Guiding vacant taxi drivers to demand locations by taxi-calling signals: A sequential binary logistic regression modeling approach and policy implications

(submitted to Transport Policy, Special Issue of Taxi Research, for publication consideration)

Abstract

Taxi-calling signals (TCSs) have appeared in many cities to reveal passenger demand at locations away from the roadside to cruising vacant taxis to reduce search times for both vacant taxi drivers and customers. This study aims to find out the factors influencing vacant taxi drivers' customer-search decisions on whether to enter or bypass recommended areas while the drivers are cruising along a road with a series of TCSs. Observational survey data were collected and analyzed to understand the travel behavior of vacant taxi drivers. A sequential binary logistic regression (SBLR) model is first proposed to examine the dynamic decision-making process of vacant taxi drivers. A simulation model and a solution procedure are then developed by adopting the intervening opportunity modeling concept to validate the SBLR model. A sensitivity analysis is consequently conducted to show that the installation of TCSs can effectively increase the number of vacant taxis entering off-road locations for picking up customers. Potential policy implications are discussed.

Keywords: sequential binary logistic regression model, simulation model, taxi-calling signal, observational survey, taxi customer-search

1. Introduction

Taxi operations are an important element in a transportation system of any city. Taxis provide flexible and personalized services and are complementary to other public transport modes with their round-the-clock and door-to-door service capabilities. Taxi services are especially useful for passengers who are physically challenged, passengers in remote areas that are underserved by other public transport modes, and passengers carrying bulky luggage. Vacant taxi drivers commonly cruise on roads to search for customers in larger cities (e.g., New York City) where most of their trips are via street hail (Schaller, 2007). In other cities, for example, Singapore, taxis are dispatched to customers' orders (Liao, 2003). The emergence of ridesourcing platforms, for instance, Uber, are bringing competition to cities in both the cruising and dispatching markets. They have significantly increased taxi-customer meeting efficiencies. In modern day operations, taxis and customers can meet each other by street hailing, call centers, e-hailing apps, taxi stands, and taxi-calling signals (TCSs). (TCSs are designed to inform cruising taxi drivers about the occurrences of passenger demand. Figure 1 shows examples of TCSs found in Hong Kong, Singapore, and Guangzhou, China. These signals are usually flashboards placed at the entrances of access roads.) Each of these means offers a distinct function and matches demand and supply differently. For example, TCSs and Uber-like apps are offering different functions. TCSs provide more immediate demand information to taxis already cruising in the vicinity, whereas Uber-like apps are sending vacant taxis to where demand is located. Locations where TCSs are installed usually are non-core areas where Uber driver don't tend to concentrate. Therefore, users of these areas shall make advance reservations. Taxis, however, can respond much more quickly to TCSs, which means the total time spent before getting in a taxi can be reduced.

(a) Hong Kong (b) Singapore (c) Guangzhou, China

Figure 1. Examples of taxi-calling signals in Asian cities

In the conventional cruising market, the time and location of customers are often unknown to vacant taxi drivers. Conversely, the whereabouts of vacant taxis are also unknown to customers. The mismatch and inequality of taxi supply and passenger demand result in unnecessary empty trips made by vacant taxi drivers and prolonged customers' waiting time and walking distance for taxi services. Meanwhile, these empty trips have become a major contributor to worsening traffic congestion, excessive vehicle emissions, and additional taxi operating costs. Therefore, understanding the travel behaviors of both vacant taxi drivers and customers and improving network efficiency have become the underlying motivations of many taxi studies.

Traditionally, taxi studies can be broadly classified into six key areas of focus, namely (1) policy and regulation, (2) drivers' customer-search behaviors, (3) customers' taxi-search behaviors, (4) passenger demand modeling and prediction, (5) taxi supply modeling and prediction, and (6) network equilibrium. The recent emergence of taxi hailing mobile applications (apps) has also become one of the focal points of many researchers. Furthermore, taxis are also used as agents in fields such as traffic flow or operational research, but these realms are beyond the scope of this study and hence will not be discussed further.

Early studies focused primarily on different types of taxi regulations and deregulations. For example, Douglas (1972) developed an aggregate demand model from a microeconomic perspective to relate taxi service quality and taxi passenger demand under price regulation. De Vany (1975) examined the capacity utilization of regulated taxi markets under competition and monopoly. Schroeter (1983) developed a model of taxi services under both fare structure and fleet size regulation. Taxi deregulation was also discussed, for instance, by Cairns and Liston-Heyes (1996), Schaller (2007), and Barret (2003, 2010). One of their key findings is that the deregulation of fare and entry may not be optimal and an equilibrium in a deregulated industry does not exist.

A simple cruising model was also discussed in the work of Cairns and Liston-Heyes (1996). Their study marked the beginning of a series of studies on modeling the customer-search behavior of vacant taxi drivers during taxi cruising. Yang and Wong (1998) used a mathematical model to depict the customer-search behavior of vacant taxi drivers but only assume that their main objective during cruising is to minimize customer search time. The perceived profitability of vacant taxi drivers is not taken into an account. Szeto et al. (2013) investigated the zonal decisions made by vacant taxi drivers in customer-search by using 24-

hour taxi global positioning systems (GPS) data. A time-dependent logit-based model was developed. They found that the rate of return, the extended concept of the perceived profitability, is critical to vacant taxi drivers' search strategies. Meanwhile, Hu et al. (2012) proposed a probabilistic dynamic programming vacant taxi routing model. However, they only assume that a vacant taxi driver at an intersection selects a link to enter to minimize their expected search time for customers. They also ignore the competition between vacant taxis and the fact that the probability of successfully meeting a taxi customer along the way increases. Wong et al. (2014a) developed a taxi customer-search model using the intervening opportunity modeling approach to modeling and analyzing local vacant taxi movements in a cell-based network and introduced the concept of the probability of success in the customer-search decision-making process. Their study considers the competition between vacant taxis and the change in the probability of success along search paths. Based on this study, Wong et al. (2015a) developed a two-stage approach taking into account both zonal and local search factors to predict the movement pattern of vacant taxis in searching for customers.

In addition to cruising, vacant taxi drivers can meet customers by going to taxi stands. However, limited studies have been carried out on modeling this search process. One of the few studies was performed by Wong et al. (2014b), who used a sequential logit approach to model a bi-level decision-making process of a vacant taxi driver about whether to go to a taxi stand after dropping off the preceding customers and whether to wait at a taxi stand until meeting a customer after arriving there.

Dispatching is another method for vacant taxis meeting their customers. Singapore used GPS to match waiting taxi customers with the nearest vacant taxis over a decade ago (Liao, 2003). In the literature, quite a number of studies are also related to taxi dispatching. For example, Wang et al. (2014) reviewed the existing dispatch algorithm adopted by Singapore and proposed a new trip-chaining method to allocate the next pick-up point in close proximity to the previous drop-off point within a reasonable time frame. This method improves the overall efficiency and cost-effectiveness of the taxi reservation and dispatch system. Maciejewski et al. (2016) proposed a real-time assignment-based dispatching strategy that improves system efficiency. Also focusing on dispatching efficiency, Miao et al. (2016) developed another real-time strategy based on GPS data with a receding horizon control framework.

Customers are simultaneously searching for vacant taxis while vacant taxi drivers are seeking customers. A customer can hail a taxi on streets or walk to a taxi stand to get a taxi. To date, very few studies have focused primarily on customers' hailing behavior. Wong et al. (2015b) as one of the few conducted a stated-preference survey to taxi customers and calibrated a nested logit model to examine the hailing preferences of customers on streets. Other studies mostly examine or model the search behaviors of both vacant taxi drivers and customers. For instance, Situ et al. (2017) proposed a parallel ant colony optimization algorithm with the region-decomposition strategy. This study provides a new solution method to the taxi-passenger matching problem.

Aside from the logit model or its discrete choice model variants, passenger demand modeling also took various other approaches. Arnott (1996) is among the first to consider uniformly distributed passenger demand over a spatially and temporally homogeneous two-dimensional city which extends infinitely far in every direction. Wong et al. (2001) extended the previously mentioned Yang and Wong's (1998) model by adding the congestion factor and the elastic demand function to model taxi movements in a congested network.

During the passenger demand modeling, it is quite common that passenger demand or its elastic function is assumed to be known. These inputs can be obtained by demand prediction. Many studies aim to predict the occurrence and emergence of passenger demand using various techniques. For example, Moreira-Matias et al. (2012) calibrated and validated a predictive model based on time series forecasting techniques and real-time data collected from 63 taxi stands in Porto, Portugal. Qian et al. (2017) applied a Gaussian condition random field model to forecast short-term demand based on historical time series data. This model can predict demand up to the next 2.5 hours using only 4 hours of past data.

The mismatch of taxi supply and passenger demand can be relieved if the time and location of vacant taxis can be predicted. A few studies, therefore, have focused on estimating the spatio-temporal distribution of vacant taxis without relying on the use of customer-search models. For example, Wu et al. (2017) optimized taxi drivers' routing choices by mining average travel speed of taxis from GPS data. The study developed a spatio-temporal trajectory model that can recommend route choices to the next pick-up point, which effectively reduce driver cruising time and passenger waiting time.

A more common approach to determine the match or mismatch between taxi supply and passenger demand is based on the concept of network equilibrium. To the best of our knowledge, Yang and Wong (1998) developed the first equilibrium taxi network model to describe how both vacant and occupied taxis circulate over a given network to search for customers and provide transportation services with a given customer origin-destination demand pattern. Subsequently, more complication network equilibrium models were developed to reflect the realism. For example, Yang et al. (2002) investigated the nature of demand and supply equilibrium in a regulated market and developed a network equilibrium model that can be used to determine a number of system performance measures, such as the utilization rate for taxi and the level of service quality, under fare structure and fleet size regulation in an either competitive or monopoly market. Yang et al. (2005) later investigated regulations in the taxi market and the nature of equilibrium by taking into account congestion externalities and adopting a realistic distance-based and delay-based taxi fare structure. Wong et al. (2008) added multiple user classes, multiple taxi modes, and customer hierarchical modal choice into the conventional single-class network equilibrium model of taxi services. The equilibria of bilateral searching and meeting between vacant taxis and customers on networks were studied by Yang et al. (2010), who proposed a meeting function in given zones where supply can meet demand therein. Unlike previous static network equilibrium models, Long et al. (2017) developed a dynamic taxi network equilibrium model to take into account the temporal variation of passenger demand.

The traditional cruising and hailing modes are now challenged by the emerging ridesourcing platforms such as the Uber and Lyft apps. Therefore, one of the evolving research areas is the taxi market with taxi hailing apps. Harding et al. (2016) discussed in depth the effects of taxi hailing apps on taxi markets, in particular, the possibility of collusion and monopoly. They also argued that regulation is needed to address the issues. Cramer and Krueger (2016) analyzed UberX in comparison with traditional taxis (without relying on taxi apps to get customers) and found that UberX drivers are more efficient as they spend more time transporting passengers. Wang et al. (2016) modeled the taxi market with a single hailing app by the aggregate and static approach and examined the impacts of taxi-hailing platform's pricing strategies on taxi market performance. Also inspired by the latest taxi-calling technologies, Zhang and Ukkusuri (2016) developed a modeling framework to determine optimal fare and fleet size strategies under stochastic demand for the taxi market with taxi apps and used New York City as a case study for the model application.

With the aforementioned studies, we now have a better understanding of how the taxi market is regulated, how taxi drivers search, how taxi customers hail, and how they correlate in terms of demand, supply, network equilibria, and taxi apps. Nevertheless, there is still serious difficulty in meeting passenger demand accurately because the timing and locations of customers may not perfectly match how vacant taxi drivers circulate on roads. In addition, some customers are further challenged by prolonged waiting times and long walking distance to locations where the chances of getting a vacant taxi are higher. More research is needed to reduce the customers' waiting and walking times and meet the passenger demand more accurately.

As aforementioned, vacant taxi drivers and waiting customers are brought together by means of street hailing (cruising), call centers (dispatching), e-hailing apps, taxi stands, and taxi-calling signals (TCSs), all of which, except for TCSs, have been studied quite extensively. TCSs are designed to inform cruising taxi drivers about the occurrences of passenger demand. They are commonly seen at the entrances of residential estates, hotels, hospitals, and commercial buildings where the actual customer pickup points are located at some distance from the roadside and where demand fluctuates. These signals are usually flashboards placed at the entrances of access roads (see Figure 1 for examples). Upon seeing a signal, vacant taxi drivers make decisions on whether to enter the access road to the TCS-recommended location or bypass it. TCSs, in theory, can also reduce the walking distance of customers to taxis. They also reduce the number of additional circulations of vacant taxis to locations where demand cannot be directly identified from the roadside. As TCS-installed locations may not be isolated, it is also common for vacant taxi drivers to find a sequence of TCSs along a stretch of a road where a series of decisions need to be made.

Among the three cities with TCSs mentioned above, Hong Kong has unique characteristics. Due to its hilly topography, many buildings and estates in Hong Kong are constructed on levels different from those of main roads. Consequently, passenger demand from these locations cannot easily be observed by vacant taxi drivers cruising on the main roads. TCSs have, therefore, become a popular tool for the users of these locations to hail taxis.

Taxis are one of the main transport modes in Hong Kong. With a fleet of 18,163 vehicles at present, the daily ridership of taxis in Hong Kong is close to 1 million, which is 8% of all motorized passenger trips and is the second highest in the world's 28 largest cities (Transport Department, 2016; Land Transport Authority of Singapore, 2011). This mode of transport, however, creates problems. Hong Kong's taxi fleet on average takes up 25% of road space (Transport Department, 1986-2009), leading to a large reduction in road capacity for other road-based transport modes such as buses and cars. Similar to other large cities, empty trips made by vacant taxis are also contributing to road congestion and roadside air pollution severely.

All taxi hailing modes exist in Hong Kong. However, due to the extra HK$5 cost for phone-booking, dispatching has a negligible impact on the taxi market (Szeto et al., 2013). Online taxi hailing apps have not gained significant market penetration in Hong Kong as Uber, the largest taxi hailing app operator, faces legal challenges and other smaller app operators cover only a limited number of vehicles (Ng, 2015). TCSs have, therefore, continued to prevail in matching supply to demand. However, little is known about the effectiveness in attracting vacant taxi drivers to the recommended location as a result of installing TCSs. It is also unclear whether any other factors may influence vacant taxi drivers' customer-search decisions to enter or bypass TCS-recommended areas.

To address the above issues, this study proposes a sequential binary logistic regression (SBLR) model to determine factors that influence vacant taxi drivers' decisions on whether to enter or bypass recommended areas while the drivers are cruising along a road with a series of TCSs. Observational survey data of Hong Kong were collected and analyzed to understand the travel behavior of vacant taxi drivers and calibrate the model. This study further develops a simulation model and proposes a solution procedure to validate the SBLR model. A sensitivity analysis is also conducted to examine the effectiveness of TCSs in increasing the number of vacant taxis entering off-road locations for picking up customers. Potential policy implications are discussed accordingly.

The contributions of this paper include the following.

- It calibrates an SBLR model to determine the influential factors in making vacant taxi drivers' decisions related to TCSs;
- It develops a simulation model and a solution procedure to obtain a solution of the percentage of vacant taxi entries and hence validate the SBLR model; and
- It demonstrates that TCSs can be effective in increasing the number of vacant taxis entering off-road locations for picking up customers, and provides taxi policy recommendations.

The remainder of this paper proceeds as follows. Section 2 explains the collection procedures of data and provides the descriptive results of the case study; Section 3 explains the research methodology. Section 4 discusses the results of model calibration and validation, examines the effectiveness of TCSs via sensitivity analysis, and presents potential taxi policy implications. A conclusion is drawn in Section 5.

2. Data

2.1 Data collection

Several hundred TCS installations can be found throughout Hong Kong. The selected study area is located along Stubbs Road in Happy Valley (a residential area) with four TCS-installed sites as shown in Figure 2. This area has a high presence of taxis in circulation (about 43% of all vehicles on the road were observed to be taxis). This area is poorly served by public transport with only two bus routes, two minibus routes (which carry a maximum of 16 seated passengers and mainly serve as feeder services), and no direct access to underground passenger rail. According to the planning regulations for a new development in this area, no additional parking space is permitted because of the congested road condition. Taxi passenger demand in the study area is consequently expected to be high. The study area consists of four sites. These four sites, from west to east, include a residential estate (Site A), a mixed-used site with residential complex and schools (Site B), another residential estate (Site C), and a hospital (Site D). Each site has an access road connecting to Stubbs Road by a priority junction with TCS installed at the entrance.

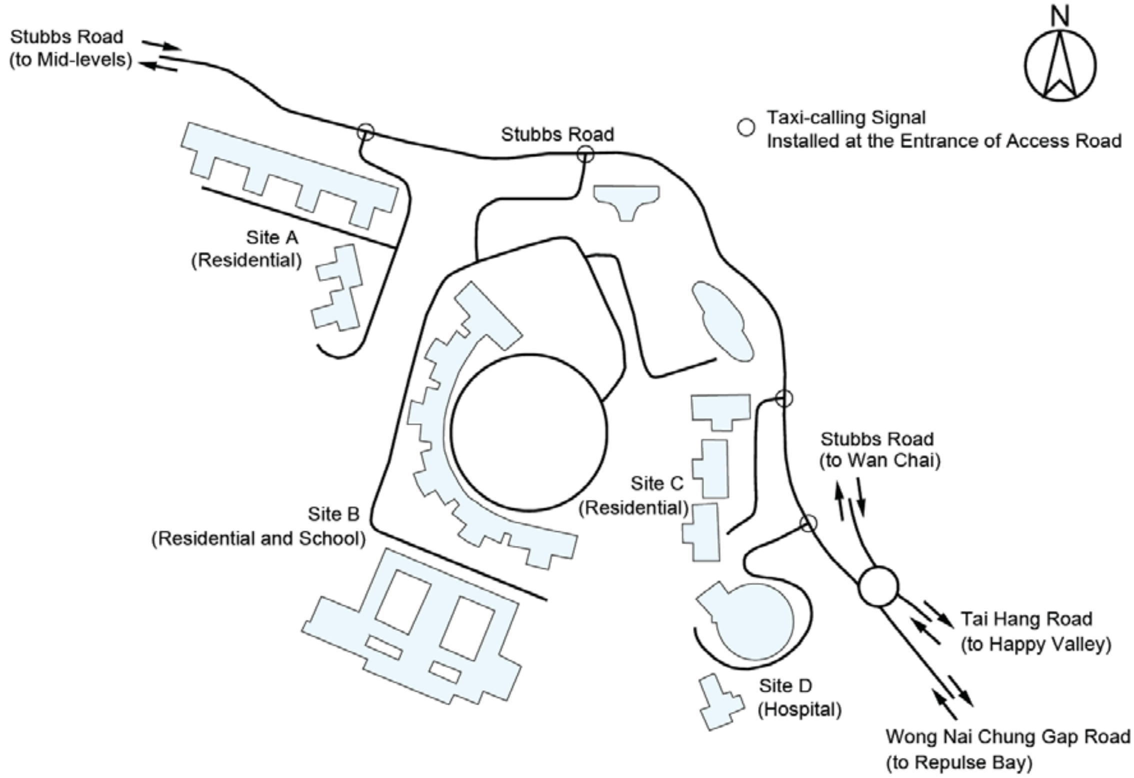

Figure 2. Illustration of the selected sites and road network in the vicinity

Data collection was conducted on six weekdays (excluded Saturday, Sunday and public holiday) in October 2015 at the entrances of the target sites during four study time periods (07:30-09:30, 09:30-11:30, 11:30-13:30, and 13:30-15:30). The items of data collected include: (1) TCS statuses; (2) time and occupancy statuses of all bypassing taxis as well as taxis entering and leaving each site; (3) number of taxis (including both vacant and occupied taxis) entering each of the sites immediately ahead of each observed vacant taxi; and (4) the travel direction of each vacant taxi. In total, 6,705 valid observations were recorded in the four sites. Data were recorded on-site by surveyors while video footages were also taken for the data cross-validation purpose.

2.2 Descriptive results

Observations show that not all the vacant taxi drivers strictly follow the signal display. Table 1 presents the vacant taxi drivers’ responses to TCSs in the study area. It is clear that an “off” signal is highly influential to vacant taxi drivers’ decisions of whether to bypass a site, whereas an “on” signal is comparatively less influential. Nevertheless, the two residential sites have shown highly positive responses to the changes of signals, suggesting that land-use alone may be embedded in vacant taxi drivers’ perceptions and hence result in their straightforward reactions. On the other hand, the attractiveness of displaying an “on” signal is much weaker in the two non-residential sites. For Site B (mixed-use) in particular, about 40% of vacant taxis bypassed the site even while the signal remained “on”. However, it received the highest percentage of vacant taxis entering the recommended site while the signal was “off”. The descriptive results suggest that signal status is not the sole factor that affects a vacant taxi driver’s decision. Differences in land-use types may have induced certain variations although their relationship is not conclusive. Other attributes are also believed to have effects on vacant taxi drivers’ decisions.

Table 1. A summary of signal statuses and the distribution of decisions of vacant taxi drivers

Signal Status	Decision of vacant taxi drivers	Frequency (Percentage)[a]			
		Site A	Site B	Site C	Site D
On	Enter	137 (96.5%)	175 (59.9%)	79 (95.2%)	83 (72.8%)
	Bypass	5 (3.5%)	117 (40.1%)	4 (4.8%)	31 (27.2%)
	Sub-total	142	292	83	114
Off	Enter	39 (2.6%)	98 (6.4%)	43 (2.8%)	41 (2.7%)
	Bypass	1479 (97.4%)	1441 (93.6%)	1468 (97.2%)	1464 (97.3%)
	Sub-total	1518	1539	1511	1505
	Total	1660	1831	1594	1619

Note: [a] The values in brackets represent the percentages of vacant taxi drivers with the corresponding decisions.

3. Research methodology

3.1 Modeling the decision-making process

Figure 3 illustrates the decision-making process of vacant taxi drivers in both eastbound and westbound traffic approaching the study area. Consider a vacant taxi driver traveling eastbound as an example (shown at the top-right corner of Figure 3). He/she first approaches Site A and upon reaching there, the TCS at the entrance of Site A is visible. A decision on whether to enter or bypass the site is then made under the influence of a number of factors including the status of the TCS at that time. If this driver decides to enter the site, the result is either a successful pickup or a failure. A successful pick-up means that the taxi becomes occupied and does not need to go to other sites ("Occupied and leave"), whereas a failure causes the driver to continue to search for customers and go to the next site. The same decision-making process is repeated at the entrance of each of the subsequent sites until a customer is met or until the driver leaves the study area. The same process also applies to a westbound taxi driver (shown at the bottom-left corner of Figure 3). For an occupied taxi driver whose destination of preceding customers is one of the four sites of this study, after dropping off the preceding customers ("Taxi customer drop-off"), the taxi becomes vacant and the driver starts searching for customers in the study area.

The decision-making process of vacant taxi drivers at locations with a sequence of TCSs is similar to the concepts of the opportunity-type model (Stouffer, 1940) and intervening opportunity model (Schneider, 1959). Stouffer proposed that a trip is made according to the relative accessibility of opportunities for achieving the objective of trip-making rather than explicitly relating to distance. Schneider enhanced this idea by adding two hypotheses about human behavior: (1) The total travel time from an origin to a destination is minimized; and (2) the probability of choosing any potential destination is a constant and independent of the order of visit. Therefore, the model assumes that a traveler makes his/her trip as short as possible and that the trip is only lengthened if the traveler fails to find an acceptable destination at a lesser distance (Heanue and Pyers, 1966).

The decision-making process cannot be modeled by simultaneous discrete choice models (e.g., multinomial logit model and nested logit model) accurately. This family of models is commonly used to predict the choice of individuals from several alternatives based on utility maximization (Train, 2009). As suggested by Schneider (1959), this family of models assumes that the choices of individuals are made with the knowledge of all given alternatives. This assumption limits the capacity of this family of models to handle behavior that individuals may be faced with one alternative at a time with no information given for the following alternatives, which is the case of vacant taxi drivers passing through our study area.

As the four sites are arranged in sequence, vacant taxi drivers can only see the TCS status of a site and decide whether they enter or bypass the site one by one. The conventional binary logistic regression model is unable to depict this sequential decision-making process. Therefore, to correctly model vacant taxi drivers' responses to a series of TCS statuses, we propose a logit-based binary sequential regression model incorporating the concept of intervening opportunities to examine the effects of various factors on vacant taxi drivers' decisions. The first model of the family of sequential discrete choice models has been developed by Amemiya (1975) and has been extended and applied in other researches (e.g., Kahn and Morimune, 1979; McCullagh and Nelder, 1989; Fu and Wilmot, 2004; Sawtelle et al., 2011) to model the sequential travel decisions. It assumes that a set of alternatives is only presented to an individual after a choice has been made from the previous set of alternatives. The fundamental assumption for this model is that vacant taxi drivers have not decided on which site to enter before reaching the study area, but rather all decisions are made on the spot at the entrance of each site based on the existence of opportunities (e.g., whether the TCS is "on") to meet their objectives (i.e., finding a customer).

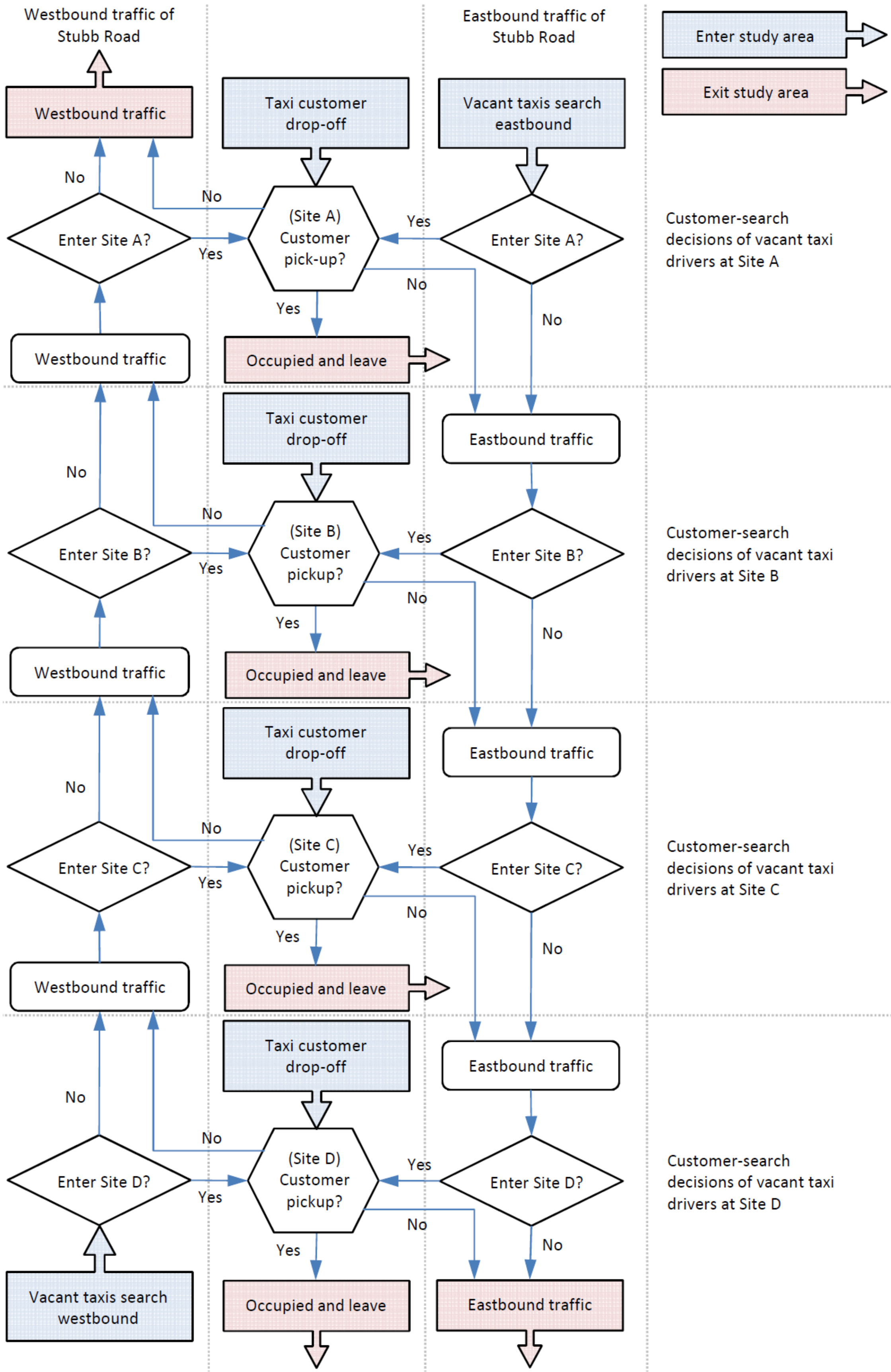

Figure 3. The decision-making process of vacant taxi drivers in the study area

3.2 *Notations*

The following notations are used throughout this paper.

s	Site index (also presents the proximity to the city center);
t	Time period index;
j	Vacant taxi driver index;
$I_{s(j)}$	Taxi-calling signal status of site *s* observed by vacant taxi driver *j*;
$V_{s(j)}$	Number of (observed) taxis ahead of vacant taxi driver *j* entering site *s*;
T	Dummy variable associated with the time period;
D	Dummy variable associated with the day of the week;
L_s	Dummy variable associated with the land-use type of site *s*;
H_s^t	Hourly (served) passenger demand of site *s* during time period *t*;
C_s^t	Time spent in site *s* for customers (in second) during time period *t*;
$W_{s(j)}^t$	Probability of vacant taxi driver *j* in westbound traffic making a left-turn into site *s* during time period *t*;
$E_{s(j)}^t$	Probability of vacant taxi driver *j* in eastbound traffic making a right-turn into site *s* during time period *t*;
$U_{s(j)}^t$	Deterministic utility capturing the factors that influence vacant taxi driver *j* entering site *s* during time period *t*;
α^{I}	Coefficient associated with the taxi-calling signal status;
α^{V}	Coefficient associated with the number of (observed) taxis ahead entering a site;
β^t	Coefficient associated with time period *t*;
β^{s}	Coefficient associated with the relative proximity of a site to the city center;
β^{D}	Coefficient associated with the day of the week;
β^{L}	Coefficient associated with the land-use type of a site;
β^{H}	Coefficient associated with the hourly passenger demand of a site;
β^{C}	Coefficient associated with the time spent in a site for picking up customers;
η	Coefficient associated with right-turn movements;
$\bar{W}_s^t$	Average probability of all vacant taxi drivers in westbound traffic making a left-turn into site *s* during time period *t*;
$\bar{E}_s^t$	Average probability of all vacant taxi drivers in eastbound traffic making a right-turn into site *s* during time period *t*;
P^t	Number of vacant taxis in westbound traffic entering the study area during time period *t*;
$\hat{P}_s^t$	Number of vacant taxis in westbound traffic at the entrance of site *s* during time period *t*;
Q^t	Number of vacant taxis in eastbound traffic entering the study area during time period *t*;
$\hat{Q}_s^t$	Number of vacant taxis in eastbound traffic at the entrance of site *s* during time period *t*;

A_s^t Availability of vacant taxis in site *s* during time period *t*;

O_s^t Number of occupied taxis entering site *s* and dropping off customers there during time period *t*;

Z_s^t Probability of successfully picking up customers in site *s* during time period *t*;

R^t Proportion of vacant taxis entering at least one of the sites during time period *t*;

R Proportion of vacant taxis entering at least one of the sites during the combined-all-four-period;

LL Log-likelihood of the estimated SBLR model;

ω Adjustment to each probability of success Z_s^t in the solution procedure;

ε Acceptable tolerance; and

k Iteration index.

3.3 Variables of discrete choice model

Based on the observational survey, the following variables are extracted and considered to be significant factors that influence the customer-search decisions of vacant taxi drivers. The variables are categorized into three groups as explained in the next three paragraphs.

Site dynamic variables: The model has two dynamic variables representing the instantaneous situation of a site when a vacant taxi driver makes his/her decision to enter or bypass the site, which includes the TCS status ($I_{s(j)}$) of site *s* and the number of taxis ahead entering site *s* observed by individual vacant taxi driver *j* ($V_{s(j)}$). When the signal is "on", it can be considered to be at least one customer waiting at the customer pickup point inside the site for taxis. Most likely, as indicated in the results presented in Table 1, the majority of vacant taxi drivers decide to enter the site. On the other hand, when the signal is "off", the vacant taxi drivers cannot judge whether there is a customer waiting there at that instant (because the customer may not inform the management office to light up the signal). The vacant taxi driver may have to decide based on other factors. The number of (observed) taxis ahead entering the site is used to investigate whether the decisions of preceding taxi drivers may affect the decisions of subsequent taxi drivers. This number is obtained by counting the number of entering taxis at a visible distance, which depends on the road alignment and the objects blocking the sightline. At least one taxi (which can be either a vacant taxi or an occupied taxi with a drop-off in the site) ahead entering the site implies that a taxi customer there will likely be picked up according to the first-come-first-served principle.

Site static variables: Six static components describing the general characteristics of the site are included in the proposed discrete choice model. The time period of the day (T) captures the changes in behaviors in association with temporal variations in the day. The location of a site (s) represents the relative proximity of that site to the city center. The day of the week (D) reflects the changes in behaviors in association with temporal variations in the week. The land use type (L_s) categorizes the exploitation of site *s* for commercial, industrial, residential, recreational, or other purposes. The hourly (served) passenger demand (H_s^t) represents the intrinsic popularity of site *s* during time period *t* and is used to reveal whether taxi drivers have had insights into their probability of successfully getting a customer. According to our observations in the sites, there have been cases where vacant taxis entered and then left the sites with no passenger, indicating that taxi supply is larger than passenger demand. It implies that the hourly *served* passenger demand is roughly equivalent to the hourly *actual* passenger demand. The time spent for searching for customers (C_s^t) is the time consumption of a vacant taxi driver entering site *s* from the entrance, waiting for and picking up customers (if any), and leaving the site during time period *t*, which represents the opportunity cost.

Right-turn movements: As a former British colony, Hong Kong follows the United Kingdom in driving on the left. For the operation of priority (uncontrolled) junctions, priority will be given to the vehicles going straight and turning left, and the right-turn vehicles have to give way. As indicated in Figure 2, westbound traffic turns left to enter the sites, which has a priority. Reversely, eastbound traffic has to give way to westbound traffic before making a right-turn into the sites. The proposed SBLR model incorporates this disadvantage of right-turn movements to explain the customer-search decisions of vacant taxi drivers from different traffic directions.

3.4 Model structure

By applying the SBLR model, the probabilities of an individual vacant taxi driver *j* entering site *s* during time period *t* are formulated as

$$W_{s(j)}^{t} = \frac{1}{1+\exp\left[-U_{s(j)}^{t}\right]}, \tag{1}$$

$$E_{s(j)}^{t} = \frac{1}{1+\exp\left[-U_{s(j)}^{t}-\eta\right]}, \text{ and} \tag{2}$$

$$U_{s(j)}^{t} = \left[\alpha^{\mathrm{I}} I_{s(j)} + \alpha^{\mathrm{V}} V_{s(j)}\right] + \left[\beta^{t} T + \beta^{\mathrm{s}} s + \beta^{\mathrm{D}} D + \beta^{\mathrm{L}} L_{s} + \beta^{\mathrm{H}} H_{s}^{t} + \beta^{\mathrm{C}} C_{s}^{t}\right], \tag{3}$$

where $W_{s(j)}^{t}$ and $E_{s(j)}^{t}$ denote the probability of an individual vacant taxi driver *j* entering site *s* during time period *t* from westbound and eastbound traffic, respectively. $U_{s(j)}^{t}$ denotes the deterministic utility capturing the factors that influence the choice of whether an individual vacant taxi driver *j* enters site *s* during time period *t*. The utility in Equation (3) consists of two parts, the dynamic components (attributes in the first square brackets) and the static components (attributes in the second square brackets). $I_{s(j)} = 0$ when the signal at the entrance of site *s* observed by vacant taxi driver *j* is "off", and 1 otherwise. $T = 0$ when the time period is 07:30-09:30, and 1 otherwise. $s = 1$ for Site D, 2 for Site C, 3 for Site B, and 4 for Site A. $D = 0$ when the day of the week is Tuesday, Wednesday or Thursday, and 1 otherwise (i.e., when the day of the week is Monday or Friday). $L_s = 0$ when the land-use type of site *s* is either mixed use or hospital, and 1 otherwise. α^{I} and α^{V} are the coefficients associated with the dynamic variables, and β^{t}, β^{s}, β^{D}, β^{L}, β^{H}, and β^{C} are the coefficients associated with the static variables. The coefficient of right-turn movements η is additionally introduced in Equation (2) to reflect the potential adverse impact of the right-turn movements to eastbound traffic.

It is worth mentioning that the static variables are the perceived values of the intrinsic characteristics of each site. We assume that vacant taxi drivers have the same perception of these characteristics and their choices can be attributed to these variables in addition to the dynamic instantaneous variables. For the purpose of this study, we assume that vacant taxi drivers do not make decisions prior to reaching the study area.

$\overline{W}_{s}^{t}$ and $\overline{E}_{s}^{t}$ represent the average probabilities of vacant taxis traveling westbound and eastbound entering site *s* during time period *t*, respectively. They are calculated as

$$\bar{W}_s^t = \frac{\sum_j W_{s(j)}^t}{\hat{P}_s^t}, \text{ and} \tag{4}$$

$$\bar{E}_s^t = \frac{\sum_j E_{s(j)}^t}{\hat{Q}_s^t}, \tag{5}$$

where $\hat{P}_s^t$ and $\hat{Q}_s^t$ denote the numbers of vacant taxis arriving at the entrance of site *s* during time period *t* from westbound and eastbound traffic, respectively. Based on both the SBLR and intervening opportunity modeling concepts, they can be expressed in Equations (6) - (8) below:

$$\hat{P}_s^t = \begin{cases} P^t & \text{, if } s = 1; \\ \hat{P}_{s-1}^t\left(1-\bar{W}_{s-1}^t\right) + \left(\hat{P}_{s-1}^t\bar{W}_{s-1}^t + \dfrac{O_{s-1}^t}{2}\right)\left(1-Z_{s-1}^t\right) & \text{, if } 1 < s \le 4, \end{cases} \text{ and} \tag{6}$$

$$Z_s^t = \min\left[\frac{H_s^t}{A_s^t}, 1\right], \tag{7}$$

where P^t denotes the number of westbound vacant taxis entering the study area during time period *t*, O_s^t denotes the number of occupied taxis entering site *s* and dropping off customers there during time period *t*, and Z_s^t denotes the probability of success in picking up customers in site *s* during time period *t*, which is defined as the hourly (served) passenger demand (H_s^t) over the availability of vacant taxis (A_s^t) subject to the condition that the probability is between zero and one inclusively. The probability of success controls the number of vacant taxis leaving the site without customers and return back to the main traffic stream to the next sites.

Equation (6) reflects a series of travel decisions to be made by the vacant taxi drivers in westbound traffic, who reach each of the sites in a sequence from $s = 1$ to 4 (i.e., from Sites D to A). For the case when $s = 1$ (i.e., for Site D), $\hat{P}_1^t = P^t$. For the case when $s = 2$ (i.e., for Site C), $\hat{P}_2^t$ equals $\hat{P}_1^t\left(1-\bar{W}_1^t\right)$, which represents the number of vacant taxis that have bypassed Site D, plus $\left(\hat{P}_1^t\bar{W}_1^t + \dfrac{O_1^t}{2}\right)\left(1-Z_1^t\right)$, which represents the total number of taxis failing to pick up customers in Site D, including the number of vacant taxis entering Site D but failing to get a customer there and the number of occupied taxis dropping off preceding customers in Site D but failing to get a customer there. (We assume that the vacant taxi drivers will travel along the original direction if they cannot reach a customer in a trial and that the occupied taxi drivers will select joining either eastbound or westbound traffic randomly after a failure in finding a customer.)

Likewise, in Equation (8), the eastbound vacant taxi drivers make a series of decisions of entering or bypassing a site, similar to those traveling westbound but in a reverse sequence of reaching each site.

$$\hat{Q}_s^t = \begin{cases} \hat{Q}_{s+1}^t\left(1-\bar{E}_{s+1}^t\right) + \left(\hat{Q}_{s+1}^t\bar{E}_{s+1}^t + \dfrac{O_{s+1}^t}{2}\right)\left(1-Z_{s+1}^t\right) & \text{, if } 1 \le s < 4; \\ Q^t & \text{, if } s = 4, \end{cases} \tag{8}$$

where Q^t denotes the number of eastbound vacant taxis entering the study area during time period t.

The availability of vacant taxis in site s during time period t (A_s^t) can then be calculated by summing the numbers of vacant taxis entered from westbound and eastbound traffic and the number of occupied taxis dropped off the preceding customers there:

$$A_s^t = \hat{P}_s^t \overline{W}_s^t + \hat{Q}_s^t \overline{E}_s^t + O_s^t . \tag{9}$$

The study objective is to find out the proportion of vacant taxis entering at least one of the sites for customers during time period t (R^t) and the (average) proportion during the combined-all-four-period (R). They can be determined as

$$R^t = \frac{\sum_s \left(\hat{P}_s^t \overline{W}_s^t + \hat{Q}_s^t \overline{E}_s^t \right)}{\sum_s \left(\hat{P}_s^t + \hat{Q}_s^t \right)}, \text{ and} \tag{10}$$

$$R = \frac{\sum_t \sum_s \left(\hat{P}_s^t \overline{W}_s^t + \hat{Q}_s^t \overline{E}_s^t \right)}{\sum_t \sum_s \left(\hat{P}_s^t + \hat{Q}_s^t \right)} . \tag{11}$$

3.5 Model calibration

In accordance with the principle of the sequential discrete choice model, vacant taxi drivers' decisions are made independently and sequentially. Decisions made at the entrance of a site can be described by a binary logistic regression model. Hence, the coefficients in Equations (2) and (3) (e.g., η, α^1, β^t, etc.) can be calibrated by using the maximum likelihood method under the objective function of

$$\text{Maximize LL} = \sum_j \sum_t \sum_s \left(\ln W_{s\,(j)}^t + \ln E_{s(j)}^t \right), \tag{12}$$

where LL is the log-likelihood of the estimated model, where $W_{s\,(j)}^t$ and $E_{s(j)}^t$ are the probabilities calculated by Equations (1) to (3).

3.6 Model validation

A simulation model is developed based on the flowchart in Figure 3 to validate the SBLR model, which estimates the proportion of vacant taxis entering at least one of the off-road locations (i.e., Sites A to D) for picking up customers during each time period t (R^t) and the proportion during the combined-all-four-period (R). The predicted proportions will then be used to compare with the observed values to show the prediction accuracy of the calibrated SBLR model.

In the simulation model, the probabilities of entering the sites under various scenarios are first estimated by applying Equations (1) to (3) with the explanatory variables and their associated coefficients calibrated in an earlier stage. All static variables take observed values on a Monday and the two dynamic variables are assigned with random values within reasonable ranges determined by the observed values. In the next step, Equations (4) to (9) are then applied for the availability of vacant taxis in each site by inputting the observed numbers of vacant taxis entering the study area from westbound and eastbound traffic, and the observed number of occupied taxis dropping off the preceding customers in each site. However, one of the variables in the equations, the probability of success in each site, cannot be pre-determined, as it is mutually dependent on the percentage of vacant taxi entries and hence the availability of vacant taxis. Therefore, a solution procedure is thus proposed to initialize and adjust the probability of success with an objective to have the predicted hourly (served) passenger demand to be mutually consistent with the observed value. The procedure is mainly based on the following:

$$Z_s^{t(k+1)} = \begin{cases} Z_s^{t(k)} - \omega & \text{, if } \dfrac{H_s^{t(k)} - H_s^t}{H_s^t} > \varepsilon; \\ Z_s^{t(k)} + \omega & \text{, if } \dfrac{H_s^t - H_s^{t(k)}}{H_s^t} > \varepsilon; \\ Z_s^{t(k)} & \text{, if } \left| \dfrac{H_s^t - H_s^{t(k)}}{H_s^t} \right| \le \varepsilon, \end{cases} \tag{13}$$

where $0 < \omega < 1$ and $\varepsilon > 0$. $H_s^{t(k)}$ and $Z_s^{t(k)}$ are the predicted hourly passenger demand and the predicted probability of success in site *s* during time period *t* in iteration *k*, respectively. If the observed hourly passenger demand is close to the predicted value in iteration *k* within the acceptable tolerance ε, no adjustment is made to the probability of success in the next iteration. Otherwise, the probability of success increases or decreases by an adjustment step ω. The updated probability of success in Equation (13) is then applied to Equations (4) to (9) for the predicted hourly (served) passenger demand. A convergent solution is achieved when the relative absolute difference between the predicted and observed hourly (served) passenger demand ($\left| \dfrac{H_s^t - H_s^{t(k)}}{H_s^t} \right|$) of all sites is less than the acceptable tolerance ε. Then, the proportions of taxis entering at least one site for picking up customers can be obtained in Equations (10) and (11).

The simulation model can be summarised by the following steps.

Step 1 – Computation of the probabilities of vacant taxi drivers entering the sites

Choose the values of $I_{s(j)}$ and $V_{s(j)}$ according to their empirical distributions, and input them as well as the values of static variables, including the observed values of H_s^t and C_s^t, into Equations (1) to (3) to compute the probabilities of vacant taxi drivers entering the sites in westbound and eastbound traffic.

Step 2 – Initialization of the solution procedure

Set ω and ε. Set iteration number $k = 1$ and initialize $Z_s^{t(k)}$.

Step 3 – Computation of the predicted hourly (served) passenger demand in the sites

Input the observed values of P^t, Q^t, and O_s^t into Equations (4) to (9) to obtain the predicted hourly (served) passenger demand in all sites *s* for all time periods *t*.

Step 4 – Convergence test

If $\max \left| \dfrac{H_s^t - H_s^{t(k)}}{H_s^t} \right| \le \varepsilon$, then stop and apply Equations (10) and (11) for determining the proportions of taxis entering at least one site for picking up customers. Otherwise, go to Step 5.

Step 5 – Adjustment of the probability of success

Adjust the probability of success according to Equation (13). Then, set $k = k + 1$ and go to Step 3.

The simulation model will be repeated for numerous times with different values of the dynamic variables for different periods and the combined-all-four-period. The predicted mean entry proportion and the associated standard deviation of each period can be obtained and recorded to compare with the corresponding observed value. The null hypothesis is that there is no difference between the observed value and predicted value. If the test statistic does not exceed the threshold value that is specified for the normal distribution at the chosen level of significance, we do not reject the hypothesis that the observed value is equal to the predicted value, and conclude that the model is accurate and robust.

4. Results and Discussion

4.1 Model calibration results

The statistical package SPSS was used to estimate the coefficients of the variables based on the maximum likelihood estimation. Table 2 below summarizes the estimated coefficients and their *t*-statistics. All the introduced variables are significant at the 5% level, and all the signs of coefficients are logical.

Table 2. Summary of estimated coefficients and their *t*-statistics

Explanatory variables	Control	Coefficients [*t*-statistics][a]
Taxi-calling signal (Signal on)	(Signal off / no signal)	5.00[b] [32.6]
Number of preceding taxis entering the site		-2.23[b] [-8.8]
Time of day (09:30-11:30)	(07:30-09:30)	0.27[c] [2.0]
Time of day (11:30-13:30)	(07:30-09:30)	0.96[b] [3.9]
Time of day (13:30-15:30)	(07:30-09:30)	1.01[b] [3.1]
Proximity to the city center		-0.63[b] [-9.8]
Day of the week (Mon / Fri)	(Tue / Wed / Thu)	-1.12[b] [-8.7]
Land-use type (Residential)	(Mixed use / Hospital)	-1.27[b] [-6.8]
Hourly passenger demand		0.08[b] [8.2]
Time spent for customer-search in the site		-0.01[b] [-2.9]
Right-turn movements	(Left-turn movements)	-2.08[b] [-14.8]

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

For the site dynamic variables, having a signal installed and showing "on" is highly effective in attracting vacant taxi drivers into a site where demand occurs, revealing that at least in the sites where signals are installed, the usefulness of these signals in directing taxis towards waiting customers is high. It is noted that the magnitude of the coefficient associated with a TCS is the highest among all other explanatory variables, which suggests the provision of the signal has the greatest influence on the vacant taxi drivers' decisions. Moreover, for a vacant taxi driver, the presence of another taxi entering the site ahead is highly discouraging. It suggests a tendency of being more conservative in vacant taxi drivers' expectations of passenger demand.

Vacant taxi drivers' customer-search decisions vary according to both day of the week and time of day. Vacant taxi drivers are less willing to enter the sites on Mondays and Fridays. In terms of time of day, with the morning peak hours of 07:30-09:30 as the baseline, vacant taxi drivers during this period are least likely to enter a site due to a higher overall demand and the associated expectation of the possibility of higher-yield customers elsewhere, which represents the selectiveness of vacant taxi drivers during the morning peak. The inclination to enter the site grows over time as the perceived probability of meeting a customer found from other places along their driving route is reduced. Thus, vacant taxi drivers are more eager to secure any possible passenger demand on the spot. Presumably, mid-afternoon (13:30-15:30) has the lowest demand as it situates just after lunch hour and before the start of an early afternoon peak. Consequently, vacant taxi drivers are more than twice as likely to enter the TCS recommended area as during the morning peak. The proximity to the city center, in addition, shows a negative impact (i.e., the closer to the city center a site is, the less likely a vacant taxi driver enters the site as the driver sees more opportunities elsewhere). This is to say that while driving away from the city center, vacant taxi drivers tend to be more active in entering.

Aside from the spatio-temporal characteristics, the model reveals that residential sites are decidedly less favorable to vacant taxi drivers than non-residential sites. In this study, the vacant taxi drivers are more responsive to TCS recommendations given at the hospital and at the site where schools are located. This behavior shows that a TCS has an unexpected benefit in serving the more vulnerable user groups. It may be due to not only taxi drivers' goodwill, but also a reflection of the anticipated certainty level associated with demand at non-residential sites. Hourly (served) passenger demand has a positive impact on drivers' entry proportion (and hence entry percentage) as higher demand relates to a greater chance of success, and such endogenous quality is influential to vacant taxi drivers. The average time spent for customer-search in each site has a slightly negative impact on vacant taxi drivers' decisions. Moreover, the model shows that the right-turn drivers are less likely to enter the sites than left-turn drivers because the right-turn movements have a lower priority to enter the sites for picking up customers as explained in Section 3.3.

4.2 Model validation results

Upon the completion of model calibration, model validation was conducted by simulating the percentages of vacant taxis entering at least one site for picking up customers. With an adjustment to the probability of success (ω) of 0.01, an acceptable tolerance (ε) of 0.05, and the predicted probability of success in all sites and time periods in the first iteration ($Z_s^{t(1)}$) of 0.30, the solution procedure provides a convergent solution. We performed 100 trials for each case. Table 3 shows the summary of validation results and provides the observed and predicted mean values of the percentage of vacant taxi entries for different periods and the combined-all-four-period. All the null hypotheses are not rejected, which implies that there is no significant difference between the observed and predicted values. Therefore, we confirm that the calibrated SBLR model is highly accurate and robust.

Table 3. Summary of validation results

Period	Percentage of vacant taxis entering at least one site for picking up customers		Two standard errors	Conclusion of hypothesis test[a]
	Observed	Predicted		
07:30-09:30	7.26%	7.31%	1.74%	Do not reject
09:30-11:30	12.55%	12.34%	3.10%	Do not reject
11:30-13:30	12.90%	12.92%	4.54%	Do not reject
13:30-15:30	7.96%	7.85%	2.28%	Do not reject
Combined all-four-period	9.23%	9.30%	2.66%	Do not reject

Note: [a] Null hypothesis test at the 95% confidence level.

4.3 Sensitivity analysis

The developed SBLR model and analysis illustrates the significance of having a signal installed and being "on" to influence the vacant taxi drivers' customer-search decisions. However, as all four selected sites are signal-enabled, the above analysis cannot evaluate the effectiveness of signal installation and operation on affecting vacant taxi drivers' decisions. Based on the simulation model, a sensitivity analysis was, thereafter, performed. We hypothesized all other 15 possible combinations of having three signals, two signals, one signal, and without signal in the four sites plus the standard scenario of having all signals in operation. A total number of 16 scenarios were tested and compared while all other variables were controlled. All 16 scenarios were simulated for the combined all-four-period for 100 repetitions. The predicted percentage of vacant taxis entering at least one site for picking up customers are plotted and shown in Figure 4.

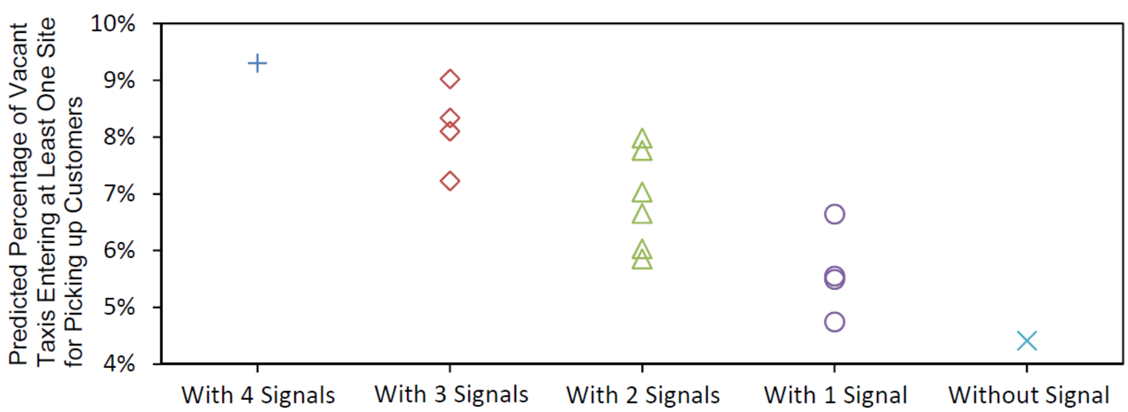

Figure 4. Variation in the percentage of vacant taxis entering at least one site for picking up customers under all scenarios of the provision of taxi-calling signals

A downward trend, dropping from 9.3% to 4.4%, is clearly discernable in the scatter plot. This means that more sites having signal provision would attract more vacant taxi drivers entering at least one of these sites. This finding supports the view that TCSs are effective in increasing the number of vacant taxis entering off-road locations for picking up customers, as vacant taxi drivers do positively respond to the TCS status. When a signal is not provided, the decision whether to enter or bypass the site simply relies on the perceived characteristics of a site, as well as vacant taxi drivers' own experience. It is natural that sites are visited by a very low number of vacant taxis when there is no signal provision. The sensitivity test shows that under such circumstance, vacant taxi drivers are more reluctant to enter the site compared with the case with signal provisions, which is also in line with our earlier findings that vacant taxi drivers tend to be more conservative towards passenger demand expectation when passenger demand is not clearly shown.

Figure 4 also demonstrates the variation in entry percentages of different numbers and locations of the TCSs provided. For the case with 3 signals, it shows the predicted percentage of vacant taxis entering at least one site for picking up customers in each of the four possible scenarios (i.e., the absence of a TCS in either one of the four sites). The percentages range from about 7% to 9%. The results indicate that the effectiveness in attracting vacant taxi drivers varies when the TCSs are installed at different locations. Two rather indicative cases are Sites B and D (mixed-use and hospital), in which the absence of signal in Site B and/or D usually causes a lower entry percentage of vacant taxis within the identified ranges. The result concurs with our findings in Table 1 and further strengthens the inference that TCSs increase the equity level of taxi services. Additionally, a site that has no TCS installed at the entrance and is situated next to another site with TCS installed at the entrance is at a disadvantage in attracting vacant taxis entering the site. Vacant taxi drivers with their objective of either maximizing profit or minimizing cost would inevitably prefer to enter a site where demand is clearly indicated.

4.4 Discussions and policy implications

The findings of this study suggest that TCSs are useful in guiding vacant taxis to remote areas, hospitals or schools where the entrances are not directly on main roads, and large residential estates with longer access roads. Usually, such areas have irregular passenger demand patterns and often do not correspond with peak/off-peak hours; TCSs serve as a proxy for customers who otherwise must stand and wait at the roadside. The government is suggested to provide adequate support to the deployment and use of TCSs. For example, consolidated multi-panel TCSs indicating the demand of various sites can be positioned at intersections connecting secondary access roads to a series of remote areas with the main roads, hospital roads, and access roads to topologically constrained areas. Moreover, the model results have suggested that the positioning of a site (proximity to the city center) would affect drivers' decisions. Having consolidated demand information presented to vacant taxi drivers up front could potentially remove this impact when drivers can make informed decisions ahead of time. Since a right-turn has a large negative effect on drivers, it is also recommended to place the consolidated panels at both ends of the access roads to ensure sufficient flow on the "easier" side of the road. Additionally, providing TCSs at these places can reduce unnecessary cruising of vacant taxis and hence reduce environmental impact and traffic load. It can also reduce the waiting time and improve the accessibility of taxi users in such areas. To increase the effectiveness of TCSs and the level of conformity of vacant taxis to the signals, it is further recommended to align TCSs with real-time demand volume. An example is the International Finance Centre in Hong Kong whose TCS is able to show the numbers of up to four waiting customers. This practice provides vacant taxi drivers a more accurate estimation of the probability of successful pickup when other preceding taxis are entering the TCS recommended area.

Arguably, ridesourcing platforms such as Uber can provide similar functions as the TCS. However, in addition to the legal disputes as seen in many cities including Hong Kong, the technological requirements and tipping mechanism of ridesourcing also bring social exclusion and equity issues. TCSs have been implementing to ensure the equity of taxi services as demands are met without associations with money incentives, trip destinations, and access to smartphones and certain apps. TCSs are ridesourcing platforms that have different demand-responsive mechanisms. In essence, TCSs are providing demand information to circulating supply, whereas ridesourcing platforms are providing supply information to demand. TCSs contribute to a market niche that is not fully covered by ridesourcing vehicles, and where conventional taxis are still dominating.

5. Conclusion

TCSs provide direct and on-the-spot passenger demand information to cruising vacant taxi drivers. They influence the vacant taxi drivers' customer-search decisions on whether to enter a TCS recommended area. This paper develops and calibrates an SBLR model that is considered more suitable for depicting the decision-making process in a sequence of TCS-installed sites. Nine significant variables that link with vacant taxi drivers' travel behavior are identified, in which the TCS status is found to be the most influential factor to affect vacant taxi drivers entering the sites for picking up customers. A simulation model and a solution procedure are developed and the results show that the calibrated SBLR model is highly accurate and robust. A sensitivity test is performed and the results demonstrate that TCSs are effective in directing vacant taxi drivers to demand locations that would otherwise be underserved. The study recommends TCSs to be installed for remote and hilly areas and at hospitals both to reduce unnecessary cruising and increase local driver-customer matching efficiencies. In addition, this study suggests aligning TCSs with real-time demand volume, which provides vacant taxi drivers a more accurate estimation of the probability of successful pickup when other preceding taxis are entering the TCS recommended area. In an era where ridesourcing platforms are increasingly popular, TCSs are still serving their own niche functions and should remain a simple yet effective tool for matching drivers with customers. Various tools and platforms including street hailing, taxi call centers, e-hailing apps, taxi stands, and TCSs are complementary to one another to form an all-inclusive and accessible taxi system of a city.

Acknowledgements

This research was jointly supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (HKU 17201217), a grant (201611159067) from the University Research Committee of the University of Hong Kong and a Research Postgraduate Studentship from the University of Hong Kong. We are also grateful for the comments from the five reviewers.

References

Agresti, A., 2007. An introduction to categorical data analysis, 2nd Edition. Wiley, New Jersey.

Amemiya, T., 1975. Qualitative response models. Annals of Economic and Social Measurements 4 (3), 363-372.

Amemiya, T., 1994. Studies in Econometric Theory: Collected Essays of Takeshi Amemiya. Edward Elgar, England.

Arnott, R., 1996. Taxi travel should be subsidized. Journal of Urban Economics 40 (3), 316-333.

Bacache-Beauvallet, M. and Janin, L., 2012. Taxicab license value and market regulation. Transport Policy 19 (1), 57-62.

Barret, S.D., 2003. Regulatory capture, property rights and taxi deregulation: A case study. Economic Affairs 23 (4), 34-40.

Barret, S.D., 2010. The sustained impacts of taxi deregulation. Economic Affairs 30 (1), 61-65.

Cairns, R.D. and Liston-Heyes, C., 1996. Competition and regulation in the taxi industry. Journal of Public Economics 59 (1), 1-15.

Carteni, A., Cascetta, E., and de Luca, S., 2016. A random utility model for park & carsharing services and the pure preference for electric vehicles. Transport Policy 48, 49-59.

Cramer, J. and Krueger, A.B., 2016. Disruptive change in the taxi business: The case of Uber. The American Economic Review 106 (5), 177-182.

De Vany, A.S., 1975. Capacity utilization under alternative regulatory constraints: An analysis of taxi markets. Journal of Political Economy 83 (1), 83-94.

Douglas, G.W., 1972. Price regulation and optimal service standards: The taxicab industry. Journal of Transport Economics and Policy 6 (2), 116-127.

Fu, H. and Wilmot, C.G., 2004. A sequential logit dynamic travel demand model for hurricane evacuation. Transportation Research Record: Journal of the Transportation Research Board 1882, 19-26.

Harding, S., Kandlikar, M., and Gulati, S., 2016. Taxi apps, regulation, and the market for taxi journeys. Transportation Research Part A: Policy and Practice 88, 15-25.

Hausman, J.A. and McFadden, D., 1984. Specification tests for the multinomial logit model. Econometrica 52 (5), 1219-1240.

Heanue, K.E. and Pyers, C.E. 1966. A comparative evaluation of trip distribution procedures. Highway Research Record 114, 20-50.

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: Journal of the Transportation Research Board 2284, 81-88.

Kahn, L.M. and Morimune, K., 1979. Unions and employment stability: A sequential logit approach. International Economic Review 20 (1), 217-235.

Land Transport Authority of Singapore, 2011. Passenger Transport Mode Share in World Cities. www.lta.gov.sg/ltaacademy/doc/J11Nov-p60PassengerTransportModeShares.pdf

Liao, Z., 2003. Real-time taxi dispatching using Global Positioning Systems. Communications of the ACM – Wireless networking security 46 (5), 81-83.

Long, J.C., Szeto, W.Y., Du, J., and Wong, R.C.P., 2017. A dynamic taxi traffic assignment model: A two-level continuum transportation system approach. Transportation Research Part B: Methodological, 100, 222-254.

Maciejewski, M., Bischoff, J., and Nagel, K., 2016. An assignment-based approach to efficient real-time city-scale taxi dispatching. IEEE Intelligent Systems, 31 (1), 68-77.

McCullagh, P. and Nelder, J.A., 1989. Generalized Linear Models, 2nd Edition. Chapman & Hall/CRC, London.

Miao, F., Han, S., Lin, S., Stankovic, J.A., Zhang, D., Munir, S., Huang, H., He, T., and Pappas, G.J., 2016. Taxi dispatch with real-time sensing data in metropolitan areas: A receding horizon control approach. IEEE Transactions on Automation Science and Engineering, 13 (2), 463-478.

Moreira-Matias, L., Gama, J., Ferreira, M., and Damas, L., 2012. A predictive model for the passenger demand on a taxi network. In: 2012 15th International IEEE Conference on Intelligent Transportation Systems, Anchorage, U.S.A., September 16-19, 2012. Proceedings. New York: IEEE. 1014-1019.

Ng, K.C., 2015. Hong Kong Consumer Council to study legality of Uber services. South China Morning Post. www.scmp.com/news/hong-kong/law-crime/article/1849908/hong-kong-consumer-council-study-legality-uber-services.

Qian, X., Ukkusuri, S.V., Yang C., and Yan, F., 2017. A model for short-term taxi demand forecasting accounting for spatio-temporal correlations. In: Transportation Research Board 96th Annual Meeting, Washington DC, U.S.A., January 8-12, 2017.

Sawtelle, V., Brewe, E., and Kramer, L.H., 2011. Sequential logistic regression: A method to reveal subtlety in self-efficacy. In M. S. Plakhotnik et al. (Eds.), Proceedings of the Tenth Annual College of Education & GSN Research Conference (pp. 216-225). Miami: Florida International University.

Schaller, B., 2007. Entry controls in taxi regulation: Implications of US and Canadian experience for taxi regulation and deregulation. Transport Policy 14 (6), 490-506.

Schneider, M., 1959. Gravity models and trip distribution theory. Papers and Proceedings of the Regional Science Association V, 51-56.

Schroeter, J.R., 1983. A model of taxi service under fare structure and fleet size regulation. The Bell Journal of Economics 14 (1), 81-96.

Situ, X., Chen, W.N., Gong, Y.J., Lin, Y., Yu, W.J., Yu, Z., and Zhang, J., 2017. A parallel ant colony system based on region decomposition for taxi-passenger matching. In: 2017 IEEE Congress on Evolutionary Computation (CEC), San Sebastian, Spain, June 5-8, 2017.

Stouffer, S.A., 1940. Intervening opportunity: A theory relating mobility and distance. American Sociological Review 6, 845-867.

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 (2), 184-198.

Train, K.E., 2009. Discrete Choice Methods with Simulation, 2nd Edition. Cambridge University Press, Cambridge.

Transport Department, Hong Kong SAR Government, 1986-2009. The Level of Taxi Services, TTSD Publication Series.

Transport Department, Hong Kong SAR Government, 2016. Annual Transport Digest 2015. www.td.gov.hk/mini_site/atd/2015/index.html.

Wang, H., Cheu, R.L., and Lee, D.H., 2014. Intelligent taxi dispatch system for advance reservations. Journal of Public Transportation, 17 (3), 115-128.

Wang, X., He, F., Yang, H., and Gao, H.O., 2016. Pricing strategies for a taxi-hailing platform. Transportation Research Part E: Logistics and Transportation Review 93, 212-231.

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: Methodological 35 (9), 819-842.

Wong, K.I., Wong, S.C., Yang, H., and Wu, J.H., 2008. Modeling urban taxi services with multiple user classes and vehicle types. Transportation Research Part B: Methodological 42 (10), 985-1007.

Wong, R.C.P., Szeto, W.Y. and Wong S.C., 2015b. Behavior of taxi customers in hailing vacant taxis: A nested logit model for policy analysis. Journal of Advanced Transportation 49, 867-883.

Wong, R.C.P., Szeto, W.Y., and Wong, S.C., 2014a. A cell-based logit-opportunity taxi customer-search model. Transportation Research Part C: Emerging Technologies 48, 84-96.

Wong, R.C.P., Szeto, W.Y., and Wong, S.C., 2014b. Bi-level decisions of vacant taxi drivers traveling towards taxi stands in customer-search: Modeling methodology and policy implications, Transport Policy 33, 73-81.

Wong, R.C.P., Szeto, W.Y., and Wong, S.C., 2015a. A two-stage approach to modeling vacant taxi movements. Transportation Research Part C: Emerging Technologies 59, 147-163.

Wong, R.C.P., Szeto, W.Y., and Wong, S.C., 2015c. Sequential logit approach to modeling the customer-search decisions of taxi drivers. Asian Transport Studies 3(4), 398-415.

Wong, S.C. and Yang, H., 1998. A network equilibrium model of urban taxi services: An improved algorithm. Transportation Research Record: Journal of the Transportation Research Board 1623, 27-30.

Wu, L., Hu, S., Yin, L., Wang, Y., Chen, Z., Guo, M., Chen, H., and Zhong, X., 2017. optimizing cruising routes for taxi drivers using a spatio-temporal trajectory model. ISPRS International Journal of Geo-information, 6 (11), 373-393.

Yang, H. and Wong, S.C., 1998. A network model of urban taxi services. Transportation Research Part B: Methodological 32 (4), 235-246.

Yang, H. and Yang, T., 2011. Equilibrium properties of taxi markets with search frictions. Transportation Research Part B: Methodological 45 (4), 696-713.

Yang, H., Leung, C.W.Y., Wong, S.C., and Bell, M.G.H., 2010. Equilibria of bilateral taxi–customer searching and meeting on networks. Transportation Research Part B: Methodological 44 (8-9), 1067-1083.

Yang, H., Wong, K.I., and Wong, S.C., 2001. Modeling urban taxi services in road networks: Progress, problem and prospect. Journal of Advanced Transportation 35 (3), 237-258.

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: Methodological 36 (9), 799-819.

Yang, H., Ye, M., Tang, W.H., and Wong, S.C., 2005. Regulating taxi services in the presence of congestion externalities. Transportation Research Part A: Policy and Practice 39 (1), 17-40.

Zhang, W. and Ukkusuri, S.V., 2016. Optimal fleet size and fare setting in emerging taxi markets with stochastic demand. Computer-Aided Civil and Infrastructure Engineering 31 (9), 647-660.