



R. M. N. T. Sirisoma
Post-doctoral Fellow,
Department of Civil
Engineering, Schulich School
of Engineering, University of
Calgary, Alberta, Canada



S. C. Wong
Professor, Department of
Civil Engineering, The
University of Hong Kong,
Hong Kong SAR, China



William H. K. Lam
Professor, Department of
Civil and Structural
Engineering, The Hong Kong
Polytechnic University,
Kowloon, Hong Kong SAR,
China



Donggen Wang
Professor, Department of
Geography, Hong Kong
Baptist University, Kowloon
Tong, Hong Kong SAR,
China



Hai Yang
Professor, Department of
Civil and Environmental
Engineering, The Hong Kong
University of Science and
Technology, Kowloon, Hong
Kong SAR, China



Peng Zhang
Professor, Shanghai
Institute of Applied
Mathematics and
Mechanics, Shanghai
University, Shanghai,
China

Empirical evidence for taxi customer-search model

R. M. N. T. Sirisoma BSc(Eng), PhD, MITE, AMIE, S. C. Wong MPhil, PhD, FCILT, FHKIE, FHKSTS, MASCE, MITE, MIHT, W. H. K. Lam MSc, PhD, FHKIE, FHKSTS, FIHT, MCILT, MICE, MITE, D. Wang MA, MSc, PhD, FHKSTS, H. Yang BSc, DEng, FHKSTS and P. Zhang MSc, PhD

A multinomial logit model of urban taxi services has been developed for the study of the operational characteristics of the taxi industry, in which it is hypothesised that the customer-searching behaviour of vacant taxis follows a multinomial logit choice model. Although the model is commonly used, little empirical evidence exists to validate the choice mechanism and determine the set of important factors that affect the choice. In the present study, a stated preference survey of 400 taxi drivers was conducted to analyse the customer-searching behaviour of vacant taxis. The results explain how the considered parameters of waiting time, journey time, travel distance and toll affect the driver behaviour when searching for customers. In addition, market segmentation analysis was carried out to study the effects of driver demographics and operational characteristics on the searching behaviour. The parameters that were considered in this section of the study were the age of the driver, taxi ownership, driver experience and marital status.

NOTATION

c_{sr}	generalised travel time via the shortest route from zone s to zone r
LR	log likelihood ratio
L_R	log likelihood for the base model calibrated for the combined dataset with all market segments
L_U	sum of the log likelihoods of the sub-models calibrated for different market segments
$P_{r/s}$	probability that a vacant taxi originating in zone s will meet a customer in zone r
U^i	utility function of a taxi driver in searching for the next customer
V^i	deterministic component of the utility
w_r	waiting (search) time in zone r
x_D^i	distance for individual i
x_{JT}^i	journey time for individual i
x_T^i	toll for individual i
x_{WT}^i	waiting time for individual i
β_{JT}	journey time coefficient
β_D	distance coefficient
β_T	toll coefficient
β_{WT}	waiting time coefficient

γ	weighting that a taxi driver places on waiting compared to journey time
ε^i	random component of the utility function for individual i
θ	degree of uncertainty of customer demand and taxi services in the whole market
ν_D	value of distance
ν_{JT}	value of journey time
ν_{WT}	value of waiting time

1. INTRODUCTION

Taxi services play a vital role in the transportation system as they offer door-to-door service to passengers, and are a more comfortable and convenient transportation mode than other public modes of transportation, especially for the travel requirements of disadvantaged groups such as the disabled and elderly (Hensher, 2007; Petzall, 1995; Pucher and Renne, 2003; Su *et al.*, 2010). The main issue of taxi services is that although an occupied taxi has to take a direct route to the customer's destination, there is no guarantee that the driver of the vacant taxi can find a new customer at the destination. This leads to inefficient movement of vacant taxis, which circulate on roads in search of customers and cause traffic congestion and air pollution in surrounding areas. The taxi drivers' knowledge of the market is limited and updated only from their experience (Kim *et al.*, 2005). Traditionally, studies use an aggregate approach to investigate the taxi industry, which is subject to various types of regulation, such as entry restriction and price control, and the economic consequences of regulatory restraints (Arnott, 1996; Beesley and Glaister, 1983; Cairns and Liston-Heyes, 1996; De Vany, 1975; Douglas, 1972; Foerster and Gilbert, 1979; Frankena and Pautler, 1986; Loo *et al.*, 2007; Manski and Wright, 1976; Schroeter, 1983; Shreiber, 1975; Yang *et al.*, 2000, 2005a, 2005b), but they do not take into account the spatial structure of the taxi market.

In view of the necessity and importance of network modelling of taxi traffic, Yang and Wong (1998) made an initial attempt to characterise taxi movements in a road network for a given customer origin–destination demand pattern. They proposed a simultaneous system of equations to describe the movements of both vacant and occupied taxis, and solved the problem with a fixed-point algorithm. The model explicitly deals with the effects of taxi fleet size and the degree of taxi driver

uncertainty about customer demand and service conditions through various system performance measures such as taxi utilisation and taxi availability at equilibrium, and thus helped to provide information for government decision-making about taxi regulations. Wong and Yang (1998) reformulated the taxi service problem in networks as an optimisation problem, which led to a more efficient and convergent iterative balancing algorithm for the case without traffic congestion. Wong *et al.* (2001) extended the simple network model of urban taxi services proposed by Yang and Wong (1998) by incorporating congestion effects and customer demand elasticity and reformulating the problem as the simultaneous optimisation of two sub-problems, and then developed a solution algorithm. Wong *et al.* (2005) modelled bilateral micro-searching behaviour for urban taxi services using the absorbing Markov chain approach. Golias (2003) developed the sensitivity of the impact of taxi traffic on capacity and delay on urban roads. Lee *et al.* (2005) developed a taxi-pooling dispatching system to determine the optimum fleet size when the demand is given. Wong *et al.* (2002) developed a sensitivity-based solution algorithm for a taxi model with congested effects and elastic demand, and Wong *et al.* (2008a) further extended this method to deal with the case of multiple user classes and multiple vehicle modes.

The potential applications of the model have been demonstrated by several case studies of the urban area of Hong Kong. Yang *et al.* (2001) conducted a case study of the calibration and validation of the simple network model for the urban area of Hong Kong. Yang *et al.* (2002) then investigated the nature of demand–supply equilibrium in a regulated market for taxi services in the urban area of Hong Kong. They concentrated on the effects of alternative regulatory restraints on market equilibrium by investigating the social surplus, firm profit, and customer demand at various levels of taxi fare and fleet size in regulated, competitive, and monopoly markets. However, the model calibrations in these studies were conducted at the macroscopic level in an attempt to match the general patterns of taxi movement in the network, in which the logit-based search model of vacant taxis was assumed and taken as a prerequisite input in the network model. To ascertain whether vacant taxis actually search for customers in accordance with the mechanism in previous studies (Wong and Yang, 1998; Wong *et al.*, 2001, 2002, 2005, 2008a; Yang and Wong, 1998; Yang *et al.*, 2001, 2002), it is essential to obtain empirical evidence to validate this hypothesis of taxi driver customer-searching behaviour through choice information from taxi drivers.

This paper aims to provide empirical evidence of the taxi operational behaviour that was assumed in developed models for taxi movements in a road network. A stated preference (SP) survey was conducted to collect choice information for a set of well-designed hypothetical games from taxi drivers, and analysed the data obtained using a multinomial logit (MNL) model that is fully consistent with the search model that was adopted in previous works. The results help to validate the taxi search model and determine the sensitivity of the taxi driver's search preference. In addition, market segmentation analysis was carried out to identify the possible effects of driver experience, age, marital status, and other operational characteristics on driver search choice.

2. DATA

Data were collected in December 2000. The first part of the questionnaire included driver demographics such as age, gender, marital status, working experience, and taxi ownership status. The summary of the demographic distribution is shown in Table 1. Age and taxi driver experience were categorised into four different intervals to make it consistent with the sample size. The second part consisted of a set of SP games, in which five attributes, each with three levels, were identified. Table 2 lists the five attributes and their levels. The choice of attributes and their levels were based on initial interviews with taxi drivers regarding the factors that influence their decision on where to find their next passengers. To derive the possible non-linear effects of the variables on the drivers' behaviour, each attribute was defined by three levels, whose exact values were again based on interviews with taxi drivers.

The first attribute (ATT1) referred to the waiting time at the place at which the previous customer alighted. The other four attributes (ATT2, ATT3, ATT4 and ATT5) were used to generate different hypothetical situations in the event that the taxi driver decided to search for passengers. These attributes included the searching time and distance, possible tolls, and waiting time. The toll fee was defined as the charge that a taxi has to pay when passing through a tolled tunnel or bridge. To ensure orthogonality, the last four attributes (ATT2, ATT3, ATT4 and ATT5) were used twice; thus, a total of nine attributes were used in the experimental design. The fractional factorial design method was applied to generate choice experiments, and a total of 27 profiles was produced as shown in Table 3. Each profile contained three alternatives, which consist of a choice set. In total, 27 (9 × 3) choice sets were generated. The choice sets were randomly distributed among the respondents (taxi drivers). To ensure data quality, each respondent received five or six choice sets. As there was a total of around 400 drivers in the sample, and each driver completed five or six choice sets, the total number of observations was around 2000. Figure 1 is an example of a choice set. 'S' means stay at the place where the previous passenger alighted and

Factor	Attribute	Sample size
Age: years	20–40	105
	41–45	102
	46–50	102
	51–	91
Gender	Male	399
	Female	1
Marital status	Single	29
	Married	355
	Divorced	16
Driving experience: years	0–5	103
	6–10	103
	11–17	96
	18–	98
Taxi ownership	Self owned	90
	Rented from taxi association	71
	Rented from individual owner	232
	Other	7

Table 1. Demographic data of drivers (number of observations = 400)

Attributes	Levels
<i>Wait at the place where previous customer alighted</i>	
(ATT1) Waiting time for the next customer at the same location: min	10, 15, 20
<i>Move to another place searching for customers</i>	
(ATT2) Time of searching for customer in a different location: min	5, 10, 15
(ATT3) Distance of searching for customer in a different location: km	1, 3, 5
(ATT4) Toll fee at toll bridges and tunnels: HK\$	0, 5, 10
(ATT5) Waiting time for the next customer at the moved location: min	3, 6, 9

Table 2. Attributes and levels used for the stated preference survey

Profile	Alternatives								
	(S)		(A)			(B)			
	Wait time (ATT1)	Journey time to A (ATT2)	Distance (ATT3)	Toll (ATT4)	Wait time (ATT5)	Journey time to B (ATT2)	Distance (ATT3)	Toll (ATT4)	Wait time (ATT5)
1	10	5	1	0	3				
2	10	5	3	10	3	15	3	10	6
3	10	15	5	5	3	10	5	5	9
4	10	5	1	0	6	10	5	10	6
5	10	10	3	10	6	5	1	5	9
6	10	15	5	5	6	15	3	0	3
7	10	5	1	0	9	15	3	5	9
8	10	10	3	10	9	10	5	0	3
9	10	15	5	5	9	5	1	10	6
10	15	5	3	5	3	10	3	5	6
11	15	10	5	0	3	5	5	0	9
12	15	15	1	10	3				
13	15	5	3	5	6	15	1	0	9
14	15	10	5	0	6	10	3	10	3
15	15	15	1	10	6	5	5	5	6
16	15	5	3	5	9	5	5	10	3
17	15	10	5	0	9	15	1	5	6
18	15	15	1	10	9	10	3	0	9
19	20	5	5	10	3	15	5	10	9
20	20	10	1	5	3	10	1	5	3
21	20	15	3	0	3	5	3	0	6
22	20	5	5	10	6	5	3	5	3
23	20	10	1	5	6	15	2	0	6
24	20	15	3	0	6	10	1	10	9
25	20	5	5	10	9	10	1	0	6
26	20	10	1	5	9	5	3	10	9
27	20	15	3	0	9	15	5	5	3

Table 3. Profiles generated for the stated preference survey

S: Waiting for the next passenger for 10 min in the same district.

A: Journey time to A is 15 min, distance is 5 km, toll fee paid is HK\$5, and waiting time for the next passenger at A is 9 min.

B: Journey time to B is 5 min, distance is 1 km, toll fee paid is HK\$10, and waiting time for the next passenger at B is 6 min.

Your choice: S A B

Figure 1. Example of a choice set

wait for the next passenger; 'A' and 'B' are the two alternative choices of going to another place to find the next passenger. Respondents were asked to choose their preferred solution among the three alternative situations. Interested readers are

referred to Louviere *et al.* (2000) and Wang and Li (2005) for more details and recent advances concerning the experimental design of SP methods.

The questionnaires were distributed to taxi drivers at taxi stands and through taxi companies in three regions of Hong Kong: the New Territories, Hong Kong Island and Kowloon. Both day-shift and night-shift drivers were targeted. A total of 400 taxi drivers participated in the survey. There were around 40 000 taxi drivers in Hong Kong in 2001. The selected sample of 400 represents 1% of the total population of taxi drivers in Hong Kong. In the sample of 400 drivers only one female driver was interviewed. The reason for having only one female driver in the sample is that the proportion of female drivers was very low in 2000; this is also true now. Drivers were asked to complete the survey form after the end of each drop-off irrespective of the time of day. Each survey

form consisted of five or six different games for the driver to play.

3. METHOD

The logit-based search model that was used in the network analysis is based on the assumption that each driver tries to minimise the individual search time for customers and the expected search time is a random variable because of the random arrival of customers. The probability that a vacant taxi originating in zone s will meet a customer in zone r is specified by the following logit form

$$1 \quad P_{r/s} = \frac{\exp[-\theta(c_{sr} + w_r)]}{\sum_m \exp[-\theta(c_{sm} + w_m)]}$$

where c_{sr} is the generalised travel time via the shortest route from zone s to zone r , w_r is the waiting (search) time in zone r , and θ reflects the degree of uncertainty of customer demand and taxi services in the whole market. This parameter represents the general perception of taxi drivers concerning the uncertainty of finding customers in relation to their knowledge and familiarisation with the potential customer demand locations and general network conditions. The case of $r = s$ represents the probability that a taxi that takes a customer to zone s , searches for and meets the next customer in the same zone, and c_{ss} defines the intrazonal travel time.

The choice of decision-makers depends on their likes and dislikes of a set of independent parameters that vary for each decision-maker. These choices are expected to represent the realistic travel patterns of decision-makers and represent the alternatives. The concept of utility is used and the observed utility is defined as a linear combination of variables (Ortuzar and Willumsen, 1996). In this paper, the utility function of a taxi driver in searching for the next customer can be specified as

$$2 \quad U^i = V^i + \varepsilon^i$$

where V^i is the deterministic component of the utility and ε^i is the random component of the utility function for individual i . The choice set and alternatives are discussed in the previous section and listed in Table 3, and an example is shown in Figure 1. The deterministic component of this model can be expressed as

$$3 \quad V^i = \beta_{JT}x_{JT}^i + \beta_Dx_D^i + \beta_Tx_T^i + \beta_{WT}x_{WT}^i$$

where the observable components x_{JT}^i , x_D^i , x_T^i and x_{WT}^i are, respectively, the journey time, distance, toll, and waiting time for individual i ; and β_{JT} , β_D , β_T and β_{WT} are the respective coefficients. Assuming that the difference between the random terms of two alternatives is independently and identically (IID) Gumbel distributed (Domencich and McFadden, 1975), the probability of a taxi driver choosing a zone (alternative) in searching for the next customer will follow the logit form in Equation 1. The significance of the model-fitting results will

provide empirical evidence of the validity of such a search model.

From utility function (Equation 3), we can obtain the following values of time and distance

$$4 \quad v_{JT} = \frac{\beta_{JT}}{\beta_T}$$

$$5 \quad v_D = \frac{\beta_D}{\beta_T}$$

$$6 \quad v_{WT} = \frac{\beta_{WT}}{\beta_T}$$

where v_{JT} is the value of journey time, v_D is the value of distance, and v_{WT} is the value of waiting time. The weighting that a taxi driver places on waiting compared to journey time can be determined by

$$7 \quad \gamma = \frac{\beta_{WT}}{\beta_{JT}}$$

which reflects the taxi driver's preference for waiting over moving to find a customer.

As driver behaviour varies with driver age, experience and marital status (Kim *et al.*, 2005), the different groups of drivers will cruise in dissimilar patterns and directions when they search for customers for their taxis. The behaviour pattern of drivers needs to be understood as unnecessary cruising causes traffic congestions on the road as well as reducing the income of the taxi driver. Therefore, the developed logit model is used to explain the driver behaviour, which is a function of their personal and operational characteristics. The model was run for different groups, and the coefficients, log likelihood ratios, and significance of the variables were obtained. The Watson and Westin pooling test approach was used for different market segments to analyse the sensitivity of the driver characteristics (Loo *et al.*, 2006; Watson and Westin, 1975; Wong *et al.*, 2008b).

The test was based on the log likelihood ratio, which was calculated by

$$8 \quad LR = -2(L_R - L_U)$$

where L_R is the log likelihood for the base model that is calibrated for the combined dataset with all market segments, and 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, when the test statistics exceed the threshold value that is specified for the chi-squared distribution at a chosen level of significance (Wong *et al.*, 2008b). The degree of freedom was calculated as the difference between the sum of the number of

parameters of the sub-models and the number of parameters of the base model.

4. RESULTS AND DISCUSSION

4.1. Results of multinomial logit model

Table 4 shows the results for the MNL model of the customer-searching behaviour of vacant taxis. The model is significant at the 1% level, which shows that Equation 1 is able to describe the search behaviour of vacant taxis. Journey time, toll, and waiting time were significant factors in customer searching at the 1% level. Distance was significant at the 5% level. From the model, the values of journey time and waiting time were HK\$26/h and HK\$34/h, respectively. According to information from the Transport Department of Hong Kong, the monthly taxi rental cost of a rentee-driver per shift in 2000 was about HK\$7653, and the number of operating days of a rentee-driver per month was about 27.3. With 12 h per shift, the rental was about HK\$23/h, which can be considered the time opportunity cost of a taxi driver, and is consistent with the values of the journey time and waiting time perceived by the drivers in the present study.

The results show that after dropping off a customer, taxi drivers preferred to cruise on the roads rather than wait to find the next customer, with the wait time weighting 31% higher than the journey time, meaning that 1 min of wait time was equivalent to 1.31 min (γ) of journey time.

Variables	Coefficient	t-statistic
Journey time: min	-0.0816	-9.804*
Distance: km	-0.0471	-2.143†
Toll: HK\$	-0.1886	-17.431*
Waiting time: min	-0.1072	-12.559*
Log likelihood function		-1835.518*
Number of observations		2000
ν_{JT}		HK\$26/h
ν_D		HK\$0.25/km
ν_{WT}		HK\$34/h
γ		1.31

* Significant at the 1% level.
† Significant at the 5% level.

Table 4. Results of the MNL model

The value of distance was HK\$0.25/km, but was significant only at the 5% level. Based on information from the Transport Department, the average daily cost of fuel for a taxi was about HK\$115 and the average daily operating mileage was about 318 km, resulting in an average fuel cost of HK\$0.36/km. Apparently, taxi drivers are willing to trade one-third of the fuel cost for the opportunity to find a customer away from their current location. However, it should be remembered that drivers may not be fully aware of their fuel costs for a given travel segment, and they may also consider other marginal costs when travelling. Therefore, this trade-off needs further investigation in a future study.

4.2. Results of market segmentation analysis

Table 5 shows the results of the market segmentation analysis. 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. The likelihood ratios were calculated on the basis of Equation 8. The age of driver and driving experience were approximately equally distributed into four different groups. The age and marital status of taxi drivers and taxi ownership status were significant at the 1% level, whereas the driving experience of a taxi driver was significant at the 5% level. Table 6 shows the detailed results of each of these market segments with the significant differences. Although the distance factor was significant in the base model, it was insignificant in most of the results in the market segments. It is doubtful that the significance of the distance factor in the base model is derived only by chance, and that taxi drivers do not consider travel distance as an important factor in searching for customers. The proximity factor in the searching mechanism is largely governed by the journey time, which reflects the opportunity cost.

4.3. Taxi ownership

Of the taxis drivers who were interviewed, around 75% rented either from a taxi association or from an individual taxi owner, of which 58% rented from individual taxi owners, and 22% owned their taxis. Table 6 shows that taxi drivers who rented their taxis from individual owners showed the highest disutility for journey time, toll, and waiting time for finding customers, probably because many of these night-shift drivers drove a taxi as a part-time job in addition to their regular day job. Therefore, these drivers were the most cost sensitive. Self-

Attribute	L_R	L_U	Degree of freedom	Likelihood ratio
Personal characteristics				
Age of the driver	-1835.158	-1825.469	4	19.378*
Driving experience	-1835.158	-1818.189	4	33.398†
Marital status	-1835.158	-1814.770	3	40.776*
Operational characteristics				
Service area	-1835.158	-1803.019	7	64.278
Taxi ownership	-1835.158	-1824.873	4	20.570*
Travel distance (km)	-1835.158	-1782.006	4	106.303

* Significant at the 1% level.
† Significant at the 5% level.

Table 5. Results of the market segmentation analysis

Taxi ownership								
Variables	Rented from a taxi association		Rented from an individual owner		Self owned		Other	
	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic
Journey time	-0.059	-3.094*	-0.095	-8.426*	-0.075	-4.406*	0.012	0.155
Distance	-0.011	-0.225	-0.052	-1.749	-0.083	-1.804	0.109	0.529
Toll	-0.157	-6.622*	-0.210	-13.996*	-0.157	-7.219*	-0.375	-3.352*
Waiting time	-0.066	-3.458*	-0.128	-11.054*	-0.099	-5.608*	0.002	0.035
ν_{JT}	HK\$22/h		HK\$27/h		HK\$29/h		N.A.	
ν_D	N.A.		N.A.		N.A.		N.A.	
ν_{WT}	HK\$25/h		HK\$37/h		HK\$38/h		N.A.	
γ	1.12		1.35		1.32		N.A.	
Experience								
Variables	0-5 years		6-10 years		11-17 years		> 17 years	
	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic
Journey time	-0.083	-5.368*	-0.085	-5.072*	-0.095	-5.187*	-0.064	-3.799*
Distance	-0.023	-0.548	-0.050	-1.144	-0.060	-1.261	-0.058	-1.294
Toll	-0.134	-7.081*	-0.206	-9.122*	-0.225	-9.233*	-0.214	-9.398*
Waiting time	-0.101	-6.167*	-0.110	-6.380*	-0.115	-6.304*	-0.107	-6.240*
ν_{JT}	HK\$37/h		HK\$25/h		HK\$25/h		HK\$18/h	
ν_D	N.A.		N.A.		N.A.		N.A.	
ν_{WT}	HK\$45/h		HK\$32/h		HK\$31/H		HK\$30/h	
γ	1.22		1.29		1.21		1.67	
Marital status								
Variables	Married		Divorced		Single			
	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic		
Journey time	-0.081	-9.083*	-0.080	-2.085†	-0.115	-3.492*		
Distance	-0.066	-2.805*	0.026	0.230	0.089	1.064		
Toll	-0.187	-16.207*	-0.217	-3.935*	-0.216	-4.899*		
Waiting time	-0.106	-11.731*	-0.269	-4.350*	-0.094	-2.941*		
ν_{JT}	HK\$26/h		HK\$22/h		HK\$32/h			
ν_D	HK\$0.35/km		N.A.		N.A.			
ν_{WT}	HK\$34/h		HK\$74/h		HK\$26/h			
γ	1.22		3.36		0.82			
Age of the driver								
Variables	20-40 years old		41-45 years old		46-50 years old		> 50 years old	
	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic
Journey time	-0.079	-5.076*	-0.100	-5.762*	-0.073	-4.535*	-0.079	-4.293*
Distance	-0.029	-0.699	-0.012	-0.272	-0.079	-1.829	-0.076	-1.594
Toll	-0.155	-7.901*	-0.215	-9.378*	-0.192	-8.978*	-0.204	-8.521*
Waiting time	-0.096	-5.907*	-0.113	-6.452*	-0.131	-7.474*	-0.096	-5.393*
ν_{JT}	HK\$31/h		HK\$28/h		HK\$23/h		HK\$23/h	
ν_D	N.A.		N.A.		N.A.		N.A.	
ν_{WT}	HK\$37/h		HK\$32/h		HK\$41/h		HK\$28/h	
γ	1.22		1.13		1.79		1.22	

* Significant at the 1% level.

† Significant at the 5% level.

Table 6. Results of the MNL models calibrated for each sub-group

owned taxi drivers have to bear mortgage and maintenance costs, and are thus more sensitive to costs than those taxi drivers who rent taxis from taxi associations, although less cost sensitive in comparison with the drivers who rent taxis from individual owners. The drivers who rent their taxis from taxi associations gave the least waiting time cost per hour. Only 12% of these drivers cruised to search for customers, whereas more than 30% of the drivers from the other two groups cruised to search for customers.

4.4. Driver experience

In this study, the driving experience was defined as the number of years of working experience as a taxi driver. Experienced drivers with 11–17 years of driving experience gave the highest coefficient or disutility for searching for customers. The least experienced drivers with less than 5 years of working experience gave the lowest disutility for distance, toll, and waiting time. The most experienced drivers were the group least concerned about the customer-searching time.

Inexperienced drivers had the highest values for both journey and waiting time costs in comparison with the other groups. As these drivers were new to the market, they needed to work harder to find customers. In contrast, the most experienced drivers with more than 17 years of experience gave the lowest journey time cost (HK\$18/h), which means that 67% of the most experienced drivers moved to a different place once they had dropped off their previous customers.

4.5. Driver marital status

According to the survey data, 89% of the drivers were married and only 7% were single. Married drivers gave the highest disutility for search distance whereas divorced drivers gave the highest disutility for both toll and waiting time. Married drivers gave the lowest coefficient on the toll to be paid. The waiting time cost of divorced drivers was significantly higher than that of the other groups. However, the waiting time cost of single drivers was lower than their journey time cost. The results showed that to find their next customers, 100% of divorced drivers moved out of the previous customer drop-off area and 100% of single drivers waited. Notwithstanding the high level of significance at 1% as shown in Table 5, it is not clear why marital status affects drivers' behaviour, and there may possibly be some confounding factors that might exist but not be included in the dataset. Careful examination of this interesting factor is needed in future research.

4.6. Age group of the driver

Regarding age group, those in the 41–45 years age group gave the highest coefficient for toll and journey time, and those in the 46–50 years age group gave the highest disutility for the searching distance and waiting time. Drivers below 40 years of age gave the lowest coefficient for toll and waiting time, which suggests that young drivers are less concerned about the toll to be paid, and do not mind waiting a little longer for a customer. The two groups of middle-aged drivers were more concerned about the searching time, distance, and toll to be paid. Drivers under 40 years of age and over 50 years of age gave the highest and lowest journey time costs, respectively, although drivers in both groups waited for a new customer after dropping off the previous customer. Drivers between 46 and 50 years of age showed the highest mobility in searching for

customers, as their journey time cost was very low in comparison with their waiting time cost per time unit.

In the foregoing analyses, the effects of the time of day or network congestion state were omitted as they were not explicitly specified in the SP survey. However, the values of the utility coefficients are expected to differentiate, depending on the prevailing network conditions such as the reliability of stated travel times. Therefore, the values of the coefficients presented correspond to drivers' general perception of the average daily network conditions. In addition, there are also other interesting factors, such as the number of hours in the shift, number of hours left to finish the shift, number of customers so far in the shift, and type of operating area (e.g. residential and business), that are not considered in this study, but could be interesting topics for future investigation. This, however, poses a challenge because of the large number of combinations in terms of variables and their attributes involved in the design of the SP survey.

5. CONCLUSIONS

The results of this study demonstrate the validity of the MNL choice model for the customer-searching behaviour of taxi drivers, which has been commonly used in network equilibrium models of urban taxi services (Wong and Yang, 1998; Wong *et al.*, 2001, 2002, 2005, 2008a; Yang and Wong, 1998; Yang *et al.*, 2001, 2002), based on the stated preference survey results of a survey that was conducted of 400 taxi drivers. The journey time, toll, and waiting time were found to be significant factors in the choice decision at the 1% level, whereas distance was significant at the 5% level. However, the market segmentation results revealed that distance was not significant in most of the market segments. The values of journey time and waiting time were found to be HK\$26/h and HK\$34/h, respectively, which reflects well the opportunity cost of the taxi operation. The taxi drivers' weighted waiting time cost 30% higher than journey time cost. The market segmentation results also showed that driver age, experience, marital status, and taxi ownership status affect the customer-searching behaviour of vacant taxis. These findings will be useful for the application of the network equilibrium models of urban taxi services, in which there are more realistic representations of different groups of taxi drivers with different search behaviours.

ACKNOWLEDGEMENTS

The work described in this paper was jointly supported by grants from the Hong Kong Research Grants Council of the Hong Kong Special Administrative Region (project numbers: POLYU5202/06E, HKUST6212/07E), the University of Hong Kong (10207394), the National Natural Science Foundation of China (70629001), and the National Basic Research Program of China (2006CB705500).

REFERENCES

- Arnott R (1996) Taxi travel should be subsidized. *Journal of Urban Economics* 40(3): 316–333.
- Beesley ME and Glaister S (1983) Information for regulation: the case of taxis. *The Economic Journal* 93(371): 594–615.
- Cairns RD and Liston-Heyes C (1996) Competition and regulation in the taxi industry. *Journal of Public Economics* 59(1): 1–15.

- De Vany AS (1975) Capacity utilization under alternative regulatory constraints: an analysis of taxi markets. *Journal of Political Economy* 83(1): 83–94.
- Domencich T and McFadden D (1975) *Urban Travel Demand: A Behavioural Analysis*. North-Holland, Amsterdam.
- Douglas GW (1972) Price regulation and optimal service standards: the taxicab industry. *Journal of Transport Economics and Policy* 6(2): 116–127.
- Foerster JF and Gilbert G (1979) Taxicab deregulation: economic consequences and regulatory choices. *Transportation* 8(4): 371–387.
- Frankena MW and Pautler PA (1986) Taxicab regulation: an economic analysis. *Journal of Research in Law and Economics* 9: 129–165.
- Golias JC (2003) Examining sensitivity of impact of taxi traffic on capacity and delays at urban road sections. *ASCE Journal of Transportation Engineering* 129(3): 286–291.
- Hensher DA (2007) Some insights into the key influences on trip-chaining activity and public transport use of seniors and the elderly. *International Journal of Sustainable Transportation* 1(1): 53–68.
- Kim H, Oh JS and Jayakrishnan R (2005) Effect of taxi information system on efficiency and quality of taxi services. *Transportation Research Record* 1903: 96–104.
- Lee KT, Lin DJ and Wu PJ (2005) Planning and design of a taxipooling dispatching system. *Transportation Research Record* 1903: 86–95.
- Loo BPY, Wong SC and Hau TD (2006) Introducing alternative fuel vehicles in Hong Kong: views from the public light bus industry. *Transportation* 33(6): 605–619.
- Loo BPY, Leung BSY, Wong SC and 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(4): 249–266.
- Louviere JJ, Hensher DA and Swait J (2000) *Stated Choice Methods: Analysis and Applications*. Cambridge University Press, Cambridge, UK.
- Manski CF and Wright JD (1976) Nature of equilibrium in the market for taxi services. *Transportation Research Record* 619: 296–306.
- Ortuzar JD and Willumsen LG (1996) *Modelling Transport*, 2nd edn. Wiley, New York.
- Petzall J (1995) The design of entrances of taxis for elderly and disabled passengers – an experimental study. *Applied Ergonomics* 26(5): 343–352.
- Pucher J and Renne JL (2003) Socioeconomics of urban travel: evidence from the 2001 NHTS. *Transportation Quarterly* 57(3): 49–77.
- Schroeter JR (1983) A model of taxi service under fare structure and fleet size regulation. *Bell Journal of Economics* 14(1): 81–96.
- Shreiber C (1975) The economic reasons for price and entry regulation of taxicabs. *Journal of Transport Economics and Policy* 9(3): 268–279.
- Su F, Schmocker JD and Bell MGH (2010) Mobility scooters on load – a scheme complementing the existing special transport services in London. *International Journal of Sustainable Transportation* 4(2): 95–111.
- Wang D and Li P (2005) Does uniform design really work in stated choice modeling? A simulation study. *Transportmetrica* 1(3): 209–211.
- Watson PL and Westin RB (1975) Transferability of disaggregate mode choice models. *Regional Science and Urban Economics* 5(2): 227–249.
- Wong KI, Wong SC and Yang H (2001) Modeling urban taxi services in congested road networks with elastic demand. *Transportation Research Part B* 35(9): 819–842.
- Wong KI, Wong SC, Yang H and Tong CO (2002) A sensitivity-based solution algorithm for the network model of urban taxi services. *Proceedings of the 15th International Symposium on Transportation and Traffic Theory* (Taylor MAP (ed.)). Elsevier, Amsterdam, the Netherlands, pp. 23–42.
- Wong KI, Wong SC, Bell MGH 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(1): 81–104.
- Wong KI, Wong SC, Yang H and Wu JH (2008a) Modeling urban taxi services with multiple user classes and vehicle modes. *Transportation Research Part B* 42(10): 985–1007.
- Wong SC and Yang H (1998) Network model of urban taxi services: improved algorithm. *Transportation Research Record* 1623: 27–30.
- Wong SC, Wong CW and Sze NN (2008b) Attitudes of public light bus drivers to penalties to combat red light violations in Hong Kong. *Transport Policy* 15(1): 43–54.
- Yang H and Wong SC (1998) A network equilibrium model of urban taxi services. *Transportation Research Part B* 32(4): 235–246.
- Yang H, Lau YW, Wong SC and Lo HK (2000) A macroscopic taxi model for passenger demand taxi utilization and level of service. *Transportation* 27(3): 317–340.
- Yang H, Wong KI and Wong SC (2001) Modeling urban taxi services in road networks: progress, problem and prospect. *Journal of Advanced Transportation* 35(3): 237–258.
- Yang H, Wong SC and Wong KI (2002) Demand-supply equilibrium of taxi services in a network under competition and regulation. *Transportation Research Part B* 36(9): 799–819.
- Yang H, Ye M, Tang WH and Wong SC (2005a) Regulating taxi services in the presence of congestion externality. *Transportation Research Part A* 39(1): 17–40.
- Yang H, Ye M, Tang WH and Wong SC (2005b) A multi-period dynamic model of taxi services with endogenous service intensity. *Operations Research* 53(3): 501–515.

What do you think?

To discuss this paper, please email up to 500 words to the editor at journals@ice.org.uk. Your contribution will be forwarded to the author(s) for a reply and, if considered appropriate by the editorial panel, will be published as discussion in a future issue of the journal.

Proceedings journals rely entirely on contributions sent in by civil engineering professionals, academics and students. Papers should be 2000–5000 words long (briefing papers should be 1000–2000 words long), with adequate illustrations and references. You can submit your paper online via www.icevirtuallibrary.com/content/journals, where you will also find detailed author guidelines.