Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 00 (2018) 000–000

Transportation Research Procedia

www.elsevier.com/locate/procedia

World Conference on Transport Research - WCTR 2019 Mumbai 26-31 May 2019

A tour-based mode choice models for commuters in Indonesia

Faza Fawzan Bastarianto[a,b], Muhammad Zudhy Irawan[a,*], Charisma Choudhury[b], David Palma[b], Imam Muthohar[a]

[a]*Department of Civil and Environmental Engineering Universitas Gadjah Mada, Jln. Grafika No.2 Bulaksumur, 55281 Yogyakarta, Indonesia*
[b]*Institute for Transport Studies University of Leeds, 34-40 University Road, LS2 9JT Leeds, United Kingdom*

Abstract

Sustainable planning and policy strategies in transport require understanding the relationship between tour types and mode, and investigating the factors that influence commuters' mode and tour type choices. Therefore, in this study we focus tour-based mode choice of commuters between Bekasi and Jakarta. A secondary data containing 24-hour daily activity patterns of 420 commuters including their trip characteristics, socio-economic and socio-demographic background was used. We analyze this data using three different logit model structures: multinomial logit (MNL), nested logit (NL), and cross-nested logit (CNL); and compare them to comprehend which model structure is the most appropriate for this problem. Among other results, we find that gender and income significantly influences commuter's choice of mode, and that reducing travel time and cost can increase the ridership of public transport. Furthermore, the NL showed significant improvement when grouping the alternatives based on tour types, and the CNL model is considered the best model structure as it captured significant correlation among alternatives within the nest and across the nests.

Keywords: Tour-based mode choice; Commute; Revealed preference; Nested logit; Cross-nested logit

1. Introduction

As the capital city of Indonesia, Jakarta is a big attraction to the cities surrounding it, such as Bogor, Depok, Tangerang, and Bekasi. Thus it generates trips which are growing every year due to increasing urbanization. Bekasi, which is a city located on the east side of Jakarta, contributes the largest number of commuters to Jakarta consisting of about 14.8% (Statistics of DKI Jakarta Province, 2015). As many commuters travel to Jakarta from Bekasi, there are some problems between these two cities such as severe traffic congestion and lack of high quality public transport serving the two cities. Therefore, there is a need for robust mode choice models to investigate the behavior of commuter between these two cities.

* Corresponding author. Tel.: +62-878-386-90515.
E-mail address: zudhyirawan@ugm.ac.id

Commuters use different modes for their trips, such as a private car, motorcycle, bus, KRL commuter line train, minibus, etc. The tours themselves have many variations such as hwh, hwh+, hw+wh, hw+wh+, hw++wh, hsh, hs+sh, hs++sh. Description of these kinds of tour types are for the tour types that have "w" and "s" representing tour type including at least one work (w) activity and one school (s) activity, respectively. The symbol (+) itself represents that the tour includes one additional stop for another activity (Bowman and Ben-Akiva, 2000), and the symbol "h" represents home. Understanding the relationship between tour types and modes is vital for sustainable planning and policy strategies aimed at private vehicles reduction and public transport promotion. Thus, tour-based mode choice of the commuters specifically in developing country will be the emphasis of this study.

Tour-based mode choice has been extensively studied in developed countries up to today. However, there are fewer studies of travel behaviour specifically in tour-based mode choice in developing countries (Dissanayake and Morikawa, 2010). The reason for this might be related to a lack of data relating to either survey of household travel or traffic pattern.

Concerning methodology and model structure, to date, many studies have used the multinomial logit (MNL), nested logit (NL), and cross-nested logit (CNL) models. The MNL model is a traditional logit model, which is widely used in transport research due to its simplicity and reliability. The NL model is an extension of the MNL model, where the choices can be structured in different levels and nests, each nest grouping correlated alternatives, in such a way that each alternative can only belong to one nest. Finally, the CNL model is a generalization of the NL model, allowing alternatives to belong to more than one nest (Train, 2009). Hess et al., (2012) point out that the CNL model can capture more correlation patterns of alternatives rather than a three-level NL structure, in a multi-dimensional choice process. However, there is an absence of studies comparing different logit model structures specifically in developing countries.

Therefore, the main aim of this investigation is to develop logit model structures of tour-based mode choice model for commuters in developing countries, specifically from Bekasi to Jakarta. The study focuses on analysing the relationship between the tour types and mode choice of the commuters, and investigating the attributes that influence the commuters on choosing their mode and tour type. The mode choice will be based on characteristics of the respondents.

This study will also explore different logit model structures between simple choice model or MNL, NL, and advanced choice model or CNL to see which structure best estimates the relationship between the tour types and mode choice.

2. Literature review

2.1. Tour-based model

Tour-based modelling approach assumes a trip chain as the basic decision unit for an individual. Trip chains are such that start at individuals' home, take them to one or more activity locations, and finish at the starting point (Bowman and Ben-Akiva, 2000). The tour-based model approach can be used to develop more complex disaggregate choice modelling such as activity-based model and tour-based mode choice model. This study, however, uses a tour-based mode choice model for analysing the data. The tour-based mode choice model can give us a better understanding of the relationship of individual decision to travel in daily activity with a particular mode used (Miller et al., 2005). Some factors are influencing an individual to choose travel mode such as; built environment factors, socio-economic factors, attitudinal factors, and trip chain (Hu et al., 2018).

In terms of modelling a large class of models with the set of alternatives which can be partitioned into subsets, a NL specification is a common model structure used especially for tour-based mode choice since it can effectively model multidimensional choice processes in a natural hierarchy (Bowman and Ben-Akiva, 2000), unlike other discrete choice models. Since this model has two-levels of modelling which are the marginal probabilities (upper model), and the conditional probabilities (lower models) which subsequently will be incorporated into a nest (Train, 2009). In the case of the tour-based mode choice model, the upper levels are the tour type, and the lower levels are travel modes used.

2.2. Travel behavior studies in Jakarta, Indonesia

Several studies about travel behaviour in Jakarta and more related developing countries have been undertaken by researchers in recent years. These kind of researches are critical for stakeholders to determine their policies before implementing them, whether these result in advantages or disadvantages. Even though conducting this researches is essential, there are a few studies of tour-based mode choice particularly in developing countries. Most of these studies that have been undertaken, are focusing on mode choice (Siddique and Choudhury, 2017; Santoso and Tsunokawa, 2005; Irawan et al., 2017), travel behaviour (Irawan et al., 2017; Hyodo et al., 2005), and activity-travel patterns (Dharmowijoyo et al., 2017).

However, there are several studies which have been done mainly in Jakarta. Most of the studies used data collected from The National Planning and Development Agency of the Republic of Indonesia (Bappenas) and Japan International Cooperation Agency (JICA) which conducted a household travel survey from 28th of June 2002 to 8th of November 2002. This data aims to study an integrated transportation master plan (SITRAMP) 2004 over Jakarta Metropolitan Area (JMA). Still, there are deficient studies with other data sources to investigate further of disaggregate choice model in Jakarta.

The study which has been done by Yagi and Mohammadian (2008) using SITRAMP 2004 data focused on joint models of home-based tour to build an activity-based modelling framework for Jakarta. They used a NL model to develop four different activity types which are work, school, maintenance, and discretionary. While on the other side, travel modes are divided into eight most commonly used mode in the region. The overall results of the model show that the model structure, choice alternatives, and critical variables significantly differ with other activity-based models in the developed countries. In Jakarta, they found that shared ride is often associated with people who employ chauffeurs, meaning that this mode is common among high-income people. Income has a significant influence on mode choice. For instance, utilities of a private vehicle and taxi increase as income rises while utilities of the motorcycle, public transport, and non-motorized transport increase as income declines. Gender, age, and status of individuals directly increase the utilities of specific travel modes. Male respondents have higher of utilities of private vehicles and females have higher utilities of public transport and taxi. In addition, full-time workers have greater utilities of private modes and homemakers have greater utilities of non-motorized transport.

Using the same SITRAMP 2004 data, Dharmowijoyo et al. (2016) examined the variability in travellers' activity-travel patterns in JMA. The explanatory variables are daily constraints, land use, road network conditions, and resources. The study looks at the influences of these variables on activity-travel patterns such as travel mode choices, a number of trips, trip chains, departure time, and total daily travel time interact with each other. The trips are made by workers, students, and non-workers. The results show that some variables influencing activity-travel patterns in developing countries are similar to developed countries, but others differ. The variables such as individual, household, transport network, and land use significantly influence workers and students' activity-travel patterns, with the exception of the departure time variable. They are more likely to use motorised modes to travel and have lesser daily trips for those who have high-income than other income levels. Their activity-travel patterns are also more predictable especially during weekdays and more flexible during weekend days. On the other hand, because non-workers have much more flexible time, they are not significantly influenced by departure time, travel time spent, and mode shares. Furthermore, older non-workers have daily trips with more trip chains using motorised modes than others.

3. Methodology

The relationship between tour type and mode of travel has been mostly modelled using logit models such as MNL (Lekshmi et al., 2016) and NL (Ho and Mulley, 2013). Meanwhile, one of the advanced logit models such as CNL has rarely been used in modelling this relationship. The simplicity of MNL was one consideration that this specific model is selected, following NL which could estimate the correlation among different alternatives within specific groups or nests. Finally, one of the most advanced logit model, CNL, is used to estimate the correlation among alternatives in different groups or nests.

We tested a single NL model structure, with two levels, where the upper level captures the tour types, while the lower level represents a joint choice of mode choice and tour type. Meanwhile, the CNL structure, has four nests based on tour types (hwh and hw+wh) and vehicle types (private vehicles and public transport). Additionally, we estimated a simpler MNL model without any nest. Fig. 1. To Fig. 3. present the structures of these models.

The tour types from the data are divided into eight types as previously stated in the introduction. However, there are only two groups with high variations of tour type and mode choice that will be analysed in this study which are hwh and hw+wh.

In regard to mode choice, the private vehicles in this study are limited to private car and motorcycle as these are the primary private vehicles used in the context of Jakarta Metropolitan Area. Meanwhile, public transport are divided into bus and KRL commuter line train since these two modes are the modes with the highest occupancy of commuters.

The interaction between tour type and mode choice generates eight alternatives or set of choices which are described in Table 1.

Table 1. Set of choices of alternatives in the model.

Set of choices	Description
Alternative 1 (HWH_MC)	Decision maker takes hwh tour type and chooses motorcycle as a mode of travel to commute
Alternative 2 (HWH_CAR)	Decision maker takes hwh tour type and chooses a car as a mode of travel to commute
Alternative 3 (HWH_BUS)	Decision maker takes hwh tour type and chooses a public transport bus as a mode of travel to commute
Alternative 4 (HWH_KRL)	Decision maker takes hwh tour type and chooses KRL commuter line train as a mode of travel to commute
Alternative 5 (HW+WH_MC)	Decision maker takes hw+wh tour type and chooses motorcycle as a mode of travel to commute
Alternative 6 (HW+WH_CAR)	Decision maker takes hw+wh tour type and chooses a car as a mode of travel to commute
Alternative 7 (HW+WH_BUS)	Decision maker takes hw+wh tour type and chooses a public transport bus as a mode of travel to commute
Alternative 8 (HW+WH_KRL)	Decision maker takes hw+wh tour type and chooses KRL commuter line train as a mode of travel to commute

The utility function of each alternative is different since varying parameters are included in each function. Each parameter is assumed to have significant effect on the value of utility. The utility function for all alternatives is presented in Equation (1) to Equation (8) below.

$$U_{HWH\_MC} = ASC_{HWH} + ASC_{MC} + \beta_{TC} * TC_{MC} + \beta_{TT} * TT_{MC} + \beta INC_{L\,MC} * LOW\ INCOME \quad (1)$$

$$U_{HWH\_CAR} = ASC_{HWH} + ASC_{CAR} + \beta_{TC} * TC_{CAR} + \beta_{TT} * TT_{CAR} + \beta INC_{H\,CAR} * HIGH\ INCOME \quad (2)$$

$$U_{HWH\_BUS} = ASC_{HWH} + ASC_{BUS} + \beta_{TC} * TC_{BUS} + \beta_{TT} * TT_{BUS} + \beta GEN_{BUS} * FEMALE \quad (3)$$

$$U_{HWH\_KRL} = ASC_{HWH} + ASC_{KRL} + \beta_{TC} * TC_{KRL} + \beta_{TT} * TT_{KRL} + \beta GEN_{KRL} * FEMALE \quad (4)$$

$$U_{HW+WH\_MC} = ASC_{MC} + \beta_{TC} * TC_{MC} + \beta_{TT} * TT_{MC} + \beta INC_{L\,MC} * (INC_{MC} == 1) + \beta_{PURPOSE} * PART\ TIME\ WORKER \quad (5)$$

$$U_{HW+WH\_CAR} = ASC_{CAR} + \beta_{TC} * TC_{CAR} + \beta_{TT} * TT_{CAR} + \beta INC_{H\,CAR} * HIGH\ INCOME + \beta_{PURPOSE} * PART\ TIME\ WORKER \quad (6)$$

$$U_{HW+WH\_BUS} = ASC_{BUS} + \beta_{TC} * TC_{BUS} + \beta_{TT} * TT_{BUS} + \beta GEN_{BUS} * LOW\ INCOME + \beta_{PURPOSE} * PART\ TIME\ WORKER \quad (7)$$

$$U_{HW+WH\_KRL} = \beta_{TC} * TC_{KRL} + \beta_{TT} * TT_{KRL} + \beta GEN_{KRL} * LOW\ INCOME + \beta_{PURPOSE} * PART\ TIME\ WORKER \quad (8)$$

The equations above show the utility function for all alternatives. The β_{TC}, β_{TT}, $\beta INC_{L\ MC}$, $\beta INC_{H\ CAR}$, βGEN, $\beta_{PURPOSE}$ are the dummy variables of mode and socio-demographic of the commuters related to travel cost, travel time, low income for motorcycle users, high income for car users, gender, and travel purpose. These dummy variables will explicitly indicate the behavioural trend of commuters in choosing their mode choice and tour types. The alternative specific constant (ASC) of each alternative shows the average effect on the utility of all factors outside of the observed parameters (Train, 2009). The presence of alternative specific constant in each function eliminates the error term or unobserved portion of utility ε in the utility function. Moreover, a variety of alternative specific dummies are included in the model to investigate the household travel behaviour and attitudes such as income, gender, and travel purpose of the commuters. Meanwhile, the differences between three logit models will be described and drawn in the figures below.

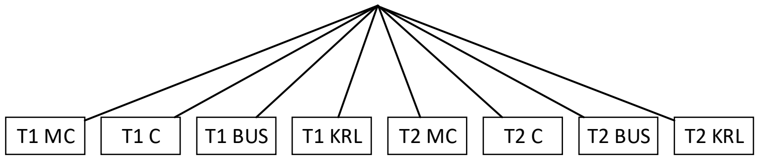

Fig. 1. Multinomial logit model structure.

The MNL model structure is drawn in Fig. 1. above. There are eight alternatives to the model which have the same level of choices. Where T represents the tour type of the commuters which is divided into T1 and T2 to represent hwh and hw+wh, while MC, C, BUS, and KRL are motorcycle, car, bus, commuter line train respectively.

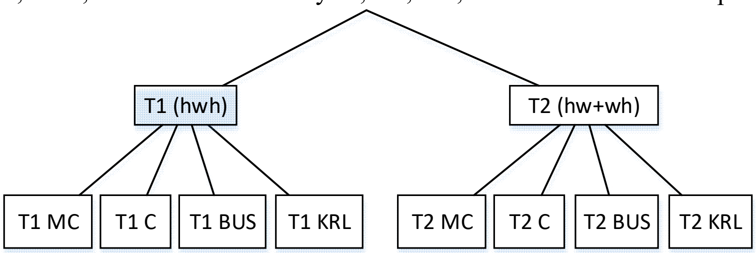

Fig. 2. Nested logit model structure.

Fig. 2. Nested logit model structure.Fig. 2. illustrates the model structure of NL that is deployed in this study. This model structure is a two level NL where the upper level is partitioned of tour type into two nests and the lower level is the joint choice of tour types and travel mode of the commuters. Thus, by using a NL model, the correlation between these two tour types will be explored in this model.

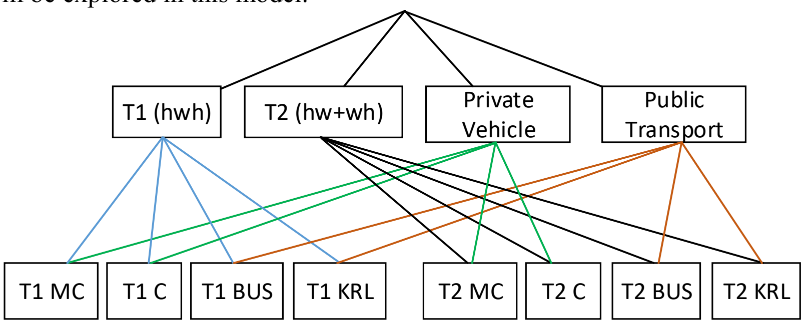

Fig. 3. Cross-nested logit model structure.

Fig. 3. above demonstrates that the CNL model is similar with NL, but the alternatives in the lower level could be the member of one or more nest in the upper level. In this model, the upper level consists of four nests: hwh, hw+wh, private vehicle, and public transport whereas the lower level includes joint choice of tour types and modes. Each joint choice belong to two nests both tour type and vehicle type. However, this structure can be drawn in NL structure but it requires either two different NL structures based on tour types and vehicle types or combined nest in three level NL structure based on the study conducted by Hess et al. (2012). Therefore, by using CNL structure from this model, the correlation pattern among alternatives across the nest can be more easily apprehended.

4. Data description

4.1. Study area

The capital city of Indonesia, Jakarta, has a population of 10.277.628 which is around 3.97% of the total Indonesian inhabitants (Statistics of DKI Jakarta Province, 2017). The administrative, industrial, and commercial activity in Indonesia is centralised in Jakarta. It makes a rapid development in Jakarta where there are plentiful constructions sites, works on infrastructure facilities, as well as a fleet of public transport to increase the mobility of the people inside and toward Jakarta. Therefore, it attracts many people in Indonesia to work and live there for permanent residence. However, for the people who live in the cities surrounding Jakarta, they are choosing to be as commuters. In terms of origin-destination (OD) of the commuters, Bekasi city is set to be the origin whereas Special Region of Jakarta is chosen to be the destination, including the municipalities within those two cities.

Fig. 4. Location map of Jakarta Metropolitan Area (JMA). Source – Google Maps, 2018.

There are several modes of transport which are provided to commute from Bekasi to Jakarta such as a private car, motorcycles, bus, and KRL commuter line train. According to the survey from Statistics of DKI Jakarta Province (2015), they pointed out that the number of commuters from the outskirts of Jakarta is 1.38 million which is comprising the cities of Bekasi, Depok, Bogor, and Tangerang. The modal shift consists of 58 percent of commuters using motorcycles, 12.8 percent of commuters using cars, and 27 percent using public transport such as buses and KRL commuter line train. The highest number of commuters was contributed by Bekasi which is about 14.8 percent of the total number of commuters.

4.2. Daily activity patterns survey data

The survey was held in 2016 in the period from February to March for the commuters from Bekasi to Jakarta. The survey was to collect the revealed preference (RP) and stated preference (SP) data which was obtained at the stations and offices both in Bekasi and Jakarta. Random sampling technique was used as the basic method of the selection of the respondents.

The survey was conducted by Gadjah Mada University with the grant from the Ministry of Research, Technology and Higher Education of the Republic of Indonesia. The survey gained three kinds of data which were socio-demographic and travel diary as revealed preference data and stated preferences data. These data are useful to analyse complex travel behaviour and forecasting travel demand for the planning project of the new LRT system from Bekasi to East Jakarta. However, the stated preference data was not analysed in this study.

Table 2. Sample profiles in the study.

Variables	Percentage/Mean	
Socio-demographic characteristics at individual level		
Male (%)	69.52	
Female (%)	30.48	
Full-time worker (%)	98.10	
Part-time worker (%)	1.90	
Household characteristics		
Average monthly income ('000 IDR)	6171.43	
Low income < 4 million IDR (%)	34.52	
Medium income 4 - 6 million IDR (%)	28.10	
High income > 6 million IDR (%)	37.38	
Vehicle ownership (%)	89.76	
Car ownership (%)	16.67	
Tour types of commuters		
HWH - Motorcycle (%)	13.33	
HWH - Car (%)	2.86	
HWH - Bus (%)	11.67	
HWH - KRL (%)	4.05	
HW+WH - Motorcycle (%)	38.33	
HW+WH - Car (%)	9.05	
HW+WH - Bus (%)	10.95	
HW+WH - KRL (%)	9.76	
Population, employment, and geography situation		
Variables	Bekasi City	Jakarta
Population (million)	2.803	10.937
Survey area (km2)	210.49	662.33
Population density (person/km2)	11413	15517.38
Total number of commuters	204240	
Number of commuters surveyed (N)	420	

Source – Secondary data, 2017; Statistics of DKI Jakarta Province, 2015; Statistics of Bekasi Municipality, 2017; Statistics of DKI Jakarta Province, 2017

The survey collected the data on socio-demographic which is relevant to tour-based mode choice. The characteristics of socio-demographic of the commuters include monthly income, gender, and also the type of profession of the commuters. Meanhile the tour type characteristics of the commuters were grouped into eight groups which included two tour types with four different travel modes. The summary of socio-demographic and tour types of the commuters from Bekasi to Jakarta is shown in Table 2 above.

Based on the data above, it can be examined that the total number of commuters from Bekasi to Jakarta is large compared to the total number of population of Bekasi City. There is around 10% of the total population of Bekasi City becoming a commuter. The sample size of commuters surveyed in this study is 420 commuters from 501 initial sample size. This number can be assumed to represent the total number of commuters by using Slovin Sampling Theory.

According to Table 2 above, it shows that the percentage of monthly income of the commuters is equally varied from low income to high income. In addition, the average monthly income of the commuters was around six million Indonesian Rupiah. Thus, due to this financial limitations, generally, vehicle ownership is limited to own one car and two or more motorcycles for each household since it is not very common to have more than one car per household in developing country (Dissanayake and Morikawa, 2002).

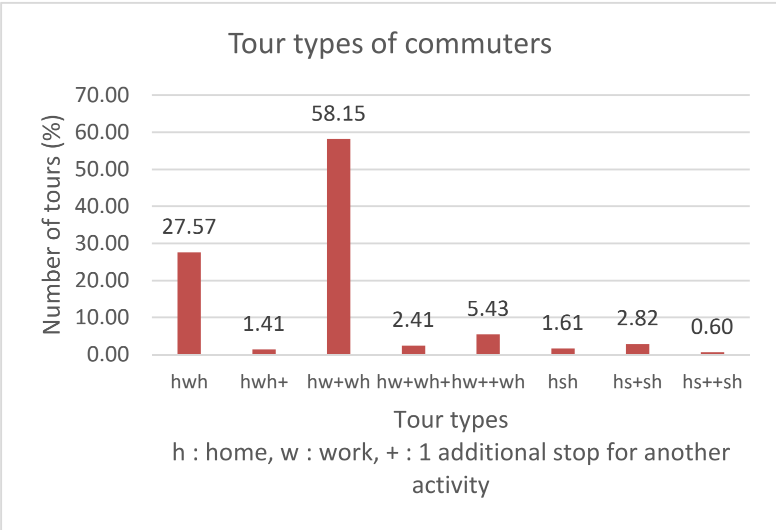

Fig. 5. (a) Initial tour type of commuters

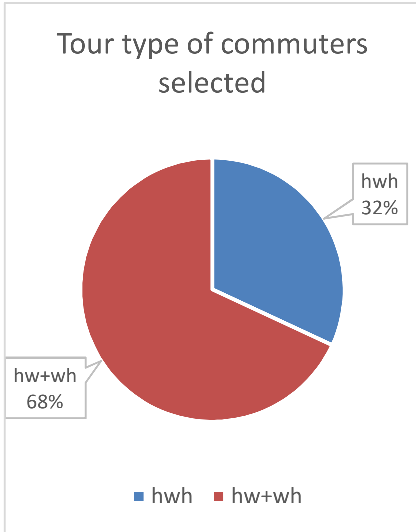

Fig. 5. (b) Tour type selected

The initial data of tour type is shown in Fig. 5. (a). There are eight tour types which include tour purpose to workplace and school. Furthermore, the tour is varied and not limited to only one destination since there will be one or more additional stop for another activity, such as lunch, meeting, shopping, and other destinations within the tour. From the Fig. 5. (b) it can be seen that the tour type hw+wh and hwh are the two most dominant tour type of the commuter. Thus, the tour type of commuters is simplified into two types from eight types collected based on initial data which will be analysed in this study. Consequently, there is no any tour type for school purpose used in this study due to the limitation of its data. The proportion of these two types are 32% and 68% for hwh and hw+wh. The tour type hw+wh is the most dominant tour type since most of the commuters from Bekasi to Jakarta were the workers who have an additional trip purpose in the workplace. Usually they would take trips for lunch during the afternoon break. Meanwhile, the hwh tours is in the second rank of number of tours, since there are several workplaces or offices which are integrated with many facilities such as shopping mall and food court so that commuters do not have to take an extra a trip to other places during the break time.

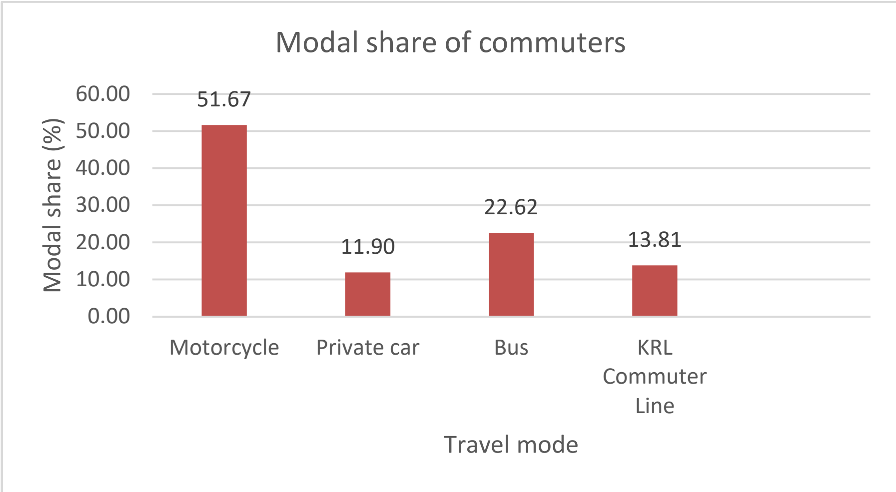

Fig. 6. Modal share of commuters from Bekasi to Jakarta

According to the data collection in Fig. 6., the modal share of commuters from Bekasi to Jakarta is dominated by motorcycle making 51.67%, followed by public transport second, and third which is bus and KRL train with the modal share about 22.62% and 13.81%. Lastly, car vehicle becomes the last choice for the commuters with the percentage of 11.90%. This initial analysis is in line with the study conducted by Yagi et al. (2013) which found that the motorcycle is the most dominant mode used in many urban areas in developing countries such as Jakarta and there was less modal share of public transport.

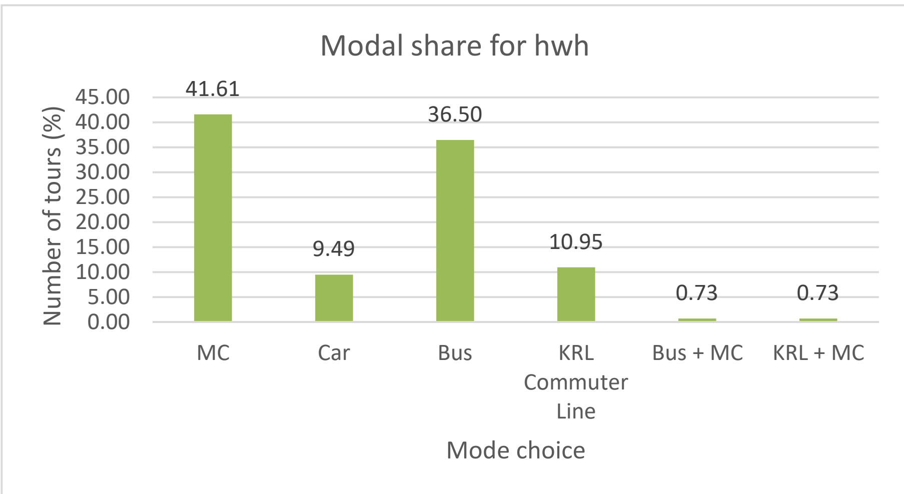

Fig. 6. Modal share for hwh tour type

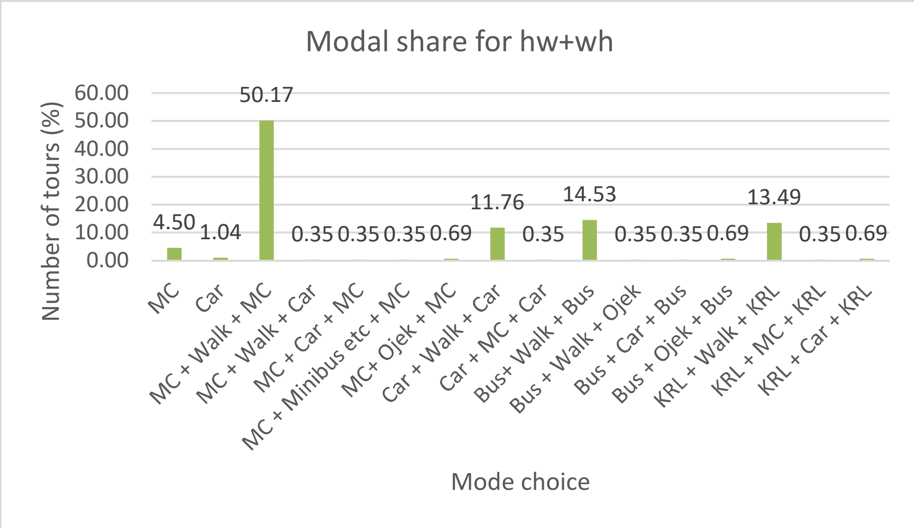

Fig. 7. Modal share for hw+wh tour type

In addition to the data above, the modal share for hwh and hw+wh tour type is drawn in Fig. 7. and Fig. 8. respectively. From those figures, it implies that both motorcycle and bus are the most chosen mode for the commuters when they are going to work and going back home with no additional trip within the workplaces. The modal share for motorcycle and bus in this hwh tour type is 41.61% and 36.50%. On the other hand, almost 50.17% of the commuter chose motorcycle as the primary mode for the tour type going to work and going back home with at least one additional trip in the workplaces. This number is the highest modal share compared to other modes in hw+wh tour type.

5. Analysis and discussion

This section of the paper will briefly discuss the results of tour-based mode choice model. The models estimated are MNL, NL, and CNL for which estimated parameter, robust standard error, and robust t-stat values are presented and discussed. Moreover, the model comparison will be based both on informal and formal tests, including signs and magnitudes of parameter estimates, log likelihood ratio test, goodness-of-fit tests, and t-statistics.

5.1. Multinomial logit model

The result from parameter estimation of MNL model structure in Fig. 1. is presented in Table 3. By analysing the informal test on the sign of the parameter estimation, the results show that several utility coefficients have a positive impact to the utility of mode choice for particular tour type as these value have positive sign. However, there are also some coefficients that have negative impact on the utility as these values have negative sign. These signs will imply whether the coefficients would either increase or decrease the propensity of commuters to choose the mode based on their preferred tour type. See Equation (1) to Equation (8) for utility functions which were used and with clarification of each estimated parameter that will be evaluated in this model.

According to the Table 3, the car, KRL, and bus are the most preferred modes for the commuters from Bekasi to Jakarta since the alternative specific constant for these modes are positive and statistically significant. Also, the parameter magnitude of these modes are vast compared to another modes such as motorcycle which is smaller. On the other hand, the commuters are less likely to take motorcycle due to the small value of parameter magnitude of the mode specific constants for this mode, and also the value is found to be insignificant.

The estimated dummy parameter of commuters with low income is positive in sign and statistically significant indicating that the low income commuters have a higher preference to choose motorcycle whereas the commuters with high income prefer to drive car to work since the estimated parameter is also positive in sign and significant.

Meanwhile, the female commuter dummy shows positive sign and is statistically significant which implies that they are more likely to take public transport for commuting, where the magnitude of parameter of female commuter dummy who travel by bus is higher than the female who takes KRL, indicating that the female has the tendency to choose a bus rather than KRL for commuting trip.

Moreover, the occupation of the commuters' is also taken into accound and analysed by using dummy parameters. The corresponding dummy parameter in the utility function is positive in result, which explains that the commuters who have a part-time job are more likely to have flexible tour type. However, this attribute is found not to be statistically significant at 10% error tolerance level.

Therefore, based on the findings above it could be argued that the income and gender influence the probability of the commuters to choose the travel mode. However, besides all of the attributes above that contribute positive impact to the utility, there are two differing attributes that have negative in sign and statistically significant which are travel time and travel cost. Thus, these attributes impact on decreasing the propensity of the utility of the commuters in choosing their mode. The estimated value of travel cost is found to be higher than the travel time. For this reason, the commuters have a higher preference to choose the cheapest and which is then followed by the fastest mode.

In terms of summary statistics of the total 420 observations, the results show that there are 11 parameters estimated and eight alternatives evaluated. In addition, the goodness-of-fit of the models such as the rho square and adjusted rho square values are 0.31 and 0.29 respectively which means that these values are substantially high. Therefore, it can be stated from the goodness-of-fit test that the model has a good fit to the data.

5.2. Nested logit model

As stated earlier, the NL model that will be analysed in this study is NL structure based on tour types (See Fig. 2. Nested logit model structure.). The initial analysis for this model was eliminating some parameters in previous MNL model and adding several parameters related to NL model. Some deleted parameters were alternative specific constant for hwh (ASC_HWH) and hw+wh (ASC_HW+WH) which then was added with the scale parameter of λ both for nest hwh and hw+wh. Consequently, some nest-specific parameters (β value) for both nest are divided in order to inspect the effect of particular attributes in a different nest. These nest-specific parameters include βINC, βGEN, and βPURPOSE. Nonetheless, the generic parameters such as βCOST and βTIME were kept alongside the nests.

Within this initial analysis, only the alternative specific constants for KRL was fixed to be one whereas the other new parameters will be estimated. Result shown that the value of λ_{hwh} and λ_{hw+wh} were greater than one which is inconsistent with the utility maximizing behaviour since the value of λ should be in the specific range between zero and one (Train, 2009). In order to make sure that the value of both λ is between zero and one, the value of λ_{hw+wh} was fixed to 0.99 while keeping λ_{hwh} free, we made this assumption based on the hypothesis that correlation among alternatives in nest hw+wh was bigger due to the more complex tour. Next, the scale parameter for nest hw+wh was set to be 0.99 (λ_{hw+wh}=0.99) while for nest hwh was free. The second approach worked better, and is then reported in Table 3.

The results in Table 3 shows that after fixing the parameter of λ_{hw+wh}, both λ_{hwh} and λ_{hw+wh} are in the range between zero and one which is consistent with the utility maximizing behaviour theory. The value of λ_{hwh} is found to be 0.55 and statistically significant. Therefore, it can be argued that there is correlation among alternatives in nest hwh.

Some findings are found to be similar with MNL model results, such as the alternative specific constants value for car, bus, and KRL which were significantly positive. This finding indicates that these mode are more likely to be chosen by the commuters whereas the motorcycle is unlikely to be chosen by the commuters since its parameter has small magnitude value and is found to be statistically insignificant at 90% of confidence level. Furthermore, travel cost and travel time also indicate similar commuters' behaviour to choose their preferred mode. The coefficients of travel cost and time are negative in sign and statistically significant.

Other results obtained from the NL model reveal that the commuters with low income with hw+wh tour type have a higher preference to choose motorcycle rather than the commuters with hwh tour type because of the estimated parameter obtained for low income commuters in hw+wh being more significant than shown in hwh. Subsequently, the commuters with high income have a higher preference to choose car as the main mode to commute in hw+wh tours rather than in hwh tours since both estimated parameters are statistically positive. This finding is sensible when it is assumed that the commuters who have private vehicles tend to have complex tour type particularly in their work place rather than having fixed tour type such as hwh.

Other findings of this model indicate that there are different behaviours of commuters who take different tour type. The female commuters are more likely to take bus in hwh tours than hw+wh tours as the estimated parameter is found to be positive and statistically significant for hwh tours, unlike hw+wh tours which is insignificant. However, the female commuters tend to choose KRL rather than bus when they have hw+wh tour type as the estimated parameter for female to take KRL is significantly positive for nest hw+wh.

In terms of commuters' occupations, the result shows similar finding with the MNL model result. The commuters who have a part-time job shows the tendency to have hw+wh tour type due to the flexibility for them to make a trip. However, this parameter is found to be insignificant statistically at 90% confidence level.

The statistics summary of 420 observations show that there are 15 parameters obtained from eight alternatives. The result of goodness-of-fit test was shown to be higher than the MNL model since the rho square, and adjusted rho square are increased to 0.35 and 0.33 respectively. Hence, this result implies that this NL model has a good fit to the data. Moreover, the likelihood ratio test between NL and MNL has a result of 64.48 which in regard to chi-square statistic, this output value is greater than the critical value of three degree of freedom, and 90% of confidence level used in this study. Thus, this indicates that the parameter estimations (See Equation (1) to Equation (8)) in this model considerably different from zero.

5.3. Cross-nested logit model

The CNL model provides a higher degree of flexibility in capturing wider correlation patterns among alternatives than the two previous models that have been used (Wen and Koppelman, 2001). The restricting aspect of grouping some alternatives in one nest in NL model can interrupt in capturing the correlation of alternatives that may belong to more than one nest. Since each of the alternatives belong to more than one nest as shown in Fig. 3., this model structure is appropriate and efficient to be used in the analyses rather than using two NL model structure based on tour types and vehicle types or three level of NL structure.

Therefore, in this model it is assumed that each joint choice alternatives of commuters will be allocated in two nests according to tour type based and vehicle type based. The upper level represents the tour types and vehicle types, while the lower level consists of the joint choice alternatives. The estimation results for CNL model are shown in Table 3. The final values of λ and α shown in the Table 3 and Appendix A have been transformed using logit transformation equation, as during estimation an appropriate logit transformation was used to avoid these values going outside the [0, 1] interval.

According to the findings, it can be said that the CNL model can categorize the correlation among alternatives in different nest since all of α values were shown below than one, with the exception the value of α for alternative six to nest private vehicle which has the scale parameter of about one. While the highest correlation between alternative and nest is apprehended on the commuters who select car in hw+wh tours with nest hw+wh, followed by the commuters who ride motorcycle in hwh and hw+wh tours with the nest private vehicles, and the commuters who take bus and KRL in hw+wh tours with the nest public transport.

In terms of correlation among alternatives in each nest, the λ values show that these values are less than one meaning that the correlation among alternatives exist. The λ values for hwh tours, hw+wh tours, private vehicle, and public transport are 0.99, 0.22, 0.92, and 0.00 respectively. Moreover, by using the t-test, the results indicates that all λ values are different from one. The most apparent correlation among alternatives is found on the nest public transport meanwhile the lowest correlation is on the nest hwh.

On the other hand, the modes car, bus, and KRL still become the most preferred mode for the commuters whereas the motorcycle is less preferred. It is found that the parameter magnitude of car is the highest among the other modes, followed by bus and then KRL.

Consistent with the previous result, the generic parameters of travel cost and travel time have reasonable significance and correct signs. Based on the magnitude value of this generic parameters, the commuters tend to choose the cheapest mode rather than the fastest mode as the estimated parameters which are shown to be -0.23 and -0.01 for travel cost and travel time correspondingly.

Most of the dummy parameters for commuters' income show reasonable significance. The commuters' income in this model also show the same interpretation with the previous model. The low income commuters have preference to ride motorcycle whereas high income commuters have a higher preference to drive car.

Contrary with the results of the previous model for female commuter dummy, the dummy parameter is found to be significantly positive for female commuters who take public transport in hwh tours, whereas it is insignificant for hw+wh tours. This indicates that the female commuters have the tendency to make use of public transport in simple tour rather than complex tour. Moreover, the female commuter is found to be more likely to choose bus rather than KRL, as the magnitude value for bus is higher than KRL.

Similar to the findings in the two previous model, the part-time worker are less likely to make simple tour since the parameters is found to be positive in complex tour but it is not significant. Therefore, this logit model also could not capture significant result for this parameter.

Summary of statistical analysis of this model shows that there are 26 parameters obtained from eight alternatives. The goodness of fit and final log likelihood are improved compared to the two previous model. The rho square and adjusted rho square are found to be 0.37 and 0.3 respectively while the final log likelihood is -478.28. Thus, the likelihood ratio compared to the NL is found to be 30.53 in which statistically significant at 10% tolerance of error and 11 degree of freedom.

5.4. Model comparison

In three logit models above, the utility parameter of mode and tour type for the commuters are estimated by using individual and socio-demographic basis to investigate the commuter preferences for travel modes when taking their tour type. This section will present the comparison of similar attributes used at least for two of three logit models that have been used in this study. The comparison of three models performance above is presented in Table 3.

As shown in Table 3, it can be seen that the three logit models used: MNL, NL, and CNL mostly yield similar results concerning the alternative specific constants and estimated parameters having similar signs and magnitude. Furthermore, the estimated parameters of an individual basis and socio-demographic also have similar significance level in terms of statistical analysis.

However, some parameters were found to have different significant results associated with the dummy female commuters. This can be found by estimating nest-specific parameters of the dummy female commuter in two different nest in the NL and CNL model. This is indicating that NL and CNL model provides a better estimation for a particular group of alternatives.

Dummy part time commuter was found not to be significant for all model. This might be because the data provided little evidence that the part-time commuter tend to have complex tour in their commute trip since the amount of part-time worker is only 1.90% from the total sample.

The generic parameters and alternative specific constants are used in the same utility functions for three models. Thus, these parameters are easy to compare based on the significant results. The CNL model provides the highest t-test value compared to the other model, indicating that CNL is more appropriate in evaluating the estimated parameters than MNL and NL.

Based on the findings, it was shown that the alternative specific constants for motorcycle has the lowest magnitude compared to other modes. However, from the initial descriptive analysis, this mode was found to be most commonly used by the commuters. This is because the travel cost and travel time for motorcycle were found to be the lowest compared to other modes based on the data source. Therefore, the utility of motorcycle is greater than other modes.

Going further, all of the λ parameters in CNL model are significantly different from one another. It implies that this model is also more robust to capture the increasing correlation inside the nests. Likewise, α parameters have reasonable significance which specifies that correlation among alternatives in different nests could be captured.

Table 3. Summary of model performance for MNL, NL, and CNL.

Utility coefficient	MNL			NL			CNL		
	Estimate	**Robust std. error**	**Robust t-stat**	**Estimate**	**Robust std. error**	**Robust t-stat**	**Estimate**	**Robust std. error**	**Robust t-stat**
Alternative-specific constants									
Motorcycle	0.69*	0.43	1.58	0.57*	0.35	1.6	0.24*	0.19	1.23
Car	4.45	1.02	4.34	4.08	0.87	4.71	4.97	0.77	6.39
Bus	2.34	0.47	4.94	2.21	0.39	5.63	2.24	0.31	7.3
KRL	1.00 (fixed)	NA	NA	1.00 (fixed)	NA	NA	1.00 (fixed)	NA	NA
Scale parameters									
λ_{hwh}				0.55	0.09	6.01	0.99	0.00	2.01
λ_{hw+wh}				0.99 (fixed)	NA	NA	0.22	0.16	4.83
Household characteristics									
Dummy: low income for motorcycle users	5.38	1.02	5.27						
Dummy: high income for car users	0.43	0.18	2.35						
hwh									
Dummy: low income for motorcycle users				6.59	0.95	3.76	3.53	1.33	2.65
Dummy: high income for car users				0.5	0.15	1.75	0.76	0.21	3.56
Dummy: female for bus users				1.42	0.28	5.08	0.76	0.21	3.58
Dummy: female for KRL users				0.40*	0.35	1.13	0.20	0.08	2.48
hw+wh									
Dummy: low income for motorcycle users				5.23	0.94	5.58	5.20	1.32	3.94
Dummy: high income for car users				0.58	0.16	3.72	0.75	0.22	3.4
Dummy: part-time worker, travel purpose	1.21*	1.07	1.13	0.97*	0.95	1.02	1.01*	1.02	0.99
Dummy: female for bus users				0.19*	0.4	0.5	-0.09*	0.29	-0.30
Dummy: female for KRL users				0.63	0.38	1.65	0.28*	0.26	1.05
Travel characteristics									
Travel cost (Rupiah)	-0.18	0.1	-1.79	-0.2	0.08	-2.39	-0.23	0.06	-3.77
Travel time (min)	-0.02	0	-4.41	-0.01	0	-3.7	-0.01	0.00	-3.41
Dummy: female for bus users	1.34	0.32	4.17						
Dummy: female for KRL users	0.71	0.38	1.89						
Value of travel time (Rp/hr)		1,381.29			627.86			544.14	
Summary statistics									
Number of observations		420			420			420	
Number of alternatives		8			8			8	
Number of parameters		12			15			26	
Rho square		0.31			0.35			0.37	
Adj. Rho square		0.29			0.33			0.34	
Null Log likelihood		-761.2			-761.2			-761.20	
Final Log likelihood		-525.78			-493.54			-478.28	
Likelihood ratio					64.48			30.53	
Chi-square statistic (for LR test)					6.25 (3, 0.10)			17.28 (11, 0.10)	

Based on the comparison table above it can be seen that the values of travel time (VTT) are different among the three models. The MNL has the highest VTT about 1,381.29 IDR/hour while the CNL has the lowest VTT around 544.14 IDR/hour. These values are sensible in the Indonesian context.

Regarding the formal test of the result, the CNL model shows better improvement compared to the MNL and NL model. The final log likelihood value increases from about -525.78 in the MNL to -493.54 in the NL and then improved to -478.28 in the CNL model. Likewise, the goodness-of-fit test which includes rho square and adjusted rho square also shows better result compared to the values obtained in both MNL and NL model since these values in CNL model are found to be substantially the highest. Furthermore, by using the likelihood ratio test, the CNL model is significantly better at 90% and even at 99% confidence level. According to this study, therefore, the CNL model is argued to be the optimum model among the tested ones, to be used in the estimation of parameters among multi alternatives in nesting structures that may belong to other nests. Moreover, NL is also considered to give significant improvement of final log likelihood for nesting the choices based on tour types, indicating that NL model is also appropriate to evaluate nesting structure in tour-based mode choice model.

6. Conclusion

This study attempted to present the discrete choice models of different logit models analysis using MNL, NL, and CNL models. The models were specified to estimate the relationship between tour type and mode choice and influence attributes of the commuters from Bekasi to Jakarta by considering the travel patterns and socio-demographic of the respondents. The estimation of these models, used a set of secondary data obtained from the travel diary of commuters. This set of data allowed these models to be compared in formal and informal tests for each model to evaluate the estimated parameters that offer significant results whether there is an existence of correlation between the alternatives.

From the analysis and discussions above, it can be argued that commuters' characteristics can influence the mode choice. According to the results from three different models, the female commuters have higher rate of using public transport such as bus and KRL which is consistent with the findings by Yagi and Mohammadian (2008). Increasing income has positive effect for the commuters on choosing car for their daily commute trip whereas low income push the commuters to have higher preference to ride motorcycle. This finding is consistent by the study conducted by Soltani (2017) in developing countries. Moreover, travel time and travel cost were found to contribute negative propensity toward the utilities.

The three logit models show reasonable result for estimation. However, the NL model showed significant improvement in final log likelihood when grouping the joint choices based on tour types. Furthermore, the CNL model is considered the better model structure as it showed the most powerful model to capture the correlation among alternatives within nest and across the nests. Some of the estimated parameters were also found to have higher significance level compared to the two other models. This CNL model also has the highest final log likelihood and the most sufficient model structure that fits the data in terms of goodness-of-fit test.

Acknowledgements

I would like to express my gratitude to the Ministry of Research, Technology and Higher Education of the Republic of Indonesia for the survey data (Grant number: 170/UN1/DITLIT/DIT-LIT/LT/2018). I would also like to thank Bashir Haque, Dian Azmi, Ahmida Rizki Aulia, and Wida Yuliar Rezika as the discussion partner for this paper. Finally, I would like to give the highest thank to the Indonesia Endowment Fund for Education (LPDP) for sponsoring my postgraduate study.

Appendix A. Scale parameter of CNL

Scale parameters	Estimate	Robust std. error	Robust t-stat
λ_{hwh}	0.99	0.00	2.01
λ_{hw+wh}	0.22	0.16	4.83
$\lambda_{private\ vehicle}$	0.92*	0.06	1.37
$\lambda_{public\ transport}$	0.00	0.00	341.66
α_{alt1_hwh}	0.99*	0.01	0.02
$\alpha_{alt1\_private\ vehicle}$	0.01*		
α_{alt2_hwh}	0.12	0.04	20.51
$\alpha_{alt2\_private\ vehicle}$	0.88		
α_{alt3_hwh}	0.59	0.08	5.27
$\alpha_{alt3\ public\ transport}$	0.41		
α_{alt4_hwh}	0.75	0.07	3.71
$\alpha_{alt4\ public\ transport}$	0.25		
α_{alt5_hw+wh}	0.99	0.00	4.98
$\alpha_{alt5\_private\ vehicle}$	0.01		
$\alpha_{alt6\ hw+wh}$	0.00	0.00	43818.49
$\alpha_{alt6\_private\ vehicle}$	1.00		
α_{alt7_hw+wh}	0.99	0.00	2.79
$\alpha_{alt7\_public\ transport}$	0.01		
α_{alt8_hw+wh}	0.98*	0.65	0.02
$\alpha_{alt8\ public\ transport}$	0.02*		

References

Bowman, J.L., Ben-Akiva, M.E., 2000. Activity-based disaggregate travel demand model system with activity schedules. Transp. Res. Part A Policy Pract. 35, 1–28. https://doi.org/10.1016/S0965-8564(99)00043-9

Dharmowijoyo, D.B.E., Susilo, Y.O., Karlström, A., 2017. Analysing the complexity of day-to-day individual activity-travel patterns using a multidimensional sequence alignment model: A case study in the Bandung Metropolitan Area, Indonesia. J. Transp. Geogr. 64, 1–12. https://doi.org/10.1016/j.jtrangeo.2017.08.001

Dharmowijoyo, D.B.E., Susilo, Y.O., Karlström, A., 2016. Day-to-day variability in travellers' activity-travel patterns in the Jakarta metropolitan area. Transportation (Amst). 43, 601–621. https://doi.org/10.1007/s11116-015-9591-4

Dissanayake, D., Morikawa, T., 2010. Investigating household vehicle ownership, mode choice and trip sharing decisions using a combined revealed preference/stated preference Nested Logit model: case study in Bangkok Metropolitan Region. J. Transp. Geogr. https://doi.org/10.1016/j.jtrangeo.2009.07.003

Dissanayake, D., Morikawa, T., 2002. Household Travel Behavior in Developing Countries: Nested Logit Model of Vehicle Ownership, Mode Choice, and Trip Chaining. Transp. Res. Rec. 1805, 45–52. https://doi.org/10.3141/1805-06

Hess, S., Fowler, M., Adler, T., Bahreinian, A., 2012. A joint model for vehicle type and fuel type choice: Evidence from a cross-nested logit study. Transportation (Amst). 39, 593–625. https://doi.org/10.1007/s11116-011-9366-5

Ho, C., Mulley, C., 2013. Tour-based mode choice of joint household travel patterns on weekend and weekday. Transportation (Amst). 40, 789–811. https://doi.org/10.1007/s11116-013-9479-0

Hu, H., Xu, J., Shen, Q., Shi, F., Chen, Y., 2018. Travel mode choices in small cities of China: A case study of Changting. Transp. Res. Part D Transp. Environ. 59, 361–374. https://doi.org/10.1016/J.TRD.2018.01.013

Hyodo, T.T., Jr, C.M.M., Fujiwara, A., Soehodho, S., 2005. Urban Travel Behaviour Characteristics of 13 Cities based on Household Interview Survey data. J. East. Asia Soc. Transp. Stud. 6, 23–38.

Irawan, M.Z., Belgiawan, P.F., Widyaparaga, A., Deendarlianto, X., Budiman, A., Muthohar, I., Sopha, B.M., 2017. A market share analysis for hybrid cars in Indonesia. Case Stud. Transp. Policy 6, 336–341. https://doi.org/10.1016/j.cstp.2017.09.003

Irawan, M.Z., Putri, M.K., Belgiawan, P.F., Dwitasari, R., 2017. Analyzing Commuters' Behavior on Egress Trip from Railway Stations in Yogyakarta, Indonesia. Open Transp. J. 11, 53–66. https://doi.org/10.2174/1874447801711010053

Lekshmi, G.R.A., Landge, V.S., Kumar, V.S.S., 2016. Activity Based Travel Demand Modeling of Thiruvananthapuram Urban Area. Transp. Res. Procedia 17, 498–505. https://doi.org/10.1016/J.TRPRO.2016.11.100

Miller, E.J., Roorda, M.J., Carrasco, J.A., 2005. A tour-based model of travel mode choice. Transportation (Amst). 32, 399–422. https://doi.org/10.1007/s11116-004-7962-3

Santoso, D.S., Tsunokawa, K., 2005. Spatial transferability and updating analysis of mode choice models in developing countries. Transp. Plan. Technol. 28, 341–358. https://doi.org/10.1080/03081060500319694

Siddique, M.A.B., Choudhury, C.F., 2017. Modelling the Behavioural Response to Congestion Pricing in Dhaka, Bangladesh. Transp. Dev. Econ. 3, 23. https://doi.org/10.1007/s40890-017-0052-y

Soltani, A., 2017. Social and urban form determinants of vehicle ownership; evidence from a developing country. Transp. Res. Part A Policy

Pract. 96, 90–100. https://doi.org/10.1016/j.tra.2016.12.010

Statistics of Bekasi Municipality, 2017. Bekasi Municipality in Figures. Statistics of Bekasi Municipality, Bekasi.

Statistics of DKI Jakarta Province, 2017. Jakarta in Figures. Statistics of DKI Jakarta Province, Jakarta.

Statistics of DKI Jakarta Province, 2015. Commuter of DKI Jakarta Province. Jakarta.

Train, K.E., 2009. Discrete choice methods with simulation, 2nd Ed. ed. Cambridge University Press, New York. https://doi.org/10.1016/S0898-1221(04)90100-9

Wen, C.-H., Koppelman, F.S., 2001. The generalized nested logit model. Transp. Res. Part B Methodol. 35, 627–641. https://doi.org/10.1016/S0191-2615(00)00045-X

Yagi, S., Mohammadian, A., 2008. Joint Models of Home-Based Tour Mode and Destination Choices: Applications to a Developing Country. Transp. Res. Rec. J. Transp. Res. Board 2076, 29–40. https://doi.org/10.3141/2076-04

Yagi, S., Nobel, D., Kawaguchi, H., 2013. Time Series Comparison of Joint Mode and Destination Choice Models in Changing Transportation Environment. Transp. Res. Rec. J. Transp. Res. Board 2394, 100–110. https://doi.org/10.3141/2394-13