

World Conference on Transport Research - WCTR 2016 Shanghai. 10-15 July 2016

Comparative Analyses of Taxi Operations at the Airport

Yuxiong Ji^a, Yixuan Cao^a, Yuchuan Du^{a*}, H. Michael Zhang^{b,a}

^a Key Laboratory of Road and Traffic Engineering of the Ministry of Education, Tongji University, Shanghai 201804, China

^b Department of Civil and Environmental Engineering, University of California at Davis, Davis, CA 95616, USA

Abstract

We propose statistical methodologies to investigate taxi operations at the airport. Based on Automatic Vehicle Location (AVL) data of nearly 9,000 taxis in five weekdays, we apply the proposed methodologies to evaluate taxi operations at Shanghai's Hongqiao international airport. The empirical studies statistically compare the operations of taxis serving the airport and urban areas and identify the periods in which the taxi profits are relatively stable. Possible explanations for the empirical results are provided. The findings would offer a better understanding of the mechanism that drives vacant-taxi drivers to serve the airport.

© 2017 The Authors. Published by Elsevier B.V.

Peer-review under responsibility of WORLD CONFERENCE ON TRANSPORT RESEARCH SOCIETY.

Keywords: Taxis; Automatic Vehicle Location (AVL); Airport; Customer Search; Profit; Vacant time

1. Introduction

Taxis play an important role in maintaining the sustainability of urban transportation systems, especially in cities where buying and maintaining a car is expensive. Taxis provide door-to-door service and have better accessibility than conventional mass transit systems such as bus and subway. On the one hand, taxis can be used for accessing and egressing from transit stations. On the other hand, taxis could be substitutes of mass transit systems. During the night when mass transit systems are not operating, many people have to rely on taxis service [1].

Many cities have employed Automatic Vehicle Location (AVL) systems to track the locations of taxis in real time. With AVL, taxi operators can locate taxis in real time and dispatch nearby taxis when customers call for taxi service. The archived AVL data, on the other hand, are useful for performance analyses and operations optimization.

The taxi AVL data have been widely used in various applications, such as inferring traffic conditions on road networks [2-4], analyzing human mobility [5, 6], and assessing taxi operations performance [1, 7-11]. For example, Liu [7] systematically studied taxi drivers' behaviors using daily taxi AVL data. They found that taxi drivers have

* Corresponding author. Tel.: +86-21-69585712; fax: +86-21-69585712.

E-mail address: ycdu@tongji.edu.cn

different strategic operations patterns, resulting in large variations in their income. Zhang, Peng and Sun [8] proposed a comprehensive framework to evaluate taxi operations from different perspectives. Amat [9] developed a methodology to assess the efficiency and profitability of the taxi system in Barcelona. Pele and Morency [10] analyzed taxi operations in Montreal. They also evaluated taxi operations under the influence of various factors, such as weather and holidays. Qian, Zhan and Ukkusuri [11] developed a two-step clustering algorithm to classify taxi trips and then analyzed the taxi travel patterns in New York City and some hotspot areas.

Most of existing literatures focused on developing methodologies to evaluate taxi operations in a large area. The methodologies are macroscopic and deterministic. As a result, the analyses tend to miss detailed features of taxi operations in some small, but important areas, such as airport, transportation hub and Central Business District (CBD).

Airport is a main generator and attractor of taxi trips. The efficiency of taxi operations is critical in determining the overall level of service of the ground transportation systems at the airport. Understanding taxi travel patterns is helpful for improving the quality of taxi service at the airport and for increasing the income of taxi drivers. However, Empirical studies on taxi operations at the airport are limited. Pele and Morency [10] and Qian Zhang and Ukkusuri [11] presented some preliminary findings about taxi operations at the airport. Further investigations are necessary to comprehend taxi drivers' behaviors at the airport.

In China, many taxi drivers are willing to travel to the airport empty to pick up customers, even though the airport is distant from the urban areas and the waiting times at the airport are long. For example, at Hong Kong's main airport, taxis need to wait about 4 hours upon their arrival before they get their chances to pick up customers [12]. The average taxi waiting time at Shanghai's Hongqiao airport was 80 minutes in the daytime and 40 minutes in the evening in 2010. The willingness to serve the airport customers and the long waiting times at the airport make it interesting to investigate taxi operations at the airport.

Early literatures assumed that vacant-taxi drivers tend to minimize their search times when searching for customers [13, 14]. Obviously, this assumption cannot explain why taxi drivers are willing to endure the long waiting times at the airport. To capture the customer-search behaviors of taxis serving the airport customers, Wong et al. [15] assumed that taxi drivers are willing to serve the airport customers because they perceive higher profits at the airport than in the urban areas. Similarly, Yang et al. [12] assumed that taxi drivers intend to maximize their "profit per unit time" when searching for customers. The profitability based explanation of the behaviors of taxi drivers who serve the airport seems plausible. Nevertheless, no empirical evidence has been collected to verify (or refute) such an explanation.

In this study, we propose statistical methodologies to investigate taxi operations at the airport. Based on AVL data of nearly 9,000 taxis in five weekdays, we apply the proposed methodologies to evaluate taxi operations at Shanghai's Hongqiao international airport. The results would offer a better understanding of the mechanism that drives vacant-taxi drivers to serve the airport.

The rest of this paper starts with the dataset used in the empirical study. Then measures of taxi operations performance and the statistical methodologies for analyzing taxi data are introduced, followed by the results of analyses. Finally, conclusion and discussion are presented.

Nomenclature

c^r	rental cost per hour of operation
c^f	fuel cost per kilometer
d_i^v	driving distance for the i^{th} vacant taxi trip
d_i^o	driving distance for the i^{th} occupied taxi trip
t_i^v	time for the i^{th} vacant taxi trip
t_i^o	time for the i^{th} occupied taxi trip
F_i	fare collected for the i^{th} occupied trip
OO	both the origins of the occupied trips and the destinations of their preceding trips are the airport
OV	origins of the occupied trips are outside of the airport, but the destinations of their preceding occupied trips are the airport
OA	both the origins of the occupied trips and the destinations of their preceding trips are outside of the airport

P_i	profit for the i^{th} search cycle
RR_i	rate of return for the i^{th} search cycle, measuring profit per hour of operation
VO	occupied trips originating from the airport, but the destinations of their preceding occupied trips are outside of the airport

2. The Data

This study focuses on taxi operations at Shanghai's Hongqiao international airport. Hongqiao airport is one of the busiest airports in China. The annual flights are about 234,942 sorties and the annual passenger throughput is about 33.8 million [16]. Hongqiao airport lies in the west of Shanghai and is 13 kilometers away from Shanghai's city center. Hongqiao airport is connected to the freeways and urban expressways through elevated roads for quick access and egress.

Hongqiao airport adopts a taxi staging policy. A staging area is designated for vacant taxis to wait for customers. There are more than 50 lanes in the staging area. Operators in the staging area dispatch taxis lane by lane on a first-in-first-out (FIFO) basis. After joining the queue, taxi drivers could shut off the engine and take a break, which is an important incentive for taxi drivers to serve the airport customers.

The AVL data used in this paper were provided by a major taxi company in Shanghai. The AVL systems record taxi information every 20 seconds when taxis are vacant and every one minute when taxis are occupied. Each record includes transaction ID, longitude, latitude, car ID, time stamp, and loading status etc. The loading status equals one when the taxi is occupied and equals zero when it is vacant. A transition from vacant status to occupied status indicates that the taxi picks up passengers and a transition from occupied status to vacant status indicates that the taxi drops off passengers. As a result, we can identify the time and location of the starting and ending points of each vacant and occupied trip.

The occupied taxi trips are classified into four groups, depending on the origins of the occupied taxi trips and the destinations of the preceding occupied taxi trips: (1) VO (Vacant-in-Occupied-out) trips: the occupied trips originating from the airport, but the destinations of their preceding occupied trips are outside of the airport; (2) OO (Occupied-in-Occupied-out) trips: both the origins of the occupied trips and the destinations of their preceding occupied trips are the airport; (3) OV (Occupied-in-Vacant-out) trips: the origins of the occupied trips are outside of the airport, but the destinations of their preceding occupied trips are the airport; (4) OA (Outside Airport) trips: both the origins of the occupied trips and the destinations of their preceding occupied trips are outside of the airport. Different taxi trip groups are illustrated in Figure 1, in which Hongqiao airport is represented by the area between two horizontal dashed point lines.

Figure 1 also illustrates some basic features of taxi operations. In Figure 1, a circle represents the time and location when a taxi drops off passengers and a rectangle represents the time and location when a taxi picks up passengers. The green dashed and blue solid lines represent the vacant and occupied trips, respectively. The vacant time, vacant distance, profit time, profit distance and origin and destination of each trip can be derived from the raw taxi AVL data.

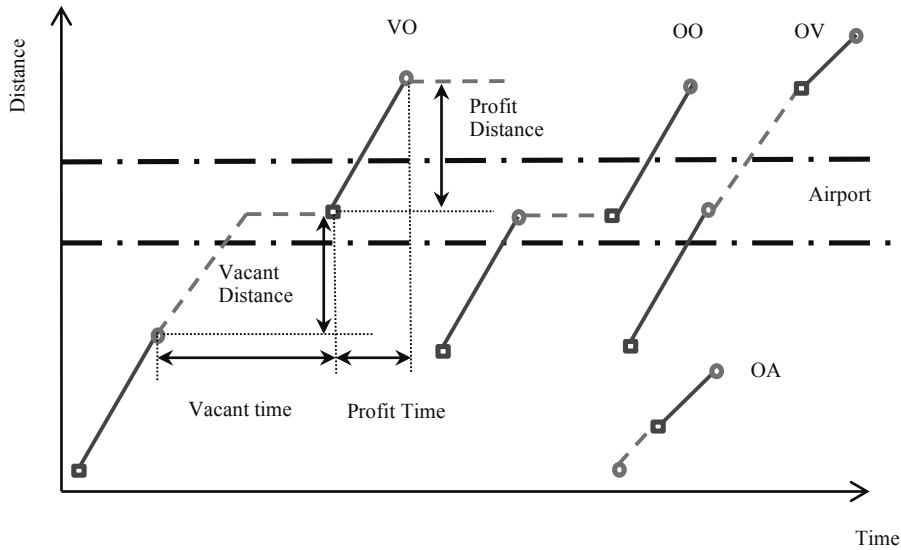


Fig. 1. Classifications of taxi trips and illustration of basic features of taxi operations.

Taxi AVL data for five weekdays between 8am and 11pm in 2010 are used in the analysis. The number of taxis and number of occupied trips in each day are summarized in Table 1. On average, the raw data includes 8,678 taxis and 342,970 occupied trips between 8am and 11pm in a day. Nevertheless, some obvious errors are identified. Records with errors as listed below are excluded from analyses: (1) taxis have records with the same time stamps, but different latitudes and longitudes; (2) profit distance is less than 100 meters; (3) profit time is less than 1 minute; (4) profit time is more than 1 hour, but profit distance is less than 5 km. Eventually, 82% of the original number of taxis and 75% of the original number of occupied trips are remained for analyses.

Table 1. Numbers of taxis and numbers of occupied trips in the raw dataset and in the remaining dataset.

Date	Raw dataset		Remaining dataset	
	Number of taxis	Number of profit trips	Number of taxis	Number of profit trips
23-Mar	8,963	306,374	6,689	208,845
15-Apr	8,620	351,896	7,264	269,840
21-Apr	8,584	351,709	7,201	268,417
9-Jun	8,641	353,508	7,157	268,595
2-Jul	8,584	351,364	7,242	275,461
Average	8,678	342,970	7,111	258,232

2.1. Fare, operational cost and profit

The fares taxi drivers collected are estimated based on the distances of occupied trips. The taxi fare in Shanghai is distance based. During the period between 5am and 11pm, the starting fee for the first 3 km was 12 yuan (1 yuan = 0.15 US dollar in 2010). The fee per km was 2.4 yuan for distance over 3 km and was 3.6 yuan for distance over 10 km. During the period between 11pm and 5am, the starting fee for the first 3 km was 16 yuan. The fee per km was 3.1 yuan for distance over 3 km and was 4.7 yuan for distance over 10 km.

Taxi operational cost includes rental and fuel costs. The monthly rental cost for a taxi in Shanghai was about 8,500 yuan in 2010. A taxi was operated by two drivers, and each driver operated 15 days per month and about 18 hours per day. As a result, the rental cost per hour of operation was approximately 15.7 yuan. In addition, the liquefied petroleum price was 6.43 yuan per liter in 2010. A taxi consumes approximately 10 liters per 100 kilometers travelled. As a result, the fuel cost per kilometer was approximately 0.64 yuan.

Since each occupied taxi trip is related to one vacant taxi trip immediately preceding it, we refer a pair of occupied and vacant taxi trips as a search cycle for convenience. The starting time of a search cycle is indicated by the starting time of the vacant trip. The sum of the times of the vacant and occupied trips is referred to as cycle time. The profit for a search cycle is calculated by subtracting the operational cost from the fare. The profit, P_i , for the i^{th} search cycle is given by:

$$P_i = F_i - c^r \times (t_i^v + t_i^o) - c^f \times (d_i^v + d_i^o) \quad (1)$$

where F_i is the fare collected for the i^{th} occupied trip. c^r and c^f are rental cost per hour of operation and fuel cost per kilometer, respectively. t_i^v and t_i^o represent the times for the i^{th} vacant and occupied trips, respectively, and d_i^v and d_i^o represent the distances for the i^{th} vacant and occupied trips, respectively.

3. Methodologies

3.1. Measures of taxi operation performance

To obtain higher income, taxi drivers would like to have shorter vacant time and longer profit distance. In this study, the profit taxi drivers perceive is measured by profit per hour of operation, which is referred to as Rate of Return (RR). Specifically, we define the rate of return, RR_i , for the i^{th} search cycle as follows:

$$RR_i = \frac{P_i}{t_i^v + t_i^o} \quad (2)$$

Higher RR values indicate that taxi drivers earn more.

3.2. Nonparametric statistical methodologies

The Wilcoxon Rank Sum test [17] is used to evaluate whether different taxi trips have the same operations characteristics. The Wilcoxon Rank Sum test compares the distributions of two variables that are independent of each other. The null hypothesis H_0 is that the distributions of both variables are equal. As a result, the probability that an observation of one variable (X) is greater than an observation of the second variable (Y) equals 0.5. The null and alternative hypotheses are represented by:

$$H_0 : X = Y \text{ vs. } H_1 : X \neq Y \quad (3)$$

Or

$$H_0 : \Delta = 0 \text{ vs. } H_1 : \Delta \neq 0 \quad (4)$$

where $\Delta = Y - X$, which is called a location parameter. Δ measures the difference in the location of X and Y distribution.

The statistic for the Wilcoxon Rank Sum test is based on the ranks of the Y (or X) variables in the pooled ranking of all the X and Y observations. Assume that there are m and n numbers of observations for variables X and Y , respectively. Order the $N = m + n$ observations from the least to the greatest. Let R_j denotes the rank of y_j in this ordering dataset, the test statistic W is computed as follows:

$$W = \sum_{j=1}^n R_j \quad (5)$$

The statistic W equals the sum of the ranks of the Y observations. Obviously, the minimum and maximum values of W are $n \times (n+1)/2$ and $n \times (2m+n+1)/2$, respectively. The null hypothesis is rejected if $W \geq W_{\alpha/2, m, n}$ or $W \leq n \times (m+n+1) - W_{\alpha/2, m, n}$, where $W_{\alpha/2, m, n}$ is the upper $\alpha/2$ quantile of W under H_0 given the sample sizes of m and n for the X and Y variables, respectively.

If the sample sizes of m and n are large, large sample approximation can be used. Specifically, the test statistic W^* as defined below follows standard normal distribution:

$$W^* = \frac{W - E(W)}{\sqrt{\text{var}(W)}} = \frac{W - \frac{n \times (m+n+1)}{2}}{\sqrt{\frac{m \times n \times (m+n+1)}{12}}} \sim N(0,1) \quad (6)$$

The null hypothesis is rejected if $W^* \geq Z_{\alpha/2}$ or $W^* \leq -Z_{\alpha/2}$ where $Z_{\alpha/2}$ is the upper $\alpha/2$ quantile of standard normal

distribution.

The Wilcoxon Rank Sum test is more robust in the presence of outliers than the t-test [17]. It has greater efficiency than the traditional t-test on variables that are not Normally distributed, and it is almost as efficient as the t-test on variables that are Normally distributed.

Point estimate and confidence interval of the location parameter Δ can be obtained using the following procedure.

- (1) Compute all the $n \times m$ possible difference $y_j - x_i$, and then order the $n \times m$ differences from the least to the greatest, the k^{th} difference is denoted by $U^{(k)}$;
- (2) Compute $C_\alpha = n \times (2m + n + 1) / 2 - W_{\alpha/2, m, n}$
- (3) The point estimate of the location parameter Δ equals the median of U .
- (4) The $(1 - \alpha)$ confidence interval for Δ is $(U^{(C_\alpha)}, U^{(m \times n + 1 - C_\alpha)})$.

To compare the distributions of more than two variables, Kruskal-Wallis test [17] can be used. Kruskal-Wallis test is a generalization of the Wilcoxon Rank Sum test. Suppose there are k variables, the null hypothesis is that all variables are equal and the alternative hypothesis is that not all variables are equal.

The statistic for Kruskal-Wallis test is based on the weighted ranks of all variables in the pooled ranking of the observations of all variables. Assume that there are k number of variables and that there are n_i numbers of observations for variables z_i . Order the $N = n_1 + \dots + n_k$ observations from the least to the greatest. Let z_{ij} denotes the j^{th} observation of variable z_i and let R_{ij} denotes the rank of z_{ij} in the ordering dataset, the average rank R_i for variable z_i is given by:

$$R_i = \frac{\sum_{j=1}^{n_i} R_{ij}}{n_i} \quad (7)$$

If all variables are close, then each R_i should be close to the average of the ranks of all the N observations $R_{..}$, which is given by:

$$R_{..} = \frac{\sum_{i=1}^N N}{N} = \frac{N+1}{2} \quad (8)$$

The test statistic H is proportion to the weighted sum of the square difference between the average rank for each variable and the average rank for all observations, with the weights given by the corresponding sample size:

$$H = \left[\sum_{i=1}^k n_i \times (R_i - R_{..})^2 \right] \times \frac{12}{N \times (N+1)} \quad (9)$$

The null hypothesis is rejected if $H \geq h(\alpha, k, (n_1, \dots, n_k))$, where $h(\alpha, k, (n_1, \dots, n_k))$ is the upper α quantile of H under H_0 . If the sample sizes of m and n are large, large sample approximation can be used.

If the null hypothesis is rejected, it indicates that at least one of the variables is different from at least one of the others. Pairwise comparisons can be carried out to identify which pair of variables are different [17].

4. Results of analyses

Fig. 2 summarizes the number of taxi trips leaving Hongqiao airport by time of day. Data are aggregated into one-hour contiguous elemental periods from 8am to 11pm. The horizontal axis represents the starting time of each elemental period. Figures 2(a) and 2(b) presents the number and proportion of VO, OO and OV trips by time of day. The numbers and proportions of VO taxi trips are higher after 7pm than those before 7pm. This result indicates that more vacant-taxi drivers are willing to go to the airport to pick up passengers in the night time. In addition, the proportions of OV trips are higher than those of OO trips before 6pm, but are close to or lower than those of OO trips after 6 pm, suggesting that more taxi drivers who transport passengers to the airport would wait at the airport for their next customers in the night time.

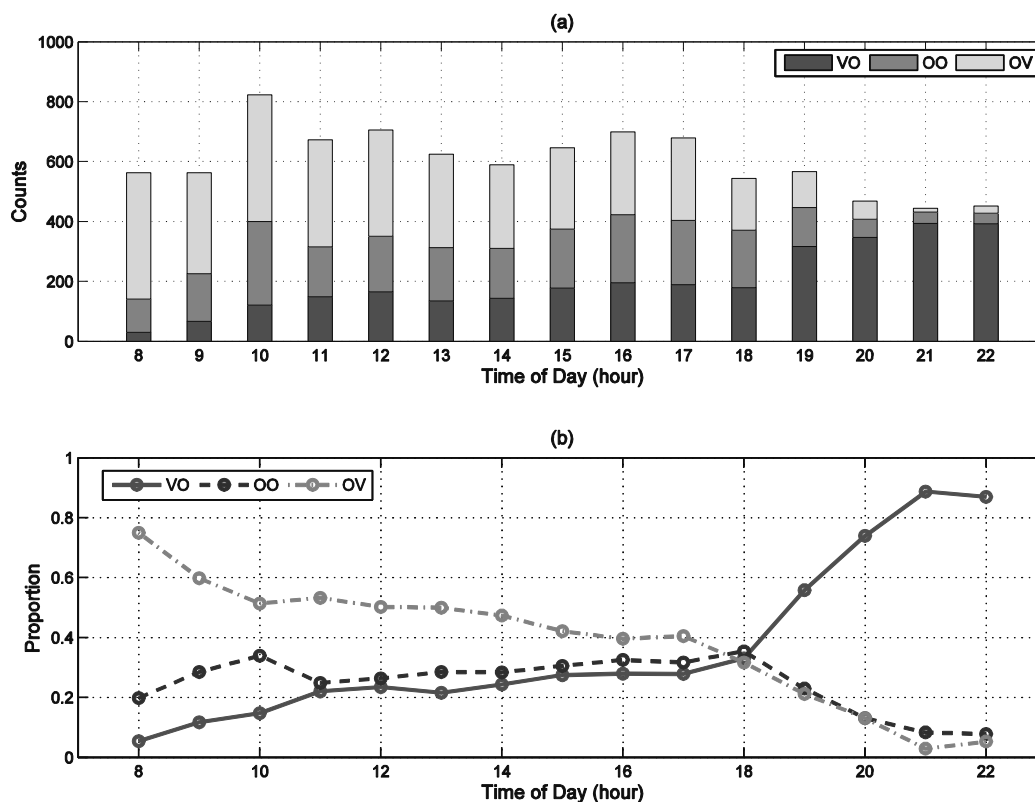


Fig. 2. Summary statistics of taxi trips leaving Hongqiao T2 by time of day: (a) Number of taxi trips; (b) Proportion of VO, OO and OV trips.

Fig. 3 presents the median values of the vacant time, profit speed, fare and rate of return for four types of taxis by time of day. In terms of the airport taxi trips (VO and OO taxi trips), The vacant times in the night time are about 40 and 20 minutes shorter than those in the daytime for the OO and VO taxi trips, respectively. The fares VO and OO taxi drivers collected are about 55 yuan, which are relatively stable across the time of day. For both the VO and OO taxi trips, the rates of return in the night time are higher than those in the day time, possibly due to the shorter vacant times in the night time.

In terms of the OV and OA taxi trips, the vacant times are relatively stable across the day. The vacant times for the OV taxi trips are about 15 minutes longer than those for the OA taxi trips, reflecting the time it took OV taxis to drive empty from the airport to the areas where they picked up customers. The fares OV and OA taxi drivers collected are about 15 yuan, which are relatively stable across the time of day. The rates of return for both the OV and OA taxi trips in the night time are relatively higher than those in the day time.

As far as the profit speeds are concerned, the profit speeds in the night time are higher than those in the daytime, thanks to the better traffic condition at night. For example, for the VO and OO taxi trips, the profit speeds in the night time are about 15 km/hr higher than those in the daytime. Since the airport is connected with the freeways and the urban expressways with elevated roads, the profit speeds of VO and OO taxis trips are higher than those of OV and OA taxi trips.

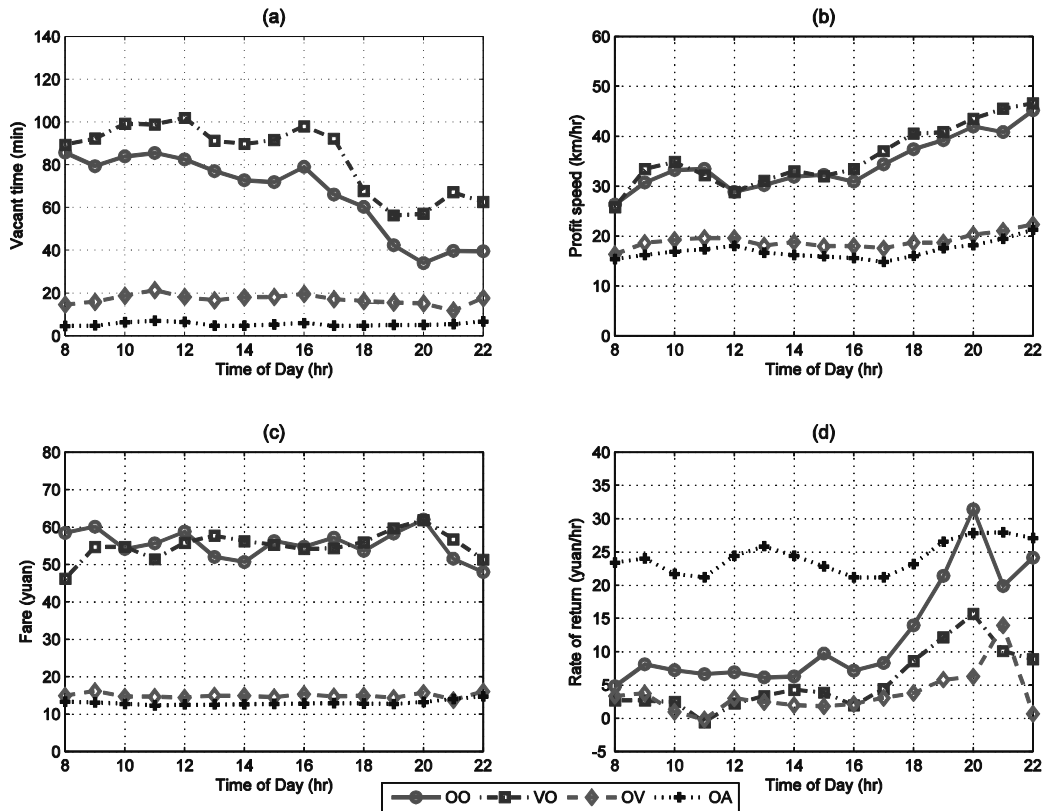


Fig. 3. Taxi operations measures by time of day: (a) Vacant time (min); (b) Profit speed (km/hr); (c) Fare (yuan); (d) Rate of return (yuan/hr).

4.1. Comparisons of taxi operations of different types of taxi trips

Fig. 4 provides 95% confidence intervals for the differences of the vacant time and rate of return by time of day. The confidence intervals are produced using the Wilcoxon Rank Sum test method. In Fig. 4, the vertical lines represent the lower and upper bounds of the 95% confidence intervals. The circle points inside the lines represent the median values. The two plots in the first row of Fig.4 compare the operations of OO and OV taxi trips. The OO and OV taxis dropped off customers at the airport and they need to decide whether they should wait at the airport or go somewhere else to pick up customers. The two plots in the second row of Fig. 4 compare the operations of VO and OA taxi trips. The VO and OA taxis dropped off customers at areas other than the airport and they may wonder whether they should go to the airport or somewhere else for their next customers.

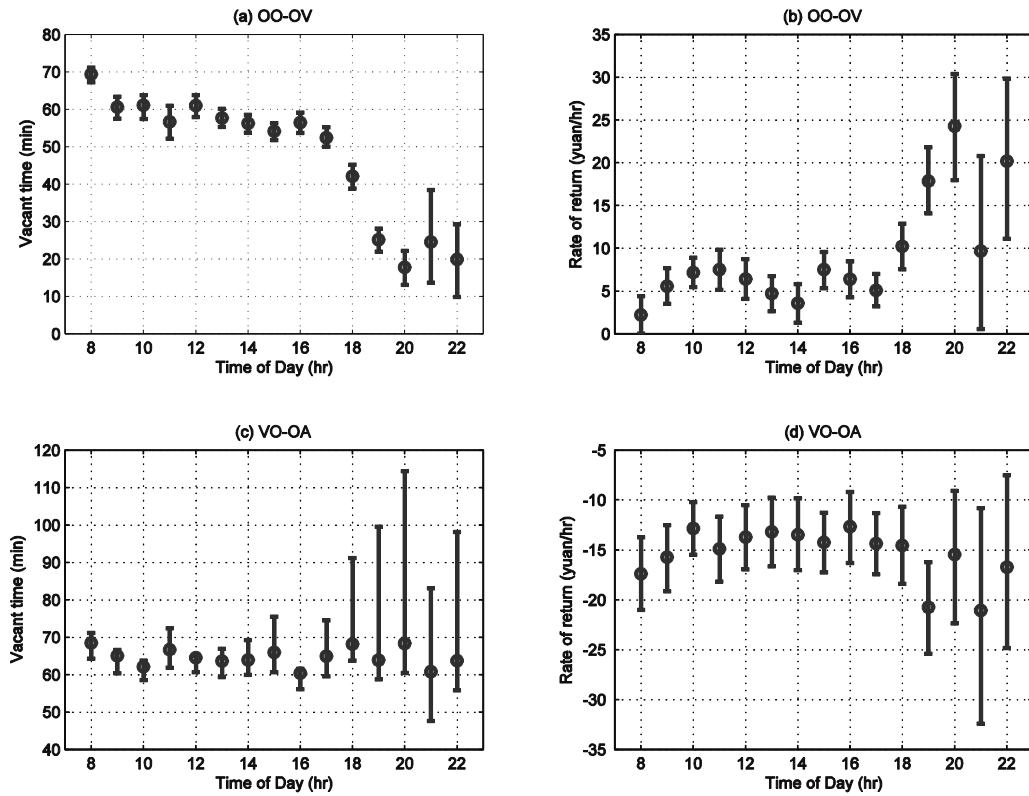


Fig. 4. Comparisons of the vacant time and rate of return by time of day: (a) Difference between the vacant times of OO and OV taxi trips; (b) Difference between the rates of return of OO and OV taxi trips; (c) Difference between the vacant times of VO and OA taxi trips; (d) Difference between the rates of return of VO and OA taxi trips.

Fig. 4(a) demonstrates that the vacant times of OO taxi trips are significantly higher than those of OV taxi trips. The vacant times of OO taxi trips are more than 40 minutes and are about 20 minutes longer than those of OV taxi trips before and after 6pm, respectively. The much longer vacant times at the airport in the day time could possibly explain the high proportion of OV taxi trips as shown in Fig. 2.

Even though the vacant times of OO taxi trips are longer, the rates of return of OO taxi trips are significantly higher than those of OV taxi trips, as revealed in Fig. 4(b). The rates of return of OO taxi trips are about 5 yuan/hr higher than those of OV taxi trips before 6pm, but are above 10 yuan/hr higher after 6pm.

Fig. 4 (c) shows that VO taxi trips have to wait much longer for customers than the OA taxi trips. The vacant times of VO taxi trips are about 60 minutes longer than those of OA taxi trips. Even though VO taxi trips could collect higher fare than OA taxi trips (see Fig. 3(c)), the rates of return of VO taxi trips are significantly lower than those of OA taxi trips, due to the much longer vacant times at the airport. Their differences are about 15 yuan/hr and 18 yuan/hr before and after 6 pm, respectively. The significantly higher vacant times and lower rates of return suggest that the assumption of either vacant time minimization or profit maximization may not be sufficient to explain the mechanism that drive taxi drivers to travel to the airport empty to pick up customers.

4.2. Effect of time of day on airport taxi trips

As presented in Figure 3, the rates of return vary by the time of day. To identify the periods in which the rates of return are relatively stable for taxis that serve the airport customers, Kruskal-Wallis test is used to compare the rates of return of OO taxi trips in 15 elemental periods from 8am to 11pm. The p value of Kruskal-Wallis test is close to zero, suggesting that the rates of return in some elemental periods are different. Then Pairwise comparisons are carried out to identify the pairs of elemental periods that are different. The comparison results are presented in Fig.

5. In Fig. 5, black cells indicate that the elemental periods in the corresponding rows and columns are not significantly different and white cells indicate that the elemental periods in the corresponding rows and columns are significantly different.

Based on the results in Figure 5, the elemental periods could be approximately classified into three groups. Group one includes 10 elemental periods from 8am to 6pm. Group two includes one elemental period from 6pm to 7pm and Group three includes four elemental periods from 7pm to 11pm. All elemental periods in Groups one or three are not significantly different. All elemental periods in Group three are significantly different from those in Group one. Based on the above classification, the median rates of return for the OO taxi trips in Groups one to three are 7.2, 14.0 and 24.7 RMB/hr, respectively. The results suggest that the time periods after 7pm are more profitable than the other periods.

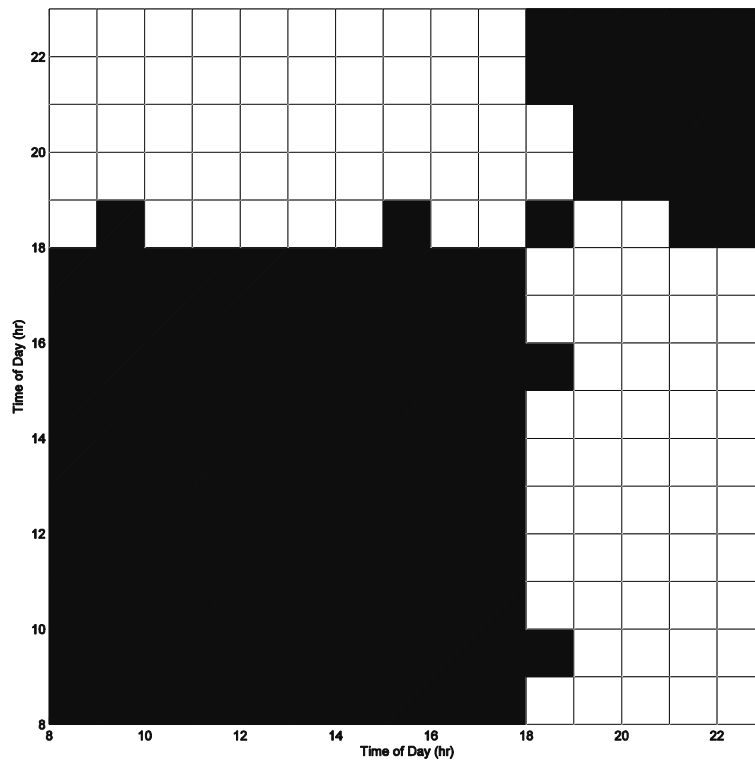


Fig. 5. Pairwise comparisons of the rates of return for the OO taxi trips in 15 elemental periods.

5. Conclusion and discussion

This paper proposed statistical methodologies to investigate taxi operations and applied the proposed methodologies to evaluate taxi operations in Shanghai's Hongqiao international airport. The analyses resulted in some interesting findings. For taxis that dropped off customers at the airport, the taxis that waited at the airport for customers had to bear longer waiting times, but they could have higher profits. For taxis that dropped off customers at areas other than the airport, the taxis that travelled to the airport empty for customers had longer vacant times and lower profits. The findings suggest that the vacant time minimization or profit maximization assumptions in the literature may not be able to capture the customer-search behaviours for taxis that serve the airport customers. In addition, the periods in which the profits are relatively homogeneous for taxis serving the airport are identified. The results suggest that taxis serving the airport customers could have significantly higher profits in the night time than in the daytime.

Given that taxis that travel to the airport empty for customers have significantly lower profits, it is surprising that many taxis are still willing to take non-profit trips to the airport. Such phenomenon may be explained from the following aspects. Firstly, taxi drivers may consider that waiting in the airport is more comfortable than driving and searching for passengers on the streets. Secondly, taxi drivers may think the longer vacant times in the airport could be compensated by the higher fare and, thirdly, taxi drivers do not like the uncertainty about the vacant time before picking up passengers and about the fare they can earn when they drive taxis in the urban areas.

Airport operators should make airport taxi trips profitable for taxi drivers to avoid taxi shortages. Understanding taxi travel patterns at the airport could help airport operators manage taxi operations more effectively. It is also valuable for taxi drivers. For example, obtaining real-time taxi operations at the airport could help taxi drivers to decide whether they should serve the airport customers.

Like all empirical studies, the results presented in this study are specific to the conditions of the airport under investigations. Nevertheless, our study adds to a rare body of literature that use large amount of taxi AVL data to analyze taxi operations at the airport. In addition, our analyses reveal certain limitations of existing taxi customer-search models. Moreover, the methodologies proposed in this study are general and could be applied for analyses at other airports. Comparing taxi drivers' behaviours at different airports and identifying key factors affecting taxi drivers' behaviours are worth pursuing in the future.

Acknowledgement

This research was supported by Science and Technology Commission of Shanghai Municipality (Grant no. 15DZ1204402). The work of the first author was supported by Program for Changjiang Scholars and Innovative Research Team in University.

References

- [1]. Austin, A.B. The Taxicab as Public Transportation in Boston M.S. Thesis. Massachusetts Institute of Technology: Massachusetts, 2011.
- [2]. Fu, H.W. and L.C. Qu. An Evaluation Method for the Estimated Travel Speed and Travel Time by Taxi Floating Car Data. *Applied Mechanics and Materials* 2012; **178-181**: 2805-2810.
- [3]. Wang, Z. Using Floating Cars to Measure Travel Time Delay: How Accurate Is the Method? *Transportation Research Record: Journal of the Transportation Research Board* 2004; **1870**: 84-93.
- [4]. Zhao, N., L. , et al. Analysis of Traffic Flow Characteristics on Ring Road Expressways in Beijing: Using Floating Car Data and Remote Traffic Microwave Sensor Data. *Transportation Research Record: Journal of the Transportation Research Board* 2009; **2124**: 178-185.
- [5]. Veloso, M., S. Phithakkitnukoon, and C. Bento. Urban Mobility Study Using Taxi Traces. in *the 2011 international workshop on Trajectory data mining and analysis*. Beijing, China, 2011.
- [6]. Li, X., et al. Prediction of Urban Human Mobility Using Large-Scale Taxi Traces and Its Applications. *Frontiers of Computer Science* 2012; **6**(1): 111-121.
- [7]. Liu, L., C. Andris, and C. Ratti. Uncovering Cabdrivers' Behavior Patterns from Their Digital Traces. *Computer, Environment and Urban Systems* 2010; **34**: 541-548.
- [8]. Zhang, D., Z. Peng, and D. Sun. A Comprehensive Taxi Assessment Index Using Floating Car Data. in *the 93rd Transportation Research Board*. Washington D.C., 2014.
- [9]. Amat, C., J. Ortigosa, and M. Estrada. Assessment of the Taxi Sector Efficiency and Profitability Based on Continuous Monitoring and Methodology to Review Fares. in *the 93rd Transportation Research Board*. Washington D.C., 2014.
- [10]. Pele, N. and C. Morency. When, Where and How Taxis Are Used in Montreal. in *the 93rd Transportation Research Board*. Washington D.C., 2014.
- [11]. Qian, X., X. Zhan, and S.V. Ukkusuri. Characterizing Urban Dynamics Using Large Scale Taxicab Data. in *the 93rd Transportation Research Board*. Washington D.C., 2014.
- [12]. Yang, H., et al. Nonlinear Pricing of Taxi Services. *Transportation Research Part A: Policy and Practice* 2010; **44**(5): 337-348.
- [13]. Wong, K.I., S.C. Wong, and H. Yang. Modeling Urban Taxi Services in Congested Road Networks with Elastic Demand. *Transportation Research Part B: Methodological* 2001; **35**(9): 819-842.
- [14]. Yang, H. and S.C. Wong. A Network Model of Urban Taxi Services. *Transportation Research Part B: Methodological* 1998; **32**(4): 235-246.
- [15]. Wong, K.I., et al. The Effect of Perceived Profitability on the Level of Taxi Service in Remote Areas. *Journal of the Eastern Asia Society for Transportation Studies* 2003; **5**: 79-95.
- [16]. Shanghai Airport Authority. <http://en.shairport.com/>. [July 10 2014].
- [17]. Hollander, M. and D.A. Wolfe. *Nonparametric Statistical Methods*. 2 ed. John Wiley & Sons, Inc: Danvers, MA, 1999.