Factors influencing ride-sourcing usage: An ordered hybrid choice model

Prawira F. Belgiawan[a*], Tri Basuki Joewono[b], Muhammad Zudhy Irawan[c]

[a] School of Business and Management,
Bandung Institute of Technology,
Jl. Ganesha No. 10, Bandung, 40132, Indonesia
fajar.belgiawan@sbm-itb.ac.id

[b]Department of Civil Engineering
Parahyangan Catholic University
Jl. Ciumbuleuit 94 Bandung, 40141, Indonesia
P: +62-22-2032655; F: +62-22-2031110
vftribas@unpar.ac.id

[c]Department of Civil and Environmental Engineering,
Universitas Gadjah Mada
Jl. Grafika No.2, Yogyakarta, 55281, Indonesia
zudhyirawan@ugm.ac.id

* Corresponding Author

Factors influencing ride-sourcing usage: An ordered hybrid choice model

Abstract
The emergence of ride-sourcing has substantially changed transportation services in the last decade. In a developing country such as Indonesia, where buses and rail system is not extensive enough, ride-sourcing has a great impact on paratransit and other transportation modes. This situation raises the question on whether ride-sourcing could weaken the role of urban public transport in Indonesia, particularly in Bandung. Thus, the objective of the paper is to investigate factors influencing ride-sourcing usage in Bandung. We distributed a questionnaire randomly to 497 respondents in Bandung. We also enquired our respondents on their travel and waiting time both in average, as well as the cost of the used mode We also gave several questions regarding attitudes of the subjects towards their ride-sourcing, in which from the principal component analysis, we have obtained six factors: reliability, fare, comfort, availability, accessibility, and service. From these six factors, we incorporate reliability and factors in the ordered hybrid choice model (OHCM). The independent variables for HCM are travel time, waiting time, travel cost (generic), two latent attitudinal factors for ride-sourcing and socio-demographic characteristics (age, and gender). We found that two latent variables are significant in ride-sourcing choice. Cost is proven to be negative in significance as expected, while waiting time is insignificant. Travel time is significant with positive sign. Age is found to be significant while gender is not.

Keywords:
ride-sourcing, ride-hailing, attitudes towards ride-sourcing, ordered hybrid choice model, developing country.

1. Introduction

In developed countries, ride-sourcing (or ride-hailing) would have greater impact on private motor vehicle use, where Harding et al. (2016) argued that taxi app will not substitute urban transportation (bus or rail systems) due to the price difference, except for some specific circumstances (i.e. journeys with luggage or tight time constraints). However, in the developing countries including Indonesia, ride-sourcing has a great impact on paratransit (such as taxi, jitney, *ojek*, or non-motorized paratransit), where the city buses and rail systems are not adequately extensive (Irawan et al., 2019a). Irawan et al. (2019a) found that ride-sourcing (motorcycle-based) in fact is a complement rather than a substitute for public transit in Jakarta. Different results found that bus services in Yogyakarta have a significant competitor from the motorcycle-based ride-hailing (Irawan et al., 2019b).

These different phenomenas hypothetically differentiate the impact of ride-sourcing between developed and developing countries. This situation thenn raises the question of whether ride-sourcing will weaken the role of urban public transport in developing countries. Therefore, we aim to investigate factors influencing ride-sourcing usage in Bandung, Indonesia. Our results will partly answer this question as an anticipation of future urban transportation in developing countries.

2. Data collection method

The set of questionnaires was constructed based on previous works of literature concerning ride-sharing and ride-sourcing. In the first section of the questionnaire, respondents were asked on their frequency of travel using ride-sourcing within the past 30 days. The result range of answer vary from once (less than four times), couple times (4-6 times), often (7-14 times), very often (15-30 times), and always (more than 30 times). These five categories further converted into five point ordinal scale variable of frequency of using ride sourcing range from 1 (once) to 5 (always). Further this variable becomes the dependent variable for the ordered hybrid choice model.

Going futher, respondents were asked to describe the length of time that they had to wait for the driver to pick them at the stated location. Other than that, we also enquired our respondents to rate, on the time length of their travel whilst using the ride-sourcing service. Next, details on the fare of commuting using ride-sourcing was asked to the respondents. For the second section, we queried several questions regarding attitudes towards ride-sourcing that we have grouped according to factors of availability, accessibility, reliability, comfort, safety, productivity, and others. We asked the respondents to rate the extent of their agreement regarding these 26 attitudinal factors indicators on five points Likert scale (ranging from fully disagree - fully agree). Finally, we requested the sociodemographic information such as age, gender, and income from the individuals.

The survey was conducted from 24 April 2018 to 14 May 2018 with the help of 20 surveyors. After the completion of questionnaire distribution, the data were evaluated based on completeness and found that 497 sets of the questionnaire can be used for deeper analysis. The respondents reported that they usually wait less than 10 minutes before the drivers arrive to pick them up, however, some wait until 20 minutes due to congested traffic. For the travel time, on average, they traveled for around 30 minutes whilst using the ride-sourcing service. Based on their perception, the travel time using this application-based service is faster around 15-30 minutes compared with their commonly used mode. The amount of cost of travel using this service is around IDR 10 to 30

thousand (equal to 0.6 up to 2 USD) per trip; though some pay more than IDR 30 K. The respondents paid mostly by cash. This fare confirms the distance that travels using ride-sourcing is short to medium, or in other words, it is a travel means inside the city. Despite that, withinthe last two years, the growth of cashless payment is ascending in line with the growing of virtual wallet.

The respondents of this study were dominated by young people with a mean age of 27.5 years old.. Around 60 percent are male with senior high school or university as their highest education. The respondents can be concluded having an occupation as a student (36%) and as a worker in a private institution or as an entrepreneur (28.6% or 16.9%). They have a working or studying time with a mean of 7.26 hours per day and around 6 days per week. This group of young individuals have a monthly income above the monthly minimum wages for the region of Bandung, i.e., IDR three million. Around 30% of the respondent can be classified as medium to high-income groups.

3. Ordered hybrid choice model of ride-sourcing

For the attitude towards ride-sourcing, some variables can not be constructed into attitudinal construct since the Cronbach alpha is below 0.6. Therefore we performed principal component analysis (PCA) to construct uncorrelated factors of attitude towards ride-sourcing. Firstly, we performed PCA with 26 indicators, then we discovered that 3 out of 26 have very low loading factor, therefore we excluded those indicators. We continued the PCA with 23 indicators. From 23 indicators, we gained six attitudinal factors. We obtained these six factors based on eigen value greater than one and varimax rotation. We labelled the six factors based on the indicators that construct the factors i.e: reliability, fare, comfort, availability, accessibility, service. From these six factors we atempted to test the factors which have significant correlation with ride-sourcing mode choice, and we traced that reliability and comfort are two factors that significantly correlate with decision to choose ride-sourcing, hence we have included these two factors in the subsequent model.

3.1 Ordered hybrid choice model

We estimated a hybrid choice model (HCM) with two ride-sourcing attitudinal factors. Hybrid choice model combines latent variable modelling with discrete choice models (Ben-Akiva et al., 2002; Walker and Ben-Akiva, 2002). Table 1 summarizes the indicator for latent variables, choice related attributes, and respondent characteristics as the explanatory variables for mode choice decision.

Table 1 Measurement of indicators, choice attributes, and respondents characteristics.

Variables		Indicators		Measurement
Latent Variables				
Z_{RI}	Ride-sourcing attitudes-reliability (RI)	I_1	faster than other modes	1 = strongly disagree; 2 = disagree; 3 = somewhat disagree; 4 = neutral; 5 = somewhat agree; 6 = agree; 7 = strongly agree
		I_2	saving time	
		I_3	more reliable travel time	
		I_4	reliable vehicle arrival	
		I_5	better travel time	
Z_{CF}	Ride sourcing attitudes-comfort (CF)	I_6	drivers comfortably driving	
		I_7	vehicle comfortably riding	
		I_8	equipped with useful facilities	
		I_9	drivers safely driving	

Variables		Indicators	Measurement
		I_{10} provide security even at night	
Attribute of alternatives			
TT	Travel time		Continuous
TC	Travel cost		Continuous
WT	Waiting time		Continuous
Respondents characteristics			
A	Age		Continuous
G	Gender		0: Male; 1: Female

General model framework can be seen in Figure 1. Ellipses represent latent factors, two attitudinal latent factors of reliability and comfort, also the utility of service. Rectangles represent observe variables which are three attribute of alternatives, travel time, travel cost, and waiting time; two personal characteristics, age, and gender; also ten latent factor indicators. Solid arrows represent structural regression relationships, while dashed arrows represent the indicators measurement relationships.

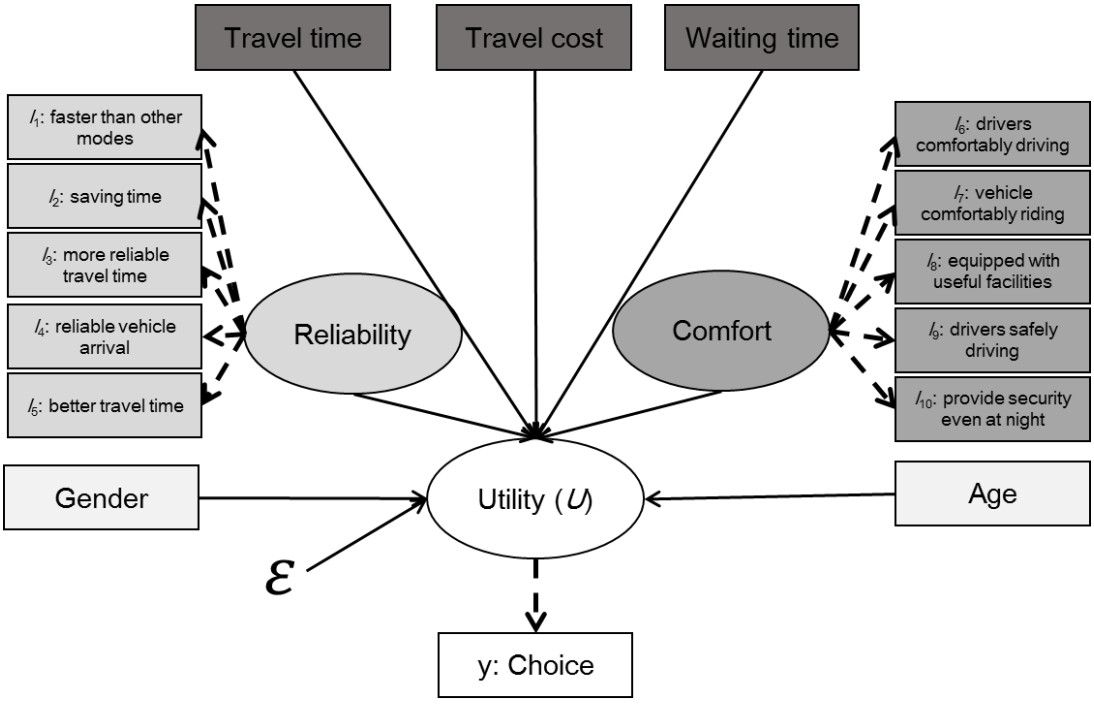

Figure 1 Ordered Hybrid choice model framework

The systematic utility function for the model can be seen in Eq.1.

$$\boldsymbol{V} = (\beta_{RI}\boldsymbol{z}_{RI} + \beta_{CF}\boldsymbol{z}_{CF} + \beta_{TT}\boldsymbol{x}_{TT} + \beta_{TC}\boldsymbol{x}_{TC} + \beta_{WT}\boldsymbol{x}_{WT} + \beta_{A}\boldsymbol{x}_{A} + \beta_{G}\boldsymbol{x}_{G}) \qquad (1)$$

Where $\boldsymbol{\beta}$ are the estimated parameters. The total utility refers to the latent response of frequency of using ride-sourcing. All latent variables are assumed normally distributed across the population,

i.e. $\mathbf{z}_{\mathrm{RI}} \sim \mathrm{N}(0, \boldsymbol{\sigma}_{\mathrm{RI}}^2)$, and $\mathbf{z}_{\mathrm{CF}} \sim \mathrm{N}(0, \boldsymbol{\sigma}_{\mathrm{CF}}^2)$ are uncorrelated with each other[1]. The $\boldsymbol{\sigma}^2$ terms are the variances of the latent variables which are to be estimated. Further, we assumed that the error term of the utility underlying the dependent variable ride-sourcing frequency is independently and identically distributed with: $\varepsilon \sim \text{Logistic}\left(0, {\pi^2}/{3}\right)$.

The utility is the sum of the systematic utility and the error term as shown in (Eq. 2), and the measurement model for car purchase intention is given by (Eq. 3) where the τ parameters are thresholds to be estimated.

$$\boldsymbol{U} = \boldsymbol{V} + \boldsymbol{\varepsilon} \tag{2}$$

$$\boldsymbol{y} = \begin{Bmatrix} 1\ (once)\ if\ -\infty \leq \boldsymbol{U} \leq \tau_1 \\ 2\ (couple\ times)\ if\ \tau_1 \leq \boldsymbol{U} \leq \tau_2 \\ 3\ (often)\ if\ \tau_2 \leq \boldsymbol{U} \leq \tau_3 \\ 4\ (very\ often)\ if\ \tau_3 \leq \boldsymbol{U} \leq \tau_4 \\ 5\ (always)\ if\ \tau_4 \leq \boldsymbol{U} \leq \infty \end{Bmatrix} \tag{3}$$

The psychometric indicators for the four latent variables are treated as continuous variables and modelled as follows:

$$\boldsymbol{I}_r = \lambda_r Z_{RI} + \boldsymbol{\upsilon}_r \text{ with } r = 1, \ldots, 5 \tag{4}$$

$$\boldsymbol{I}_r = \lambda_r Z_{CF} + \boldsymbol{\upsilon}_r \text{ with } r = 6, \ldots, 10 \tag{5}$$

where: $\upsilon_r \sim \mathrm{N}\left(0, \sigma_{\upsilon_r}^2\right), r = 1, \ldots, 10$, where the variances $\sigma_{\upsilon_r}^2$ are to be estimated. The error terms in (Eqns. 3-4) are assumed uncorrelated. The identification of latent variables is ensured by fixing $\lambda_1 = 1$ (for latent variable reliability); $\lambda_6 = 1$ (for latent variable comfort).

The joint probability of the choice and psychometric indicators is given by:

$$P(\boldsymbol{y}, \boldsymbol{I}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P(y|\boldsymbol{X}, \boldsymbol{z}_{RI}, \boldsymbol{z}_{CF})\ g_1(\boldsymbol{I}_1|\boldsymbol{z}_{RI}) \ldots g_5(\boldsymbol{I}_5|\boldsymbol{z}_{\boldsymbol{z}_{RI}})\ g_6(\boldsymbol{I}_6|\boldsymbol{z}_{\mathrm{CF}}) \ldots g_{10}(\boldsymbol{I}_{10}|\boldsymbol{z}_{CF}) \\ f_{RI}(\boldsymbol{z}_{RI}) f_{CF}(\boldsymbol{z}_{CF}) d\boldsymbol{z}_{RI} d\boldsymbol{z}_{CF} \tag{6}$$

where $\mathrm{P}(y|\boldsymbol{X}, \boldsymbol{z}_{RI}, \boldsymbol{z}_{CF})$is an ordinal logit model; That is, the probability that *y* takes level *m* is given as follows:

[1] In line with most of the literature on hybrid choice models (e.g. Belgiawan et al., 2017; Ben-Akiva et al, 2002b; Bolduc & Alvarez-Daziano, 2010), it is assumed that the latent constructs are orthogonal/uncorrelated with each other.

$$P(y = m) = F(\tau_m - V) - F(\tau_{m-1} - V) \quad (7)$$

where F is the cumulative distribution function of the logistic distribution, $\boldsymbol{g}$ are the probability sdensity functions for the indicators $\boldsymbol{I}$ and $\boldsymbol{f}$ the probability density functions of the latent constructs.

3.2 Results and Discussion

We estimated the OHCM using simulated maximum likelihood estimation with 10,000 Halton draws in Python Biogeme (Bierlaire, 2016). We also have tested the model with different number of draws as well as different starting values to ensure convergence and stability of the results. The results is presented in Table 3. The number of observation is equal to the number of sample, 497. We have 31 estimated parameters and the model fit is adequate, with adjusted R^2 0.253. We presented the parameter estimate of ride-sourcing attributes, respondents characteristics, latent attitudes toward ride-sourcing and latent attitudes of standard deviation.

It can be seen age significantly influences frequency of using ride-sourcing, the older the person, the less likely he/she use ride-sourcing. There is no effect on gender toward frequency of using ride-sourcing. The waiting time has no significant effect. Travel time has positive significant effect which is not as expected. Travel cost is negative significant showing that the more expensive ride-sourcing, the less likely people will use it.

Table 2 Ordered hybrid choice model results

Category	Name	Value	Robust t-test	Category	Name	Value	Robust t-test
Ride-sourcing Attributes	Travel Time	0.32	2.41	Indicators Lambda and Sigma	Lambda I_4	0.94	11.36
	Travel Cost	-0.21	-2.28		Sigma I_4	0.55	21.48
	Waiting Time	0.04	0.36		Lambda I_5	0.89	11.21
Latent Attitudes	Reliability	1.56	4.70		Sigma I_5	0.49	22.86
	Comfort	-1.20	-3.99		Lambda I_6	1.00	
Latent attitudes st. dev	Sigma Reliability	0.47	12.24		Sigma I_6	0.43	17.59
	Sigma Comfort	0.45	14.28		Lambda I_7	0.97	14.37
Respondents characteristics	Age	-0.05	-4.87		Sigma I_7	0.44	20.81
	Gender	-0.08	-0.48		Lambda I_8	0.95	10.70
Indicators Lambda and Sigma	Lambda I_1	1.00			Sigma I_8	0.53	22.88
	Sigma I_1	0.6	21.29		Lambda I_9	0.83	9.64
	Lambda I_2	0.97	10.57		Sigma I_9	0.47	21.02
	Sigma I_2	0.51	19.82		Lambda I_{10}	0.91	9.43
	Lambda I_3	1.02	10.48		Sigma I_{10}	0.66	25.78
	Sigma I_3	0.50	22.86				
	Sample Size		497				
	Estimated parameters		31				
	Halton draws number		10,000				
	Log-likelihood (L(0))		-6744.79				
	Log-likelihood (L(1))		-5005.94				
	McFadden R^2		0.258				
	McFadden adjusted R^2		0.253				

4. Conclusion

This investigation presents a study exploring determinant factors in using ride-sourcing by users in Bandung, Indonesia. Using ordered hybrid choice model, two latent variables are incorporated into the model including users' and ride-sourcing attributes. The model is able to obtain the explanation of the reasons behind the decision to frequently use ride-sourcing.

In the light of specific condition of the traffic and public transport services in Bandung, the latent variable of comfort are found to negatively influence the mode choice. People tend to compensate the comfort with the travel time and reliability. They are willing to endure service that have less comfort but they can gain a faster travel time in exchange, especially for ride-sourcing.

Being present in the majority of Indonesian cities, ride-sourcing remains, in theory, illegal as motorcycle is not classified as public transport. To accommodate users' dynamic needs, and recognizing the positive contribution of technology in providing transport services, laws and regulations need to be upgraded. At the same time, public safety must be a priority for regulators (Susilo and Joewono, 2017). Meanwhile, in the on going discussion of the most recent regulation, motorcycle may be classified as paratransit. Their service can be managed as a feeder, especially operated in residential area to supply the collector in road services.

To anticipate the coming future, it is notable to refer to Cohen (2018) suggestions about the role of government, namely the government needing to clearly understand every possible impact of a given development and the government should anticipate coming developments for what they are and respond accordingly. This demands a change of perspective and way of taking and managing action to be more anticipative and to be well planned instead of responding the after effect.

References

Ben-Akiva, M., Mcfadden, D., Train, K., Walker, J., Bhat, C.R., Bierlaire, M., Bolduc, D., Boersch-Supan, A., Brownstone, D., Bunch, D.S., Daly, A., De Palma, A., Gopinath, D., Karlstrom, A., Munizaga, M. a., 2002. Hybrid Choice Models: Progress and Challenges. Mark. Lett. 13, 163–175. https://doi.org/10.1023/A:1020254301302

Bierlaire, M., 2016. Estimating Choice Models with Latent Variables with PythonBiogeme.

Cohen, T., 2018. Being ready for the next Uber: can local government reinvent itself? Eur. Transp. Res. Rev. 10. https://doi.org/10.1186/s12544-018-0330-8

Harding, S., Kandlikar, M., Gulati, S., 2016. Taxi apps, regulation, and the market for taxi journeys. Transp. Res. Part A Policy Pract. 88, 15–25. https://doi.org/10.1016/j.tra.2016.03.009

Irawan, M.Z., Belgiawan, P.F., Joewono, T.B., Simanjuntak, N.I.M., 2019a. Do motorcycle-based ride-hailing apps threaten bus ridership? A hybrid choice modeling approach with latent variables. Public Transp. https://doi.org/10.1007/s12469-019-00217-w

Irawan, M.Z., Belgiawan, P.F., Tarigan, A.K.M., Wijanarko, F., 2019b. To compete or not compete: exploring the relationships between motorcycle-based ride-sourcing, motorcycle taxis, and public transport in the Jakarta metropolitan area. Transportation (Amst). https://doi.org/10.1007/s11116-019-10019-5

Susilo, Y.O., Joewono, T.B., 2017. Indonesia. pp. 107–126. https://doi.org/10.1007/978-3-319-43851-1_6

Walker, J.L., Ben-Akiva, M., 2002. Generalized random utility model. Math. Soc. Sci. 43, 303–343. https://doi.org/10.1016/S0165-4896(02)00023-9