BI-LEVEL DECISIONS OF VACANT TAXI DRIVERS TRAVELING TOWARDS TAXI STANDS IN CUSTOMER-SEARCH: MODELING METHODOLOGY AND POLICY IMPLICATIONS

R.C.P. WONG [a], W.Y. SZETO [a,*] and S.C. WONG [a]

[a] Department of Civil Engineering, The University of Hong Kong, Hong Kong
[*]E-mail: ceszeto@hku.hk
Phone: (852) 28578552
Fax: (852) 25595337

ABSTRACT

This study adopts the sequential logit approach to modeling bi-level decisions of vacant taxi drivers in customer-search. The first level decision is about whether the drivers will travel to one of the nearest taxi stands after dropping off their customers and the second level decision is on whether the drivers will join the queue at the nearest taxi stand once they have arrived there. A stated preference survey was conducted to interview 258 urban taxi drivers about their choices of the two level decisions. The statistical test shows that search districts, travel distance from the customer's drop-off location to the designated taxi stand, the congestion level on the way of cruising, as well as the preference for traveling towards taxi stands are found to be the significant factors of the first level decision. This study also confirms that the queue lengths of both taxis and passengers at taxi stands, the expected customer-search distance after leaving taxi stands, and the preference of vacant taxi drivers for staying at taxi stands are found to be significantly influence the second level decision. The likelihood ratio tests for market segmentation analysis demonstrate the variations in preferences of taxi drivers operated in different taxi shifts and service areas. Some policy implications on introducing more taxi stands and improving the utilization rates of taxi stands are also discussed. We believe that the proposed sequential logit modeling approach, findings, and discussions are useful for developing micro-simulation models in terms of evaluating the performance of road traffic networks with taxi services and for developing simulation-based optimization models to answer policy questions related to taxi services.

Key words: sequential logit model; stated preference survey; behavior of vacant taxi drivers; taxi customer-search; taxi stands

1. INTRODUCTION

Taxis provide door-to-door transportation services. However, taxis tend to circulate around in searching for customers but these circulation activities occupy spare road spaces, which could worsen traffic congestion and air pollution problems. Moreover, because there is always a problem of mismatch between demand and supply of taxis over time and space, the taxi system is often found to be inefficient and the customers dissatisfy with the service. Furthermore, poorly designed taxi bays with insufficient capacities will cause long taxi queues extending onto nearby roadways and lead to the obstruction of local traffic. Advanced technologies and taxi regulation policies are established to tackle these problems. Studies which relate to advanced technologies (e.g., Lee and Cheng, 2008; Conway et al., 2012; da Costa and de Neufville, 2012) are aimed to increase the customer satisfaction or to improve overall system efficiency from the perspective of taxi operators. Studies which provide insights on taxi regulatory policies (e.g., Fernández et al., 2006; Manski and Wright, 1976; Schaller, 1998; Yang et al., 2005) are focused on examining the consequences on implementation of such policies. However, these studies have been developed based on an idealized market of conventional economic analysis in which the spatial structure of the market has been 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. Later, the model was improved by Wong et al. (2001, 2008), Kim et al. (2005), Yang et al. (2010a,b; 2012a,b), Yang and Yang (2011), and Hu et al. (2012) to capture congestion effects, multiple user classes, multiple taxi modes, customer hierarchical modal choice, day-to-day learning processes, stochastic travel time, taxi search behavior of taxi customers, and search frictions between vacant taxis and taxi customers. Solution algorithms were also developed by Wong and Yang (1998) and Wong et al. (2001, 2002, 2003, 2005, 2008) for solving the improved models. In these existing taxi network models, multinomial logit models were used to describe the search behavior of vacant taxi drivers. Yet, the logit models have never been calibrated and validated.

To calibrate and validate the multinomial logit models, Sirisoma et al. (2010) performed a stated preference (SP) survey. Wong et al. (2013) calibrated and validated the logit models based on the global positioning system (GPS) data obtained from 460 urban taxis to predict the drivers' strategic zonal choice for searching customers in both peak and off-peak periods. Szeto et al. (2013) further extended the consideration to every hour in a day, and Wong et al. (2014) extended Wong et al. (2013) by adopting the sequential logit approach to modeling sequential customer-search decisions of the vacant taxi drivers on finding customers at intermediate zones while heading to their designated zones. These behavioral studies of vacant taxi drivers are focused on zonal choice. These studies are useful in terms of transportation planning and policy evaluation purposes by developing a better taxi network models to capture the spatial interaction between zones, such as determining the zonal level of service and an optimal taxi fare structure. Unfortunately, these models cannot be used to develop micro-simulation-based models and simulation-based optimization models for depicting traffic flow including taxi flow on local streets, managing local traffic, evaluating the utilization rates of taxi stands and the capacities of taxi bays, and determining the optimal number and locations of taxi stands in each zone taking into the consideration of utilization rates. It is because some choice decisions of the vacant taxi drivers, such as traveling to a taxi stand, cannot be modeled in the zone choice model. Therefore, more vacant taxi driver decisions are required to model for these applications.

In fact, the vacant taxi drivers tend to have at least two choices after dropping off their customers, namely 1) cruising on streets and 2) traveling to and then waiting at a taxi stand. These aggregated decisions affect the road congestion level, roadside air quality, and utilization rates of taxi stands. More taxi drivers preferring waiting at taxi stands implies fewer vacant taxis circulating on streets, leading to better roadside air quality, less road congestion, and higher utilization rates of taxi stands. It is therefore important to understand the preferences of the vacant taxi drivers for these choices. Nevertheless, to the best of our knowledge, no study has conducted to examine the vacant taxi drivers' preferences until

Kitamura and Yoshii (2005) adopted GPS data to develop a single-level discrete choice model to illustrate their preferences. However, due to the limitations of the collected GPS data, their model could not take into account of the numbers of passengers and vacant taxis that stand in queues at the taxi stands, the districts of where taxi stands locate at, the congestion level on the way to their destination, and the expected customer-search distance after leaving the taxi stands. All these factors will intensely influence the decisions of vacant taxi drivers on whether to travel or stay at taxi stands in customer-search.

This study adopts a sequential logit approach to develop a model to formulate the vacant taxi drivers' preferences for traveling towards taxi stands for searching customers into two levels. The first level decision is on whether the drivers will travel to one of the nearest taxi stands after dropping off their customers and the second level decision is on whether the drivers will join the queue at the taxi stand once they have arrived there. To calibrate the model for future use in the simulation studies, this study has conducted a SP questionnaire survey. A total of 258 urban taxi drivers were interviewed throughout the survey. Each urban taxi driver was invited to report his/her choices under four hypothetical scenarios and hence more than 1,000 observations were obtained. This study also investigates whether the drivers' search strategies change across different market segments in terms of taxi shift and service area. This was accomplished by developing individual models and undertaking Watson and Westin pooling test (Watson and Westin, 1975). The paper also discusses the potential taxi policy implications on enhancing the utilization rates of taxi stands and on introducing additional taxi stands in different taxi shifts and service regions. Since there is a surcharge for telephone bookings, and it is not difficult for a passenger to catch a vacant taxi on any urban roadsides or at taxi stands in Hong Kong, which is similar to other metropolitans with dense business, commercial and entertainment areas, telephone bookings are not common in Hong Kong. Therefore, the impact on how the telephone bookings may help on taxi drivers' decisions is excluded in this study.

The contributions of this paper include the following:

- Proposing a sequential logit modeling approach to modeling bi-level decisions of vacant taxi drivers in customer-search. The sequential logit approach is capable of describing the realistic situation actually encountered by taxi drivers. The model structure should be more suitable to describe the choice behavior of the taxi drivers traveling towards taxi stands in finding customers;
- Determining the factors that affect the choices of vacant taxi drivers on traveling towards the nearest taxi stand to search for customers and those vacant taxi drivers who wait at taxi stands;
- Discussing the potential implications of implementing the proposed taxi policies, such as introducing more full-time, day-time, and night-time taxi stands in different service regions, and implementing policies to increase the utilization rates of taxi stands, and;
- Providing a sub-model for developing micro-simulation models to evaluate the performance of road traffic networks with taxi services as well as developing simulation-based optimization models to answer policy questions related to taxi services.

The remainder of this paper proceeds as follows. Section 2 describes the data collection method, the demographical distribution of interviewed urban taxi drivers, and the details of drivers' customer-search decisions. Section 3 presents the model structure and the methodology of market segmentation analysis. Section 4 discusses the results and policy implication, and finally, Section 5 concludes the paper.

2. DATA

2.1 Data Collection

At present, there are 18,138 taxis operating in Hong Kong territory, including 15,250 urban taxis, 2,838 New Territories taxis and 50 Lantau taxis. These taxis serve about 1 million occupied trips a day. Urban taxis operate at most areas in Hong Kong (Hong Kong Island, Kowloon Peninsula and the New Territories); New Territories taxis mainly operate in the north-eastern part and north-western part of New Territories; and Lantau taxis are only permitted to operate on Lantau Island and Chep Lap Kok (the airport) only. Figure 1 shows the district map of Hong Kong and the associated permitted operating areas for different types of taxis.

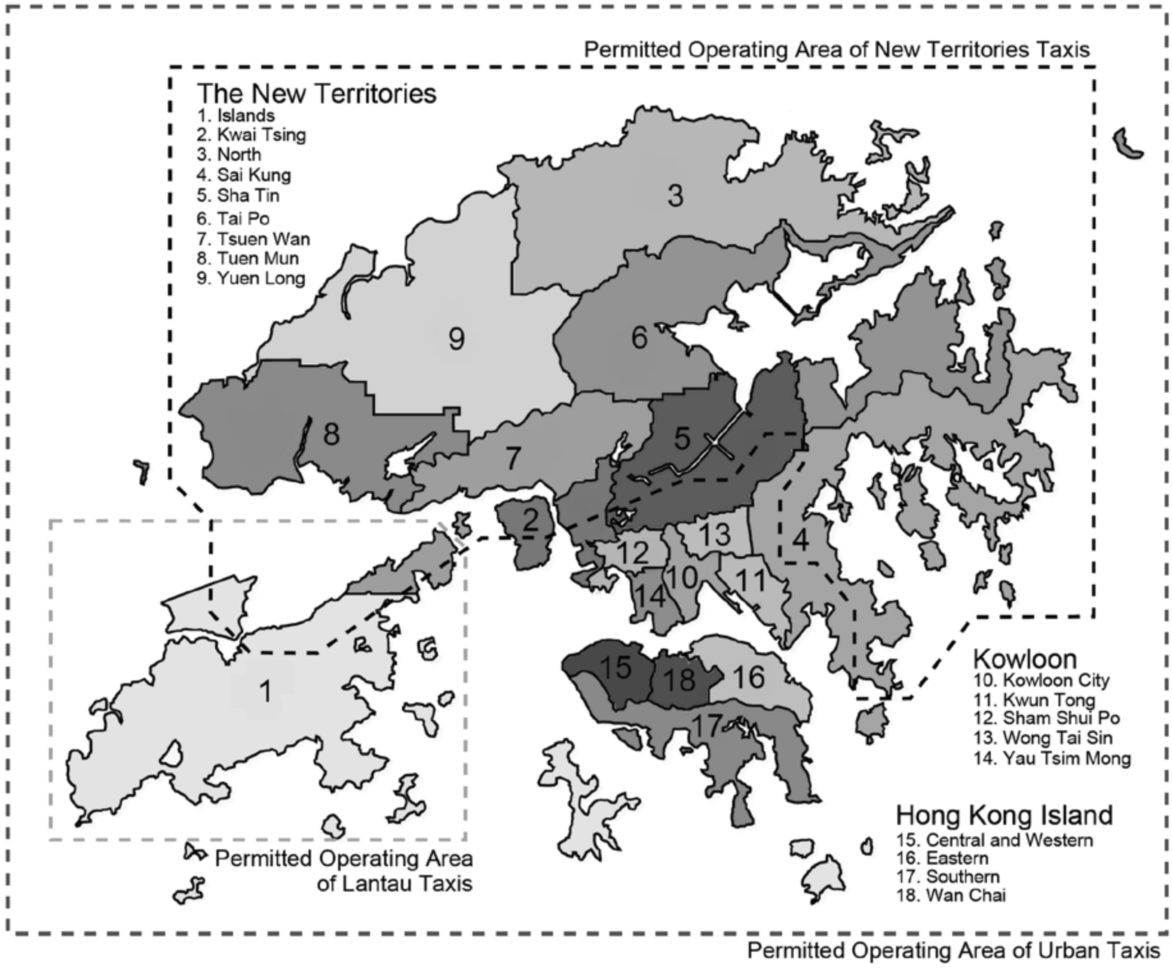

Figure 1. District Map and Permitted Operating Areas of Taxis

The Hong Kong Transport Department has set up 468 roadside taxi stands in total of which 79 are in Hong Kong Island, 107 are in Kowloon Peninsula, 252 are in New Territories and the remaining 30 roadside taxi stands are located at Lantau Island. Most of these taxi stands are ordinary on-street taxi stands specifically providing services for one (sometimes two) type of taxi. Less than 10% of these stands are cross-harbor taxi stands that serve the passengers with their destinations at the opposite side of the Victoria Harbor. Amongst all these on-street taxi stands, a few of them operates in either day-time or night-time only. This arrangement prevents further worsening the traffic congestion problem at the urban areas at peak hours due to one traffic lane being occupied by taxis in queue in front of an on-street taxi stand, and caters the variations in passenger demand with spatio-temporal considerations. Other than the above mentioned ordinary roadside taxi stands, the study area also covers the taxi waiting points at hotels, shopping malls, major commercial buildings, and residential estates that are operated under private operations which offer similar services as the ordinary taxi stands. However, on-street taxi pick-up/drop-off points are excluded in this study as taxi drivers are prohibited from waiting for passengers at these pick-up/drop-off points where taxi drivers must leave immediately after dropping off the preceding customer unless a passenger is waiting at that spot.

The SP questionnaire survey was conducted from January to March 2012. The face-to-face interview questionnaire surveys were taken place at numbers of selected commercial and residential areas in Hong Kong, Kowloon Peninsula, the New Territories, and Hong Kong International Airport in Lantau Island. At the survey, we have interviewed the urban taxi drivers at roadside taxi stands, taxi waiting points, gas stations, popular shift changing places, and along roadsides at both day time and night time. The questionnaires with return envelopes were distributed to those taxi drivers who did not have sufficient time to complete the questionnaires during the face-to-face interview. To ensure that the surveys were widely distributed to different types of taxi drivers, hundreds of questionnaires were also sent to more than 10 taxi associations to have them allocated the questionnaires to their members. In total, 258 responses from taxi drivers were collected, where about half of them were collected during face-to-face interviews and the remaining questionnaires were collected by mail. The overall response rate was estimated to be about 18% in this study.

2.2 Demographical Distribution of Interviewed Taxi Drivers

The questionnaire has two sections. One section collects the driver’s demographics such as his/her gender, age, education level, taxi ownership, main taxi shift period, service area, weekly working days, taxi driving experience, and average income (excluding the rental cost, fuel cost of liquefied petroleum gas (LPG), and toll) in one shift. The summary of the demographic distribution is summarized in Figure 2 below.

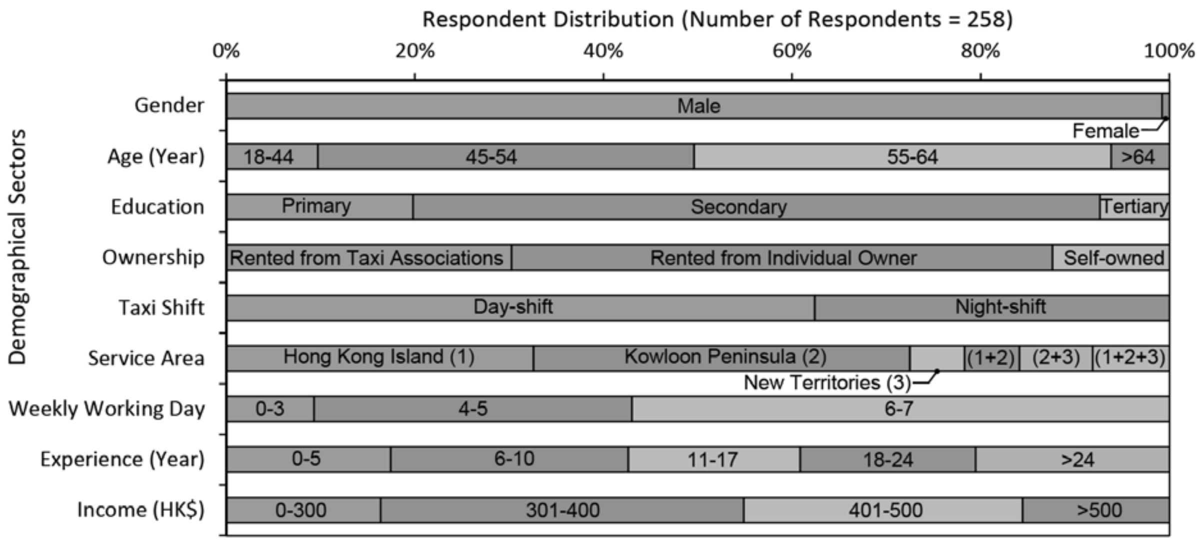

Figure 2. Demographical Distribution of Interviewed Taxi Drivers

From the above figure, it is observed that most of the interviewed taxi drivers were male and only 1% was female, which matched with our observations of this industry. The age of taxi drivers were categorized into four groups: 18-44, 45-54, 55-64, and older than 64. The 45-64 age group was shown to dominate the overall taxi industry. Most of the taxi drivers had completed or at education level. More than 80% of the drivers rented their taxis from either taxi associations or individual owners. Slightly more than half of the drivers operated in day-shift. Majority of the respondents identified themselves as either Hong Kong Island taxi or Kowloon Peninsula taxi drivers, while 20% of them were operating for multiple service areas. Most of the taxi drivers were full-time drivers for not less than 4 days a week. Taxi driving experiences were evenly spread over each group, with a mean of 16 years. The reported income per shift of the interviewed taxi drivers was mostly ranged from HK$ 301 to HK$500.

Other than responding personal particulars of the interviewed taxi drivers, they were also requested to report their frequencies of traveling towards taxi stands per shift in search of customers. As shown in Figure 3 below, it provided a general idea regarding the utilization rates of taxi stands by urban taxi drivers.

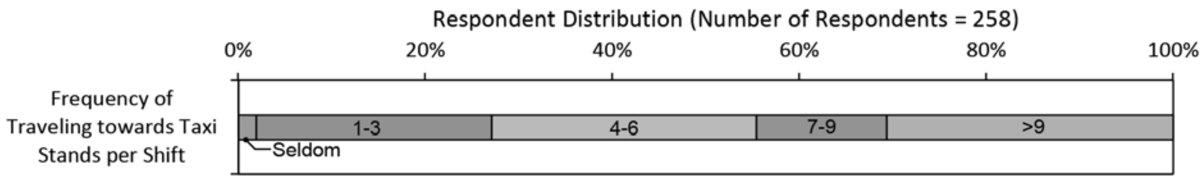

Figure 3. Frequency of Traveling towards Taxi Stands per Shift

It is noted that more than half of the interviewed taxi drivers approached to taxi stands in customer-search less than 6 times in one shift. Considering that each taxi takes about 25 occupied trips per shift as informed by the taxi drivers, less than one-fourth of their trips were involving taxi stands. The findings agreed with the observations on the low utilization rates of taxi stands in the urban areas.

We have also asked the urban taxi drivers an open-ended question about the drawbacks of waiting for customers at taxi stands. Some drivers commented that some taxi stands in the urban area were located at inconvenient locations for taxi passengers and consequently the passengers preferred waiting at roadsides rather than walking to a nearby taxi stand. It is noted that during the off-peak period and after mid-night period when the overall passenger demand was low, a lot of vacant taxis tended to cruise around and hence it could be easy for the passengers to catch taxis along roadsides. It gave less motivation for passengers to wait at taxi stands, and the customer-wait time at taxi stands became unacceptably long. Other reasons are also found: 1) Some taxi stands do not provide enough space for the vacant taxi drivers to join the taxi queues; 2) taxi stands are located at the congested areas; 3) travel distance to the taxi stand is relatively long even going to the nearest taxi stand; and 4) some designs of the taxi stands confine taxis in one traffic lane which make them unable to leave after they join the queue.

2.3 Customer-search Decisions

The other section of the questionnaire collects the decisions of the vacant taxi drivers who made in customer-search. Figure 4 illustrates the decision tree of the vacant taxi drivers in customer-search.

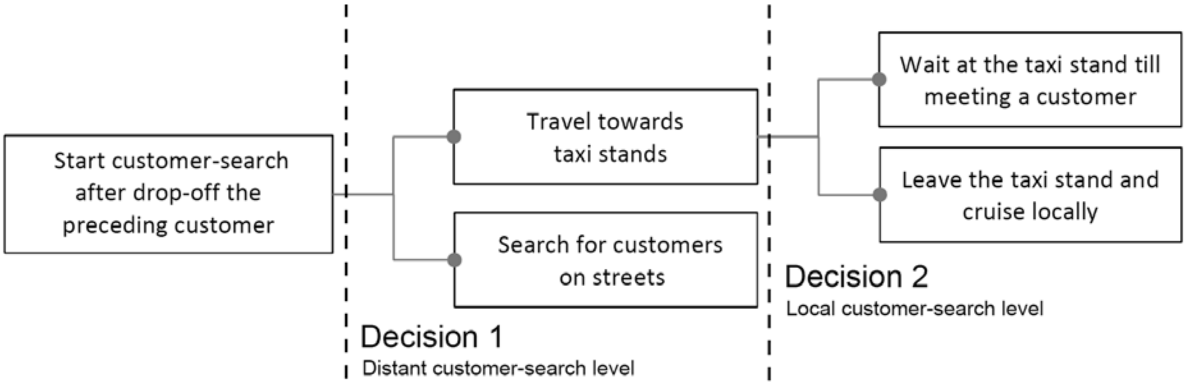

Figure 4. Bi-level Decisions of Vacant Taxi Drivers in Customer-search

In a normal operation (when the driver does not receive any telephone order for a dispatch service) after a vacant taxi driver drops off the preceding customer, the drivers have two options to search for customers: 1) Heading to a taxi stand; and 2) cruising on streets for customers. It is considered that after dropping off the customers and before reaching a nearby taxi stand, only limited road network information is available to this taxi driver. For example, the driver only has the information about the surrounding traffic condition (say, about 1 km), the required travel distance to concerned taxi stand, and the districts which the taxi stand located at. The actual conditions of the concerned taxi stand such as the queue lengths of taxis and passengers are somehow unpredictable to this vacant taxi driver before he/she actually arrives there. Once the driver reaches the taxi stand, the observed taxi competition and passenger demand at the taxi stand may influence the decision of this vacant taxi driver whether to wait at or leave the taxi stand. Because of the uncertainty of taxi stand conditions, it is assumed that taxi drivers made bi-level decisions sequentially in searching for customers at taxi stands as in Figure 4. Hence, the SP questionnaire survey asked the customer-search preferences in two levels:

Upper Level (or distant customer-search level): In the distant customer-search level, it was assumed that the interviewed taxi drivers had dropped-off their preceding customer and started searching for the next. The conditions of the nearby taxi stands were unknown. They made their choices based on three attributes, namely the search district, distance from the taxi stand to the drop-off location, and the congestion level on the way. These three attributes were used to generate different hypothetical situations for the taxi drivers to decide on traveling to a nearby taxi stand or cruise locally. In order to evaluate the preferences of the drivers for traveling towards nearby taxi stands in different conditions, two different options were designed for heading to these taxi stands.

Lower Level (or local customer-search level): Once the interviewed vacant taxi drivers arrived at a taxi stand, the queue lengths of both taxis and passengers could be observed. Hence, in this level, the drivers decide either waiting at the taxi stand until they meet a customer, or leaving the taxi stands based on the queue lengths of both taxis and passengers, and the opportunity cost of leaving the taxi stands to meet a customer in certain distance. These three attributes were used to generate different hypothetical situations for the local search decision. For the taxi drivers who decided to wait at the taxi stands, the expected customer-search distance was obviously equal to zero. On the other hand, if the taxi drivers decided to leave the taxi stands and cruise locally, then the numbers of taxi and passenger in queues at the expected roadside pickup location were considered to be equal to zero.

Table 1 tabulates the attributes and their levels for both distant and local decisions. Except the attributes about the search district and the locations of taxi stands, all the other quantitative attributes were designed at three levels to capture the possible non-linear effects of the attributes about the urban taxi drivers' response. The fractional factorial design method was applied to generate choice experiments, and a total of 27 profiles were produced. The choice sets were divided randomly into 7 sets of questionnaires and then distributed among taxi drivers.

Table 1. Attributes and Levels Used for the Stated Preference Survey

Distant Decisions	Attributes	Levels
Travel towards taxi stand A (D1) / taxi stand B (D2)	District (land use) where taxi stands locate	Commercial (-1), Residential (1)
	Number of blocks[a] apart from taxi stands	3, 6, 9
	Congestion level along the way to taxi stands	No (0), Slight (1), Heavy (2)
Search for customers along roadsides (D3)	District (land use) of customer-search	Commercial (-1), Residential (1)
	Number of blocks[a] required to travel in customer-search	4, 8, 12
	Congestion level along the way of search	No (0), Slight (1), Heavy (2)
Local Decisions	**Attributes**	**Levels**
Wait at the taxi stand till meeting a customer (L1)	Number of taxis in queue	0, 5, 10
	Number of passengers in queue	0, 3, 6
Leave the taxi stand and cruise locally (L2)	Number of blocks[a] traveled in local search	2, 4, 6

Note: [a] The travel distance of one block in urban areas is assumed to be equal to 100 m.

The interviewed taxi drivers were requested to select the most preferable choice among the choice set. An example of the choice set of distant decisions asked in the questionnaire is shown as below:

Distant decision (D1): Travel towards taxi stand A in a commercial district where is 9 blocks away from the current location and it is anticipated that there will be slight traffic congestion along the way;

Distant decision (D2): Travel towards taxi stand B in a residential district where is 3 blocks away from the current location and it is anticipated that there will be heavy traffic congestion along the way; and

Distant decision (D3): Do not travel towards any taxi stand. Search for customers in a residential district within 12 blocks away from the current location and it is anticipated that no congestion along the way.

An example of the choice set of local decisions in the questionnaire is shown as below:

Local decision (L1): Wait at a taxi stand until meeting a customer, given that currently there is 5 vacant taxis and 3 gangs of customers in the queue; and

Local decision (L2): Leave the taxi stand. Cruise locally within 6 blocks nearby the taxi stand for customers.

To ensure the data quality for model calibration, each interviewed taxi driver was requested to report their choice decisions in 4 hypothetical games at each decision level, and hence more than 1,000 observations were collected from the 258 urban taxi drivers in this study.

3. METHOD

3.1 Sequential Logit-based Search Model

Instead of using the multinomial (simultaneous) logit model that assumes the vacant taxi drivers make their decisions at the first instance, this study adopted the sequential logit model in depicting their sequential choice decisions. An important difference between sequential logit and multinomial logit models is that: while the multinomial logit model can be derived under the assumption that an individual chooses the alternative by considering the entire choice set simultaneously, the sequential logit model is derived under the assumption that an individual's choice process consists of some sequential and independent choices. Several researches have been conducted on the comparison between simultaneous and sequential choice models, and have discussed the models' properties (Gudishala and Wilmot, 2011; Kahn and Morimune, 1979; Nagakura and Kobayashi, 2009; Ophem and Schram, 1997). The sequential logit approach is capable of providing a similar situation actually encountered by the vacant taxi drivers, and the model structure should be more suitable to describe their choice behavior on whether to travel towards taxi stands in finding customers, and if so, whether to stay there to wait for their customers.

As it is assumed that the vacant taxi drivers made their bi-level decisions separately and sequentially at each level, we calibrated the sequential logit model in two parts based on the data collected in the SP questionnaire survey. The resultant model was used to validate the significances of the variables on affecting the sequential customer-search strategy at each level. The sequential logit model can be expressed as the following form (McFadden, 1974):

$$P_{iq}^{\mathrm{d}} = \frac{\exp\left(V_{iq}^{\mathrm{d}}\right)}{\sum_{m} \exp\left(V_{mq}^{\mathrm{d}}\right)} \text{ and } P_{jq}^{\mathrm{l}} = \frac{\exp\left(V_{jq}^{\mathrm{l}}\right)}{\sum_{n} \exp\left(V_{nq}^{\mathrm{l}}\right)}, \tag{1}$$

where P_{iq}^{d} and P_{jq}^{l} are the probabilities that an individual vacant taxi driver chooses *i* and *j* as his/her distant and local decisions towards taxi stands in customer-search trip *q*, respectively. V_{iq}^{d} and V_{iq}^{l} are the deterministic utilities that capture the factors influencing the distant and local decisions of vacant taxi driver *i* in customer-search trip *q*, respectively. According to Equation (1), the higher utility of a choice implies a higher probability of that choice being selected by the driver.

The utility for the distant decision is modeled as follows:

$$\text{Distant decision:} \qquad V_{iq}^{\mathrm{d}} = \beta_{\mathrm{S}} S_i + \beta_{\mathrm{D}}^{\mathrm{d}} D_i^{\mathrm{d}} + \beta_{\mathrm{C}} C_i + \beta_{\mathrm{M}}^{\mathrm{d}} M_i^{\mathrm{d}}, \tag{2}$$

where S_i, D_i^{d}, C_i, and M_i^{d} are, respectively, the search district, the distant travel distance (which can be the search distance along the roadside or the travel distance towards the concerned taxis stand from the last customers' drop-off point), the congestion level on the way of search, and the dummy variable of distant decision option i . β_{S}, $\beta_{\mathrm{D}}^{\mathrm{d}}$, β_{C}, and $\beta_{\mathrm{M}}^{\mathrm{d}}$ are the respective coefficients. The distant decision dummy variable (M_i^{d}) is introduced into the utility function to determine the general perception of taxi drivers, other than the first three attributes, using the taxi stands for customer-search. The distant decision dummy variable is equal to one if the choice is finding customers at taxi stands; otherwise, it is equal to minus one.

The utility for the local decision in Equation (1) is modeled as follows:

$$\text{Local decision:} \qquad V_{jq}^{\mathrm{l}} = \beta_{\mathrm{T}} T_j + \beta_{\mathrm{P}} P_j + \beta_{\mathrm{D}}^{\mathrm{l}} D_j^{\mathrm{l}} + \beta_{\mathrm{M}}^{\mathrm{l}} M_j^{\mathrm{l}}, \tag{3}$$

where T_j, P_j, D_j^{l}, and M_j^{l} represent the number of taxis in queue, the number of passengers in queue, the local travel distance (which is the expected customer-search distance after leaving the taxi stand), and the dummy variable of local decision option *j*, respectively . β_{T}, β_{P}, $\beta_{\mathrm{D}}^{\mathrm{l}}$, and $\beta_{\mathrm{M}}^{\mathrm{l}}$ represent the respective coefficients. The local decision dummy variable (M_j^{l}) is designed similarly to the distant decision dummy variable. The local decision dummy variable is equal to one if the choice is waiting for customers at taxi stands; otherwise, it is equal to minus one.

It is important to clarify that the attributes in the above utility functions are indeed the perceived values to a taxi driver for the associated decision making towards taxi stands in customer-search. Every taxi driver is assumed to have the same perception to the above attributes. Therefore, for simplicity, the subscript *q* is omitted in each attribute.

3.2 Watson and Westin Pooling Test

Model comparisons were also conducted in this study to demonstrate the changes in customer search preferences of the taxi drivers in different taxi shifts and service areas. The Watson and Westin pooling test (Watson and Westin, 1975) was applied in the comparison. The test is based on the log likelihood ratio LR, which can be calculated as:

$$\mathrm{LR} = -2(L_{\mathrm{R}} - L_{\mathrm{U}}), \quad (4)$$

where L_{R} is the log likelihood for the base model calibrated for the combined dataset with all market segments (either in terms of taxi shift or service area, but not both simultaneously). L_{U} is the sum of the log likelihoods of the sub-models calibrated for different market segments. The null hypothesis that there is no intervention in market segmentation is rejected, as the test statistic exceeds the threshold value that is specified for the chi-squared distribution at the chosen level of significance. The degree of freedom is calculated as the difference between the number of variables of the base model and the sum of the number of individual sub-models.

4. RESULTS AND DISCUSSION

4.1 Results of the Base Model

A typical modeling software NLOGIT was adopted in this study. The software NLOGIT uses the maximum likelihood estimation to determine the coefficient of each variable in the choice model at each decision level. The coefficient of each variable at each level is tabulated in Table 2. All the coefficients are not equal to zero at the 1% significance level. Hence, the search district, distant and local travel distances, the congestion level on the way of search, the numbers of taxis and passengers in queues at taxi stands, and the perceptions on traveling towards and waiting at the taxi stands are the significant factors that affect the decisions of the vacant taxi drivers traveling towards taxi stands and staying there in customer-search.

Table 2. Results of the Base Model

Decision Levels	Explanatory variables	Coefficients [t-statistics][a]
Distant	Search district	-0.22[b] [-5.74]
	Distant travel distance	-0.09[b] [-5.46]
	Congestion level	-0.98[b] [-17.85]
	Distant decision dummy	-0.14[b] [-3.66]
Local	Taxi queue	-0.31[b] [-13.53]
	Passenger queue	0.42[b] [11.08]
	Local travel distance	-0.14[b] [-2.77]
	Local decision dummy	0.40[b] [3.43]

Note: [a] The values in brackets represent the t-statistics of the explanatory variables.
[b] Parameters are significant at the 1% level.

For the distant customer-search level, the search district provides a negative coefficient, which suggests the taxi drivers prefer to search customers at a commercial district (-1) in general. The coefficients of the distant travel distance and the congestion level are also negative, which implies that the longer travel distance and a more congested road condition on the way to a taxi stand will result in fewer taxi drivers going to the taxi stand as it is less attractive. It is also found that the coefficient of the distant decision dummy variable is negative as well, it explains that the general perception of the vacant taxi drivers prefer to search along roadsides to travel towards a taxi stand. This finding is reasonable because taxi drivers normally need to spend additional time (generally not equal to zero) in waiting for customers at taxi stands.

For the local decisions, the model indicates that a long taxi queue at a taxi stand would discourage the vacant taxi drivers to wait at that taxi stand for customers. Conversely, a long passenger queue at a taxi stand would attract more drivers to travel there for customers. If the required local travel distance in customer-search after leaving the taxi stand is shorter, more drivers may decide to cruise locally instead. Furthermore, it is also noted that the coefficient of the local decision dummy variable is positive, which indicates that taxi drivers tend to stay at the taxi stand once they have arrived there.

One interesting finding worth to mention is that the coefficients of the numbers of taxis and passengers in queues at taxi stands are quite different as shown in Table 2. We expect that taxi drivers would only concern about the net number of passengers (i.e., the number of passengers minus the number of vacant taxis) in queue to make their decisions on either waiting or leaving the concerned taxi stands. However, the results indicate that the taxi drivers would give a higher weight to the passenger queue than the taxi queue when they made their choices. This result can be explained by the perceptions of the vacant taxi drivers on the expected wait time at taxi stands. A taxi stand with passengers in queue normally implies that the passenger arrival rate is larger than the service rate of passengers. This is a good sign to the vacant taxi drivers which lead them to have a preference for waiting at those taxi stands with passenger queues.

To sum up, the model results of both the distant and local customer-search strategies are logical and meet our expectations.

4.2 Model Results for Different Market Segments

Table 3 shows the results of the market segmentation analysis. The market segments were defined by 2 taxi shifts, namely day-shift and night shift. The market segment analysis was done at both distant and local decision levels. Table 3 also gives the log likelihood values of the base model and the sum of the individual models for different market segments. Both of which are used to calculate the log likelihood ratios. The degree of freedom for each market segmentation analysis is determined as the difference between the sum of the number of parameters associated with each segment and the number of parameters of the combined model.

Table 3. Results of the Market Segmentation Analysis

Market segments	Decision levels	L_R	L_U	Degree of freedom	LR	Critical value	Conclusion
Taxi shift	Distant	-910.10	-894.97	4	30.26	13.28[a]	Reject
	Local	-480.40	-476.97	4	6.86	9.49[b]	Do not Reject
Service area	Distant	-1,166.02	-1,153.26	8	25.52	20.09[a]	Reject
	Local	-623.64	-615.75	8	15.78	15.51[b]	Reject

Note: [a] Parameters are significant at the 1% level.
[b] Parameters are significant at the 5% level.

For the taxi shift segmentation analysis, the log likelihood ratio at the distant decision level is significantly higher than the chi-square critical value at the 1% significance level. Reversely, the log likelihood ratio at the local decision level is lower than the chi-square critical value. The null hypothesis that there are no interventions between different taxi shifts is rejected at the distant decision level only, but it is not rejected at the local decision level. The result implies that only the distant customer-search strategies are found to be different in various taxi shifts.

For the service area segmentation analysis, the log likelihood ratios at both the distant and local decision levels are higher than the associated critical values. Therefore, the null hypotheses that there are no interventions in service areas at both decision levels are rejected. As a result, we conclude that both decision levels vary from the others and cannot be pooled together.

It is important to clarify that about 20% of the interviewed taxi drivers reported that they have multiple service areas as mentioned in Section 2.2. In this case, their search behavior could not be categorized for a single service area. Therefore, we treat their choice decisions as independent observations for each of the involved areas. As a result, the log likelihoods of the base model (L_R) for service areas are different from those for taxi shifts at both distant and local decision levels.

Table 4 presents the detailed results of the differences in drivers' search behavior in various shifts at the distant decision level. The coefficients in each individual model are found similar to those in the base model shown in Table 2, except that the coefficient of the distant decision dummy variable is found to be insignificant at the 5% level. The results suggest that the day-shift taxi drivers might not have a clear preference for searching customers along roadsides over traveling towards taxi stands in the distant customer-search level. It can be explained by the higher passenger demand at taxi stands in day-time, but this is not the case for the night-shift drivers. Also, the coefficient of the congestion level indicates that the day-shift drivers have a higher concern on cruising in a congested area for customers; this is probably because the road networks are normally highly congested at the urban areas in day-time, and the traffic congestion would affect their turnaround time, which in turn would affect their profit. Furthermore, the night-shift drivers have a stronger preference for traveling to the commercial districts in searching for customers because most of the taxi passenger demands are generated at commercial districts during night-time.

Table 4. Model Results for Different Market Segments

<table>
<tr><th rowspan="2">Decision levels</th><th rowspan="2">Explanatory variables</th><th colspan="3">Coefficients [t-statistics][a]</th></tr>
<tr><th>Day-shift</th><th colspan="2">Night-shift</th></tr>
<tr><td rowspan="4">Distant</td><td>Search district</td><td>-0.10[c] [-2.05]</td><td colspan="2">-0.41[b] [-6.54]</td></tr>
<tr><td>Distant travel distance</td><td>-0.09[b] [-4.11]</td><td colspan="2">-0.10[b] [-3.62]</td></tr>
<tr><td>Congestion level</td><td>-1.08[b] [-15.13]</td><td colspan="2">-0.83[b] [-9.45]</td></tr>
<tr><td>Distant decision dummy</td><td>-0.08 [-1.53]</td><td colspan="2">-0.25[b] [-4.08]</td></tr>
<tr><th>Decision levels</th><th>Explanatory variables</th><th>Hong Kong Island</th><th>Kowloon Peninsula</th><th>New Territories</th></tr>
<tr><td rowspan="3">Distant</td><td>Search district</td><td>-0.32[b] [-5.86]</td><td>-0.14[b] [-2.73]</td><td>-0.15 [-1.78]</td></tr>
<tr><td>Distant travel distance</td><td>-0.05[c] [-2.24]</td><td>-0.14[b] [-5.94]</td><td>-0.11[b] [-2.99]</td></tr>
<tr><td>Congestion level</td><td>-0.84[b] [-10.99]</td><td>-1.18[b] [-15.10]</td><td>-1.04[b] [-8.66]</td></tr>
</table>

	Distant decision dummy	-0.16[b] [-3.07]	-0.17[b] [-3.37]	-0.13 [-1.58]
Local	Taxi queue	-0.28[b] [-8.93]	-0.34[b] [-10.71]	-0.26[b] [-5.56]
	Passenger queue	0.37[b] [7.16]	0.50[b] [9.67]	0.33[b] [4.16]
	Local travel distance	-0.14[c] [-2.03]	-0.15[b] [-2.21]	-0.10 [-0.88]
	Local decision dummy	0.27 [1.70]	0.42[b] [2.64]	0.65[c] [2.50]

Note: [a] The values in brackets represent the t-statistics of the explanatory variables.
[b] Parameters are significant at the 1% level.
[c] Parameters are significant at the 5% level.

Table 4 also provides an idea of the variations in customer-search strategies of the vacant taxi drivers in different service regions at both decision levels. At the distant customer-search level, the Hong Kong Island taxi drivers have weighted the search district the most, whereas the Kowloon Peninsula taxi drivers have considered the distant travel distance and the congestion level to be more important. This could be due to the fact that the passenger demand is normally highly concentrated at commercial districts on the Hong Kong Island. On the other hand, because there is less concentrated passenger demand on Kowloon Peninsula, the taxi drivers prefer to reach a customer in a shorter distance as fast as possible on the way of their search. The New Territories taxi drivers are mainly concerned about the travel distance and the congestion level at the distant level. However, these drivers did not show a clear preference for traveling toward taxi stands in customer-search at their distant decision level. It may be due to the fact that the expected amount of time spent on cruising along the roadside to meet a customer is about the same as that on waiting at the taxi stands for a customer. Moreover, the search district is not a significant factor that affects the search choice; it is because the passenger demand is about evenly distributed in each district.

At the local decision level, it is noticed that the individual models for different service regions are similar. The result suggests that the taxi drivers perform similarly, except that the taxi drivers in Hong Kong Island do not show a clear preference for staying at taxi stands until they have picked up a customer. This is because the high passenger demands in Hong Kong Island and leads to a higher chance to meet a customer in a short travel distance away from the taxi stands. This observation is opposite to those New Territories drivers who have a very strong preference for waiting at the taxi stands and also do have not much concern about the required local travel distance after leaving the taxi stands. Again, it is because the chance of meeting a customer after leaving the taxi stands is not significantly higher than staying there. Hence, it is better for them to stay at the taxi stands to wait for customers and save fuel cost.

4.3 Discussion on Policy Implications

As mentioned previously, the distant decision dummy variable is expected to give a negative coefficient; it is because the taxi drivers require additional wait time (generally not equal to zero) at taxi stands. The coefficient's magnitude of this dummy variable also reflects the conceptual degree of dissatisfaction of the vacant taxi drivers about traveling to taxi stands, including but not limited to the reasons reported by the taxi drivers as discussed in Section 2.2. The small absolute value of the coefficient of the distant decision dummy variable in each sub-model suggests that adjusting the value of one of the other three attributes will significantly affect the decision choices of vacant taxi drivers in searching for customers. Hence, each of these sub-models can be used to examine the effect of the policies that affect the value of one of the three attributes. For example, introducing more taxi stands would reduce the required travel distance to reach the nearby taxi stands. According to the sub-model findings, the shorter travel distance to reach the nearby taxi stands will increase the probability of taxi drivers traveling toward the taxi stands in customer-search. Hence, we can conclude that introducing more taxi stands can encourage more taxi drivers traveling toward taxi stands in customer-search in general.

According to the response of the open-ended questions and the model results, the results suggest that the taxi stands are strongly not recommended to be located at heavily congested road sections. The utilization of those taxi stands will be low and it is not cost-effective. These results further imply that when developing the simulation-based taxi models or simulation-optimization taxi models, it is necessary to take into account the background traffic that causes the traffic congestion.

The aggregate model also indicates that most of taxi drivers have a strong preference for searching customers at commercial areas in general because of having a higher passenger demand in those areas, even though there are no taxi stands. In order to achieve a better distribution of vacant taxis for providing better quality of services to all passengers in different regions, it is recommended to set up more taxi stands at the residential areas. These taxi stands may adopt advance technology to collect taxi passenger demand information and provide the information online to attract vacant taxi drivers to pick up customers there.

The model results about various taxi shifts as discussed in Section 4.2 also revealed that there is a significant variation in customer-search behavior between taxi drivers operate in different shifts. The night-shift taxi drivers have little motivation on traveling towards taxi stands for searching the next customers, and they prefer to circulate within the commercial areas. The findings offer a great support of establishing the part-time taxi stands to cater the variations in taxi drivers' customer-search preferences and spatio-temporal passenger demand. Considering the coefficients' magnitudes of search district and distant search dummy variables, more day-time only taxi stands at the residential areas and night-time only taxi stands at the commercial areas are recommended to be built in Hong Kong and other cities with similar characteristics.

The service area segmentation analysis shows that the sub-model of Hong Kong Island has a smaller absolute value of the coefficient of the distant travel distance variable than the other two sub-models. It implies that providing more taxi stands (consequently decreasing the required travel distance to the nearest taxi stand) on Hong Kong Island will improve less local usage of taxi stands than providing more taxi stands at Kowloon Peninsula and the New Territories.

The coefficient of the local decision dummy variable in each sub-model is small and positive, which indicates that under current arrangement and operation, the vacant taxi drivers do not show a strong preference to wait at the taxi stands until picking up a new customer there. In making local customer-search decisions, it is noted that the drivers also consider the values of the other three attributes, including the number of passengers in queues at taxi stands that implies the passenger arrival rate there. In order to enhance the usage of taxi stands by the taxi drivers, the most effective way is to attract more passenger demand arriving at taxi stands which is supported by the calibrated lower level sub-model. To achieve this objective, it is recommended to set up an appropriate direction sign system that guides passengers walking towards their nearest taxi stands. This system will enhance the passenger arrival rates at taxi stands and potentially prolong the expected local search distances after the drivers leaving the taxi stands. Both consequences will encourage more taxis waiting at taxi stands and will reduce cruising activities on the streets.

5. CONCLUSION

This study analyzes the preferences of vacant taxi drivers for traveling towards taxi stands in finding customers or staying at the taxi stands to wait for customers. The vacant taxi drivers are considered to have bi-level decisions, namely a distant customer-search level decision and a local customer-search level decision. The distant decision is made after the drivers drop off the preceding customers, whereas the local decision is made when they arrive at a taxi stand. A stated preference survey was conducted to interview 258 urban taxi drivers about their distant and local choice decisions under different hypothetical games. The data obtained is used to develop a sequential logit model for the analysis.

By carrying out statistical tests, it is discovered that the search district, the distant travel distance, the congestion levels on the way of search as well as the distant decision dummy variable (the preference of the drivers for traveling towards taxi stands) are found to be the significant factors that affect the distant decisions of drivers after they drop off the preceding customer. In particular, it is found that the longer travel distance and a more congested road condition on the way to a taxi stand will result in fewer taxi drivers going to the taxi stand, and that the vacant taxi drivers tends to search along roadsides to traveling towards a taxi stand.

This paper also confirms that the queue lengths of both taxis and passengers at taxi stands, the expected customer-search distance after leaving taxi stands, and the local decision dummy variable (the preference of taxi drivers for staying at a taxi stand) significantly influence the vacant taxi drivers' local decisions on joining the queues at taxi stands once they arrive there. Particularly, a long passenger queue and a short taxi queue at a taxi stand probably attract the vacant taxi drivers to wait there for their customers. We also find that the taxi drivers tend to stay at the taxi stand once they arrive there. The results also illustrate that the taxi drivers give a higher weight to the passenger queue than the taxi queue when they make their choices.

The likelihood ratio tests of market segmentation analysis demonstrate the variations in the preferences of vacant taxi drivers operate in different taxi shifts and service areas. The distant customer-search strategies are found to be different in various taxi shifts; however there is insufficient evidence to show that the local search strategies are different in various taxi shifts. It is also observed that the search strategies at each decision level are different in various service areas.

This paper also discusses the potential taxi policy implications on enhancing the utilization rates of taxi stands and introducing additional taxi stands in different taxi shifts and service regions. To sum up, the model, findings, and discussion on the policy implications presented in this paper are believed to be useful for developing the micro-simulation models to evaluate the performance of road traffic networks with taxi services as well as for developing the simulation-based optimization models to answer policy questions related to taxi services.

ACKNOWLEDGEMENTS

This research was jointly supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (HKU 716913E), two grants (201111159056 and 201209160014) from the University Research Committee, a Research Postgraduate Studentship and Faculty of Engineering Top-up Grant from the University of Hong Kong, and a grant from National Natural Science Foundation of China (71271183). The authors also wish to thank Michelle Po Yee TAI, a student of Civil Engineering Department, The University of Hong Kong for her assistance for data collection and analysis. The authors are also grateful to the reviewer for his/her constructive comments.

REFERENCES

Conway, A., Kamga, C., Yazici, A. and Singhal, A. (2012) Challenges in managing centralized taxi dispatching at high-volume airports: Case study of John F. Kennedy International Airport. Transportation Research Record 2300: 83-90.

da Costa, D.C.T. and de Neufville, R. (2012) Designing efficient taxi pickup operations at airports. Transportation Research Record, 2300: 91-99.

Fernández, J.E., De Cea, J. and Briones, J. (2006) A diagrammatic analysis of the market for cruising taxis. Transportation Research Part E, 42: 498-526.

Gudishala, R. and Wilmot, C. (2011) A Comparison of time-dependent sequential logit and nested logit for modeling hurricane evacuation demand. Transportation Research Record, 2312: 134–140.

Hu, X., Gao, S., Chiu, Y.C. and Lin, D.Y. (2012) Modeling routing behavior for vacant taxicabs in urban traffic networks. Transportation Research Record, 2284: 81-88.

Kahn, L.M. and Morimune, K. (1979) Unions and employment stability: a sequential logit approach. International Economic Review, 20: 217-236.

Kim, H., Oh, J.S. and Jayakrishnan, R. (2005) Effect of taxi information system on efficiency and quality of taxi services. Transportation Research Record, 1903: 96-104.

Kitamura, R. and Yoshii, T. (2005) Rationality and heterogeneity in taxi driver decisions: Application of a stochastic-process model of taxi behavior. In Mahmassani, H. (ed.) Transportation and Traffic Theory: Flow, Dynamics, and Human Interaction. Elsevier Science, 609-628.

Lee, W.C. and Cheng, B.W. (2008) Incorporating e-technology to advantage in a greener taxi industry and its impact on driving performance and safety. Transportation Planning and Technology, 31: 569-588.

Manski, C.F. and 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, 105-142.

Nagakura, D. and Kobayashi, M. (2009) Testing the sequential logit model against the nested logit model. The Japanese Economic Review, 60: 345-361.

Ophem, H. and Schram, A. (1997) Sequential and multinomial logit: A nested model. Empirical Economics, 22: 131-152.

Schaller, B. (1998) Issues in fare policy: Case of the New York taxi industry. Transportation Research Record, 1618: 139-142.

Sirisoma, R.M.N.T., Wong, S.C., Lam, W.H.K., Wang, D., Yang, H. and 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., and Yang, H. 2013. A time-dependent logit-based taxi customer-search model, International Journal of Urban Sciences, 17: 184-198.

Watson, P.L. and Westin, R.B. (1975) Transferability of disaggregate mode choice models. Regional Science and Urban Economics, 5: 227-249.

Wong, K.I., Wong, S.C., Bell, M.G.H. and 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. and 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. and 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. Amsterdam: Elsevier Science, 23-42.

Wong, K.I., Wong, S.C., Yang, H. and 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., Yang, H. and 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. and Wong, S.C. (2014) Sequential logit approach to modeling customer-searching decisions of taxi drivers. Asian Transport Studies, in press.

Wong, R.C.P., Szeto, W.Y., Wong, S.C. and Yang, H. (2013) Modeling multi-period customer-searching behavior of taxi drivers. Transportmetrica B, doi:10.1080/21680566.2013.869187.

Wong, S.C. and Yang, H. (1998) Network model of urban taxi services: improved algorithm. Transportation Research Record, 1623: 27-30.

Yang, H., Fung, C.S., Wong, K.I. and Wong, S.C. (2010a) Nonlinear pricing of taxi services. Transportation Research Part A, 44: 337-348.

Yang, H., Leung, C.W.Y., Wong, S.C. and 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, S.C. and 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. and Yang, T. (2011) Equilibrium properties of taxi markets with search frictions. Transportation Research Part B, 45: 619–742.

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

Yang, T., Yang, H. and Wong, S.C. (2012a) Taxi services with search frictions and congestion externalities. Journal of Advanced Transportation, doi:10.1002/atr.1210.

Yang, T., Yang, H., Wong, S.C. and Sze, N.N. (2012b) Returns to scale in the production of taxi services: an empirical analysis. Transportmetrica A, doi:10.1080/23249935.2013.794174.