of the first stage model is the zonal distribution of vacant taxis, and influences to the probability of success in each cell in each zone. Because a zone with more vacant taxis circulating has a lower probability of success meeting a customer there, s_y in cell $y \in X_j$ in zone $j \in I$ is assumed to be inversely proportional to Q_j^M obtained from the first stage model. Moreover, s_y is assumed to be proportional to the observed probability of success in cell $y \in X_j$, s_y^{O}, and the observed zonal distribution of vacant taxis to zone $j \in I$, Q_j^{O}. Therefore, we have

$$s_y = s_y^{\mathrm{O}} \left(\delta_j \frac{Q_j^{\mathrm{O}}}{Q_j^M} \right), \ \forall y \in X_j, \ j \in I \ , \tag{15}$$

where δ_j is the proportionality constant associated with zone $j \in I$ and $\delta_j \frac{Q_j^{\mathrm{O}}}{Q_j}$ can be interpreted as a balancing factor to ensure the models in the two stages to be consistent to each other.

The observed probability s_y^{O} in Equation (15) is defined as the passenger demand generated from cell $y \in X_j$ over the availability of vacant taxis in that cell, subject to the condition that the probability is between zero and one inclusively. Mathematically, it can be expressed as

$$s_y^{O} = \min\left[\frac{O_y}{A_y}, 1\right], \ \forall y \in X_j, j \in I, \tag{16}$$

where O_y denotes the number of occupied trips starting in cell $y \in X_j$ and A_y denotes the number of vacant taxis appeared in cell $y \in X_j$. This probability is cell specific to allow capturing the variation of the probability of successfully meeting a customer in different cells.

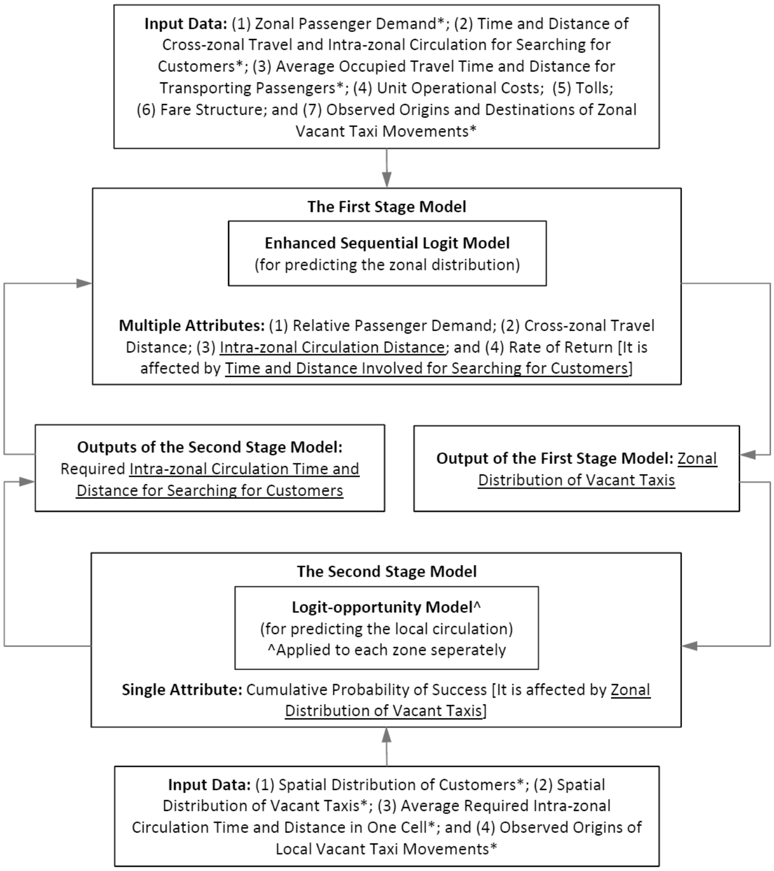

Fig. 1. Flowchart of the proposed two-stage taxi customer-search model

Equations (15) and (16) are used to calculate the cumulative probability of success in each cell. The cumulative probability of success in each cell is the unique attribute for the second stage model as stipulated in Section 2.2. Together with the inputs of (1) a spatial distribution of customers; (2) a spatial distribution of vacant taxis; (3) average required intra-zonal circulation time and distance in one cell; and (4) the observed origins of local vacant taxi movements, the second stage model estimates the required intra-zonal circulation time and distance for customer-search in each zone. They are part of the multiple attributes of the first stage model. Hence, these two sub-models influence each other.

2.4. Iterative solution procedure

Based on the preceding flowchart, we introduce an iterative solution procedure for the two-stage model to obtain a convergent solution of a vacant taxi trip distribution and also the balancing factor. The estimated circulation distance and time in each zone are determined in the second stage model to update the inputs to the first stage zonal distribution model. It is noted that a zone that requires a shorter distance and time finding the next customer attracts more vacant taxis going there, and hence reduces the overall probability of successfully being hired. Therefore, we propose applying an adjustment factor for zone $j \in I$ to the probability of success in each cell in that zone under different conditions to achieve an equilibrium zonal vacant taxi distribution and the required circulation distance and time. The adjustment factor can be expressed as follows:

$$\omega_j^{(k)} = \begin{cases} (1-\phi) & \text{, if } Q_j^{M\ (k)} - Q_j^{M\ (k-1)} > \varepsilon; \\ (1+\phi) & \text{, if } Q_j^{M\ (k-1)} - Q_j^{M\ (k)} > \varepsilon; \\ 1 & \text{, if } \left| Q_j^{M\ (k)} - Q_j^{M\ (k-1)} \right| \le \varepsilon, \end{cases} \quad \forall j \in I \text{ and } s_y^{(k)} = \omega_j^{(k)} s_y^{(k-1)}, \ \forall y \in X_j, j \in I, \tag{17}$$

where $0 < \phi < 1$, $\varepsilon > 0$. $\omega_j^{(k)}$ is the adjustment factor for zone $j \in I$ in iteration *k*. $Q_j^{M\ (k)}$ and $s_y^{(k)}$ are, respectively, the predicted number of trips ending in zone $j \in I$ during a customer-search and the corresponding estimated probability of success for cell $y \in X_j$ in iteration k. If the total number of trips assigned to a zone is close to those predicted in the preceding iteration $k-1$ within an acceptable tolerance ε, no adjustment is made to the probability of success in each cell in that zone. Otherwise, the reduction factor $(1-\phi)$ and the expansion factor $(1+\phi)$ are applied to the probability of success because in such cases, the total number of assigned trips is obviously higher and smaller than those predicted, respectively.

Recursive Equation (17) is used to update the probability of success in Equation (8) during the solution procedure until a convergent solution is obtained. It is important to clarify that the probability of success should always be controlled by the upper limit of 100% in order to prevent the violation of the implicit model condition of $0 \le S_y^l \le 1$.

The solution procedure can be summarized as follows:

Step 1 – Initialization and parameter setting

Initialize $s_y^{(0)}$, $Q_j^{M\ (0)}$, $D_j^{c\ (0)}$ and $T_j^{c\ (0)}$. Set ϕ and ε. Set the iteration number $k = 1$.

Step 2 – Compute the zonal vacant taxi distribution

Apply the zonal distribution model in Equations (1)-(7) based on the circulation distance $D_j^{c\ (k-1)}$ and time $T_j^{c\ (k-1)}$ in iteration $k-1$ to obtain an initial solution $Q_j^{M\ (k)}$ for all zones $j \in I$.

Step 3 – Convergence test

If $\max \left| Q_j^{M\ (k)} - Q_j^{M\ (k-1)} \right| \le \varepsilon$, then stop; otherwise, go to Step 4.

Step 4 – Adjust the probability of success

Adjust the probability of success $s_y^{(k)}$ according to Equations (15) and (16).

Step 5 – Update the circulation distance and time

Use the adjusted probability of success to compute the updated circulation distance $D_j^{c\ (k)}$ and time $T_j^{c\ (k)}$ using Equations (8)-(14). Then, set $k = k+1$ and go to Step 2.

After finishing the proposed solution procedure, the balancing factor for zone $j \in I$ introduced in Equation (15) can be calculated as the product of the reduction and expansion factors in all iterations:

$$\delta_j \frac{Q_j^{O}}{Q_j^{M}} = \prod_{k=1}^{k^{\max}} \omega_j^{(k)}, \forall j \in I, \tag{18}$$

where $k^{\max}$ denotes the number of iterations required to achieve a convergent solution of a zonal vacant taxi distribution (i.e., $\max\left|Q_j^{M\ (k)} - Q_j^{M\ (k-1)}\right| \le \varepsilon$).

2.5. Model calibration

Calibration - First Stage Zone Choice Model: According to the ESL model structure, decisions are made independently and sequentially. Each zonal search decision is indeed described by an MNL model. The zone choice made by a driver in a particular zone is independent to the decision sequence. Hence, the variable coefficients (β^{E} , β^{t} , β^{c} and β^{R}) can be calibrated under the maximum log-likelihood principle using the following objective function:

$$\text{Maximize } P_{\log} = \sum_{i \in I} \sum_{j \in H_i \cup \{i\}} \ln P_q(j \mid i), \tag{19}$$

where $P_{\log}$ is the log-likelihood of the estimated ESL model with $M = 1$, in which $P_q(j \mid i)$ is the probability calculated from Equation (5).

The determination of the *optimal* number of zonal search decisions ($\hat{M}$) is based on a RMSE evaluation. The RMSE for a particular number of search-decisions is defined as the square root of the average of the sum of the square differences between the estimated vacant taxi zonal distributions (formed by Q_j^M for all zones $j \in I$) and that observed from the GPS data. The lower the RMSE, the better the fit of the model to the taxi trip data. Hence, the optimal number of zonal search decisions is the value of *M* that gives the lowest RMSE. The objective can be formulated as

$$\text{Minimize RMSE} = \sqrt{\frac{\sum_{m=1}^{N} \left(Q_m^{\mathrm{O}} - Q_m^{M}\right)^2}{N}}, \tag{20}$$

where Q_m^{O} is the observed number of vacant taxis terminating in zone $m \in I$ during customer-searches and *N* is the number of zones. By finding the minimum RMSE value, we can obtain the optimal number of zonal search decisions ($\hat{M}$).

Calibration – Second Stage Local Customer-search Model: The coefficient α^L can be calibrated under the maximum log-likelihood principle by solving the following optimization problem for each value of *L*:

$$\text{Maximize } V_{\log}^{L} = \sum_{j \in I} \sum_{x \in X_j} \sum_{y \in Z_x} \ln V_w^L(y \mid x), \tag{21}$$

where $V_{\log}^{L}$ is the log-likelihood of the estimated logit-opportunity model for a given number of local search decisions anticipated *L*, in which $V_w^L(y \mid x)$ is the probability calculated from Equation (10).

The optimal number of local search decisions anticipated ($\hat{L}$) is determined based on the Bayesian information criterion (BIC) evaluation. The model with the lowest BIC value is the most preferable as it has the best fit to the taxi trip data or involves the fewest explanatory variables, or both. Therefore, the optimization problem can be formulated as

$$\text{Minimize BIC} = \frac{-2V_{\log}^{L} + b\ln \sum_{j\in I} W_j^{\mathrm{T}}}{\sum_{j\in I} W_j^{\mathrm{T}}}, \tag{22}$$

where b is the number of parameters in the second stage model. By solving Problem (22), we can obtain the optimal number of local search decisions anticipated ($\hat{L}$).

3. Data and network

The two-stage model was calibrated with the survey data collected from GPS devices in 460 urban taxis, which provided field survey data on the daily activities of the taxis in Hong Kong. Using satellite communication, the database recorded the locations in terms of longitude and latitude, travel speed, and occupied status of the taxis at 30-second intervals. The 460 urban taxis studied represented approximately 3% of the entire population and provided an adequate sample size to represent the travel behavior of vacant taxi drivers.

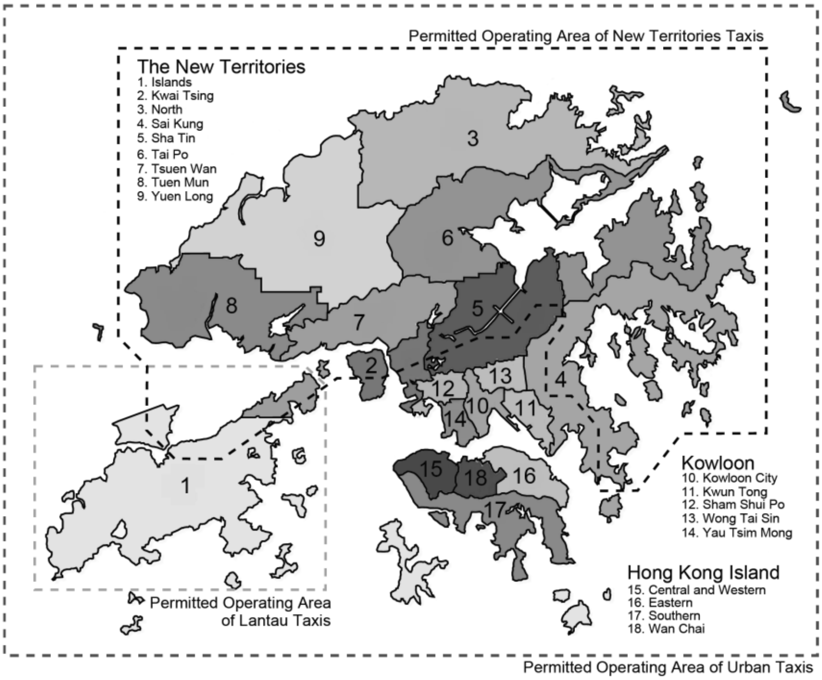

Fig. 2. District map and permitted operating areas of taxis

The GPS survey data were collected in the week spanning August 16-23, 2009. The data of taxi operations during the morning peak (07:30 to 09:30) were extracted for the model calibration, and trips occurring outside this period were excluded from the data analysis. This data set was used in the previous studies (Wong et al., 2014a, 2015a) for developing the MNL model and ESL model, respectively, and also used in this paper to facilitate our model comparison in Section 4.

The study area, Hong Kong territory, is divided into 18 zones according to administrative districts as shown in Figure 2. The districts were set based on population size, but each district also had a certain degree of homogeneity of land use. Nine of the zones are in rural areas (the New Territories, Zones 1 to 9) and the remaining zones are in urban areas (Kowloon and Hong Kong Island, Zones 10 to 18) as shown in Figure 2. Urban taxis are permitted to operate in all the districts of Hong Kong; New Territories taxis are only allowed to cruise for customers in Tuen Mun, Yuen Long, Tai Po, and North District, as well as parts of the Sha Tin, Tsuen Wan, Kwai Tsing, Sai Kung, and Islands Districts; Lantau taxis are only authorized to operate in Islands District and part of Tsuen Wan District.

To calibrate the first stage logit-based choice model, the attributes such as the average values of travel distance and time, revenue, and operational cost for trips between each origin-destination (OD) pair were determined. OD pairs with less than five trips were considered insufficient to give a representative average value, and were excluded. After the data extraction, about 11,200 vacant trips and 11,600 occupied trips from these 460 taxis were identified for model calibration.

To form a cell-based network for each zone for the second stage local customer-search model, each zone was divided into many identical squares, in which the length of each square equals 200 m. There are imaginary bi-directional links connected between each pair of adjacent cells, representing taxi movements between these cells. The vacant taxi trajectories extracted from the GPS data disclosed the cells and links involved during a customer-search process. In the case if these vacant trips were operating at high travel speeds and thus bypass the adjacent cells in our GPS records, the search trajectories of the cells that were passed through were estimated by the interpolation method. Cells and links with no recorded taxi activity were considered invalid and were excluded. The whole week (7 days a week, 24 hours a day) of GPS trip data was extracted to develop a cell-based urban taxi operating road network, in which the 2-hour (07:30 to 09:30) data in each day of the week was used for the model calibration and examining the local taxi customer-search behavior during the morning peak period.

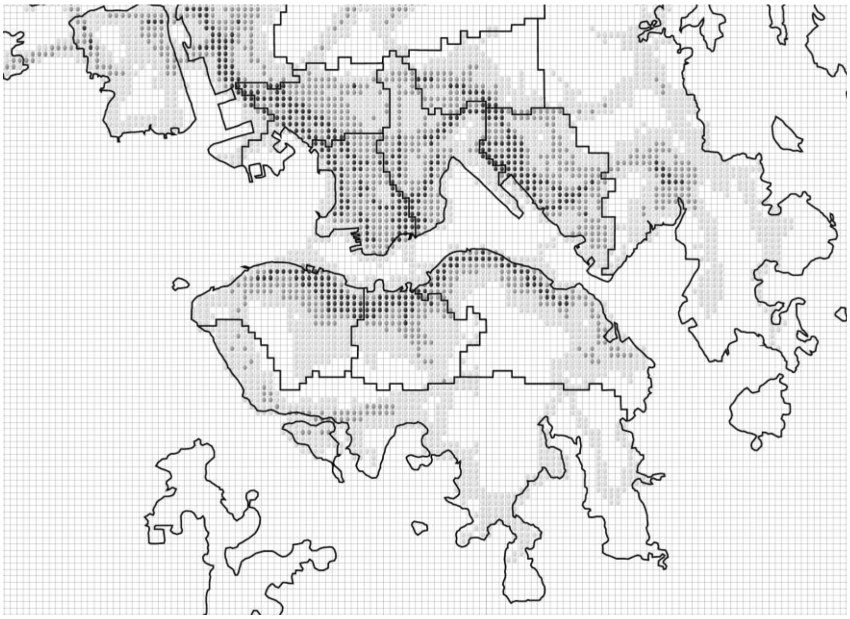

Fig. 3. The cell-based taxi operating road network for the zones in Hong Kong Island and Kowloon Peninsula

Figure 3 shows the cell-based urban taxi operating road network for the zones in Hong Kong Island and Kowloon Peninsula. The shaded cells represent the valid cells in the developed network. The cells without taxi service activities are left blank and were excluded from further analysis. Darker dots in the cells represent more trips involved in those cells. Based on the darkness of cell distribution indicated in this figure, it is noticed that the sampled urban taxis mostly concentrated in the urban areas of Hong Kong Island and Kowloon Peninsula in search of customers. This could happen because the urban taxis charge a higher fare and are less competitive to operate in rural areas than New Territories and Lantau taxis (Wong et al., 2014a). Therefore, the developed cell-based urban taxi operating road network mainly covers the urban areas.

Note that the proposed modeling methodology could be applied to any study area using different sizes of zones and cells. However, in general, the sizes of the zones and cells should be designed carefully based on the required level of modeling accuracy and the detail of taxi movements and the information available to the modeler. Some problems can be expected if the zones and cells are either too big or too small. If the sizes were too big, all differing data collected would be aggregated, making a useful analysis and conclusion impossible. An unnecessarily large zone or cell would also create concentrated traffic loading in particular cells or zones of the network, thereby distorting the overall traffic pattern. Hence, the taxi movements and operation predicted could not be useful for planning if the sizes are too large. If the sizes were too small, the relevant data collected would be statistically unreliable, and the number of samples in each zone and cell would be insufficient to provide representative means on the model explanatory variables. Moreover, some of the information may be unavailable if the sizes are too small because it is too expensive to collect.

4. Results and discussion

4.1. Convergence of the solution procedure

The iterative solution procedure in Section 2.4 was applied to solve the two-stage taxi customer-search model. With an adjustment coefficient (ϕ) of 0.05 and an acceptable tolerance (ε) of 0.05, the proposed iterative solution procedure provides a convergent solution of a zonal vacant taxi distribution as shown in Figure 4. The maximum absolute deviation $\max\left|Q_j^{M\ (k)} - Q_j^{M\ (k-1)}\right|$ starts from around 16 and declines to less than 0.05 (defined as the acceptable level) after 382 iterations.

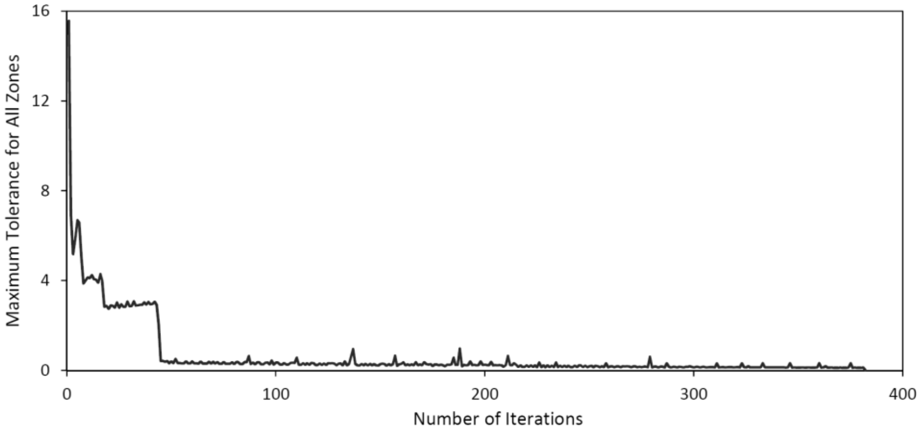

Fig 4. Convergence of the proposed iterative solution procedure

4.2. Coefficients of the two sub-models

The coefficients of the two sub-models were calibrated separately. Using the unit fuel cost (c^{f}) of HK$0.46 per kilometer and the unit rental cost (c^{r}) of HK$0.50 per minute (Wong et al., 2014a), we calibrated the model and performed t-tests on all of the explanatory variables in the first stage trip distribution model. The calibration results are tabulated in the upper part of Table 1. The RMSE value reached the minimum of 100 when the optimal number of search decisions to adjacent zones ($\hat{M}$) equaled 1. All the corresponding explanatory variables were significant at the 1% level during the morning peak period. Moreover, the coefficients for relative passenger demand and the rate of return were positive, and all other coefficients were negative. All these results make sense because a vacant taxi driver prefers to travel to a high demand zone and circulate over a short distance before reaching the next customer. Moreover, the largest coefficient (and strongest statistical significance) is given by the rate of return. This result agrees with our expectation because the rate of return is what taxi drivers aim to maximize as a business decision.

Table 1. Coefficients of the two sub-models and their t-statistics for the morning peak period

Explanatory variables	Coefficients[a] [t-statistics][b]
The first stage zone choice model (The optimal number of search decisions to adjacent zones, $\hat{M}$ = 1)	
Relative passenger demand (%)	0.08 [17.2]
Cross-zonal travel distance (km)	-0.19 [-14.7]
Intra-zonal circulation distance (km)	-0.13 [-4.7]
Rate of return (HK$/min)	2.38 [43.2]
The second stage local customer-search model (The optimal number of local search decisions in cell-based network, $\hat{L}$ = 14)	
Cumulative probability of success (%)	4.55 [57.1]

Note: [a] All parameters are significant at the 1% level.
[b] The value in a pair of brackets represents the *t*-statistics of the explanatory variables.

The lower part of Table 1 further tabulates the calibration results for the second stage model. The BIC value reached the minimum of 82.3 when the optimal number of local search decisions anticipated ($\hat{L}$) equaled 14 and the associated calibrated coefficient ($\alpha^{\hat{L}}$) was 4.55. The cumulative probability of success was confirmed as a significant factor influencing the local decisions of taxi drivers in the second stage customer-search at the 1% level. It is logical that the calibrated coefficient was positive, which implies that a cell with a higher probability of successfully meeting the next customer had a greater attraction to the vacant taxi drivers.

4.3. Model comparison

Figure 5 demonstrates the actual distribution of vacant taxis (based on about 11,200 trips by the concerned 460 urban taxis) deduced from the GPS data, and the estimated distributions based on the MNL model (Wong et al., 2014a), the ESL model (Wong et al., 2015a), and our proposed two-stage model. The MNL model assumes that drivers consider all 18 zones in their zonal choice decisions while the other two models assume that drivers only consider neighbor zones and the current zone in their zonal choice decisions. The RMSEs are also provided in this figure. The RMSEs of the distribution estimated by the MNL, ESL, and proposed models, were 107, 100, and 47, respectively. It means that our proposed model formulation fits the observed data better than the other two, which agrees with our expectation. In particular, a small reduction of RMSE from 107 to 100 means that the RMSE can be slightly improved by capturing sequential customer-search decisions of vacant taxi drivers on finding customers at intermediate zones while heading to their designated zones in the modeling framework. This improvement is because according to Wong et al. (2015a) (which used the same set of data), over 95% of the drivers met their next customers in the current zone or a neighbor zone because of high passenger demand during the morning peak period, and only limited drivers traveled to non-neighbor zones through expressways for customer-search. The implication is that it is good enough to capture neighbor zone choices for the peak period modeling. This better performance is also the reason why our proposed model adopts the ESL approach for the first stage model. The large reduction from 100 to 47 means that the RMSE can be largely improved by further modeling the fact that as vacant taxi drivers search for customers, the probability of successfully meeting a taxi customer along the way increases.

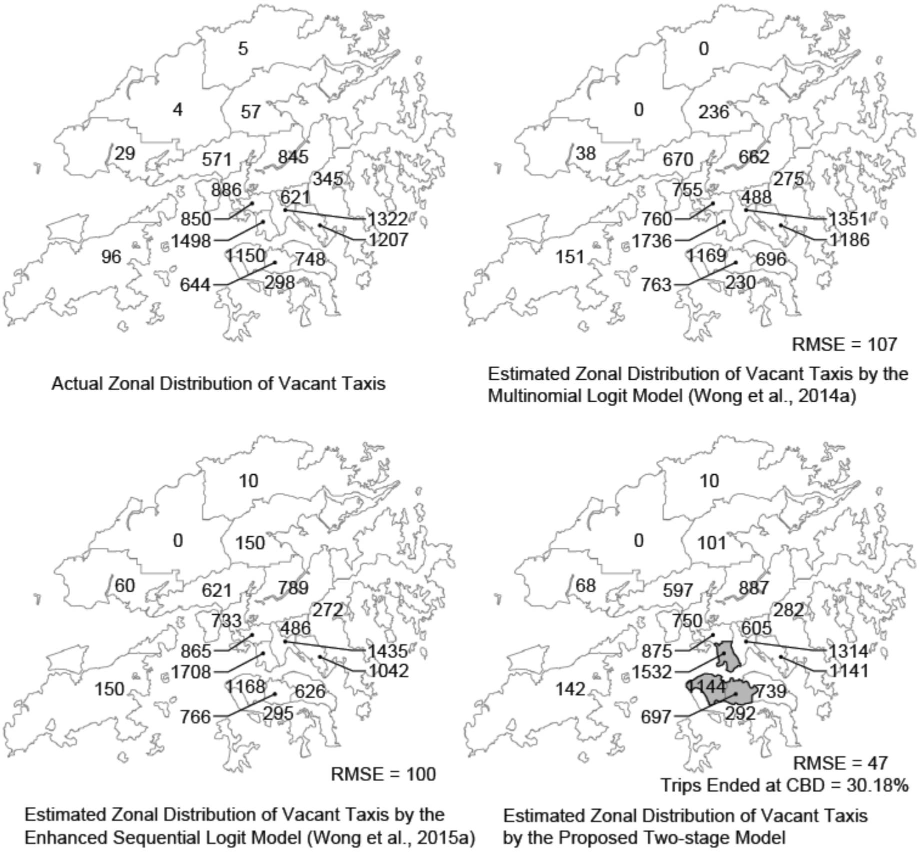

Fig 5. Actual and various estimated zonal distributions of vacant taxis during the morning peak period

4.4. Effects of the solution procedure parameters

To test the possibility of convergence of the iterative solution procedure and the uniqueness of the solution, we attempted inputting different initial (zonal vacant taxi distribution) solutions into the procedure. The procedure achieves the same convergent solution to all these scenarios, but the numbers of iterations required are different which depends on the similarity between the inputted initial solution and the convergent solution. In addition, from Table 2, we found that the adjustment coefficient (ϕ) and the acceptable tolerance (ε) affect the model results and the speed of convergence. We tested five combinations of ϕ and ε , where different solutions (as reflected by the RMSE) are obtained under various settings of these parameters, because these parameters control the set of acceptable solutions. Especially, when we adjusted ϕ to 0.5 and maintained ε at 0.05, the solution procedure cannot stop because the adjustment rate per iteration is too large to satisfy the requirement of acceptable tolerance.

Table 2. Effects of the adjustment coefficient and the acceptable tolerance

Combinations	Adjustment coefficient (ϕ)	Acceptable tolerance (ε)	Number of iterations	RMSE
1	0.05	0.50	367	55.01
2	0.05	0.10	371	47.43
3 (base case)	0.05	0.05	382	47.08
4	0.10	0.05	191	47.12
5	0.50	0.05	cannot stop	not applicable

4.5. Sensitivity analyses on potential taxi regulation policies

Figure 5 shows that a high proportion of vacant taxi trips (30.18%) ended in the highly congested CBDs (including Western and Central District, Wan Chai District, and Yau Tsim Mong District, which are shaded in the figure). These trips worsened the traffic congestion and air pollution problems there. Furthermore, the figure also shows that the taxi service quality to taxi customers at the non-CBDs was poor since the taxi supply to those areas was limited. It is therefore important to develop taxi regulation policies to achieve a more even distribution of vacant taxis to all zones and cause a lower proportion of vacant taxi trips ended in the CBDs in customer-search. The proposed two-stage model and the solution procedure can be used for this purpose. We considered two types of taxi policies: entry restrictions (adjusting the fleet size) and price controls (adjusting the fare).

Figure 6 presents the consequences due to adjusting the fleet size (from 50% to 150% of the current level) and fare (from 50% to 150% of the current level) during the morning peak period. The fare structure was assumed to remain the same but all charges, including the initial charge, addition charges, and unit charges for waiting time and travel distance, were adjusted according to the same percentage. (e.g., all the charges were discounted by 20% if the percentage equals 80%.) The solid lines represent the estimated proportion of trips ended in the CBDs. From this figure, it is noted that lowering the fare might help reduce the proportion of trips ended in the CBDs. However, the effect was not guaranteed. For example, if the fleet size remained unchanged, the proportion of vacant taxi trips ended in the CBDs was anticipated decreasing from 30.18% (the current situation) to 30.16% when the fare was reduced by 10%. However, if the fare was further reduced to 50% of the current level, the proportion would climb up to 30.35%. Nevertheless, the effect was minimal.

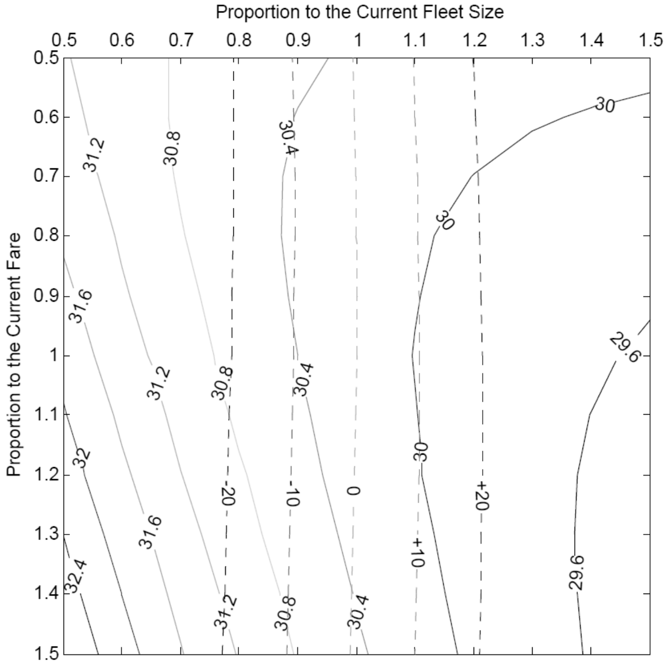

Fig 6. The estimated proportion of trips and variation in the number of vacant taxis ended in the CBDs under various fleet sizes and fares

If the fleet size was increased but the fare remained the same, the proportion of vacant taxi trips ended in the CBDs would decrease, and the proportion ended in the non-CBDs (i.e., one minus the proportion ended in the CBDs) would increase, meaning that vacant taxis would be more evenly distributed among all the districts and the supplies of vacant taxis among these districts would be closer. Increasing fleet size could increase the taxi supplies and hence the service quality at the non-CBDs, as the proportions ended in the non-CBDs increased with the total fleet size. However, a lower proportion of trips ended in the CBDs after increasing fleet size does not necessary mean that fewer vacant taxis ended there. The resultant number of taxis ended in those districts depends on the percentage decrease of the estimated proportion and the percentage increase of the fleet size. It could happen that the resultant number of vacant taxis going those districts increased. For example, if the fleet size was increased by 50% and the fare remained unchanged, the estimated proportion of trips ended in the CBDs would be decreased from 30.18% to 29.55% or equivalently the proportion had a 2.09 percentage reduction. This percentage reduction was much less than the percentage increases of the fleet size of 50%. Therefore, an increase of about 47% of vacant taxis was expected to circulate in those congested areas.

If the fare was also increased in addition to the fleet size, the proportion ended in the CBDs would be further reduced compared with the case of just increasing the fleet size, implying that the supplies among districts would be more even and the supplies to the non-CBDs would be further increased. In the extreme case, if both the fleet size and fare were increased by 50%, the estimated proportion of trips ended in the CBDs would decline to 29.44%, almost the minimum percentage in the ranges of fare and fleet size considered. However, about 46% more vacant taxis would circulate in the congested areas, and hence the traffic congestion and air pollution problems at the CBDs could not be relieved in this way.

To forecast the actual effect to the number of vacant taxis at the CBDs, the dash lines representing the variation in the number of vacant taxis at the CBDs (from +20% to -20% of the current situation) are also presented in Figure 6. These lines can be used as the taxi policy targets of reducing the number of vacant taxis at the CBDs to certain levels. For example, if the policy target was to drop down the number of vacant taxis at the CBDs by 10%, the fleet size should at least be decreased by about 11% while the fare remained the same.

From the above sensitivity analysis, it is found that adjusting the fare could not effectively minimize the number of vacant taxis at the CBDs. A more effective taxi policy that specifically deals with the vacant taxi trips entering the CBDs was hence necessary implemented. For this purpose, we considered the surcharge imposed specifically to the trips entering the CBDs, and predicted the effects of this policy on the vacant taxi distribution. It is expected that the operational cost to vacant taxis travel across zones to the CBDs would be increased and hence the attractiveness of the associated zone choice for customer-search would be reduced after imposing the surcharge.

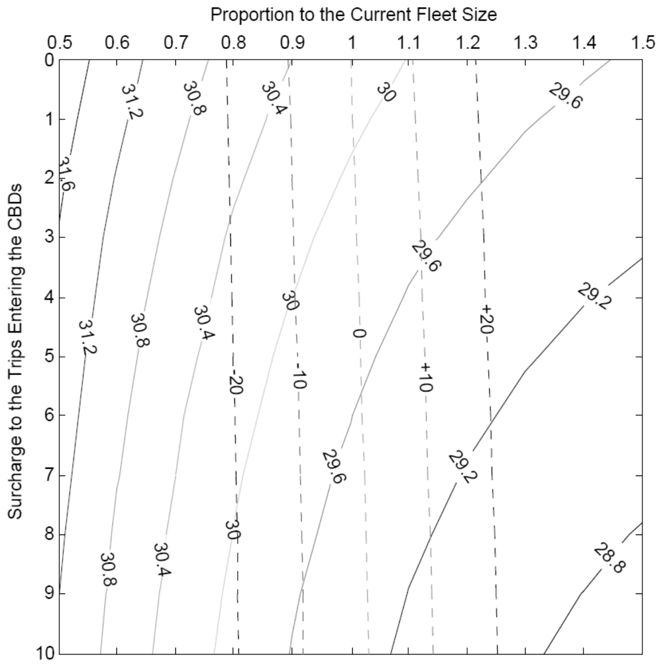

Fig 7. The estimated proportion of trips and variation in the number of vacant taxis ended in the CBDs under various fleet sizes and surcharges imposed to trips entering the CBDs

Figure 7 illustrates the effects on the estimated proportion of trips ended in the CBDs with surcharge (from none to HK$10) imposed to the trips entering the CBDs. The solid lines represent the estimated proportion of trips ended in the CBDs. As in Figure 6, the estimated proportion kept decreasing when increasing the fleet size only. Moreover, Figure 7 shows that the proportions also kept decreasing when increasing the surcharge only. Hence, the proportion decreased when both the fleet size and surcharge were increased. For example, when the fleet size remained unchanged, the proportion of vacant taxis ended in the CBDs was anticipated decreasing from 30.18% (the current situation) to 30.06% while imposing HK$1 surcharge to the trips entering the CBDs. The estimated proportion could even reach 29.50% if the fleet size was further increased to 150%.

By comparing Figures 6 and 7, we found that imposing HK$1 surcharge to the trips entering the CBDs could reduce the proportions ended in those districts more than increasing the fare by 10% (which is equivalent to at least

HK$1.8 increase in fare for the initial charge of HK$18). In general, we found that imposing the surcharge was more effective than adjusting the fare on influencing the proportion of vacant taxi trips ended in the CBDs. This observation is expected because the surcharge was imposed specifically to the trips entering the CBDs to increase the operational cost of the trips and reduce the attractiveness of the districts, but the fare increment was not.

The dash lines representing the variation in the number of vacant taxis at the CBDs (from +20% to -20% of the current situation) are also presented in Figure 7. If the taxi policy target was to drop down the number of vacant taxis at the CBDs by 10% (same as the previous numerical example), the fleet size should only be decreased by about 8% while imposing a HK$10 surcharge to the trips entering the CBDs. Moreover, if the fleet size were further reduced, the surcharge required could be smaller.

We also noticed that the dash lines representing the variation in the number of taxis at the CBDs in Figures 6 and 7 were almost vertical. It implies that either solely adjusting the fare or imposing the surcharge to the trips entering the CBDs had limited effects to achieve our policy target as compared with directly decreasing the fleet size. However, as in the situations of Hong Kong and other international cities, permanent licenses had already been issued to the urban taxi operators. The fleet size could only be increased but unlikely be reduced, unless purchasing the licenses back from the licensees. Therefore, to more effectively control the trip distribution of vacant taxis, additional supporting measures should also be investigated, such as strategically improving the meeting efficiency between vacant taxis and taxi customers at the non-CBDs by introducing more taxi stands there to shorten the circulation distances and times as suggested by Wong et al. (2014c, 2015b). Other taxi regulation measures (e.g., increase the initial charge and decrease the distance charge, or control the operational areas of taxis) may be effective. The cost-effectiveness analysis is left to future studies.

This study found that reducing the number of taxi licenses can effectively limit the excessive vacant taxi movements at the CBDs, but the Hong Kong Government has no flexibility to adjust this number. Therefore, rather than issuing more the permanent licenses in the future, it is suggested issuing temporary taxi licenses (e.g., Macau, a twin city of Hong Kong, issued taxi licenses with validity periods of 8-year and 10-year, and after which no renewal would be allowed). The city government could evaluate the conditions of passenger demand and number of vehicles on urban road networks to decide whether issuing new licenses after the expirations. This approach can reserve flexibility to the city government to control the taxi market to balance the demand and supply, and prevent the traffic congestion and air pollution problems at the CBDs to reach an unacceptable level.

5. Conclusion

This paper proposed a two-stage model to predict vacant taxi movements in search of customers. The first stage model was formulated as an ESL model, in which the vacant taxi trip distribution to adjacent zones was modeled as functions of the relative passenger demand, the cross-zonal travel distance, the intra-zonal circulation distance, and the rate of return. The second stage model is a cell-based logit-opportunity model, which captures the local customer-search movements of vacant taxi drivers within a single zone. These two sub-models were designed to influence each other, and hence an iterative solution procedure was introduced to obtain a convergent vacant taxi trip distribution. The two-stage model was calibrated by the GPS data of 460 Hong Kong urban taxis. The results show that all of the explanatory variables in both the first and second stage models were significant at the 1% level, and the proposed model offered a great improvement in terms of goodness of fit over the existing taxi-customer search models, which gave that the lowest RMSE value of 47 as compared with the observed trip distribution. The sensitivity analyses also illustrated the model capabilities to predict the changes in vacant taxi trip distribution under the implementations of various taxi policies. The effects on the fleet size and fare to the proportion of trips ended in the CBDs for customer-search were also explained. Potential taxi policies were investigated and discussed to achieve the taxi policy target on achieving a more even distribution of vacant taxis to all zones and cause a lower proportion of vacant taxi trips ended in the CBDs in customer-search. The proposed model allows us developing simulation-based models and simulation-based optimization models for depicting and managing taxi flow on local streets in future research.

Acknowledgements

This research was jointly supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (HKU 716913E), a grant (201411159063) from the University Research Committee, a Research Postgraduate Studentship from the University of Hong Kong, and a grant from National Natural Science Foundation of China (71271183). We would like to express our sincere thanks to Concord

Pacific Satellite Technologies Limited and Motion Power Media Limited for their kind provision of the GPS data from the taxis used for completing this study. We are also grateful to the comments from the two reviewers.

References

Arnott, R., 1996. Taxi Travel Should Be Subsidized. Journal of Urban Economics 40, 316-333.

Beesley, M.E., Glaister, S., 1983. Information for Regulation: The Case of Taxis. The Economic Journal 93, 594-615.

Cairns, R.D., Liston-Heyes, C., 1996. Competition and Regulation in the Taxi Industry. Journal of Public Economics 59, 1-15.

De Vany, A.S., 1975. Capacity Utilization under Alternative Regulatory Constraints: An Analysis of Taxi Markets. Journal of Political Economy 83, 83-94.

Douglas, G.W., 1972. Price Regulation and Optimal Service Standards: the Taxicab Industry. Journal of Transport Economics and Policy 6, 116-127.

Evans, S.P., 1976. Deviation and Analysis of Some Models for Combining Trip Distribution and Assignment. Transportation Research 10, 37-57.

Fernández, J.E., De Cea, J., Briones, J., 2006. A Diagrammatic Analysis of the Market for Cruising Taxis. Transportation Research Part E 42, 498-526.

Flores-Guri, D., 2003. An Economic Analysis of Regulated Taxicab Markets. Review of Industrial Organization 23, 255-266.

Florian, M., Nguyen, S., Ferland, J., 1975. On the Combined Distribution Assignment of Traffic. Transportation Science 9, 43-53.

Hu, X., Gao, S., Chiu, Y.C., Lin, D.Y., 2012. Modeling Routing Behavior for Vacant Taxi Cabs in Urban Traffic Networks. Proceedings of the 91st Annual meeting of the Transportation Research Board, National Research Council, Washington, D.C., United States.

Kim, H., Oh, J.S., Jayakrishnan, R., 2005. Effect of Taxi Information System on Efficiency and Quality of Taxi Services. Transportation Research Record 1903, 96–104.

Loo, B.P.Y., Leung, B.S.Y., Wong, S.C., Yang, H., 2007. Taxi License Premiums in Hong Kong: Can Their Fluctuations be Explained by Taxi as a Mode of Public Transport? International Journal of Sustainable Transportation 1, 249-266.

Manski, C.F., Wright, J.D., 1976. Nature of Equilibrium in the Market for Taxi Services. Transportation Research Record 619, 296-306.

McFadden, D., 1974. Conditional logit analysis of qualitative choice behavior. In: Zarembka, P. (ed.), Frontiers in Econometrics. New York: Academic Press, pp. 105-142.

Moore, A.T., Balaker, T., 2006. Do Economists Reach a Conclusion on Taxi Deregulation? Economic Journal Watch 3, 109-132.

Schroeter, J.R., 1983. A Model of Taxi Service under Fare Structure and Fleet Size Regulation. Journal of Economics 14, 81-96.

Sirisoma, R.M.N.T., Wong, S.C., Lam, W.H.K., Wang, D., Yang, H., Zhang, P., 2010. Empirical Evidence for Taxi Customer-search Model in Hong Kong. Proceedings of the Institution of Civil Engineers, Transport 163, 203-210.

Szeto, W.Y., Wong, R.C.P., Wong, S.C., Yang, H., 2013. A Time-dependent Logit-based Taxi Customer-search Model. International Journal of Urban Sciences 17, 184-198.

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, 819-842.

Wong, K.I., Wong, S.C., Yang, H., Tong, C.O., 2002. A Sensitivity-Based Solution Algorithm for the Network Model of Urban Taxi Services. In: Taylor, M.A.P. (ed.), Proceedings of the 15th International Symposium on Transportation and Traffic Theory. Elsevier Science, pp. 23-42.

Wong, K.I., Wong, S.C., Yang, H., Tong, C.O., 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, 81-104.

Wong, K.I., Wong, S.C., Yang, H., Wu, J.H., 2008. Modeling Urban Taxi Services with Multiple User Classes and Vehicle Modes. Transportation Research Part B 42, 985-1007.

Wong, R.C.P., Szeto, W.Y., Wong, S.C., Yang, H., 2014a. Modeling Multi-period Taxi Customer-searching Behavior. Transportmetrica B 2, 40-59.

Wong, R.C.P., Szeto, W.Y., Wong, S.C., 2014b. A Cell-based Logit-opportunity Taxi Customer-search Model. Transportation Research Part C 48, 84-96.

Wong, R.C.P., Szeto, W.Y., Wong, S.C., 2014c. Bi-level Decisions of Vacant Taxi Drivers Traveling towards Taxi Stands in Customer-search: Modeling Methodology and Policy Implications. Transport Policy 33, 73-81.

Wong, R.C.P., Szeto, W.Y., Wong, S.C., 2015a. Sequential Logit Approach to Modeling the Customer-search Decisions of Taxi Drivers. Asian Transport Studies, in press.

Wong, R.C.P., Szeto, W.Y., Wong, S.C., 2015b. Behavior of Taxi Customers for Hailing Vacant Taxis in Urban Areas: A nested logit model for policy analysis. Journal of Advanced Transportation, accepted.

Wong, S.C., Yang, H., 1998. Network Model of Urban Taxi Services: Improved Algorithm. Transportation Research Record 1623, 27-30.

Xu, J.M., Wong, S.C., Yang, H., Tong, C.O., 1999. Modeling the Level of Urban Taxi Services Using a Neural Network. ASCE Journal of Transportation Engineering 125, 213-216.

Yang, H., Fung, C.S., Wong, K.I., Wong, S.C., 2010a. Nonlinear Pricing of Taxi Services. Transportation Research Part A 44, 337-348.

Yang, H., Lau, Y.W., Wong, S.C., Lo, H.K., 2000. A Macroscopic Taxi Model for Passenger Demand, Taxi Utilization and Level of Services. Transportation 27, 317-340.

Yang, H., Leung, C.W.Y., Wong, S.C., Bell, M.G.H., 2010b. Equilibria of Bilateral Taxi-customer-searching and Meeting on Networks. Transportation Research Part B 44, 1067-1083.

Yang, H. and Wong, S.C. 1998. A Network Model of Urban Taxi Services. Transportation Research Part B 32, 235-246.

Yang, H., Wong, K.I., Wong, S.C., 2001. Modeling Urban Taxi Services in Road Networks: Progress, Problem and Prospect. Journal of Advanced Transportation 35, 237-258.

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, 799-819.

Yang, H., Yang, T., 2011. Equilibrium Properties of Taxi Markets with Search Frictions. Transportation Research Part B 45, 696-713.

Yang, H., Ye, M., Tang, W.H.C., Wong, S.C., 2005a. Regulating Taxi Services in the Presence of Congestion Externality. Transportation Research Part A 39, 17-40.

Yang, H., Ye, M., Tang, W.H.C., Wong, S.C., 2005b. A Multi-period Dynamic Model of Taxi Services with Endogenous Service Intensity. Operations Research 53, 501-515.
Yang, T., Yang, H., Wong, S.C., 2012. Taxi Services with Search Frictions and Congestion Externalities. Journal of Advanced Transportation 48, 575-587.