Techno-Science Research Journal 8 (2020) 140-149

Content list available at ITC

Techno-Science Research Journal

Journal Homepage: www.ric.itc.edu.kh

Techno-Science Research Journal

Utilization of Ride-Hailing Services in Asian Developing Cities: Phnom Penh Case Study

Veng Kheang Phun[1,3*], Hironori Kato[2], Tetsuo Yai[3]

[1] *Transport Study Unit, Institute of Technology of Cambodia, Phnom Penh, Cambodia*
[2] *Department of Civil Engineering, The University of Tokyo, Tokyo, Japan*
[3] *Civil and Environmental Engineering Department, Tokyo Institute of Technology, Yokohama, Japan*

Received: 18 May 2020; Accepted: 07 August 2020; Available online: xxxx

Abstract: *The growing popularity of transport services via ride-hailing apps (RHAs) has posed major debates for transport researchers and planners in many cities worldwide. In Asia, there remains limited information and data about the characteristics of RHA users. This paper investigated the factors affecting users' utilization rate of RHA services. Zero-Inflated Ordered Probit Model (ZIOPM) was applied with data collected from 1,108 citizens (87.5% aged 12-29) in Phnom Penh, December 4-11, 2018, as a case study. It was found that approximately 55.0% of respondents were identified as non-users of RHAs. Three aspects that majority of users (78.6%) liked about RHAs were easy to find a ride, lower fare, and safety. Users traveled via RHAs for various trip purposes including home (18.1%), shopping (18.1%), social/private (15.7%), and work/school (24.5%). Modes that users used to travel via RHAs were auto-rickshaws (Indian Bajajs: 82.5% and Remorks: 6.9%), Taxis (5.5%), and motorcycles (2.9%), cars (0.8%), and others (1.4%). Modeling results showed that individuals with a higher income, higher daily travel cost, and longer time of using a smartphone were more likely to use RHAs more frequent than their counterparts. Those with older age and a motorcycle were likely to use RHAs less frequently. Results further showed that higher educated individuals were 6.4% more likely to use RHAs, while higher-income individuals were 4.3% less likely to use RHAs. The predicted probability of the ZIOPM was estimated to be within ±3.0%, compared with the actual probabilities of frequency of using RHAs. The findings from this study added further knowledge about characteristics of RHA users and about factors influencing on utilization rate of transport services via RHAs; especially among young citizens, who would be the potential RHA users.*

Keywords: Emerging Technology, Phnom Penh; Ride-hailing App; Ordered Probit Model

1. INTRODUCTION

The growing popularity transport services via ride-hailing apps (RHAs) has greatly influenced the way people travel (Schmitz et al., 2016). RHAs have been developed as an online-platform to connect customers and transport operators, facilitate their matching, and allow them to communicate more efficiently. Since the introduction of Uber in 2009, RHAs have experienced significant growth in adoption worldwide (He et al., 2018). The growing number of RHAs (e.g., Grab, PassApp, Didi) and RHA adoption among citizens have also seen in Asia. In Phnom Penh, the transport services via RHAs have started since 2016 (Phun et al., 2018).

Traditionally, citizens hail an empty-cruising LAMAT (= Locally Adapted, Modified, and Advanced Transport) on-streets or look for it at a particular pick-up station along the streets (e.g., near markets and intersections). The term "LAMAT" is a new concept and is proposed by Phun and Yai (2016). LAMAT is defined as the indigenous public transport modes that are locally, adapted, modified, and advanced for a certain transport service in a particular city or region. The term "LAMAT" has been proposed to replace the term "paratransit" in Asia, because the concept of paratransit in developing countries is quite different from that in developed countries. LAMAT includes all intermediate public modes between private transport and mass transit system, ranging from non-motorized two-wheelers (bicycle taxis) up to motorized four-wheelers (minibuses), with a maximum seating of about 25. Instead of hailing LAMAT on-streets, citizens can now easily hail for LAMAT with the press of a button on RHAs in their smartphones, and GPS takes care of the locations. RHAs notably help citizens to overcome the previous information barriers caused by spatial deviation between LAMAT users and drivers—and are thus believed to be a powerful instrument for improving efficiency of LAMAT market (Wang et al., 2016). RHAs improve the quality of on-demand transport services such as safety, convenience, and

* Corresponding author: Veng Kheang Phun
E-mail: kheang@itc.edu.kh

seamless experience (Adriano and Su, 2017)—and in particular, shorter waiting time could be an attractive feature of RHAs (He and Shen, 2015).

The advent of RHAs have created significant debates in many cities worldwide on various issues, including how RHAs should be regulated, their safety implications, their impacts on existing transport operators, and how they influence travel behavior (Clewlow and Mishra, 2017). Several studies assessed the impacts of RHAs on traditional transport services (e.g., Harding et al., 2016; Sadowsky and Nelson, 2017; Tan et al., 2017), but a few investigated the adoption and utilization of RHAs. For example, Clewlow and Mishra (2017) examined the differences between users and non-users of RHAs in major USA cities. Based on fundamental descriptive statistics, the authors found that 21.0% of adults personally used RHAs, an additional 9.0% used RHAs with friends but have not installed RHAs themselves. They also found that college-educated and affluent Americans have adopted RHAs at double the rate of less educated and lower income populations. Further, Lim et al., (2018) investigated behavioral adoption of RHAs in Malaysia, using multiple linear regression. Similarly, Peng et al. (2014) developed attitude-intention-behavior models to understand the adoption mechanism of RHAs in China. These studies were based on subjective responses to examine the causal relationship between behavioral adoption and other latent variables, including perceived usefulness, subjective norm, and perceived playfulness of RHAs. However, there are limited information and data about characteristics of RHA users as well as factors that have influence on their utilization rate. This poses significant challenges for transport researchers and planners to address the growing issues related to rapid penetration of RHA services.

The objective of this study is to fill the above gaps by exploring the characteristics of RHA users and to investigate the factors affecting their utilization rate. The factors are investigated in a Zero-Inflated Ordered Probit Model using survey data collected from 1,108 citizens in Phnom Penh, as a case study. The information and findings from this study, among the few previous studies, are expected to add further knowledge about users of RHA services in Asia. The information also serves as a basis for relevant authorities and transport planners to discuss suitable regulations and planning for the growing demand for ride-hailing services in Asia.

2. METHODOLOGY

2.1 Data Collection

2.1.1 Interview survey

A questionnaire-based interview survey was conducted with general citizens in Phnom Penh, December 4-11, 2018. The questionnaire was first written in English, and was later translated into Khmer. A pilot test was performed to ensure the consistency between the English and Khmer versions. The questionnaire contains five parts. Part 1 asked citizens about their general trips. Part 2 first asked them about the RHAs in Phnom Penh. Part 3 asked them about their viewpoints on mass transit system. Part 4 asked them about a stated preference choice. And Part 5 asked about their personal information. The perception questionnaire items were based on a 5-point scale (1: very unlikely, 2: unlikely, 3: neither, 4: likely, and 5: very likely). This study mainly analyzed the data from part 1, 2, and 5.

A sample must be representative of the selected population. The good sample is not only the representative, but also adequacy, and avoiding bias (Molugaram and Rao, 2017). The sample size $n > 30$ is known as large sample, and would be statistically sufficient for fundamental analyses (e.g., group mean comparison tests). For modeling, however, the sample size should be determined systematically. In this study, we determine the sample size using the 95% confidence level and the marginal error of 3%. With the population of approximately 1.8 million in Phnom Penh in 2017 (CSES, 2017), the sample sized was estimated to be 1,067.

Eleven surveyors, who were trained to fully understand and administer the questionnaire, visited several locations around Phnom Penh such as schools, markets, terminals, bus stops, and general public places. We adopted a simple random sampling technique; i.e., the surveyors did not request every citizen they saw to join the survey. Instead, they first observed and then verbally confirmed whether a targeted random respondent would join the survey. The surveyors requested approximately 1,390 citizens, but only 1,139 respondents voluntarily participated in the survey. Citizens rejected our requests because they were busy with school exam or their daily activities. Some were not in the good mood, and others did not like the survey. Respondents were recruited with an incentive gift. On average, each respondent took approximately 20 minutes to answer the questionnaire. After screening the information, only 1,108 sample were usable for further analyses.

2.1.2 Characteristics of respondents

The characteristics of the interviewed respondents are reported in Table 1. The majority were male (51.0%), single (84.4%), without a driving license (85.0%), young citizens aged between 12-29 (87.5%), students (67.0%), low-income individuals (less than 200 USD/month, 71.3%), and high-educated (senior high school or higher, 88.9%). Majority also owned a motorcycle (74.5%), and experienced traffic accidents at least once (54.7%). In sum, majority of the interviewed respondents were single, young, and high-educated students, who had monthly income level less than 200 USD. Young individuals were found to be users of

RHAs in major USA cities (Clewlow and Mishra, 2017). It should be noted that, as majority of respondents were young college students, the sampling may be biased and should focus more on young citizens who would be potential RHA users in the future. On the other hand, as can be seen in Table 1, there were 30.1% of respondents with age below 20. One may argue that the sample include many too young respondents who are not economically independent. We checked out sample data and found that only 8.3% of respondents with age below 20 had no income. In other words, 9.7% of respondents with no income were 20 years old or over. The source of their income could be from their parents or doing a part-time job. Therefore, we include all respondents for further analyses.

Table 1. Characteristics of respondents (N = 1,108)

Variable	Percentage	Variable	Percentage
Male	51.0%	*Education level*	
Single	84.4%	Never study	0.5%
With a driving license	15.0%	Grade 1-6	3.0%
Age		Grade 7-9	5.7%
12 to 19	30.1%	Grade 10-12	11.1%
20 to 29	57.4%	Bachelor	75.5%
30 to 39	6.7%	Master/PhD	2.3%
40 to 85	5.1%	Others	1.4%
Missing	0.7%	Missing	0.5%
Occupation		*Own a vehicle*	
Own business	5.0%	No	17.8%
Staffs/Official	20.2%	Motorcycle	74.5%
Students	67.0%	Car	4.3%
Others	7.6%	Bicycle	2.3%
Missing	0.2%	Others	1.0%
Income level (USD/month)		Missing	0.1%
No income	18.0%	*Experiences of traffic accidents*	
1-100	30.6%	Never	45.3%
101-200	22.7%	One	23.3%
201-400	20.8%	2 to 3	23.8%
> 400	7.6%	4 to 20	5.5%
Missing	0.3%	Missing	2.1%

1USD ≈ 4,000 KHR

Figure 1 shows the frequency of using a transport mode. The modes that were used almost every day by the respondents were own vehicles such as motorcycle and car (71.8%). Three modes were rarely used (≤1x a month): auto-rickshaws such as Remorks or Bajajs (62.3%), Motodop (55.8%), and public bus (44.9%). The modes that were never be used by the respondents were public bus (41.6%), Motodop (37.1%), and auto-rickshaws (27.3%).

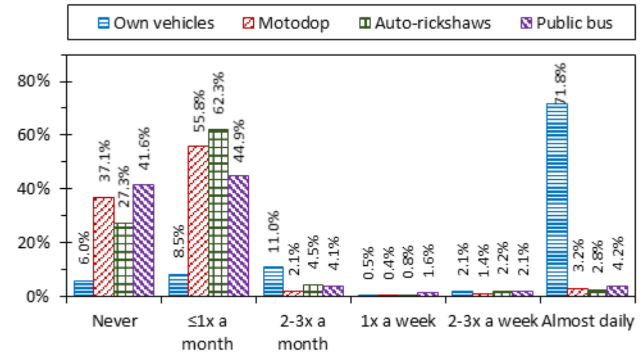

Figure 1. Frequency of using a transport mode (N = 1,108)

2.1.3 Characteristics of RHA users

Most of the interviewed respondents possessed a smartphone (91.7%), of which 49.6% were Android devices. Figure 2 shows the adoption and utilization of RHAs in Phnom Penh. It was found that 36.8% of the respondents have personally used and installed RHAs on their smartphones, and an additional 8.3% have used RHAs with friends but have not installed RHAs on their smartphones. The adoption rate of RHAs in Phnom Penh (45.1%) appeared to be higher than that in major USA cities (30.0%) (Clewlow and Mishra, 2017). On the other hand, other 46.5% have heard of RHAs, but they have not used RHAs. Some 8.4% have never heard of RHAs. Approximately 54.9% of the respondents have not used RHAs. About 35.9% reported that they rarely use RHAs, once or less per month. Some 5.1% use RHAs from 2 two 3 times per month, while a low rate (4.2%) use RHAs on a weekly to almost daily basis. In sum, 45.1% among all respondents (i.e., N = 500) were identified as users of RHAs.

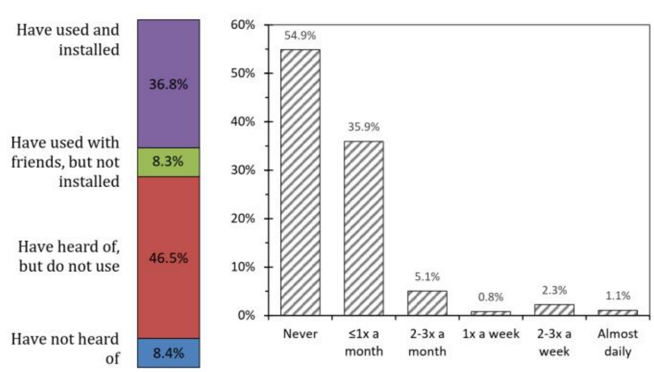

Figure 2. Adoption and utilization of RHAs among respondents (N = 1,108)

In part 2 of the questionnaire, respondents who utilized RHAs were further requested to report their trip characteristics via RHAs. About 24.3% of the respondents installed one RHA, 9.6% installed two RHAs, and 3.0% installed from 3 to 10 RHAs. Users may installed multiple RHAs in order to travel by a different RHA when one RHA was not available. Figure 3 shows the transport modes that users traveled via RHAs. The 500 RHA users provided n = 510 responses. Majority of the RHA users traveled by auto-rickshaws (82.5% by Bajajs and 6.9% by Remorks). This is reasonable because Bajajs, the motorized three-wheelers imported from India, has just gained its popular transport services following the advent of RHAs in Phnom Penh (Phun et al., 2018). The number of Bajajs registered quickly increased from 3,232 in February to 14,338 in November 2018. Only few Remorks, the two-wheeled carriage pulled by a motorcycle, have registered themselves to operate with RHAs (Phun et al., 2020). Some 5.5% used RHAs to travel by Taxis, and 3.7% traveled by motorcycles and cars. From this data, we could divide the respondents into two main groups: Non-users (54.9%) vs. users of RHAs (45.1%).

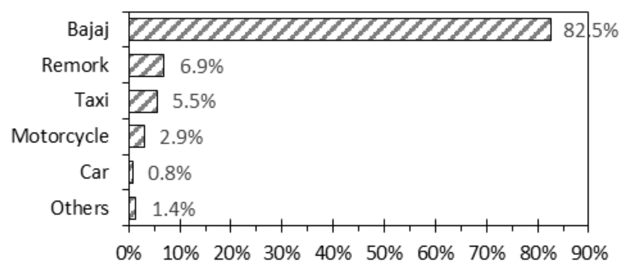

Figure 3. Travel modes via RHAs (n = 510)

Figure 4 shows the trip purposes of RHA users. The data revealed that RHAs were used for various trip purposes including home (18.1%), shopping activities (18.1%), social/private (15.7%), schools (15.3%), work (9.2%), medical centers (3.5%), and business (0.7%). A number of users (19.5%) used RHAs for other trip purposes including connective transport to intracity/intercity bus stop/terminal.

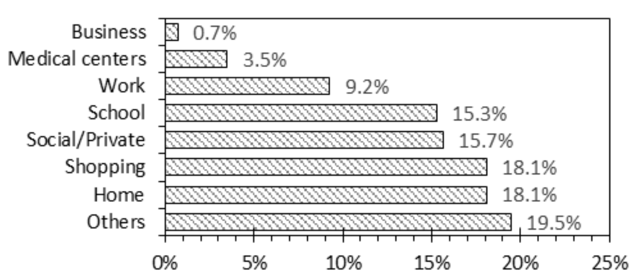

Figure 4. Trip purposes of RHA users (n = 575)

Trips for home/work/school/business accounted for 43.3%, while trips for leisure activities (i.e., shopping, social, or private) accounted for 33.8%.

Respondents were requested to freely describe up to three aspects about what they liked about the transport services via RHAs. 1,271 mixed responses were received and were later categorized as shown in Figure 5. The top four aspects that users liked about RHAs were easy to find a ride (33.1%), lower fare (23.9%), safety (21.6%), and not to drive by themselves (17.3%). Some 3.5% reported that they used RHAs because they did not have other travel options.

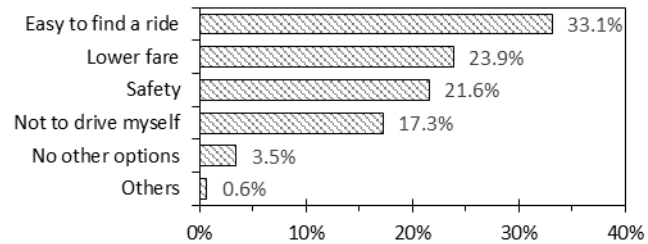

Figure 5. Aspects that users liked about transport services via RHAs (n = 1,271)

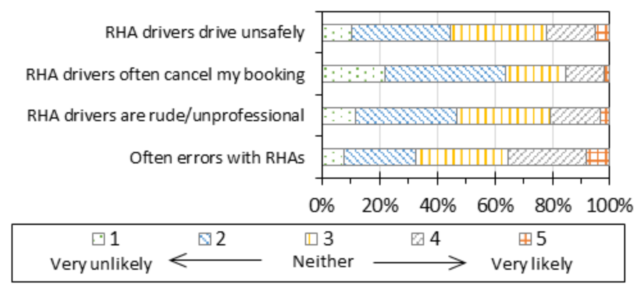

Figure 6. Perceived negative aspects of transport services via RHAs (N = 500)

Respondents were also asked to evaluate on four subjective questions about negative aspects that they experienced when using transport services via RHAs. They evaluated based on the 5-point scale (1: Very unlikely, 2: Unlikely, 3: Neither, 4: Likely, and 5: Very likely). Their responses were reported in Figure 6. Minority (22.0%) of the users (scores of 4 and 5) reported that the RHA drivers exhibited unsafe driving. Some 15.4% of the users seemed to experienced cancellation of their bookings via RHAs. 20.8% thought that the drivers operated via RHAs tended to be rude or unprofessional. 35.2% were likely to experience the errors within RHAs (e.g., errors in RHAs, internet, and digital map). In sum, only about a quarter (average of 23.4%) of RHA users appeared to have negative perception about unsafe driving behavior, booking cancellation, rude/unprofessional drivers, and errors within RHAs.

In major USA cities, it was found that early RHA adopters tend to be younger, more educated, and have higher income than the rest of population (Clewlow and Mishra, 2017). In this study, the data showed similar findings.

2.2 Zero-Inflated Ordered Probit Model

2.2.1 Model description

Zero-Inflated Ordered Probit Model (ZIOPM) fits a model for a discrete ordered response with a high fraction of zeros—this is called "zero inflation" because the lower-end zeros are overly dominant (STATA, 2019). In the context of ZIOPM, zero is an actual 0 value or the lowest outcome category. Like the Ordered Probit Model (OPM), the actual values taken by the ordered response variable are irrelevant. The inflation is assumed to occur in the lowest value to ensure that shifting the levels of the ordered response variable by a constant will not affect the estimated parameters in the model, which is common in OPM.

ZIOPM accounts for the zero inflation by assuming that the zero-valued responses come from both an OPM and a Probit Model (PM), allowing potentially different sets of covariates for each model (Harris and Zhao, 2007). In the traditional OPM, all observations with zero-valued responses are treated as a homogenous group. In contrast, ZIOPM assumes that zeros could occur in the data as members of two latent (unobservable) groups: nonparticipation group and participation group. Nonparticipation group always has the outcome 0 as the only possible value. Participation group, in addition to 0, has the other values (1, 2, 3, …) as the outcomes. The result of having two groups is an inflation in the proportion of zero-valued observations in data.

ZIOPM has been lately applied in several studies (e.g., Bagozzi et al., 2015; Fountas and Anastasopoulos, 2018; Jiang et al., 2017; Kelly and Anderson, 2008), including the adoption of new building technologies (e.g., Ganguly et al., 2010). The classic application of ZIOPM is the study of tobacco use in Harris and Zhao (2007). ZIOPM was developed using two-stage decision process. An individual must decide whether to participate in an activity (e.g., smoking) and, conditional participating, must decide on the level of participation (e.g., smoking amount/frequency), which also includes zero participation. The first decision is a binary choice and is modeled using a PM, while the second decision is an ordered choice and is modeled using an OPM. To account for the excess of zeros, Harris and Zhao (2007) allowed for zero observations to happen in two ways: (1) as a realization of the PM (i.e., nonparticipant) and (2) as a realization of OPM when the binary random variable in the PM is 1 (i.e., participation with zero activity). Further, Xu et al. (2019) developed ZIOPM to examine the probability of multiple secondary crashes after the occurrence of one primary crash with real time traffic flow, geometric, weather, and primary crash characteristics. As most crashes (> 90.0%) do not lead to any secondary crashes, there are excess zeros frequency of secondary crashes in the dataset. It was believed that, given the first crash information, ZIOPM can effectively model the frequency of secondary crashes with excess zero values. ZIOPM is also expected to produce a better model fitness and less bias than OPM. ZIOPM can be considered as the combination of the binary PM and OPM. The parameters in these models are estimated simultaneously in a ZIOPM. In line with this situation, it is also plausible that ZIOPM can be applied to model the frequency of using ride hailing services. In other words, ZIOPM can be used to investigate the factors affecting utilization frequency of transport services via RHAs.

2.2.2 ZIOPM for frequency of using RHAs

ZIOPM is considered suitable for modeling the frequency of using RHAs because, as can be seen in Figure 2, the frequency of using RHAs contains excess lower-end values—i.e. there are numerous respondents (54.9%) who have never used RHAs (i.e., non-users of RHAs). We believe that the observation for respondents who have never used RHAs is inflated. Similar to Harris and Zhao (2007) and Xu et al. (2019), our assumption is that individual must decide whether to use RHAs and must decide on the usage level of RHAs, which also includes zero usage. This assumption is based on the questions in Part 2 of the questionnaire. First, respondents were asked "Have you ever heard of ride-hailing apps (e.g., Uber, Grab, PassApp, iTsumo, WeGo)?" They should select one of the four answers: (1) Have not heard of, (2) Have heard of, but do not use, (3) Have used with friends, but not installed, and (4) Have used and installed. If respondents chose answer (3) or (4), it means they would be users of RHAs and they would further be asked to provide the frequency of using RHAs (never, 1x a month or less, 2-3x a month, 1x a week, 2-3x a week, and Almost daily). These frequency responses are treated as the dependent variable in this study. The first decision is a binary choice and is modeled using a PM, while the later decision is an ordered choice and is modeled using an OPM. To accommodate the excess lower-end values in the dataset of frequency of using RHAs, therefore, ZIOPM that combine the binary PM and OPM is developed and employed in this study. The parameters in binary PM and OPM will be estimated simultaneously in the ZIOPM.

2.2.3 ZIOPM formulation

Let Y denotes a discrete ordered response with levels coded as 0, 1, 2, …, *H* (STATA, 2019). For notional simplicity, the zeros are assumed to be inflated, but the following derivation maybe adapted to accommodate inflation in the lowest outcome category (i.e., never use an RHA). Harris and Zhao (2007) derived the ZIOPM in two steps. First, the group membership (users vs. non-users of RHAs) can be modeled using PM. Let $s_j = 1$ if the *j*th individual belongs to the group of RHA users or let $s_j = 0$ otherwise. With the PM, the probability of using an RHA is given by:

$$\Pr(s_j = 1/\boldsymbol{z}_j) = \Phi(\mathbf{z}_j\gamma) \qquad \text{(Eq.1)}$$

where $\mathbf{z}_j$ is a vector of covariates that determines group membership, γ is a vector of coefficients to be estimated, and $\Phi(.)$ is the standard normal distribution function. By conditioning on $s_j = 1$, the levels of using RHAs $\tilde{y}_j$ are modeled using an OPM; these levels may also include 0. The corresponding probabilities are given by

$$\Pr(\tilde{y}_j = h|\mathbf{x}_j, s_j = 1) = \Phi(k_h - \mathbf{x}_j\beta) - \Phi(k_{h-1} - \mathbf{x}_j\beta) \qquad \text{(Eq.2)}$$

where $h = 0, 1, 2, \dots, H;\ k_{h-1} = -\infty; k_H = +\infty; and\ \mathbf{x}_j$ is a vector of covariates that could be different from $\mathbf{z}_j$. k_h are boundary parameters that need to be estimated in addition to the coefficients vector β.

The intercept β_0 is set equal to 0 in Eq. (2) for identification. Note that s_j and $\tilde{y}_j$ are both unobservable in terms of the zeros. The observed response variable is $y_j = s_j\tilde{y}_j$. Thus, the zero outcome occurs when $s_j = 0$ (the individual is not a user of RHAs) or occurs when $s_j = 1$ and $\tilde{y}_j = 0$ (the individual is a user of RHAs with zero activity). To observe a positive y_j, it is a joint requirement that $s_j = 1$ and $\tilde{y}_j > 0$.

The distribution of Y is given by:

$$\Pr(Y) = \begin{cases} \Pr(y_j = 0|\mathbf{z}_j, \mathbf{x}_j) \\ \Pr(y_j = h|\mathbf{z}_j, \mathbf{x}_j) \end{cases}$$

$$\Pr(Y) = \begin{cases} \Pr(s_j = 0|\mathbf{z}_j) + \Pr(s_j 1|\mathbf{z}_j)\Pr(\tilde{y}_j = 0|\mathbf{x}_j, s_j = 1) \\ \Pr(s_j = 1|\mathbf{z}_j)\Pr(\tilde{y}_j = h|\mathbf{x}_j, s_j = 1) \end{cases} \qquad \text{(Eq.3)}$$

where $h = 1, 2, \dots, H$. The probability of zero outcome has been inflated because it is the sum of the probability of zero activity from the OPM and the probability of non-users of RHAs from the PM. By substituting Eq. (1) and Eq. (2) in Eq. (3), we get

$$\Pr(Y) = \begin{cases} \Pr(y_j = 0|\mathbf{z}_j, \mathbf{x}_j) \\ \Pr(y_j = h|\mathbf{z}_j, \mathbf{x}_j) \\ \Pr(y_j = H|\mathbf{z}_j, \mathbf{x}_j) \end{cases}$$

$$\Pr(Y) = \begin{cases} \{1 - \Phi(\mathbf{z}_j\gamma)\} + \Phi(\mathbf{z}_j\gamma)\Phi(k_0 - \mathbf{x}_j\beta) \\ \Phi(\mathbf{z}_j\gamma)\{\Phi(k_h - \mathbf{x}_j\beta)\Phi(k_{h-1} - \mathbf{x}_j\beta)\} \\ \Phi(\mathbf{z}_j\gamma)\{1 - \Phi(k_{H-1} - \mathbf{x}_j\beta)\} \end{cases} \qquad \text{(Eq.4)}$$

where $h = 1, 2, \dots, H - 1$. The corresponding log-likelihood function is

$$\ln L = \sum_{j=1}^{N} w_j \sum_{h=0}^{H} I(y_j = h) \ln\{\Pr(y_j = h|\mathbf{z}_j, \mathbf{x}_j)\} \qquad \text{(Eq.5)}$$

where w_j is an optional weight for the *j*th observation (not included in this study due to the unavailable data for defining w_j) and $I(y_j = h) = 1$ if $y_j = h$ and 0 otherwise. Noted that the choice between the ZIOPM and the OPM cannot be made using a likelihood-ratio test because the two hypotheses are not nested in the usual sense of parameter restrictions (STATA, 2019). The restriction $\gamma = 0$ does not eliminate the inflation effect; it makes the group membership probabilities both equal to 0.5. What is needed to remove the inflation effect is $\mathbf{z}_j\gamma \to \infty$, which cannot be imposed. In particular, Vuong test has been suggested for testing between two non-nested models (Greene, 2003). For the standard formulas of OPM and PM, see Downward et al. (2011) and Phun et al. (2015).

Table 2 reports the summary statistics of variables for the models. The dependent variable is the frequency of using RHAs. Since the actual values taken by the ordered response variable are irrelevant, we label the frequency of using

Table 2. Summary statistics and descriptions of variables for the models

Variable	All (*N* = 1,108)		Users (*N* = 500)		Non-users (*N* = 608)		Label description
	Mean	SD	Mean	SD	Mean	SD	
Frequency of using RHA service	1.63	0.93	2.39	0.93	1	0	1: Never, 2: 1x a month or less, 3: 2-3x a month, 4:1x a week, 5: 2-3x a week, and 6: Almost daily
Age	23.32	8.30	22.11	6.39	24.31	9.48	Age in years
D_female	0.47	0.49	0.48	0.49	0.47	0.49	Female = 1, Male = 0
Income	2.72	1.27	2.75	1.24	2.70	1.30	1: No income, 2: 1-100, 3:101-200, 4: 201-400, 5: 401-800, 6: 801-1200, 7: 1201-1500, 8: >1,500 USD per month
Phone_years	2.80	2.11	3.26	2.19	2.42	1.97	Years of using a smartphone, equal 0 for individual without a smartphone
D_motorcycle	0.77	0.42	0.77	0.42	0.77	0.42	Own a motorcycle = 1, otherwise = 0
Daily travel cost	4.48	5.16	4.48	4.19	4.47	5.85	Average daily travel cost, in 1000 KHR
Education	4.74	0.89	4.86	0.67	4.64	1.03	1: Never study, 2: Grade 1-6, 3: Grade 7-9, 4: Grade 10-12, 5: Bachelor, 6: Master, 7: PhD, and 8: Others

1USD ≈ 4,000 KHR

RHAs as follow—1: Never, 2: 1x a month or less, 3: 2-3x a month, 4:1x a week, 5: 2-3x a week, and 6: Almost daily. The lowest outcome category "1: Never" is equivalent to zero value in the ZIOPM. Regarding the explanatory variables, results from Welch's *t*-tests showed that the age of those who adopted RHAs was 22.11, in average, significantly younger than those who did not adopted RHAs (24.31 years old) [t (1106.00) = 4.43, $p < 0.001$]. In addition, users of RHAs appeared to have a higher education level than non-users [t (1051.19) = -4.034, $p < 0.001$]. Users of RHAs also had a longer period of using a smartphone (3.26 years) than non-users (2.42 years) [t (1106.00) = -6.6906, $p < 0.001$]. There are no specific differences in terms of gender, income, motorcycle ownership, and daily travel cost ($p > 0.05$). There is no problem of multicollinearity as the highest correlation among explanatory variables is 0.39 (Age vs. Income).

3. RESULTS

Table 3 reports the estimate results of OPM and ZIOPM for comparison. We performed Vuong test for the null hypothesis that the inflate part of the model is unnecessary (Vuong, 1989). We found that the null hypothesis is rejected at 5% significant level (t = 2.43, p = 0.0076), indicating that the inflate part of the ZIOPM is necessary. Therefore, the results by ZIOPM are favored.

There are three sets of coefficients in the ZIOPM. The first set of coefficients, which is labeled as "Frequency of using RHAs", corresponds to the participation (RHA usage) levels. These coefficients could be interpreted in the same way as those in OPM. The second set of coefficients, which is labeled as "Inflate", corresponds to the equation for the participation decision (i.e., membership in the never-used RHA group). These coefficients could be interpreted in the same way as those in PM. A positive (negative) coefficient of an explanatory variable indicates that an increase (decrease) in this variable associated with an increase in the probability of the frequency of using an RHA (Xu et *al.*, 2019). It should be noted ZIOPM does not require the same variables for both participation (Frequency of using RHAs) and decision (Inflate) equations. If the same variables are included, their coefficients commonly have the opposite signs (STATA, 2019). For example, the income variable has the positive sign in the participation level equation and negative sign in the decision of participation equation. And the third set of parameters, which is labeled as "Boundary parameters", corresponds to the cut-off points.

In the participation equation, we included six explanatory variables, five of which are statistically

Table 3. Regression results

Variable	OPM		ZIOPM	
	Coefficient	*t*-stats	Coefficient	*t*-stats
Frequency of using RHAs				
Age	-0.0247 ***	-4.61	-0.0191 **	-3.10
D_female	0.1126	1.54	0.1309 *	1.70
Income	0.0792 **	2.53	0.1456 ***	3.53
Phone_years	0.1102 ***	6.48	0.1030 ***	5.53
D_motorcycle	-0.1499 *	-1.74	-0.1942 **	-2.12
Daily travel cost	0.0139 **	2.04	0.0254 **	2.14
Inflate				
Education			0.6453 ***	3.99
Income			-0.4331 ***	-3.49
Boundary parameter				
k_1	0.0812	0.55	0.2854 *	1.70
k_2	1.3291 ***	8.70	1.6001 ***	8.71
k_3	1.7420 ***	10.89	2.0305 ***	10.51
k_4	1.8444 ***	11.34	2.1367 ***	10.87
k_5	2.3213 ***	12.58	2.6258 ***	12.05
N		1108		1108
Log-likelihood		-1098.37		-1087.73
Vuong test of ZIOPM vs. OPM				*t*-stats = 2.43, p = 0.0076

$^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

significant at 5% confidential level. The estimate coefficient for age is negatively significant ($p < 0.05$), indicating that older individuals are less likely to use RHAs more often, than younger ones. This result can be explained by the increasing awareness of new technologies such as RHAs among young people. Older individuals may not be as tech-savvy as younger ones. Further, the positive and statistically significant income variable shows that higher income individuals are more likely to use RHAs more frequent than lower income individuals ($p < 0.01$). These results are consistent to that found among American adults (Clewlow and Mishra, 2017). Similarly, the years of using a smartphone has a positive coefficient, indicating that individuals who have used a smartphone for longer time are more likely to use RHAs. This is plausible because individuals with a smartphone would have more knowledge about smartphone's apps, thus making them easier to use RHAs for their cheaper and more convenient travel options around the city. Besides, the positively significant coefficient of daily travel cost explained that individuals with higher travel cost per day are more likely to use RHAs. One plausible reason is that individuals with ability to pay higher cost for daily commutes were also able to afford to use transport services via RHAs. Moreover, D_motorcycle has a negative coefficient, suggesting that individuals who owned a motorcycle were less likely to travel via RHAs. This is true because they would regularly ride their own motorcycles, rather than using RHAs. Further, D_female is positively significant at 10% level, indicating that females are more likely to user RHAs than males. One possible explanation is that females may consider more about their own safety—they tried to avoid going into the streets and hailing a transport mode.

In the participation decision equation (or Inflate), the two explanatory variables (Education and Income) are statistically significant at 1% confidential level. The positive coefficient of Education indicates that individual with a higher education level is more to use an RHA. The negative coefficient of Income indicates that individual with a higher income level is less likely to use an RHA. This is plausible because higher income individuals might prefer private vehicles (Rashedi *et al.*, 2017), rather than depending on public transport or ride-hailing services.

However, the coefficients in the Inflate part are difficult to interpret. For example, what does an Income coefficient of -0.4331 mean? As, under the ZIOPM, the sign of Income coefficient counter each other in the participation level using OPM (positive sign) and the participation decision using binary PM (negative sign), this make the corresponding result interpretation uneasy. Since the estimated coefficients in ZIOPM are not particularly informative, and as all discrete choice models, marginal effects are better to interpret (STATA, 2019). We additionally estimated the average marginal effects for Education and Income, by computing the predicted probability for non-users of RHAs—i.e., the probability of being "never use" an RHA in the participation decision model (or Inflate). The average marginal effects for Education and Income are -0.0641 (t = -4.25) and 0.0430 (t = 3.04), respectively. This means that, on average in the data, individual with a higher education level was about 6.4% less likely to be a non-user of RHAs, than those with a lower education level. In the other words, higher educated individual is 6.4% more likely to be a user of RHAs. Besides, individual with a higher income level was 4.3% (t = 3.04) more likely to be a non-user of RHAs, than those with a lower income level. In sum, the effect of Income in ZIOPM might suggest that individuals with a higher income level are about 4.3% less likely to use RHAs, but, when they become RHA users, they use RHAs more frequently, than those with a lower income level.

Further, we computed the predicted probability of the ZIOPM for frequency of using RHAs as follows: 52.3%, 38.9%, 5.1%, 0.8%, 2.1%, and 0.9% (which are close to the actual probabilities in Figure 2: 54.9%, 35.9%, 5.1%, 0.9%, 2.3%, and 1.1%, respectively) for Never, ≤1x a month, 2-3x a month, 1x a week, 2-3x a week, and Almost daily, respectively. The differences between the predicted and actual probabilities are within ±3.0%. This implies that ZIOPM performed reasonably well for identifying factors affecting the frequency of using RHAs.

The estimate marginal effects are shown in Table 4. The sum of these effects for each variable equal zero (e.g., Age: 0.0067 – 0.0037 – 0.0013 – 0.0003 – 0.0008 – 0.0005 = 0). The interpretation is said, for example, an increase in the

Table 4. Marginal effects of frequency of using RHAs

Variable	Never	≤1 a month	2-3x a month	1x a week	2-3x a week	Almost daily
Age	0.0067 **	-0.0037 **	-0.0013 **	-0.0003 **	-0.0008 **	-0.0005 **
D_female	-0.0458 *	0.0257 *	0.0092 *	0.0018	0.0056	0.0036
Income	-0.0288 **	0.0114	0.0076 **	0.0015 **	0.0050 **	0.0033 **
Phone_years	-0.0361 ***	0.0202 ***	0.0072 ***	0.0014 **	0.0044 ***	0.0028 **
D_motorcycle	0.0680 **	-0.0381 **	-0.0136 **	-0.0026 *	-0.0083 **	-0.0054 *
Education level	-0.0089 **	0.0050 **	0.0018 **	0.0003 *	0.0011 **	0.0007 *
Daily travel cost	-0.0330 **	0.0255 **	0.0039 **	0.0007 **	0.0019 **	0.0010 **

*$p < 0.1$, **$p < 0.05$, ***$p < 0.01$

individual age by one year is associated with being approximately 0.67% more likely to be a non-user of RHAs, 0.37% less likely to use RHAs at most once a month, and 0.05% less likely to use RHAs almost daily.

LAMAT services with RHAs appear to provide several benefits to the society including a lower fare and a more reliable transport services for citizens, a higher revenue for drivers who adopted RHAs, a shorter waiting/travel time via RHA's platform, and a lower air polluting emission via the suggested shortest paths in RHA's digital maps as well as a reduction in the cruising time looking for costumers along the streets (He et al., 2018; Phun et al., 2018). The growing popularity of transport services via RHAs has also been seen as threats to existing transport providers (Clewlow and Mishra, 2017; Tan et al., 2017; Phun et al., 2020)—e.g., lower demand for conventional Taxi drivers in several cities. However, whether to encourage or discourage the use of RHAs for LAMAT remains questionable and quite depends on the local context. LAMAT with RHAs may not be a sustainable transport mode. Yet, it provides better quality of transport services (e.g., safer, faster, more convenience) to citizens in needs, especially in developing cities. The relevant authorities may want to promote the LAMAT with RHAs for two conditions: (1) citizens mainly depend on LAMAT with RHAs to travel around their cities, when there are no sufficient mass transit services (bus and rail services); and (2) citizens might use LAMAT with RHAs as connective first/last-mile transport mode to complete their trips, when there are sufficient mass transit services. Phnom Penh city falls in the first condition, in which the minimal regulations (e.g., driving behavior, vehicle status) should be imposed to ensure safety for all road users. However, the advent of RHAs in Asia is still new, and the authorities may incorporate the findings in this study into the policy discussion, that is suitable to regulate the ride-hailing services in their cities.

4. CONCLUSION

This paper investigated the factors affecting users' utilization rate of RHA services in Phnom Penh. Among 1,108 interviewed respondents, there were 54.1% who experienced traveling via RHAs. Four aspects that majority of users (95.9%) liked about RHAs were easy to find a ride, lower fare, safety, and not to drive by himself/herself. These factors clearly demonstrate the favorable benefits of ride-hailing services over those of the traditional transport services (e.g., hailing for ride on streets, negotiable fare, security & safety concern, unidentifiable drivers). Users traveled via RHAs for various trip purposes including home (18.1%), shopping (18.1%), social/private (15.7%), and work/school (24.5%). Modes that users used to travel via RHAs were auto-rickshaws (Indian Bajajs: 82.5% and Remorks: 6.9%), Taxis (5.5%), and motorcycles (2.9%), cars (0.8%), and others (1.4%).

Moreover, results from ZIOPM showed that the individuals with higher daily travel cost and longer time of using a smartphone were more likely to use RHAs more frequent than their counterparts. Those with older age and a motorcycle were less likely to use RHAs. Results from the average marginal effects also showed that individuals with a higher education level were about 6.4% more likely to be users of RHAs. Moreover, under ZIOPM, individuals with a higher income level were about 4.3% less likely to use RHAs, but, when they became RHA users, they use RHAs more frequently, than those with a lower income level.

The findings from this study added further knowledge about characteristics of RHA users and about factors influencing on utilization rate of transport services via RHAs; especially among young citizens, who would be the potential RHA users. However, more studies are needed to explore how RHA services influence on travel behavior—e.g., to what extents RHA services would influence ridership of (future) mass transit services in Asia; and how older citizens view/use RHAs in the future. Finally, this study discussed only the users' preferences for the current RHA usage. The stated preference with different attributes of RHAs (e.g., travel time and travel cost) might also be an interesting future topic to extend the findings in this study.

ACKNOWLEDGEMENTS

The first author is grateful to the Nippon Foundation for its financial support via Japan Transport and Tourism Research Institute. The authors thank Mr. Chhay Horng CHHOUK and eleven anonymous students at Institute of Technology of Cambodia for their help with the data collection. The authors also thank Ms. Pharinet PHENG for help with proofreading. The contents of this paper reflect the viewpoints of the authors, who are responsible for any errors.

REFERENCES

Adriano, A.M.M., & Su, C.C.S. (2017). Out with the old, in with the new: A study on the vehicle hailing preferences of Filipino taxi riders based on participation intent. *International Journal of Real Estate Studies*, 11(1), 75−81.

Bagozzi, B.E., Hill, D.W., Moore, J.W.H., & Mukherjee, B. (2015). Modeling two types of peace: The zero-inflated ordered probit (ZiOP) model in conflict research. Journal of Conflict Resolution, 59, 728-752.

Clewlow, R.R., & Mishra, G.S. (2017). Disruptive transportation: The adoption, utilization, and impacts of ride-hailing in the United States. Institute of

Transportation Studies, University of California, Davis, Research Report—UCD-ITS-RR-17-07.

CSES (2017). Cambodian Socio-Economic Survey. Retrieved on 01 July, 2020 from https://www.nis.gov.kh/index.php/en/14-cses/12-cambodia-socio-economic-survey-reports

Downward, P., L-Lopez, F., & Rasciute, S. (2011). The Zero-Inflated ordered probit approach to modelling sports participation. Economic Modelling 28, 2469-2477.

Ganguly, I., Koebel, C.T., & Cantrel, R.A. (2010). A categorical modeling approach to analyzing new productadoption and usage in the context of the building-materials industry. Technological Forecasting and Social Change 77, 662-677.

Greene, W. (2003). Econometric Analysis, fifth ed. Prentice-Hall, Englewood Cliffs, NJ, USA.

Fountas, G., & Anastasopoulos, P.C. (2018). Analysis of accident injury-severity outcomes: The zero-inflated hierarchical ordered probit model with correlated disturbances. *Analytic Methods in Accident Research*, 20, 30-45.

Harding, S., Kandlikar, M., & Gulati, S. (2016). Taxi apps, regulation, and the market for taxi journeys. *Transportation Research Part A*, 88, 15−25.

Harris, M.N., & Zhao, X. (2007). A zero-inflated ordered probit model, with an application to modeling tobacco consumption. Journal of Econometric 141, 1073-1099.

He, F., & Shen, Z.-J. (2015). Modeling taxi services with smartphone-based e-hailing applications. *Transportation Research Part C*, 58, 93-106.

He, F., Wang, X., Lin, X., & Tang, X. (2018). Pricing and penalty/compensation strategies of a taxi-hailing platform. *Transportation Research Part C*, 86, 263−279.

Jiang, Y., House, L., Kim, H., & Percival, S.S. (2017). Zero-inflated ordered probit approach to modeling mushroom consumption in the United States. *International Food and Agribusiness Management Review*, 20(5), 655-672.

Kelley, M.E., & Anderson, S.J. (2008). Zero inflation in ordinal data: Incorporating susceptibility to response through the use of a mixture model. Statistics in Medicine 27, 3674-3688.

Molugaram, K., & Rao, G.S. (2017). *Statistical Techniques for Transportation Engineering*. Butterworth-Heinemann, United Kingdom.

Lim, K.B., Yeo, S.F., Goh, M.L., & Gan, J.A.X. (2018). A study on consumer adoption of ride-hailing apps in Malaysia. Journal of Fundamental and Applied Sciences, 10(6), 1132-1142.

Peng, L., Wang, H., He, X., Guo, D., & Lin, Y. (2014). Exploring factors affecting the user adoption of Call-taxi App. Proceedings of the 25th Australasian Conference on Information System, Auckland, New Zealand.

Phun, V.K., Masui, R., & Yai, T. (2018). Operational characteristics of paratransit services with ride-hailing apps in Asian developing cities: The Phnom Penh case. *Journal of Transportation Technologies*, 8, 291-311.

Phun, V.K., Peng, P., Masui, R., Kato, H., & Yai, T. (2020). Impact of Ride-Hailing Apps on Traditional LAMAT Services in Asian Developing Cities: The Phnom Penh Case. Asian Transport Studies (in press).

Phun, V.K., Pheng, P., & Yai. T. (2015). Using ordered pobit modeling to assess perceived bus performance in Phnom Penh. *Journal of the Eastern Asia Society for Transportation Studies*, 11, 1155-1172.

Phun, V.K., & Yai. T. (2016). State of the art of paratransit literatures in Asian developing countries. *Asian Transport Studies*, 4, 57−77.

Rashedi, Z., Mahmoud, M., Hasnine, S., & Habib, K.N. (2017). On the factors affecting the choice of regional transit for commuting in Greater Toronto and Hamilton Area: Application of an advanced RP-SP choice model. Transportation Research Part A: Policy and Practice, 105, 1-13.

Sadowsky, N., & Nelson, E. (2017). The impact of ride-hailing services on public transportation use: A discontinuity regression analysis. *Economics Department Working Paper Series*, 13.

Schmitz, C., Bartsch, S., & Meyer, A. (2016). Mobile app usage and its implications for service management—Empirical findings from German public transport. *Procedia-Social and Behavioral Sciences*, 224, 230–237.

STATA (2019). Zero-inflated ordered probit regression. Retrieved on 01 July, 2020 from https://www.stata.com/manuals/rzioprobit.pdf and https://www.stata.com/features/overview/zero-inflated-ordered-probit/

Tan, F.T. C., Tan, B., Lu, A., & Land, L. (2017). Delivering disruption in an emergent access economy: A case study of an e-hailing platform. *Communications of the Association for Information Systems*, 41, 497−516.

Vuong, Q. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. Econometrica, 57, 307–334.

Wang, X., He, F., Yan, H., & Gao, H.O. (2016). Pricing strategies for a taxi-hailing platform. *Transportation Research Part E*, 93, 212−231.

Xu, C., Xu, S., Wan, C., & Li, J. (2019). Investigating the factors affecting secondary crash frequency caused by one primary crash using zero-inflated ordered probit regression. *Physica A*, 524, 121-129.