DISS. ETH NO. 27271

ACTIVITY BASED MODEL (ABM) APPROACHES FOR SUSTAINABLE CITIES

A dissertation submitted to attain the degree of

DOCTOR OF SCIENCES of ETH ZÜRICH
(Dr. sc. ETH Zürich)

presented by

ANUGRAH ILAHI
M.Sc. in Transport System and Engineering, Universitas Gadjah Mada

born on 23 September 1989
citizen of Indonesia

accepted on the recommendation of

Prof. Dr. Kay. W. Axhausen, examiner
Prof. Dr. Abolfazl Mohammadian, co-examiner
Dr. Prawira Fajarindra Belgiawan, co-examiner

2021

For my country Indonesia, my case study Greater Jakarta, my hometown Bengkulu, and for sustainable cities.

ABSTRACT

Urban transport is a substantial issue in many countries. Congestion and public transport unreliability need to be addressed. Any new policy needs to be tested comprehensively against people's behavior. This thesis comprehensively captures travel behavior and willingness to pay for the transportation modes. This thesis further develops an agent-based model of Greater Jakarta to model all complex interactions in which individuals interact with households, activities, and facilities.

There were several steps to develop the model. Firstly, the demand was built based on the survey of JICA in 2012 and combined with the RP survey to capture the trip purposes. We used the Bayesian network and generalized raking for the population synthesis. This method can generate the travel demand that represents the population in Greater Jakarta. We created a 30 million synthetic population based on responses from around 600,000 individuals from 170,000 households. Secondly, networks and public transport lines were built as supply facilities. The data from Open Street Map was used to create the network, and the data of public transport lines were scraped from Trafi. This supply and demand-side were used for running the agent-based model using multi-agent transport simulation (MATSim).

The thesis also created a state of the art a stated preference (SP) and revealed a preference (RP) survey with 5,000 respondents. This thesis developed a model by pooling SP and RP datasets using the multinomial logit model (MNL) and mixed logit model (MXL) to see how those modes affect mode choice. The parameter estimated from the model is used to calibrate the MATSim model so that the mode share from the MATSim model matches the HTS. Utilizing parameter estimated from the choice model claims to speed up the simulation time and convergence.

There are new modes of transportation options available that have been investigated, such as on-demand transports (car-based, motorcycle-based) that exist in many countries due to the development of information communication technology (ICT). Other alternatives will be available in the coming years, such as autonomous vehicles and urban air mobility (UAM).

The results show interesting results. First, the development of the scenario for greater Jakarta required a lot of effort. The availability of data and the population's size required expensive computations starting from population

synthesis to model simulation. Second, we investigated the willingness to pay (WTP), e.g., the value of travel time savings (VTTS), value of travel time assigned to travel (VTAT), and the elasticity for all mode choice alternatives, including on-demand transport (ODT) and UAM. Third, we found that the recorded travel behavior answered several questions, such as who the users are, when the users use it, the users' purposes, and the speed at different distances and locations. Finally, we developed a scenario for policy recommendations using an agent-based model. In this case, we modeled road pricing implementation on main roads in Jakarta.

This research contributes by building an agent-based model scenario and investigates the impact of congestion charging in Greater Jakarta. This research will be a base of future research for developing more complex scenarios. This thesis advises on the improvement of urban transportation in Greater Jakarta. The model can be employed and replicated to answer several issues in other agglomerations in Indonesia, such as Greater Bandung, Greater Bali, Greater Surabaya, Greater Medan, and Greater Makassar.

ZUSAMMENFASSUNG

Der Stadtverkehr ist in vielen Ländern ein wesentliches Thema. Staus und die Unzuverlässigkeit des öffentlichen Verkehrs müssen angegangen werden. Jede neue Politik muss umfassend anhand des Verhaltens der Menschen getestet werden. Diese Arbeit erfasst umfassend das Reiseverhalten und die Zahlungsbereitschaft für die Verkehrsmittel. Diese Arbeit entwickelt außerdem ein agentenbasiertes Modell des Großraums Jakarta, um alle komplexen Interaktionen zu modellieren, bei denen Individuen mit Haushalten, Aktivitäten und Einrichtungen interagieren.

Es gab mehrere Schritte zur Entwicklung des Modells. Zunächst wurde die Nachfrage auf der Grundlage der JICA-Umfrage von 2012 erstellt und mit der RP Umfrage kombiniert, um die Fahrtzwecke zu erfassen. Wir verwendeten das Bayes'sche Netzwerk und verallgemeinertes Raking für die Bevölkerungssynthese. Diese Methode kann die Reisenachfrage generieren, die die Bevölkerung im Großraum Jakarta repräsentiert. Wir erstellten eine synthetische Bevölkerung von 30 Millionen, basierend auf den Antworten von etwa 600.000 Personen aus 170.000 Haushalten. Zweitens wurden Netze und Linien des öffentlichen Verkehrs als Versorgungseinrichtungen erstellt. Zur Erstellung des Netzwerks wurden die Daten von Open Street Map verwendet, und die Daten der öffentlichen Verkehrslinien wurden von Trafi gescraped. Diese Angebots- und Nachfrageseite wurden für die Ausführung des agentenbasierten Modells unter Verwendung der Multiagenten Transportsimulation (MATSim) verwendet.

Im Rahmen dieser Arbeit wurde auch eine State-of-the-Art-Umfrage zur angegebenen Präferenz (SP) und zur offengelegten Präferenz (RP) mit 5.000 Befragten erstellt. In dieser Arbeit wurde ein Modell entwickelt, indem SP- und RP-Datensätze unter Verwendung des multinomialen Logit-Modells (MNL) und des gemischten Logit-Modells (MXL) zusammengeführt wurden, um zu sehen, wie diese Modi die Verkehrsmittelwahl beeinflussen. Die aus dem Modell geschätzten Parameter werden zur Kalibrierung des MATSim-Modells verwendet, so dass der Verkehrsmittelanteil aus dem MATSim-Modell mit der HTS übereinstimmt. Die Verwendung der aus dem Wahlmodell geschätzten Parameter soll die Simulationszeit und die Konvergenz beschleunigen.

Es gibt neue Transportmöglichkeiten, die untersucht wurden, wie z. B. Transporte auf Abruf (mit dem Auto, mit dem Motorrad), die in vielen

Ländern aufgrund der Entwicklung der Informations- und Kommunikationstechnologie (IKT) existieren. Weitere Alternativen werden in den kommenden Jahren verfügbar sein, wie z. B. autonome Fahrzeuge und urbane Luftmobilität (UAM).

Die Ergebnisse zeigen interessante Resultate. Erstens erforderte die Entwicklung des Szenarios für den Großraum Jakarta einen hohen Aufwand. Die Verfügbarkeit von Daten und die Größe der Bevölkerung erforderten aufwendige Berechnungen, angefangen von der Bevölkerungssynthese bis zur Modellsimulation. Zweitens untersuchten wir die Zahlungsbereitschaft (WTP), z. B. den Wert der Reisezeitersparnis (VTTS), den Wert der für die Reise aufgewendeten Zeit (VTAT) und die Elastizität für alle Verkehrsmittelwahlalternativen, einschließlich On-Demand-Transport (ODT) und UAM. Drittens fanden wir heraus, dass das aufgezeichnete Reiseverhalten mehrere Fragen beantwortet, z. B. wer die Nutzer sind, wann die Nutzer es nutzen, die Zwecke der Nutzer und die Geschwindigkeit bei verschiedenen Entfernungen und Orten. Schließlich entwickelten wir ein Szenario für politische Empfehlungen unter Verwendung eines agentenbasierten Modells. In diesem Fall modellierten wir die Implementierung von Straßenbenutzungsgebühren auf Hauptstraßen in Jakarta.

Diese Forschung trägt dazu bei, indem sie ein agentenbasiertes Modellszenario erstellt und die Auswirkungen von Staugebühren im Großraum Jakarta untersucht. Diese Forschung wird eine Basis für zukünftige Forschung sein, um komplexere Szenarien zu entwickeln. Diese Arbeit gibt Ratschläge für die Verbesserung des städtischen Verkehrs im Großraum Jakarta. Das Modell kann eingesetzt und repliziert werden, um verschiedene Fragen in anderen Ballungsräumen in Indonesien zu beantworten, wie z.B. im Großraum Bandung, Großraum Bali, Großraum Surabaya, Großraum Medan und Großraum Makassar.

ACKNOWLEDGEMENTS

First of all, I would like to thank Allah S.W.T, for giving me so much grace and guidance for this life; without Allah, I will be nothing in this world who is taking care of my health, family and giving me the spirit of life. His plans are always the best.

My thanks also go to my very humble supervisor, Prof. Kay W. Axhausen, for having given me this opportunity. I did not expect that he gave me this opportunity without known each other personally. I am thankful for your belief that I could get funding from my government after failed in several scholarship applications. I never imagined that I would be part of your Ph.D. student and part of this prestigious university. I am very grateful for your support during my stay in Switzerland and for struggling with my Ph.D.

I want to thank my co-supervisor, Prof. Abolfazl Mohammadian, for giving input and other insights to improve my thesis. My thanks also go to my second co-supervisor, Dr. Prawira Fajar Belgiawan, which we have met before during his postdoctoral in IVT. I still remember in 2016 that he told me that there are plans of Allah S.W.T why we meet here. There will be something beneficial for us and our country, for sure. It was nice to cooperate with you during fieldwork in greater Jakarta and also to do research together.

I want to thank Lembaga Pengelola Dana Pendidikan (LPDP), who gave me funding for doing my P.hD for four years; without support from LPDP, it will be tough for me to get Ph.D. LPDP helps the child from a small city in Sumatera achieve the highest education at the world's best university.

I would also like to thank all colleagues in IVT ETH Zurich and admin and IT, such as Jenni, Valerie, Elisabeth, and Pieter. I remembered that Jenni helped me a lot with the administration at ETH Zurich. Milos, who helps me understand MATSim and answer my questions, Kirill, who teach me about R for the first time when I have zero experience with coding, Sebastian, who also helps me with MATSim. Taha, who did postdoc, which we have discussed many things. Basil who help me to understand the discrete-choice model. The other IVT members, such as Aoyong, Felix, Grace, Thomas, Nuan Nuan, Sergio, Stefano, and others, may interact with me. Other visiting students Kun, Yizhe, and Kalrav. Indonesian Embassy, Percikan Iman, Indonesian and Indonesian students such as Raka, Nanda and family, Azmir, Agung and family, Bram, Nova, Nui, Wahyu, Hesty, Danang, Syifa, and others.

I want to thank my family and relatives who always pray and hope for the best for me. All teachers, since I was at school in Bengkulu and University in Yogyakarta. All colleagues when I was working in GIZ, Bappenas, MoT. All friends in Indonesia.

I want to thank my parents (Mr. Rizkan and Mrs. Herawati), who always support, pray, and believe in me, especially for education. I had never imagined that I could pursue my dream to study abroad, which my mother feels very sad about when I did not get any scholarship that I applied for in 2015, but I told her that I would keep searching for scholarships and give up my dream if I already 40 years old. They are both are the best parents that can handle my limitations. Next, it goes to my brother (Rengga) and sister (Mona), who always support me; I hope we can always make our parents proud and happy.

Last but never least, I would like to extend my sincerest gratitude to my beloved and beautiful wife, Mrs. Nony Natadia Ernel. Without her love, pray, encouragement, and continuous support, this dissertation would not have existed. I still remember how we struggled immensely with the visa procedures, finding an apartment in Zurich, and support when we were sick before submitting my thesis for Ph.D. defense, especially during pandemic COVID-19. Thank you always with me.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

1.1 MOTIVATION

Agent-based modeling has been growing in recent decades with its ability to capture people's behavior. This approach is an important step towards capturing the complex interaction between people's behavior and infrastructures. On one hand, each person has different characteristics like income, age, gender, and trip purpose. On the other hand, infrastructure has several constraints, such as starting time, closing time, and access to infrastructures. This complex interaction makes the system difficult to model.

We develop a model of this type for Greater Jakarta, which has a massive population of around 30 million inhabitants and many different types of transport modes available. Greater Jakarta consists of 3 provinces with a total of 13 cities. This area, as the capital region, has a considerable impact on the Indonesian economy. Therefore, solving urban transport problems will be beneficial for the economy and quality of life in Greater Jakarta.

However, agent-based modeling has not been well adopted for Indonesia. Only in Padang, Indonesia where Taubenböck *et al.* (2013) simulate the evacuation during Tsunami. Transport infrastructures have been planned based on the needs for small scale regions with some facilities not connected to each other. For example, when we have to transfer to another transport mode and it is not in a walkable distance. In the past, we used the four-step model as a transport planning tool. Recently, there are many emerging modes of transport, like ride-hailing, car-sharing, electric car sharing, urban air mobility, which can not be modeled fully with a conventional four-step model. Therefore, this method can not help us to model complex interactions anymore.

1.2 RESEARCH GAP

Previous studies were conducted in Greater Jakarta. However, they are limited to a four-step model, which only consider aggregate origin and destination trip (Hadian, 2018), choice modeling (Belgiawan *et al.*, 2019b;

Ilahi *et al.*, 2018), and analysis of an activity diary survey (Dharmowijoyo *et al.*, 2016; Yagi and Mohammadian, 2010).

This research, therefore, is integrating the results from a choice model and activity diary survey to develop an agent-based model, which can model people's behaviour this should improve the accuracy of the model as each person is simulated based on its socio-demographic attributes, activity location, activity chain, mode choice, and facilities constraint. However, the agent-based model consume considerable amount of time to simulate and need much more detailed data-sets especially for large number of agents.

Furthermore, this study comprehensively captures the impact of emerging transport modes in Greater Jakarta such as car On-Demand transport (ODT), and motorcycle (MC) ODT, which are growing in recent years. This study also investigates urban air mobility (UAM), which can be an option for future transport. The impact of congestion charging is also investigated by usng an agent-based model. In this model, we consider for both car and motorcycle the income of the users, which previous studies did not consider (de Freitas *et al.*, 2017; de Palma and Lindsey, 2006; Kaddoura and Kickhöfer, 2014).

We conducted a Stated Preference (SP), and Revealed Preference (SP) survey to obtain the parameter estimates for a mode choice model. Further, we used the results from the RP survey to include mandatory and secondary activities in our simulation. We used MATSim (Multi-Agent Transport Simulation) as a tool for the agent-based model (Horni *et al.*, 2016).

1.3 RESEARCH OBJECTIVES

The main objectives of this research is to create an agent-based model of Greater Jakarta, which further is used as a model to investigate different transport policies. To develop the model, we did several steps from the demand and supply side. The second chapter discusses population synthesis, where we develop the demand for our model.

The third chapter describes two surveys, i.e. the revealed preference (RP) survey, and stated preference (SP) survey. It discusses the survey design, descriptive analysis, and the parameter estimate based on the multinomial logit (MNL) model and mixed logit (MXL) model.

The fourth chapter discusses the MATSim network and the public transport lines, which we have developed based on current conditions. The fifth chapter describes the model building, where we combine the results of the population synthesis with the choice model. The sixth chapter discusses the policy scenario. Finally, the last chapter discusses the conclusions and recommendations.

1.4 ORIGINAL PAPERS AND CONTRIBUTIONS

Each chapter of this thesis is based on several papers. The following paragraphs explain the part of the paper in the chapter.

- Chapter 2 is based on Ilahi and Axhausen (2019), which was done by Ilahi and updated to cover all activities of people behaviour in Greater Jakarta. Kay Axhausen provided guidance, comments, and editing.
- Chapter 3 is based on Ilahi *et al.* (2019d), which is under review. Ilahi created the survey and experimental design, did the analysis, managed the survey process, and prepared the manuscript. Axhausen provided guidance, comments, and editing. Belgiawan supervised the fieldwork, commented, and edited, and Balac provided comment and editing. The survey was funded by IVT ETH Zurich as a part of an ongoing Airbus project.
- Chapter 4 is based on Ilahi *et al.* (2019d,c); Belgiawan *et al.* (2019b) which are under review and published respectively, for which Ilahi created the survey, did the analysis, managed the survey process, and prepared the manuscript. Axhausen provided guidance, comments, and editing. Belgiawan supervised the fieldwork, comments, and editing, and Balac provided comments and editing. The survey was funded by IVT ETH Zurich as a part of an ongoing Airbus project.
- Chapter 5 and 6 are based on Ilahi *et al.* (2019a). They also based on Ilahi *et al.* (2019b), which is not published yet. Balac provided comments, Kay Axhausen provided guidance, comments, and editing.

2

POPULATION SYNTHESIS

2.1 INTRODUCTION

This chapter discusses how we conducted the population synthesis. Several methods are discussed in this chapter, the advantages and disadvantages of each method and implementation in other countries.

Many studies have mentioned that agent-based transportation models require detailed individual and household information such as socio-demographic variables, and geocoding of activity locations (Balmer *et al.*, 2006). These data can be collected from multiple data sources. To collect all those data for a population is costly in terms of time and resources. Therefore, several approaches can be used for population synthesis from limited data. One of the initial methods is iterative proportional fitting (IPF), which was first introduced by Deming and Stephan (1940) and first implemented in transport research by (Beckman *et al.*, 1996). To tackle the sample problems, Barthelemy and Toint (2013) developed a method for the synthetic population without a sample by using the available information of dis-aggregated level data and random draws. This method was shown to be able to produce a consistent synthetic population. IPF consists of several iteration steps, where each row and each column are proportionally adjusted to be equal to the marginal row, and column totals. The steps are repeated until both row and column converge or the sum of the rows and columns are relatively similar to their marginal total. However, the control total of IPF is based on marginal individual or marginal household totals. Since IPF is widely used for population synthesis, there are several efforts to improve the quality of IPF and to develop new approaches. As reviewed by Müller and Axhausen (2011); Sun and Erath (2015), the IPF has been extended by many researchers, for example, to deal with the zero-cell issue (Guo and Bhat, 2007). Ye *et al.* (2009) proposed an iterative proportional updating (IPU), Pritchard and Miller (2012) addressed memory consumption issues. Additionally, Casati *et al.* (2015); Zhu and Ferreira (2014) introduced hierarchical and multi-stage IPF procedures. IPU and Hierarchical IPF are an advanced method of IPF that consider household and individual control totals. However, if it only uses one type of control total, it is considered as IPF.

Furthermore, there are other methods, such as Combinatorial Optimization (CO) (Voas and Williamson, 2001). CO is a method that estimates the integer weight of a sample and replicates it to the desired marginal target total. This method claims to have less variance than IPF. Farooq *et al.* (2013); Saadi *et al.* (2016a) proposed a Markov Chain Monte Carlo (MCMC) method that deals with the zero-cell issue. There is an extension of MCMC, called hierarchical MCMC, that uses individual and household attributes simultaneously. Hafezi and Habib (2014) further employed fitness based synthesis (FBS), which allows considering multilevel controls by selecting variables that have maximum fitness value. However, this approach gave less accuracy when there was a more detailed classification of attributes. Sun *et al.* (2018) developed a hierarchical mixture model. This method is combining three different models i.e., probabilistic tensor factorization, multilevel latent class modeling, and rejection sampling. This approach can produce a pool of agents that consider the inter-dependencies of household and individual structure. This method is complicated and not widely applicable, though.

Several use cases employ population synthesis with different approaches. We summarize the use cases of the various methods for different locations and population sizes. As can be seen in Table 2.1, most of the locations are in developed countries and have a population under 10 million. Furthermore, there is related software that can be used for the synthetic population generation, such as, PopoSynWin, ILUTE, PopGen, FSUMTS, CEMDAP, ALBATROS , R package sms, R package synthpop, TRANSIMS, Synthia, and SMILE (Müller and Axhausen, 2011; Templ *et al.*, 2017).

Source	Location	Methods	Sample size (%)	Population in millions
Barthelemy and Toint (2013)	Belgium	randomly drawing	0	10.00
Casati *et al.* (2015)	Singapore	IPF	1	4.00
Farooq *et al.* (2013)	Brussels	MCMC	0.1	1.20
Hafezi and Habib (2014)	Canada	FBS	1	1.30
Huynh *et al.* (2013)	Sydney	CO	NA	5.00
Moeckel *et al.* (2003)	Netanya	IPF with MCMC	6	2.60
Müller and Axhausen (2011)	Switzerland	IPF	5	8.00
Pritchard and Miller (2012)	Canada	IPF With MCMC	2	3.42
Saadi *et al.* (2016a)	Brussels	MCMC,IPF	0.1	1.20
Sun and Erath (2015)	Singapore	Comparing BN, MCMC, IPF, and DI	1*	4.00
Ye *et al.* (2009)	Arizona	IPU	8	3.07
Zhang *et al.* (2017)	San Francisco	BN	6	7.00
Zhu and Ferreira (2014)	Singapore	IPF	1	4.0

*Tests from 1% to 100% of sample size

TABLE 2.1: Summary of use cases of population synthesis

As an alternative, Sun and Erath (2015); Zhang *et al.* (2017) employed Bayesian networks (BN) for the task. They claim that it can capture complex interaction and hierarchical household structure of the sample and show that it is better than other methods (i.e., MCMC, DI, and IPF). Moreover, Saadi *et al.* (2016b) used the Hidden Markov model (HMM), which is a stochastic model that learns a complex joint distribution sample structure or known as the simplest BN. However, this model has less consideration for the hierarchical household structure. Since HMM could not fit the marginal totals, Saadi *et al.* (2018) improved the method that integrated an HMM and IPF. After the combination, the model can give a quasi-perfect result.

In this paper, we utilize a BN. However, similar to an HMM, a BN is only able to generate a synthetic data and gives a similar distribution to the sample (Sun and Erath, 2015; Zhang *et al.*, 2017). Therefore, we add a similar type of integration as in Saadi *et al.* (2018). In this case, we integrated a BN and a GR multilevel IPF to fit the aggregate census data for the following reasons. First, when the sample is less than 40 %, BN is known to give better results to model inter-dependencies of individual and household attributes (Sun and Erath, 2015). Second, BN outperforms other methods in terms of the resulting square root of the mean square error (SRMSE) and becomes an excellent model to capture heterogeneity when the sample is less than 70%.

This study makes the following contributions:

- We integrated two methodological approaches: BN integrated with generalized raking (GR) multilevel IPF
- We applied the model in a developing country and to a megacity, one of a total of 31 locations worldwide of 26 of them are in the less developed regions (United Nations, 2016) that have lower data availability
- This chapter adds to the growing literature on BN for population synthesis and its application in developing countries and megacities.

The remainder of this chapter is structured as follows. Section 2.2, explains the concept of the Bayesian network, and in Section 2.3 discusses study area and framework model. In Section 2.4, we apply the approach for a population synthesis using BN and GR multilevel IPF to fit aggregate census data. Discussions and conclusions follow in Section 2.5 and 2.6.

2.2 A BAYESIAN NETWORK

The Bayesian network uses a graphical method to learn probabilities for a model (Cowell *et al.*, 2006). It consists of two parts: a directed acyclic graph (DAG) and a set of the conditional probability distribution (Sun and Erath, 2015; Hornỳ, 2014), where DAG consists of a set of correlated random variables. The variables of a graphical structure $G = (V, A)$ are represented by a node or vertex (V), and the correlation is represented by the directed edge or arc A. In the example in Figure 2.1, there are variables income, age, and gender. The directed arrows from *NodeIncome* to *NodeAge* and from *NodeIncome* to *NodeGender* indicate that *NodeAge* and *NodeGender* are linked by a conditional probability with *NodeIncome*. Therefore, the conditional probability distribution of this condition is $P[NodeAge|NodeIncome]$ and $P[NodeGender|NodeIncome]$.

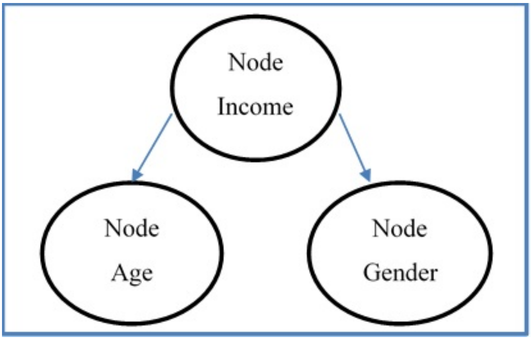


FIGURE 2.1: Example Directed acyclic graph (DAG)

2.2.1 *A learning algorithm for Bayesian networks*

There are different algorithms for learning Bayesian networks, as explained by Scutari (2010), such as the constraint-based algorithm, the score-based algorithm, or the hybrid algorithm. Each type of algorithm includes a learning algorithm. Here, we used the R package bnlearn (Scutari, 2010), which implements tabu search as part of the score-based algorithm. Tabu search, as a generic heuristic procedure, is an iterative searching procedure to obtain the best solution from complex correlation patterns (Glover and Taillard, 1993). Besides, it can handle local optima by selecting a very close solution to optimality, which can minimize the score (Scutari, 2010). It supports a whitelist and a blacklist. A blacklist means that those arcs will not be presented in the network structure, and a whitelist is a reverse of a blacklist.

2.2.2 *Network scores*

The step of measuring the candidate of the graphical structure that fits the data is important to ensure that our structure can produce a fitting synthetic population. In this process, several methods are introduced, such as maximum likelihood:

$$\mathcal{L}(G^h|D) = \max_G \sup_\Theta \mathcal{L}(G^h, \Theta|D) = \max_G (G^h, \hat{\Theta}|D) \tag{2.1}$$

Where $\mathcal{L}(G^h|D) = \max_G \sup_\Theta \mathcal{L}(G^h, \Theta|D)$ is the log-likelihood of a provided pair (G, Θ) given observation D. However, the log-likelihood is not re-presentable as explained by Sun and Erath (2015) due to over-fitting. This method will always build a fully connected DAG. Therefore, the most applicable approaches are using the Bayesian Information Criterion (BIC) (Rissanen, 1978; Schwarz *et al.*, 1978) and Akaike Information Criterion (AIC) (Rissanen, 1978; Akaike, 1974).

$$\mathrm{BIC}(G^h|D) = \log P(D|G^h, \hat{\Theta}) - \frac{d}{2} \log m \tag{2.2}$$

$$\mathrm{AIC}(G^h|D) = \log P(G^h, \Theta|D) - d \tag{2.3}$$

Where Θ, in the first equation, is the maximum likelihood estimate parameter given a hypothetical structure G^h. d is the degree of freedoms in Θ, and m is the number of observations. In contrast to maximum likelihood (2.1), BIC (2.2) and AIC (2.3) have a penalty function for the optimal likelihood $\log P(G^h, \Theta|D)$. For BIC, the penalty is $\frac{d}{2} \log m$, and for AIC is d. Using the scoring function, the best network is selected for constructing the synthetic population.

2.3 CONSTRUCTING THE POPULATION OF GREATER JAKARTA AREA

The study area is the Greater Jakarta Area or Jabodetabek, which consists of Jakarta Province, parts of West Java Province, and Banten Province. It has 31.7 million inhabitants in 2016 (BPS-Statistics, 2016a,b,c), see Table 2.2. The population data used for the synthetic population generation was partly obtained from the JAPTRAPIS study (Jabodetabek Public Transport Policy Implementation and Strategy) in 2009 (JICA, 2009; Ilahi and Axhausen, 2017).

Province	Region	Male	Female
Jakarta	South Jakarta	1,096,469	1,089,242
	East Jakarta	1,436,128	1,407,688
	Central Jakarta	457,025	457,157
	West Jakarta	1,246,288	1,217,272
	North Jakarta	867,727	879,588
Banten	Tangerang City	1,045,113	1,001,992
	Tangerang Regency	1,724,915	1,645,679
	South Tangerang City	777,713	765,496
West Java	Depok	1,061,900	1,044,200
	Bogor	532,000	515,900
	Bogor Regency	2,792,900	2,666,800
	Bekasi City	1,369,600	1,345,200
	Bekasi Regency	1,654,600	1,591,400
		Total	31,689,992

TABLE 2.2: The population of the Greater Jakarta area in 2016

There are two different types of data in the JAPTRAPIS study, which are the Household Travel Survey (HTS) and the Activity Diary Survey (ADS). However, we only use HTS data consisting of 178,954 households or 334'973 individuals for the population synthesis, which are equal to three percent of all households. In the HTS, the respondents are individuals who are going to school or office. Therefore, the aggregate synthetic population is limited to the census population of individuals who have activities for studying or working with a total of more than 20 million, as seen in Table 2.3.

The data used for this approach are from multiple sources within the HTS data. The variables of the individuals include age, gender, education, and employment status. The variables of the households are income, housing status, vehicle ownership, and address. In the activities-based model, we also need geocoding of activity locations. Since we geocode based on google API. Therefore, there are three sources of data in the model. We use the R package dplyr for the data joint. Figure 2.2 shows the framework model of data combination and integration between BN and GR. Further detail of the BN is presented in Figure 2.4.

Province	Region	Male	Female
Jakarta	South Jakarta	877,679	626,495
	East Jakarta	1,141,943	815,130
	Central Jakarta	367,092	262,034
	West Jakarta	989,250	706,136
	North Jakarta	701,639	500,837
Banten	Tangerang City	999,131	590,688
	Tangerang Regency	1,645,087	972,578
	South Tangerang City	753,194	445,290
West Java	Depok	711,791	394,497
	Bogor	354,155	196,284
	Bogor Regency	1,845,195	1,022,666
	Bekasi City	917,511	508,514
	Bekasi Regency	1,097,039	608,014
		Total	20,049,867

TABLE 2.3: The population of Jakarta greater area with the relevant activities in the census in 2016

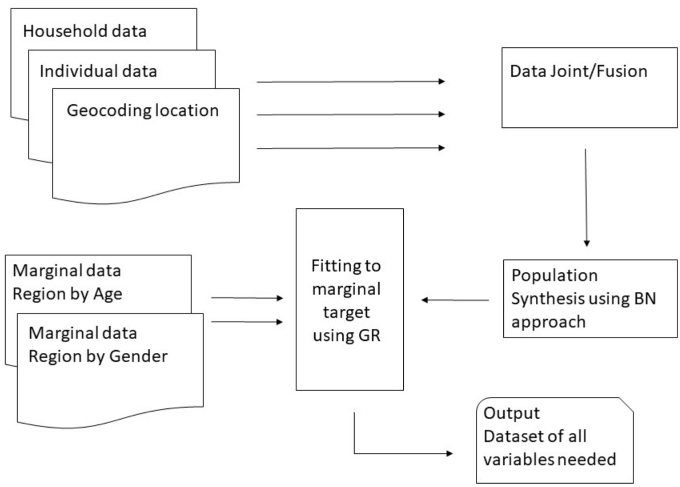


FIGURE 2.2: Framework model

2.4 MODEL ESTIMATION

2.4.1 *Bayesian network step*

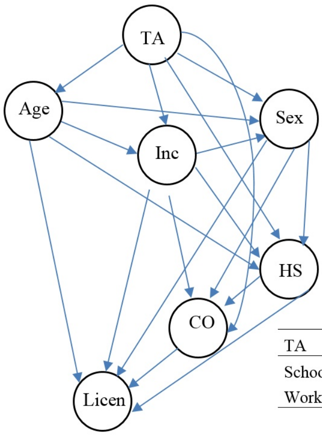
We consider seven variables by combining individual and household data for the population synthesis using BN, as presented in Table 2.4: type of activities, age, sex, income, housing, car ownership, and driving license. These seven variables are considered as the main socio-demographic variables to reduce the complexity of the network structure. However, for the variables income, housing status, and car ownership, the household data is used. The chosen network structure is based on the AIC score employing the Tabu search algorithm to learn the network structure of the BN, as implemented in the R package bnlearn (Scutari, 2010; R Core Team and others, 2013).

Variable	Definition [number of categories]	Values
TA	Type of activities of individual [2]	School; Work
Age	Age of individual [8]	<6; 6-12, 12-18; 18-24; 24-32; 32-42; 42-60, >60
Sex	Gender of individual [2]	Male; Female
Inc	Income of household in million IDR [7]	NA; <1, 1-3; 3-5; 5-8; 8-15; >15
HS	Housing status of household [2]	Owned; Rented
CO	Car ownership of household [2]	Yes; No
Licen	License of individual [4]	Motor cycle; Car; Motorcycle and Car; No License

TABLE 2.4: The population of greater Jakarta area

There are two steps in the tabu search in this scenario, without using the whitelist and blacklist G structures for the initial search and with using the whitelist and blacklist for the final search There are 256 searches in the initial search, and 64 searches for the final search. The final structure is obtained by an iterative process after the error for each arc is measured. The arc, which gives the smallest error, is included in the network using the whitelist command and the arc, which gives the highest error, is never included in the network using the blacklist command. In the final search, we found an AIC of -1679334 for the selected BN.

The model structure and the conditional probabilities of the variables can be seen in Figure 2.3. The tables accompanying Figure 2.3 present the conditional probabilities. In this case, we show sex and income variable. There is a directed arc of *NodeTA* and *NodeAge* that goes to *NodeSex*. Therefore, the conditional probability, with an income of 5-8 TA school, and age <6, is 0.43 for females and 0.57 for males. This way of interpretation is the same for other variables, and the total of conditional probability is always equal to one. After we identified the best structure, we generated the data as close as possible to the aggregate census data. In this research, the data is generated for 4 million, 8 million, 16million, to 22 million agents.



			Sex	
Inc	TA	Age	Male	Female
5-8	School	<6	0.43	0.57
5-8	Working	12-18	0.78	0.22
5-8	School	18-24	0.50	0.50
5-8	Working	24-32	0.52	0.48

	Age							
TA	<6	6-12	12-18	18-24	24-32	32-42	42-60	>60
School	0.0696	0.4695	0.3768	0.0699	0.0073	0.0037	0.0030	0.0003
Working	0.0002	0.0013	0.0181	0.1250	0.2484	0.2978	0.2810	0.0280

FIGURE 2.3: Final model structure G

We randomly joined the generated BN population to the remaining 20 variables left in HTS data, based on the available seven variables in both data sets. Complete sets, after both data are joined, will be used for the agent-based model. We obtain a joint data set consisting of 27 variables (Figure 2.4). For the extended target of 16 million agents, this joint operation takes three days and 13 hours. Moreover, for 22 million agents, this joint operation takes more than five days on the servers of the ETH Zurich computer cluster (ETH Zürich, 2016).

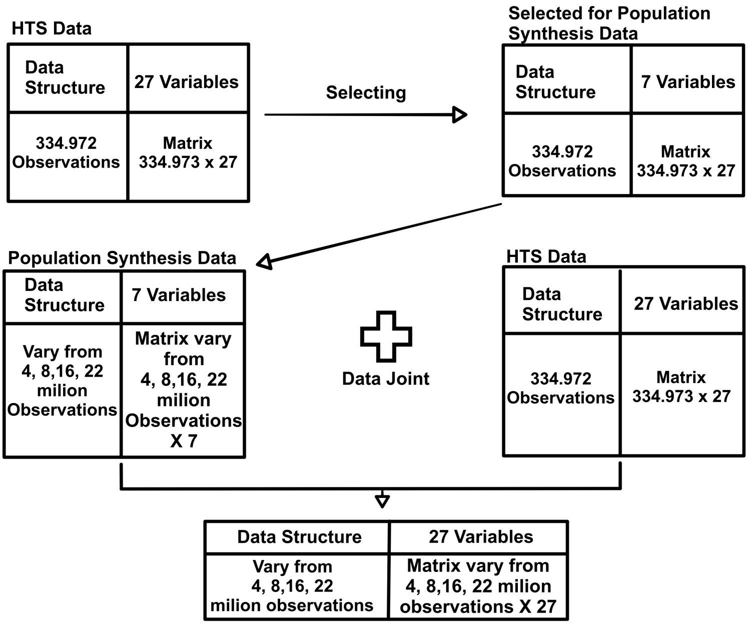


FIGURE 2.4: Population synthesis steps

To visualize the goodness of fit after the BN approach and the joint data, we compared HTS data and the final structure from BN. However, in this case, we only visualize two different types of joint distributions. Figure 2.5 shows the joint distribution of income and age, and Figure 2.6 shows the joint distribution of income and region. The left figure shows the joint distribution data from HTS. Then, the middle figure shows the joint distribution of the data from BN, and the right shows the fit of the joint distribution of both HTS and BN. We considered the region variable, which was not used as a variable in the BN approach as in Figure 2.6, to ensure that a joint data procedure gives a good fit with unselected variables. We found a similar distribution after both HTS and BN data were joined.

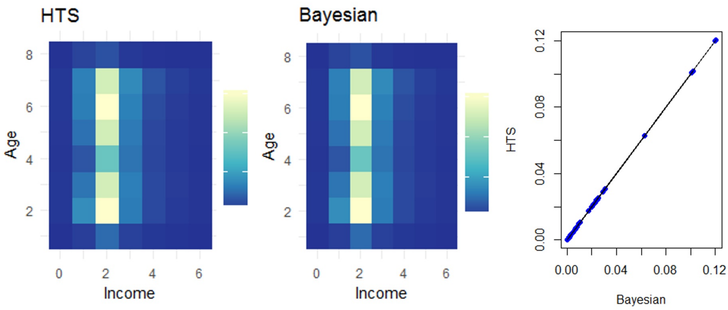


FIGURE 2.5: Joint distribution of income and age

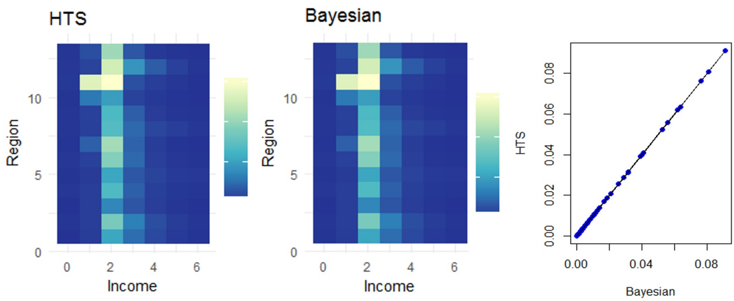


FIGURE 2.6: Joint distribution of income and region

Furthermore, as stated by Sun and Erath (2015); Zhang *et al.* (2017), the BN model can reproduce the distribution of the HTS data. That is because we can only assign to this target value, whether we estimate 4 million or 22 million. It is not based on the control total of aggregate census data in Table 2.3. Therefore, to solve this issue, we employ multilevel IPF with the R-package MultilevelIPF (Müller, 2017) to adjust the artificial sample to the aggregate census data, as presented in Table 2.3.

2.4.2 *GR Multilevel IPF for fitting against census data*

Several multilevel IPF algorithms have been implemented, such as hierarchical iterative proportional fitting (HIPF), iterative proportional updating (IPU), entropy optimization (Ent) and generalized raking (GR). Nevertheless, we use a GR algorithm that has been shown to outperform the other algorithms, which generates weights with minimal squared error and is faster. It matches the data with the aggregate control groups by estimating a weight for each observation in the sample (Müller, 2017). There are two group controls used: region and gender, and region and age.

Furthermore, the weakness of IPF is on non-integer weights (Müller, 2017; Lovelace and Ballas, 2013). Meanwhile, it requires integer weights for generating the target population. Integerisation is the process of converting the decimal weights (related to how many times each agent is replicated) to integer values. Therefore, integerisation is an important step to produce the best fit to the marginal census data. It can be done using weighted random sampling without replacement using the wrswoR package (Müller, 2016), and also using the 'truncate, replicate, sample' (TRS) method (Lovelace and Ballas, 2013). However, we employ random sampling without replacement, which consists of three steps. The first step is removing the decimal values. The second step is calculating the decimal remainders that will be used as a vector of probability weights, and the third step is implementing weighted sampling without replacement. A crank algorithm is used in this operation for faster results (Müller, 2016).

The data used, from BN for fitting using GR, differ in size from 4 million, 8 million, 16 million, to 22 million agents. Less fit result was acquired when we use data from 8 million, 16 million, to 22 million. It can be because the huge size and complex data structure create an error in implementation. Moreover, integerisation has also created an error in replication. Therefore, we used the 4 million agent data set that fit with less than 0.2 % data difference compared to the target census data. The synthetic population data fits after conducting GR multilevel IPF to the target of 20 million marginal census data and the control totals in Table 2.3, which are based on age and gender for each region. The results can be seen in Table 2.5 and Figure 2.7.

Province	Region	Male	Female
Jakarta	South Jakarta	-0.0075	-0.0054
	East Jakarta	-0.0072	-0.0015
	Central Jakarta	0.0155	0.0007
	West Jakarta	0.0123	-0.0040
	North Jakarta	0.0466	-0.0032
Banten	Tangerang City	-0.0176	0.0003
	Tangerang Regency	-0.0250	-0.0222
	South Tangerang City	-0.0356	0.0114
West Java	Depok	0.0476	-0.0223
	Bogor	0.0085	0.0056
	Bogor Regency	0.0057	0.0010
	Bekasi City	0.0440	0.0248
	Bekasi Regency	-0.0186	0.0026

TABLE 2.5: The difference (in %) between population synthesis and census data of region and gender

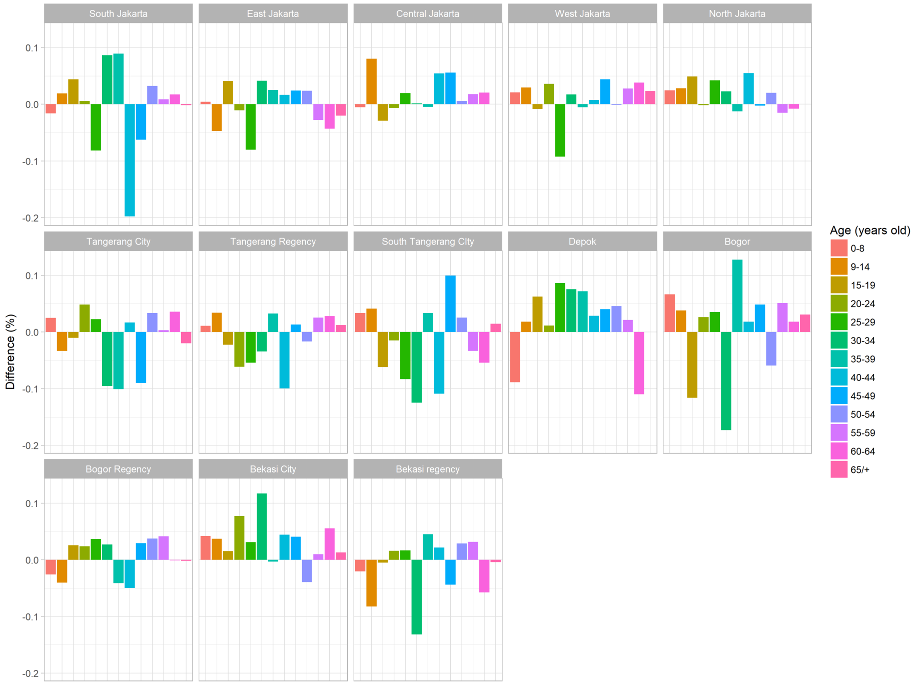


FIGURE 2.7: The difference (in %) between population synthesis and census data of region and age

2.5 IMPUTING ALL ACTIVITIES

In this section, we improved our population synthesis to get all other secondary activities, for example leisure, shop, and other. We used the data from our household travel survey (HTS) survey in 2019, which can be seen in chapter 3 for more detail. For commuting trips, such as home to work (h-w), home to school (h-s), and vice versa, are based on the data from JICA in 2012 (JICA, 2009). In this study, we used the same fully automated and customizable open-source pipeline, which has been used successfully in several places such as Switzerland, Los Angles, San Francisco, and Île-de-Fance region (Hörl and Balac, 2020).

The HTS data consists of activity chain of respondents in a day, including their socio-demographich profile, such as age, gender, household income, car ownership, and mode transport used. It also includes the exact time, when respondents do their activities.

We also have the address, where the respondents living and doing the activities. We used this address to get coordinate locations using google API. However, the address in Indonesia is complicated, instead of only street name, house number, and post code. The address also has the district number (*kecamatan*), sub-district (*kelurahan*) number, citizen number (*rukun warga* (RW)), neighbour number (*rukun tetangga* (RT)). This makes an error in our coordinate, which make some coordinate are outside of greater Jakarta. For those coordinate that outside of Greater Jakarta we deleted.

2.5.1 *Imputing primary and secondary activity locations*

The primary and secondary locations including, home, work, leisure, education, shops, others were retrieved from Open Street Map (OSM). We also assigned home locations alongside the residential roads.

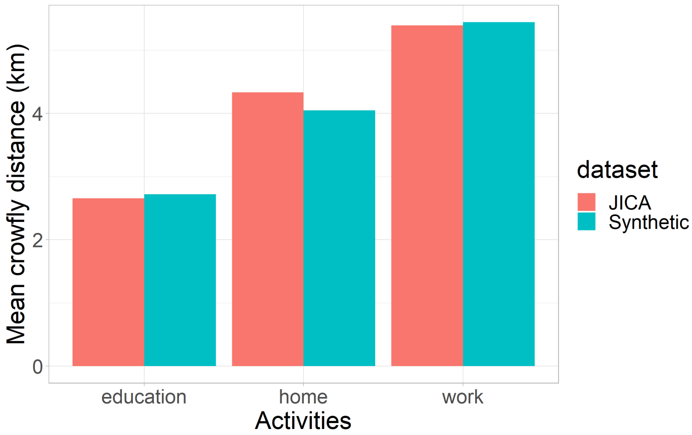
2.5.2 *Pre-processing input data*

The information gathered form HTS data is adapted to reduce the complexity. For mode transport, which HTS had 13 type of modes, are converted to seven type of modes including walk, public transport, car, motorcycle, car ODT, motorcycle ODT, and car passenger.

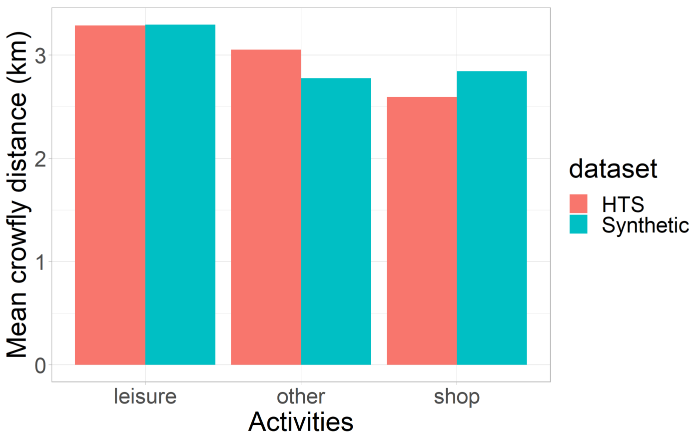
The trips on observations, which are not starting at home and end at home were filtered out. The trips purpose were merged to six categories, such as home, work, education, leisure, shop, other.

2.5.3 *Results after matching activities*

We use age, gender, and employment variables to match with activity chain from HTS survey. To clearly see that our activity are matched, we compare average crowfly distance for different activities, which shows how far the agents perform for different activities. This can be seen in Figure 2.8, and also Figure 2.9 for cumulative distributions of crow-fly distance.

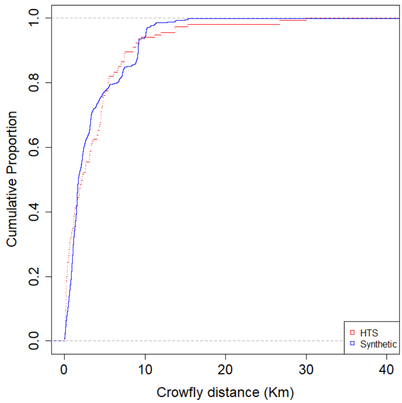


(a) Mandatory activities

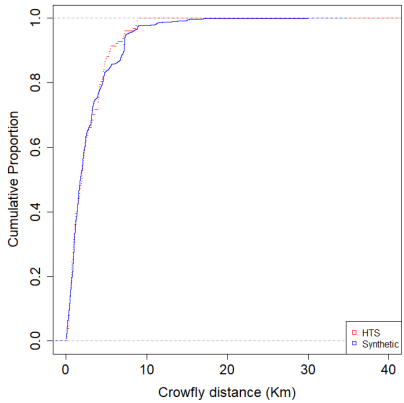


(b) Secondary activities

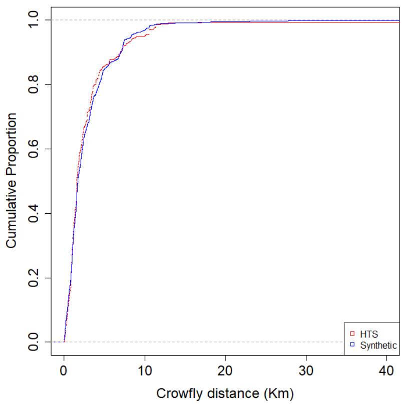
FIGURE 2.8: Comparison of mean crow-fly distance



(a) Leisure activities



(b) Other activities



(c) Shop activities

FIGURE 2.9: Distance cumulative distributions of crow-fly distance

The Figure 2.10 shows the distribution of activity chains in the population synthesis and compares it to the observed distribution obtained from the household travel survey. It shows that activity chain from population synthesis can match very well with the activity chain from household travel survey.

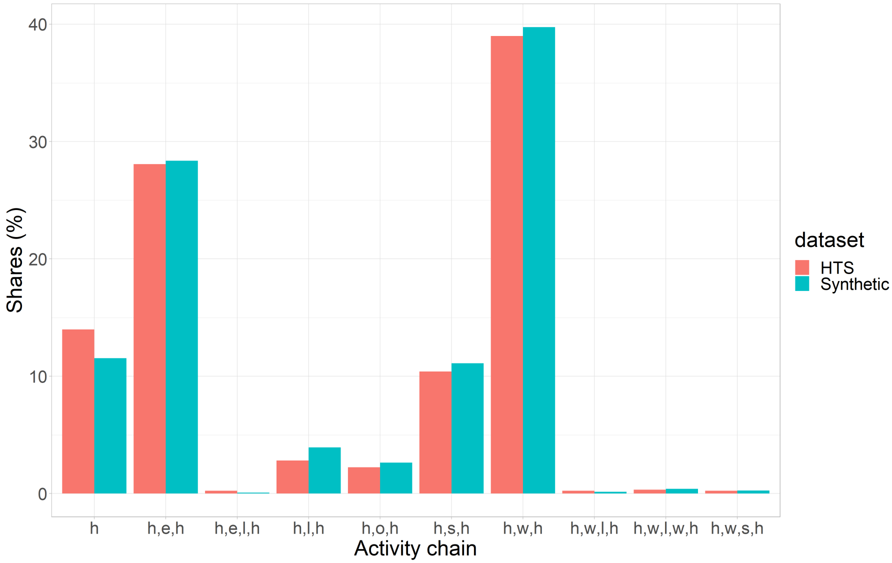


FIGURE 2.10: Activity chain distributions

2.6 DISCUSSION

Population synthesis is an important step when we develop an agent-based model. Data become an important factor in ensuring that the population synthesis gives a precise description of the agents. There are several methods available. Some of methods tried to improve the previous methods, and other methods tried to use new methods, as mentioned in Section 2.1. Each region has different difficulties since each country has different regulations regarding data issues and data availability. Less developed countries may have less accessibility and quality of the data. One of the studies generated a synthetic population without a sample to address the availability of a sample issue (Barthelemy and Toint, 2013). However, it becomes questionable at which level of disaggregation this can give quasi-perfect distribution. On the other hand, it would be a solution for the initial development of an agent-based model.

The size of the data set also influences the processing. As mentioned in Section 4, it takes more than a day to combine the different data sets. It will be an interesting topic to accelerate the combining process for huge data sets. Moreover, a BN alone cannot fit the marginal census data. However, using only GR also fails to make sure that the dependency of each variable is perfectly reproduced. Meanwhile, the combination of BN and GR is a solution to maintain the stability of the dependencies and to ensure a quasi-perfect distribution. Saadi *et al.* (2018) also combined HMM and IPF in the same spirit.

2.7 CONCLUSIONS

In our case, we found that the BN could construct a synthetic population and reproduce the HTS data. The result after data combining between population synthesis and HTS data gives well-fitting distributions. However, BN has some differences vis-à-vis marginal data from the census. Therefore, we need to fit against the marginal distribution of the census data.

The differences are addressed with GR. We made an effort to remain as close as possible to the target census data when we fit with GR. We started at 4 million, 8 million, 16 million, and 22 million agents produced by the BN approach. However, it gave different distribution against census data when we tried with 8 million, 16 million, and 22 million. Therefore, we used 4 million agents from the BN to fit with the 20 million target agents of the marginal census data/control total using GR, which is equal to 20% of the final data. Based on the results described in Section 4, we have three conclusions i.e.::

- Our results confirm that the BN approach can be used to produce large samples with well-fitting distributions, which is useful for any researcher who has a limited sample to start.
- The result from GR can fit the control totals or marginal census data.
- Integrating BN and GR can help researchers to produce data that fit to control totals or marginal census data.
- The pipeline can match the distribution of the chain of activity in the population synthesis to HTS survey.

This synthetic population is used for our further research to develop an agent-based model using Multi-Agent Transport Simulation (MATSim) (Horni *et al.*, 2016). This is the first scenario of an agent-based model for Greater Jakarta (Ilahi *et al.*, 2019a). Several variables from the synthetic population will be used in the agent-based modeling, such as age, gender, income, the coordinate of home and office, activities, license, car ownership. Furthermore, we also integrate secondary activities in our population synthesis by matching them with all activities based on HTS in 2019.

3

TRAVEL BEHAVIOUR IN GREATER JAKARTA

3.1 INTRODUCTION

This chapter discusses the travel behavior in Greater Jakarta. The objectives of this research are to understand the travel behaviors and to explore the influencing factors of mode choice. This research answers several questions for each mode of transportation i.e.: who the users are, when the users use it, the purposes of the users, the speed of different distances and locations.

We have experienced that rapid growth of Information and Communications Technology (ICT), which causes the evolution of emerging transportation modes, is inevitable. The concepts of peer production, also known as a sharing economy in the digital era, were discussed by (Benkler, 2002), and (Pepić, 2018). It connects the service, which is offered by a company, to individuals through the internet. This business model exists not only in the transportation industry but also in many sectors of industry, including hotels, restaurants, ticketing, and e-commerce.

One such alternatives is On-Demand Transport (ODT). ODT connects potential passengers and potential drivers through a smartphone app. Growing ODT systems have become popular in many countries. The conventional taxi industry's passenger volume has decreased since the ODT mode of transportation became available (Lam and Liu, 2017). The lower price and the greater convenience of using a smartphone are the main advantages of ODT compared to conventional taxi service. ODT also becomes a link to connect the last miles of a trip to its final destination, especially for motorcycle (MC) ODT, that can move faster through traffic congestion and can drive on narrow roads. These systems have also reduced the number of unemployed (AngryWorkersWorld, 2019). Driver's income was higher than the minimum wage when this system started; however, it has decreased due to the growing number of drivers (Lam and Liu, 2017). Resistance from conventional taxi drivers exists in many countries (Borowiak and Ji, 2019; Lam and Liu, 2017; Peticca-Harris *et al.*, 2018; Rogers, 2018) for various reasons, one of which is the lack of regulations imposed on ODT (Irawan *et al.*, 2019; Rogers, 2018).

Several studies have investigated ODT, both car-based (Dias *et al.*, 2017; Rayle *et al.*, 2016; Young and Farber, 2019) and MC-based (Irawan *et al.*, 2019; Medeiros *et al.*, 2018), but they are limited to the characteristics of ODT users and the effects of ODT on other modes. In a study by Dias *et al.* (2017) on the socio-demographics of respondents that use car-based ODT, it was found that ODT users tend to not only be young, well-educated, and have a higher income, but they also live in higher-density areas. In another study, Rayle *et al.* (2016) showed that the user characteristics, wait times, and trips served differed between car taxi and car-based ODT. There are few studies on MC-based ODT (see, e.g., Irawan *et al.*, 2019; Medeiros *et al.*, 2018). Irawan *et al.* (2019) showed that MC-based ODT had a positive effect on the use of public transport when the system becomes a feeder to public transport. They also found that MC taxis and MC ODT competed with each other (Irawan *et al.*, 2019; Medeiros *et al.*, 2018). Other studies found the same competition between car taxis and car ODT (Contreras and Paz, 2018; Habib, 2019).

Another alternative mode is Urban Air Mobility (UAM). There is a growing interest in solving urban transportation problems by using air mobility. Nevertheless, UAM might be suitable only for high-income users because the price is much higher than other alternative modes. UAM may eventually become a realistic alternative mode of transportation. Land transport is indeed insufficient to accommodate the demand for mobility in Greater Jakarta. Three-dimensional transport, such as UAM, may be a strategy to address such congestion. Balac *et al.* (2019a) noted that several studies were trying to measure the demand for UAM in urban areas (see, e.g., Balac *et al.*, 2019b; Fu *et al.*, 2019; Garrow *et al.*, 2017). Balac *et al.* (2019b,a) attempted to simulate UAM in the urban transport environment using an agent-based modeling approach based on the potential demand of UAM. Shaheen *et al.* (2018) measured the potential demand of UAM in several cities in the U.S. using SP experiment and attitudinal questions. In a related study, Eker *et al.* (2020b) measured individuals' perceptions regarding the potential benefits of UAM.

The developments of Information and Communications Technology (ICT), which enable the evolution of emerging transportation modes, are inevitable. Peer production concepts, also known as a sharing economy in the digital era, have been discussed by Benkler (2002) and Pepić (2018). This business models connects the service offered by a company to individuals through the internet and currently exists in the transportation industry and many other sectors, including hotels, restaurants, ticketing, and e-commerce.

There are several companies in the ODT industry, such as Uber, Lyft, Grab, and Gojek. Uber and Lyft were launched in San Francisco in 2012.

Uber focused on a black-car limousine service, while Lyft focused on a long-distance intercity carpooling named Zimride in 2007 (Henao, 2017). Uber was the big ODT player in Southeast Asia before Grab took over its business and Uber began selling its shares. This has also happened in other countries (Sothy, 2019). Currently, the local big players are Grab, which started in 2012 in Malaysia, and Gojek, which began in 2010 in Indonesia. Gojek, which was started from only a MC-based ODT in Indonesia, has expanded into other Southeast Asian markets. Gojek is backed by tech giants like Google and Tencent (Russell, 2018).

Moreover, most ODT companies expanded their businesses by providing other services, including transporting goods, and buying and delivering food. In this way, they also helped micro and small businesses to increase their sales (Harsono, 2019). Now that ODT is established, society does not want to reduce its availability. The service is very convenient; people can easily request rides anytime and anywhere and it operates as a door-to-door service with a fixed upfront price. This system is quickly growing as it meets transportation demands when conventional urban transportation modes cannot.

In the future, alternative forms of transportation will continue to emerge, including the development of electric-based or even autonomous vehicles, flying transportation, and the development of the bundling scheme of mobility as a service (MaaS). Several companies, including Airbus, Uber, and Lilium, have been investing in the development of UAM. Airbus tested the flight of UAM in Eastern Oregon Downing Downing (2019). This system may help the congested city and longer-distance travelers to minimize their travel time. The vertical take-off and landing (VTOL) aircraft, which can land and take-off vertically, may reduce land transportation infrastructure growth in the future. However, the UAM will continue to be more expensive than other transportation alternatives, as its operational cost will be higher.

UAM might face several challenges in the future, however. As mentioned by Ahmed *et al.* (2020), the sustainability of UAM involves several aspects to be considered, such as safety, training, infrastructure, environment, logistics, cybersecurity, and the human factor. Reiche *et al.* (2018) and Cohen *et al.* (2020) also discussed aspects related to the development of UAM, such as challenges, infrastructures, technology, public acceptance, and laws and regulations. Al Haddad *et al.* (2020) found that safety was the main concern for the adoptions of UAM. Moreover, Eker *et al.* (2019) found that older persons are relatively more concerned about safety issues related to UAM than younger people.

To achieve our objectives, we conducted a Revealed Preference (RP) survey in Greater Jakarta. Similarly, several studies have previously conducted RP surveys to understand travel behavior (see, for example, Axhausen (1995); Axhausen *et al.* (2002); Dharmowijoyo *et al.* (2015)). For observing the willingness to pay (WTP) from the mode choice experiment, we conducted a Stated Preference (SP) survey, which has been widely used to estimate WTP of mode choice alternatives. We estimated the model using pooled RP and SP data sets, which could give robust estimations overcoming the limitations of both data sets. There are several contributions from this study, including:

- We conducted a state-of-the-art RP and SP survey and presented its methodology with a total of 5,143 respondents, which covers 52,731 observations.
- We gained new insight into travel behavior in Greater Jakarta.

The remainder of this chapter is structured as follows. The second section describes the survey design and data collection. This is followed by the third section, which shows the descriptive statistics of the data in general. The fourth section explains the results obtained from the RP survey. The fifth sections describes the experimental designs and construction choice alternatives.

3.2 SURVEY DESIGN AND DATA COLLECTION

The survey was conducted from April to May 2019 in Greater Jakarta, which includes three provinces: West Java, Jakarta, and Banten, comprising 13 cities. The cities outside Jakarta are called Bodetabek (Bogor, Depok, Tangerang, Bekasi). We administered the survey in three waves: The first from April 1 to 13, 2019, the second from April 18 to 26, 2019, and the third from April 29 to the May 9, 2019. Due to Indonesia's 2019 presidential and parliamentary elections, the survey was paused during April 13-17, 2019. A total of 5,143 respondents were interviewed, some of whom represent complete households. To the best of our knowledge, no previous study conducted in Jakarta has used such a large sample size. After data cleaning, the survey consisted of 3,708 respondents in 952 households, 1,432 individual respondents, and 53,977 valid choice observations. The respondents' home location can be seen in Figure 3.1.

A paper-and-pencil survey was used for this study, and most of the respondents were willing to fill in the survey form with guidance from

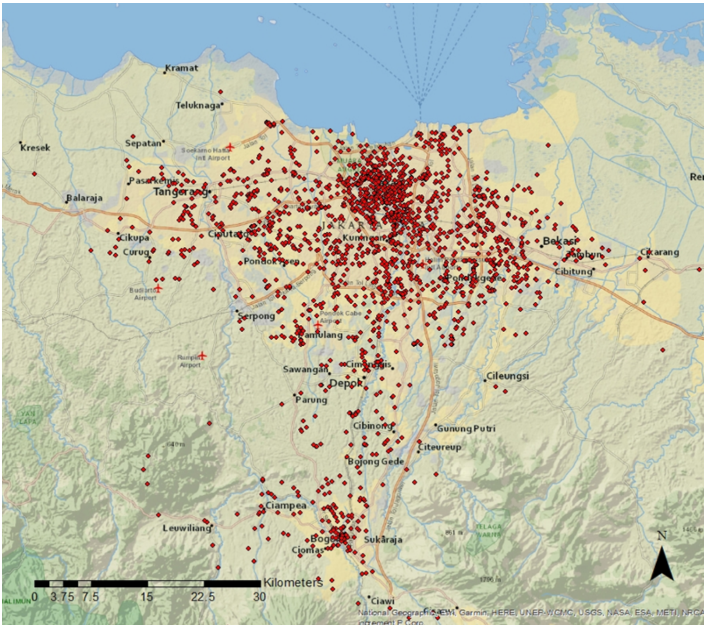


FIGURE 3.1: Home location of respondents in the study area

the surveyor. The response rate was 50% in the pre-test survey and rose to more than 80% in the main survey, as can be seen in Table 1. This response rate was considerably higher than studies in other countries, for example, Axhausen (2008) for the Swiss experiences. The response rate was high because we approached the informal subdivisions in Indonesia: local communities "Rukun Warga" (RW) and local households groups "Rukun Tetangga" (RT). Each RW consists of approximately five RT, and each RT contains between 30 and 50 household groups. This approach is used to gain respondents' trust by asking permission of the heads of the RW and RT before conducting the survey. The heads of approximately 33 RT were approached, which encompasses about 1,000 households.

Additionally, most of the surveyors lived in the study area, which made it easier for them to approach the respondents directly. A similar approach had also been used in Bandung, as briefly explained by Dharmowijoyo *et al.* (2015). To know the effort required by the respondents to answer a questionnaire, we estimated the response burden. Table 3.2 shows that the response burden for socio-demographic questions was 177 points, 737 points for the RP travel diary survey and 450 for the SP survey. The method

Survey	Location	Response rate
Pre-test	Jakarta	50%
Main survey	Jakarta	87.5%
	Bodetabek	98%

Table 3.1: Response rate of the survey

to calculate the response burden was by assigning the points to each question types. The scheme is described in Axhausen and Weis (2010) and updated in Schmid and Axhausen (2019). The primary response burden is the address, especially for the RP survey, as respondents are usually not aware of the street numbers in their neighborhood, RT, or RW of their trip destinations. Therefore, the response burden increases with the number of trips. However, the response burden in our survey was relatively low.

Question type	Observation	Response burden
Socio-demographic	5,143	177
RP Survey	37,042	737
SP Survey	20,482	450

Table 3.2: Response burden of the survey

The survey included the usual socio-demographic questions, including age, gender, income, car ownership, and the primary mode of transport. The RP travel diary of three working days consisted of detailed information on the respondent's departure time, transport mode, trip destination, and trip purpose. The survey design follows the Mobidrive protocol that was well-designed in Axhausen *et al.* (2002). The last element in the survey was a mode choice SP experiment. Respondents were asked a couple of preliminary questions before completing the SP experiment, i.e., whether their trips were in Jakarta or not, whether they were a driver or non-driver, and their travel distance. These initial questions reduced the complexity of the choice experiment. For example, in the choice experiment, private vehicle options would not be available for non-drivers. The walking mode option did not exist for trips longer than 1.5 km, and congestion/tolls price variables would not be present for the trips outside Jakarta. The survey included emerging transportation options: ODT MC (Gojek, Grab), ODT car (GoCar, Grab), and UAM. Since the respondents might not know about UAM and might never have seen it before, the surveyor gave a brief explanation of UAM using figures to illustrate its concept. The questions are summarized in Table 2. Public Transport (PT) refers to Bus, Bus Rapid Transit (BRT), train, and angkot (Microbus). An angkot is a microbus with a capacity for a maximum of 12 passengers (Cervero, 1991; Ilahi *et al.*, 2015).

Socio-demographics	Travel diary	Choice alternatives
Age	Destination	Walk
Gender	Mode transport	Car
Income	Departing time	Motorcycle
Expenditures	Arrival time	Public transport
Address	Address	Car ODT
Number of households	Trip distance	Motorcycle ODT
Vehicle ownership	Transport cost	Car Taxi
License	Frequency activity	Motorcycle Taxi
Access to private vehicles	Type of activity	UAM
Main mode		
Education		
Occupation		
Dwelling		
Working hour		

TABLE 3.3: The survey questions

3.3 DESCRIPTIVE ANALYSIS

Table 3.4 sshows the socio-demographic characteristics of the respondents. The share of male respondents in the sample was slightly higher than in the census (57.30%). About 46.90% of the respondents were younger than 34 years old, which was slightly higher than in the census. 31.70% of the sample had a university degree. More than 90% of respondents live in a single-family house and own the house; this is expected as the apartment share of the housing market is less than 2% (Yudis, 2019). However, the number of respondents who are homeowners overrepresented, as other sources suggest that it should be around 47.85% (BPS-Statistics, 2016a). House prices in Indonesia are relatively affordable, and thus ownership does not necessarily suggest that owners have high incomes. Furthermore, a house can be inherited from previous generations, depending on the size and location. The most likely homeownership arrangement is that the owners have a mortgage for the house from a bank based on the husband and wife's joint income.

Jakarta City tends to have high rise buildings in the center and then low-rise buildings towards the city's outskirts, making the city more spread out and expensive to maintain or invest in infrastructure. Around 41.60% of respondents drove in Jakarta, and 28.60% drove in the urban agglomeration. The shares of the main modes of transportation, or the most frequently used, can be seen in Table 3. The number of ODT or ride-sourcing users was substantial. We found that MC had the highest share (54.30%), followed by car (15.30%), MC ODT (10.90%), and public transport (bus, BRT, commuter rail, microbus) (9.9%).

TABLE 3.4: Descriptive statistics of the survey respondents

Variable	**Sample (%)**	**Census (%)**
Male	57.30	50.70
Female	42.70	49.30
Age categories		
Younger than 24 years old	46.90	44.09
Aged 24-29 years old	11.30	9.19
Aged 29-34 years old	6.30	9.10
Aged 34-39 years old	8.30	8.44
Aged 39-44 years old	8.80	7.39
Aged 44-49 years old	9.20	6.20
Aged 49-54 years old	5.30	5.01
Older than 54 years old	3.90	10.54
University degree	31.70	-
Owned house	92.40	47.85
Landed house	97.20	-
Has access to car	25.60	-
Has access to motorcycle	67.90	-

(To be continued)

Variable	Sample (%)	Census (%)
Driving license		
Car	5.60	-
Motorcycle	41.00	-
Car and motorcycle	23.40	-
No license	30.00	-
Working hour		
Full time	32.50	-
Half-time (30 hours)	11.60	-
Half-time (20 hours)	13.40	-
Student	29.20	-
Non worker	13.20	-
Saving (%)		
0-25	21.10	-
25-50	26.10	-
50-75	38.10	-
75-100	14.70	-
Transport cost (%)		
0-25	80.50	-
25-50	16.30	-
50-75	0.80	-
75-100	2.40	-
Type of respondents (%)		
Driver in Jakarta	41.60	-
Driver in agglomeration cities	28.60	-
Non-driver in Jakarta	11.70	-
Non-driver in agglomeration cities	18.20	-
Main mode of transport (%)		
Walk	6.70	-
Bike	0.70	-

(To be continued)

Variable	Sample (%)	Census (%)
Bus	0.60	-
BRT	1.20	-
Commuter rail	3.70	-
Microbus (angkot)	4.40	-
Car	15.30	-
Motorcycle	54.30	-
Car taxi	0.10	-
Car ODT	1.10	-
Motorcycle taxi	1.00	-
Motorcycle ODT	10.90	-

3.4 REVEALED PREFERENCE

3.4.1 *Mode choice by socio-demographics, trip purpose, and distance*

The results, as can be seen in Table 3.5, show that the mode share was different from when JICA conducted its travel diary surveys (JICA, 2009) in 2009. This was mainly due to the spread of information and communications technology (ICT) and the arrival of ODT in Greater Jakarta. In general, those who are younger than 24 dominated the use of all modes, as the number of those persons in Indonesia was dominant. Around 66.90% of the motorcycle ODT users and 34% of the car ODT users were less than 24. This number was equal to 6.5 % of all trips. The majority of car and motorcycle ODT users were female. Similar findings were reported previously in Canada, where most of the ride-hailing users were young and female (Young and Farber, 2019).

The users of PT (bus, BRT, and train) usually do not have access to a car, but they are more likely to have access to a MC. Angkot users, on the other hand, have access neither a car nor a MC. Less than 10% of the respondents are more likely to take ODT or conventional taxis. Cars were more likely to be used by full-time workers and university degree holders. Table 3.6 shows that MC was more likely to be used by persons with a monthly income less than IDR 18 million (1,250 USD)[1], and the car is more likely to be used by those with an income higher than IDR 18 million.

The mode choice of the respondents depends on the trip's purpose. Users preferred private vehicles for all trip purposes. The results show that the mode choice of public transport increases for a return home and work trip. While mode choice of ODT increases for education and leisure trip purposes. Walking was more frequently used for work, daily shopping, and religious activities. Respondents were more likely to use MC, followed by car, walk, and MC ODT for morning, afternoon, and evening trips. However, respondents were more likely to walk, followed by using MC at midday. During the night, respondents preferred to use MC, followed by a car and by walk. The chi-square test shown in Table 3.5 and Table 3.6 indicates a significant relationship between the chosen mode of transportation and the socio-demographic and trip purpose attributes of the respondents.

1 USD 1 is equal to IDR 14,400

Attributes	Walk	Bike	Bus	BRT	Train	Angkot	Car	Car Taxi	Car ODT	MC	MC Taxi	MC ODT
Row percentage	18.80	0.50	0.60	0.90	3.00	3.80	13.50	0.10	1.00	47.30	1.10	9.50
Male	59.42	59.89	64.56	34.78	74.73	29.92	65.52	54.55	18.38	66.48	19.11	34.60
Age(years)												
< 24	32.34	58.76	55.34	61.18	41.37	38.05	26.40	50.00	33.70	50.01	34.49	66.69
24-29	17.32	3.39	17.96	15.53	32.12	9.55	8.42	20.45	4.74	12.52	8.68	6.65
29-34	10.40	5.65	6.31	5.59	9.96	7.99	5.60	0.00	1.95	7.49	5.46	3.00
34-39	11.71	10.17	3.40	7.45	5.69	16.9	8.90	2.27	4.46	8.20	11.41	6.28
39-44	9.90	0.00	2.43	5.59	5.96	11.24	15.70	0.00	9.75	7.49	16.38	5.20
44-49	8.49	10.73	8.74	2.48	3.29	8.84	16.84	13.64	13.37	7.78	13.90	4.65
49-54	4.11	6.78	2.91	2.17	1.07	3.54	11.90	13.64	15.32	4.15	5.96	4.31
> 54	5.74	4.52	2.91	0.00	0.53	3.89	6.24	0.00	16.71	2.35	3.72	3.23
University degree	33.18	9.04	43.20	43.48	28.38	11.17	57.88	36.36	37.05	30.67	8.44	25.66
Owned of house	87.79	100.00	97.57	84.16	89.23	95.19	97.58	86.36	96.10	92.15	82.38	91.98
Has access to Car	15.32	3.39	21.36	20.19	11.74	5.16	85.10	6.82	29.25	16.82	7.69	14.24
Has Access to Motorcycle	61.21	18.64	66.02	53.73	85.50	26.10	55.20	70.45	45.13	87.95	22.33	33.60
Full-time	30.15	20.90	38.83	41.30	66.64	13.58	50.50	22.73	20.89	32.74	12.90	20.30
Half-time (30 hours)	21.22	12.99	11.17	13.04	9.61	7.64	13.38	27.27	5.57	12.73	4.96	7.62
Half-time (20 hours)	17.75	6.78	19.42	8.07	12.54	23.34	11.60	13.64	15.04	15.27	8.68	9.45
Student	11.77	44.63	22.82	37.27	9.88	19.80	18.92	31.82	24.51	29.73	24.32	50.39
Unemployed	19.12	14.69	7.77	0.31	1.33	35.64	5.60	4.55	33.98	9.53	49.13	12.25
Driving license												
Car	4.62	0.00	2.91	9.32	1.51	0.21	23.96	0.00	10.58	2.00	1.99	5.60
Motorcycle	42.74	16.38	20.87	23.60	66.46	18.10	7.14	29.55	12.53	60.61	7.69	41.00
Car and Motorcycle	9.41	3.95	36.41	19.88	10.05	2.76	60.98	29.55	15.60	20.91	2.98	23.40
No license	43.24	79.66	39.81	47.20	21.98	78.93	7.92	40.91	61.28	16.48	87.34	30.00
Start time												
Morning (4-9am)	10.72	0.59	0.69	1.06	3.47	4.52	13.81	0.14	0.63	53.56	1.26	9.57
Midday (9am-2pm)	49.15	0.51	0.25	0.30	0.73	4.04	7.48	0.09	1.31	26.48	1.92	7.74
Afternoon (2pm-7pm)	8.75	0.31	0.59	1.06	3.94	3.15	15.67	0.12	1.08	54.54	0.44	10.35
Evening (7pm-12am)	9.21	0.68	0.75	0.75	4.47	1.42	24.32	0.07	1.56	45.53	0.07	11.18
Night (12am-4am)	16.67	0.00	0.00	0.00	0.00	0.00	27.08	0.00	0.00	54.17	0.00	2.08

Chi-squares tests (X^2) are all significant ($p - value < 0.001$)

TABLE 3.5: Socio-demographics of the respondent by chosen mode

Attributes	Walk	Bike	Bus	BRT	Train	Angkot	Car	Car Taxi	Car ODT	MC	MC Taxi	MC ODT
Income (IDR Million)												
< 1	18.13	0.16	0.00	0.90	2.71	7.38	4.68	0.57	0.08	58.90	1.56	4.92
1-3	29.97	1.04	0.86	1.43	2.32	9.56	2.94	0.27	0.48	44.34	0.83	5.97
3-5	27.15	0.01	0.59	0.47	5.59	4.67	4.75	0.10	0.31	49.15	1.17	6.04
5-8	20.97	0.44	0.53	1.21	3.96	4.22	10.43	0.05	0.20	50.36	1.66	5.98
8-12	16.07	0.74	0.68	0.25	1.16	2.09	17.50	0.00	0.91	46.62	0.42	13.56
12-15	5.96	0.81	0.23	1.12	1.41	1.27	22.42	0.00	1.41	49.19	0.40	15.77
15-18	5.08	0.95	0.40	0.95	0.00	0.71	26.19	0.00	2.22	47.62	0.95	14.92
18-21	5.69	0.00	0.71	1.00	0.85	0.43	28.31	0.85	5.19	34.85	1.21	20.91
21-25	2.54	0.00	1.14	0.76	0.38	0.76	42.39	0.00	1.52	37.31	0.51	12.69
25-28	4.43	0.95	0.48	1.11	0.79	0.48	35.79	0.16	4.59	31.67	1.82	17.74
>28	4.90	1.18	0.00	1.76	0.20	0.39	42.55	0.00	5.29	30.39	0.39	12.94
Trip purpose												
Go home	9.99	0.52	0.58	0.88	3.44	4.34	14.84	0.12	1.09	52.29	1.18	10.73
Drop off	18.39	0.15	0.73	0.00	0.15	9.64	13.14	0.00	2.77	50.22	0.73	4.09
Work	18.54	0.16	0.76	0.96	4.68	2.44	17.07	0.12	0.35	48.34	0.54	6.04
Education	8.00	0.84	0.49	1.54	1.23	3.44	9.33	0.18	0.62	54.82	0.99	18.52
Daily shopping	32.89	1.39	0.28	0.00	0.42	9.99	4.16	0.21	1.80	33.59	6.32	8.95
Special shopping	20.46	0.53	0.00	0.00	1.06	9.17	16.58	0.18	3.35	39.68	1.23	7.76
Leisure	17.56	0.47	0.16	0.79	1.27	6.33	17.72	0.00	6.65	36.08	1.42	11.55
Religion	88.55	0.51	0.00	0.00	0.30	0.30	1.82	0.00	0.61	6.38	0.00	1.52
Other purpose	78.82	0.35	0.06	0.40	0.40	0.17	2.94	0.00	0.40	14.95	0.06	1.44

Chi-squares tests (X^2) are all significant ($p - value < 0.001$)
1 USD is equal to 14,085 IDR on 31th of July 2019

TABLE 3.6: Income, trip distance and trip purpose of the respondent by chosen mode

As depicted Figure 3.2, for short distances (0 to 1.5 km), the share of Non-Motorized Transport (NMT) was the highest; however, when the distance increased, the mode share of private vehicles also increased. NMT refers to walk and bike alternatives. Interestingly, the share of the conventional taxi was lower than ODT. Certainly, ODT is more convenient than a conventional taxi as people may have a price estimate upfront and it has a lower price and faster service. ODT was preferred, especially when compared to a conventional motorcycle taxi, where people have to negotiate the price first with the driver. MC ODT was used not only for short distances, but also for long distances. The share of MC ODT was higher than 12% when the distance was farther than 5 km.

The public transport share was more than double when the distance was more than 30 km. The reason is that the access to private vehicles is insufficient for the people living far from their activities' locations. The people living farther from activity locations have a relatively lower income level, and they cannot afford to live in places that close to CBD. Moreover, the price of public transport is lower than other modes due to the price subsidies from the government, especially for BRT and Train.

Furthermore, Figure 3.3 shows mode shares by travel time, in which the shorter the travel time, the higher the shares of NMT and ODT. However, private vehicle and public transport are reverse, and ODT is more preferred than the conventional taxi in any cases.

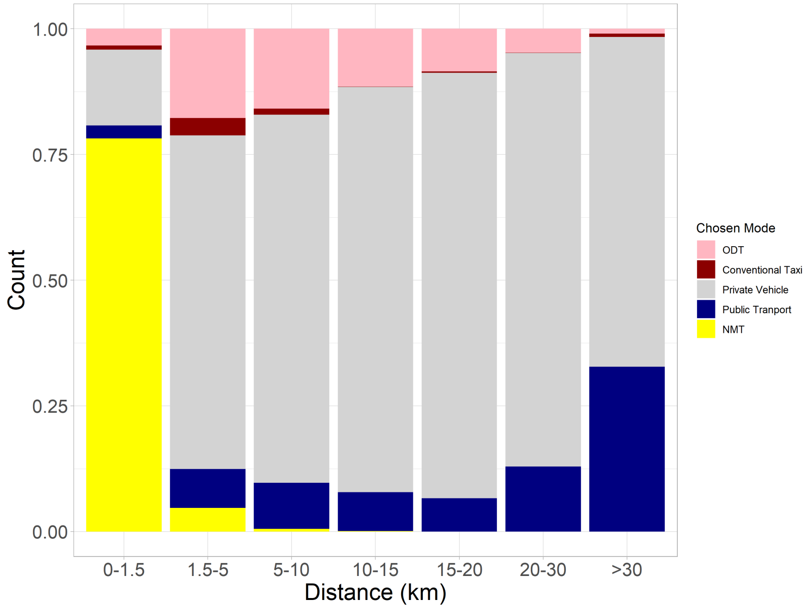


Figure 3.2: Travel distance by mode

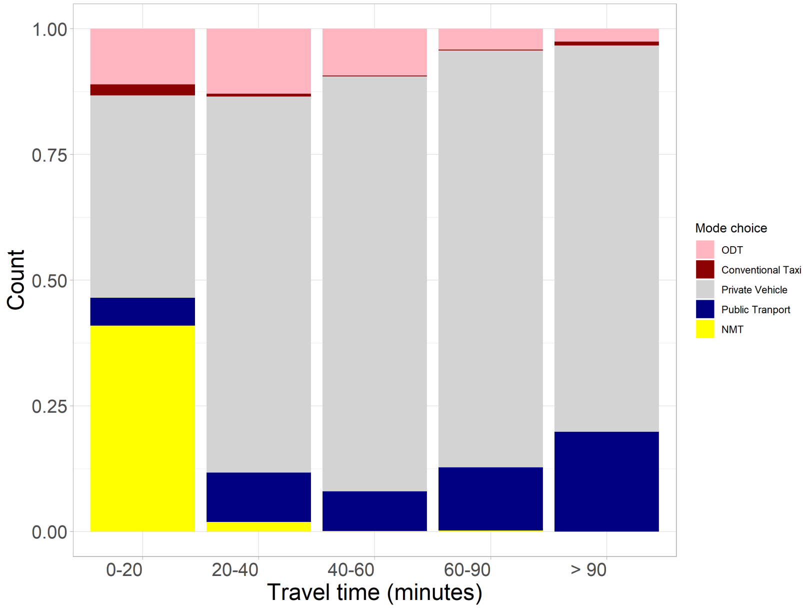


FIGURE 3.3: Travel time by mode

3.4.2 *Cost structure of the transport modes*

Figure 3 shows the price structure of the different modes of transportation, presented in a log-time and log-cost scale format. These figures are based on the travel cost and travel distance for each trip of the respondents. Blue represents car-based modes, orange represents MC-based modes, and green represents the public transport modes. It shows that public transport modes have a lower cost in general. When the distance increases, the price of public transport is lower than car-based or MC-based modes. Compared to other public transport modes, the lower the marginal price of traveling by train decreases by distance. For car-based modes, it can be observed that traveling by car is inexpensive for short-distance trips, but it becomes the most expensive mode for longer trips. Car ODT prices are higher than those of conventional car taxis as reported by the respondents. Regarding the price of MC-based modes, we can see that the price of MC taxis is the highest, followed by MC ODT and MC.

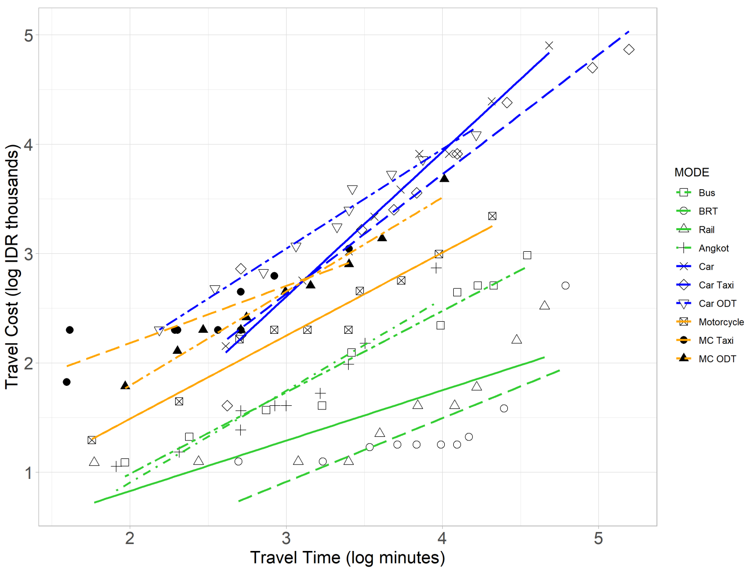


FIGURE 3.4: Travel time (log) and travel cost (log) by mode

3.4.3 *Trip chains*

Each trip had a particular purpose: work (w), to go home (h), education (e), errand (er), leisure (l), and daily and special shopping (s). Trips for religion and other purposes are combined into leisure purposes, and drop-offs are considered to be errands. Table 3.7 shows the 10 most frequent trip chains that make up 94.52% of all activity chains. The most frequent chain is home-work-home (h-w-h), followed by home-education-home (h-e-h). The result shows that these simple mandatory activity chains have a share of 55.81%, which is considerably higher than the results in other countries, Schlich *et al.* (2004) report this share to be less than 25% in the case of Mobidrive and Uppsala. Following these two chains, the home-shopping-home (h-s-h) trip chain is the third most common with 9.52%.

We also divided the trip chains based on gender and household monthly income categories. Males had a higher share for most trip chains; however, females had a higher share of home-shopping-home, home-leisure-home, and home-errand-home trip chains. It showed that transport for shopping, leisure, and errand purposes was more likely to be taken by females because they are less likely to be employed. The income categories also influenced the trip chains of respondents. The income was divided into three-levels: low income (less than IDR 5 million), medium-income (between IDR 5 and 15 million), and high income (more than IDR 15 million); medium-income respondents had higher shares for all trip chains, but lower-income respondents had a higher share for h-w-l-w-h and h-er-h trip chains.

Trip chains	Male	Female	Low Income	Medium Income	High Income	Total
h-w-h	67.60	32.40	32.37	51.14	16.49	39.79
h-e-h	54,56	45.44	25.23	52.91	21.87	26.02
h-s-h	11.00	89.00	31.44	53.11	15.45	9.52
h-w-l-w-h	77.92	22.08	59.05	40.44	0.51	7.69
h-l-h	35.23	64.77	32.72	41.72	25.56	4.91
h-w-l-h	84.62	15.38	45.16	53.11	1.74	2.64
h-er-h	33.63	66.37	47.66	39.77	12.57	2.23
h-e-l-h	58.82	41.18	41.18	47.06	11.76	0.66
h-s-w-h	96.74	3.26	26.09	73.91	0.00	0.60
h-w-s-h	55.07	44.93	37.68	47.82	0.51	0.45
Row percentage	57.20	42.80	33.30	50.30	16.40	94.52

Chi-squares tests (X^2) are all significant ($p - value < 0.001$)
1 USD is equal to 14,085 IDR on 31th of July 2019

TABLE 3.7: Trip chains by gender and income level

3.4.4 *Speed by distance, mode, and region*

The average speed of modes increased with the traveled distance, as can be seen in Figure 3.5. We calculated the speed based on respondent's reported travel time and distance. The calculation is based on the aggregate of all stages and modes used. For short distances, the speed was slower due to the frequent walking and bike use. Figure 5 presents the speed for different modes in Jakarta and its agglomeration. The speed of the transport modes was different for each mode and region. The average speed of bike, bus, BRT, angkot, MC Taxi was higher in the agglomeration; however, the average speeds of the other modes were quite similar. In general, the agglomeration had a higher speed than Jakarta due to lower traffic congestion.

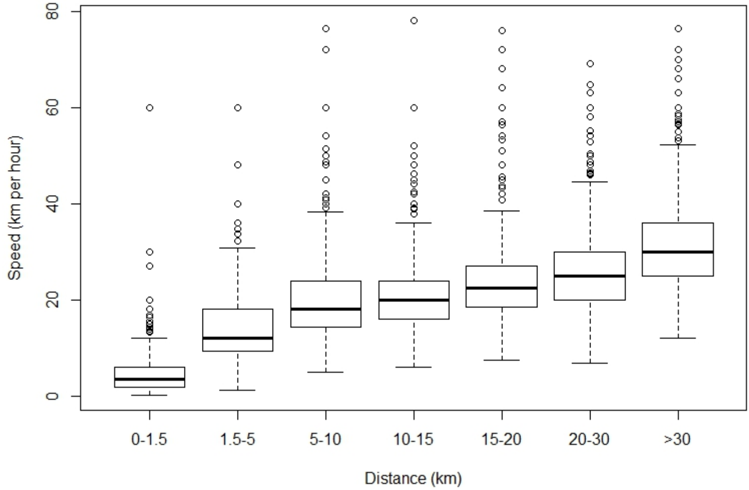


FIGURE 3.5: Average mode speed by distance

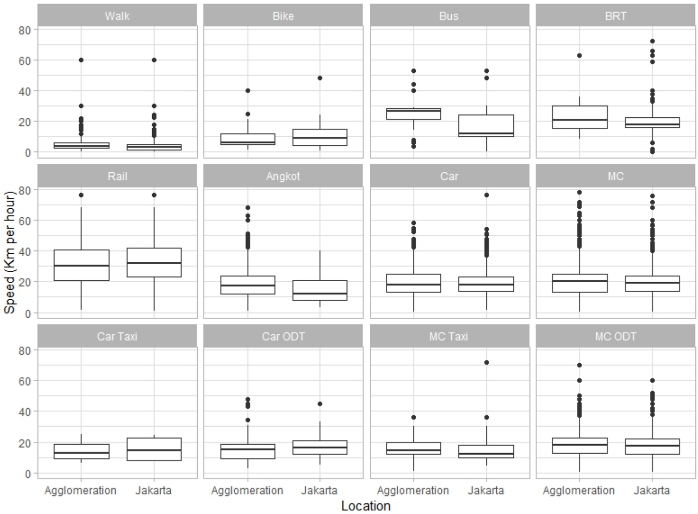


FIGURE 3.6: Speed distribution by mode and region

3.5 CONSTRUCTING CHOICE ALTERNATIVES

3.5.1 *SP data set: Experimental Designs*

We constructed stated choice experimental designs with a D-efficient design using Ngene (ChoiceMetrics, 2014). All the respondents of the RP survey, equaling 5,143 respondents, were given SP surveys. The mode choice experiment in Greater Jakarta was categorized by travel distance to the place of their daily activities, driver or non-driver, traveling inside or outside of Jakarta. The respondent received preliminary questions about these categories. The mode alternatives and variables were based on the respondent's answers to the preliminary questions. Each respondent received four-choice experiments. In total, there are 20,064 observations.

The types of experiments are shown in Table 3.8. The congestion/toll charging attribute was only available for the respondent who travelled within or to Jakarta. There are nine different modes, including walking, PT, car, MC, car Taxi, MC Taxi, MC ODT, car ODT, and UAM. Walking was only available for a distance less than 1.5 km and applicable for drivers or non-drivers. The car and MC were always available in each distance interval, but not available for non-drivers. To reduce the complexity of the choice alternatives, we assigned ODT and conventional taxi at random, meaning that for some respondents, "Taxi" appeared, and for some respondents, "ODT" appeared as a choice option. For the access time, we assume a range of 5 minutes, 10 minutes, and 15 minutes to get from the station or shelter to vertiports. The detail of the attributes can be seen in Table 3.9. We ensured that the travel time and travel cost offered produced VoT values within the range found in the paper by Belgiawan *et al.* (2019b). For car-based modes, the car VoT was the highest in the survey design, followed by conventional car taxis and car ODT. Then, for MC-based modes, the MC VoT was the highest, followed by MC taxis and MC ODT. The VoT offered for UAM was the highest compared to other modes because UAM has the fastest travel time and the highest cost. The assumptions of the cost per km of UAM in this scenario varied between USD 0.69-2.08[2]. However, if we account for the value on purchasing power parity (PPP)[3] in 2019 (OECD, 2019), the cost is around 2.1-6 USD/km. Uber Air

2 USD 1 is equal to IDR 14,400
3 USD 1 is equal to IDR 4,753

expects to have a cost of around 5.73 USD per mile/3.5 USD/km and hopes to reduce the price to 1.15 USD/km and even to 0.27 USD/km in the long-run (Dickey, 2020). Neighboring Singapore expects the cost to be around 3.75 USD per mile/2.3 USD/km or double the price of ground car raid-hailing (TheJapanTimes, 2019). Therefore, we tried to cover that range of costs in our scenarios.

Mode	0-1.5 km	1.5-5 km	5-15 km	15-25 km	>25 km	Driver	Non driver
Walk	True	False	False	False	False	True	True
PT	True	True	True	True	True	True	True
Car	True	True	True	True	True	True	False
MC	True	True	True	True	True	True	False
Car Taxi	Random	Random	Random	Random	Random	Random	Random
Car ODT	Random	Random	Random	Random	Random	Random	Random
MC Taxi	Random	Random	Random	Random	Random	Random	Random
MC ODT	Random	Random	Random	Random	Random	Random	Random
UAM	False	False	True	True	True	True	True

Table 3.8: Availability mode by distance and driving ability

Attributes	PT	Car	MC	Car Taxi	MC Taxi	Car ODT	MC ODT	Walk	UAM	
Travel cost (Thousand IDR)										
0-1.5 km	3;6;8	6;8;10	2;4;6	12;14;16	5;10;15	7;12;15	6;10;12	-	-	
1.5-5 km	6;12;17	14;18;20	9;10;13	25;30;40	15;20;25	14;25;35	15;18;22	-	-	
5-15 km	9;18;27	22;28;32	13;20;25	58;90;110	30;35;45	55;80;100	25;35;40	-	60;100;150	
15-25 km	13;30;55	35;60;75	20;30;45	75;120;160	50;65;80	72;110;145	45;55;68	-	150;200;250	
>25 km	20;40;65	62;90;115	30;40;50	110;170;250	72;90;120	105;165;220	65;80;96	-	250;300;350	
Travel time (minute)										
0-1.5 km	5;10;16	6;12;15	4;6;8	6;10;15	6;7;8	6;8;10	4;6;8	30;50;70	-	
1.5-5 km	10;20;30	10;20;30	8;15;25	10;20;25	9;15;20	10;20;30	10;15;25	-	-	
5-15 km	15;30;45	15;30;45	15;25;35	25;40;55	15;25;32	25;40;55	15;25;32	-	8;9;10	
15-25 km	30;45;60	37;60;70	25;40;50	35;55;70	25;40;50	35;55;70	25;40;50	-	10;12;15	
>25 km	35;60;90	52;90;120	27;50;70	52;75;100	35;60;70	105;165;220	35;50;70	-	13;17;23	
Transfers (minute)	0;1	-	-	-	-	-	-	-	-	-
Waiting time (minute)	5;15;30	-	-	5;10;20	5;10;20	5;10;20	5;10;20	10;15;25		
Toll/congestion charging (Thousand IDR)	-	10;15;25	5;10;15	10;15;25	5;10;15	10;15;25	5;10;15	-	-	
Accessl time (minute)	5;10;15	-	-	-	-	-	-	-	5;10;15	

1 USD is equal to 14.400 IDR on 25th of May 2019

Table 3.9: Attributes of each mode and category

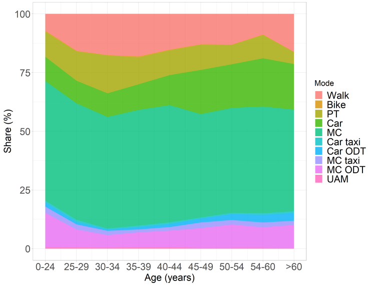
3.5.2 *RP data set: Non-Chosen Choice Alternatives*

The non-chosen transport alternatives for each trip were constructed based on information from google API (Google, 2019). We collected the coordinates of each trip origin and destination based on geocoding Google API. Then we collected the information of non-chosen alternatives: transit, driving, and walking based on Google API direction (Google, 2019). We report the travel time information of each non-chosen alternative mode available on the exact departure time reported in the survey. For bike and MC, however, there was no Google information available for travel time and limited research regarding the speed of these modes of transport are in urban settings. We thus assume that a MC's speed is 3.3 km/h faster than that of a car (Walton and Buchanan, 2012). For a bike, the speed depends on the age of the respondents. The speed of an older person was 10 km/h, while a younger person's speed was 15 km/h (City of Copenhagen, 2013; Woodcock *et al.*, 2018). We assumed each mode's travel cost based on the travel prices that exist in Greater Jakarta.

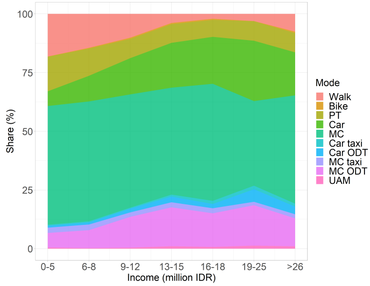
The detail of the assumptions can be seen in Table 9. There were no costs related to walking or biking. The base in the parameter travel cost indicates the travel cost when the respondent first begins to use the mode. We collected waiting time, transfer, and walking time for transit based on the Google API. There was no specific API for each different transit mode. The mode was not always available. For example, if the respondents did not have access to cars and MC, those modes would not be available as a non-chosen alternative. The waiting time for an angkot, ODT, and a conventional taxi was five minutes. The access walking time of angkot or microbus was five minutes, and transfers only occur when the trip was longer than 10 km.

3.5.3 *Description of the pooled SP and RP data set*

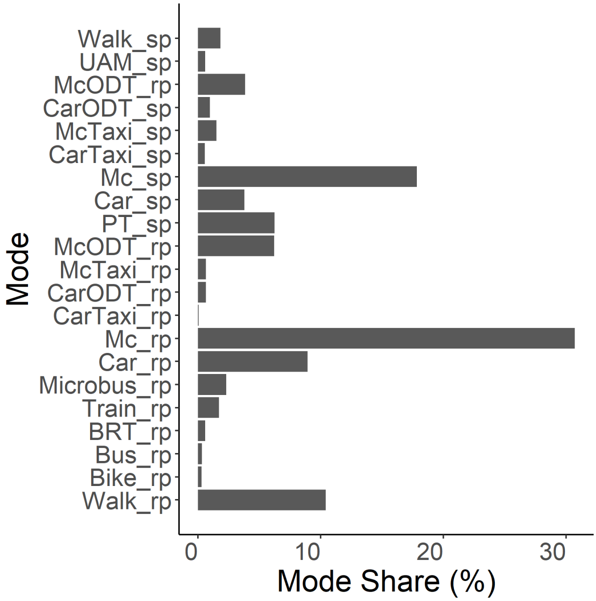
Our data set contains 52,731 observations, excluding microbus alternative. The share of MC choices from the SP and RP data set is the highest. As shown in Figure 6, the share of choice alternatives varies by age group and income group.



(a) Share by age group



(b) Share by income group



(c) Share on SP and RP data set

FIGURE 3.7: Mode shares of data set

Mode	Travel time (minutes)	Travel cost (thousand IDR/km)	Waiting time (minute)	Transfer	Walking time transit (minute)	β ODT	Availability
Walk	API Walking	-	-	-	-	-	always
Bike	$\frac{APICarDistance}{SpeedBike}$	-	-	-	-	-	always
Car	CarAPI	2.95	-	-	-	-	Access Car
Motorcycle	$\frac{APICarDistance}{APICarSpeed+3km*h^{-1}}$	0.59	-	-	-	-	Access Motorcycle
Car ODT	API Car	$10(base) + 3.5$	5	-	-	Yes	Always
Motorcycle ODT	$\frac{APICarDistance}{APICarSpeed+3km*h^{-1}}$	$10(4km) + 2.5$	5	-	-	Yes	Always
Bus	API Transit	10 per 10 km	API Transit	API Transit	API Transit	-	Has transit
BRT	API Transit	3.5	API Transit	API Transit	API Transit	-	Has transit
Train	API Transit	5.5	API Transit	API Transit	API Transit	-	Has train
Microbus (angkot)	API Car	5 per 10 km	5	$>$ 10 km	5	-	Always
Car taxi	API Car	$6(base) + 4.5$	5	-	-	-	Always
Motorcycle taxi	$\frac{APICarDistance}{APICarSpeed+3km*h^{-1}}$	$10(base) + 3$	5	-	-	-	Always

TABLE 3.10: Parameter assumptions for non-chosen mode alternatives

4

MODE CHOICE IN GREATER JAKARTA

4.1 MODELING FRAMEWORK

In this chapter, we developed a model based on a pool of stated preference (SP) and Revealed Preference (RP) data sets. We explore the demand of each choice alternatives: the willingness to pay (WTP) or value travel time savings (VTTS), value travel time assigned to travel (VTAT), and elasticity of all choice alternatives, including ODT and UAM. We conducted a stated choice experiment to gather the data and used a discrete choice model to do the analysis. The measurement of WTP: VTTS, VTAT, and elasticity all together is rarely explored by other researchers.

We employed the multinomial logit (MNL) and mixed logit (MXL) formulation for the choice modeling analysis, both of which are widely used for policy analysis. We used 1,000 Halton draws for MXL. The estimation took seven days. This paper used the R package, mixl, to estimate the model (Molloy *et al.*, 2019). The model that we presented here was based on pooled SP and RP data sets. Train (2003); Cherchi and Ortúzar (2011); Schmid *et al.* (2019) show that the pooled SP and RP data sets have a better estimation and robustness, which could improve the quality of only the SP or RP data sets. MNL assumes that the error term ε is equally Identical and Independently Distributed and that the alternatives have the same probability distribution and independence (McFadden, 1973; Train, 2003). The alternative specific constants (ASCs) are decomposed into their mean value and their standard deviation, denoted by $\eta_{i,n}$. The utility of a person n choosing alternative i in choice situation t can be seen in equation 1 for MNL and equation 2 for MXL.

$$U_{i,n,t} = ASC_i + \beta_i X_{i,n,t} + \varepsilon_{i,n,t} \tag{4.1}$$

$$U_{i,n,t} = (ASC_i + \eta_{i,n}) + \beta_i X_{i,n,t} + \varepsilon_{i,n,t} \tag{4.2}$$

There are 11 alternatives in Model 1 and 8 alternatives in Model 2 and Model 3. MC-based and car-based taxis were converted to taxi, and MC-based and car-based ODT were converted to ODT. The utility formulation for choice alternatives $i \in \{walk, bike, ..., UAM_{\text{SP}}\}$ and individual $n \in \{1, 2, .., N\}$ in choice scenario $t \in \{1, 2, ..., T\}$ can be seen in the Appendices. Travel cost has a continuous interaction term with income and travel distance. In the meantime, the travel time of UAM has a continuous interaction term with travel distance, which corresponds to elasticity λ_{Income} and $\lambda_{Distance}$ (Ilahi *et al.*, 2019c; Vrtic *et al.*, 2010; Mackie *et al.*, 2003). $Income$ refers to the household income, while $AverageIncome$ is the sample mean of income. $Distance$ is the individual trip distance, and the $AverageDistance$ is the sample mean of distance.

4.2 RESULTS

4.2.1 *Model Estimations of Pooled SP and RP*

The results for the three models are presented in Table 10, in which MC is the base category. Model 1 has 11 modes of transport alternatives presented: walking, bike, bus, BRT, train, car, MC, taxi, ODT, PT SP (public transport in SP data set), UAM. Model 2, which combines all public transport modes (Bus, BRT, Train, and PT SP) into a single PT, has eight choice alternatives: walk, bike, car, MC, taxi, ODT, PT, and UAM. For Models 1 and 2, we implemented MNL model. Model 3 has the same alternatives as in Model 2. For Model 3, we implemented MXL model.

The parameters of Model 1, established travel time for a specific alternative, generic travel costs for all choice alternatives, and generic congestion charging only for car, MC, taxi, and ODT modes of transport. Socio-demographic attributes, such as household income, age, gender, and education, were specific only for some alternatives. The models include the attribute of living in the urban agglomeration for UAM. Model 2 and Model 3 had similar parameters to Model 1 except for cost and PT travel time. Cost is a combination of travel cost and congestion charging, while PT travel time in Model 2 and Model 3 is a combination of all public transport alternatives.

In the case of Model 1, we found an insignificant result for the ASC of train, meaning that the train is not more or less preferred than a MC. However, the other choice alternatives had negative and significant ASCs, suggesting that the MC is more preferred than the alternative choices. In Model 2 and Model 3, we found that the ASCs for all choice alternatives were negative, showing that the MC is preferred more than the other transport choice alternatives.

In the models, we found that males, non-university, and older people were less likely to choose ODT, which is supported by significance of male ODT (+), university degree ODT (+), and age ODT (-). Except in Model 3, only age for UAM is significant, and being male in ODT is not significant. For UAM, the impact of being male and living outside Jakarta (in the urban agglomeration) was not statistically significant. However, the young and university degree holding respondents tended to choose UAM. It is supported from the significant of age UAM (-), and university degree UAM (+). This finding is in line with that by Eker *et al.* (2020a), that young respondents tend to choose UAM. Furthermore, variable travel costs, congestion charging, and access time of UAM were negatively significant, as expected.

The variable of travel time for all choice alternatives was negative and significant in all models, and significant except for PT (in the SP data set) and bus in Model 1. These results suggests that they did not like biking. This finding may be because the bike lane facilities do not accommodate bikers very well, and bikers have to compete for space with the motorized vehicles. Furthermore, hot weather makes it difficult to ride bikes. All models also indicate that respondents like to use cars and MC, and that residents in Jakarta prefer to use private vehicles. However, for UAM, from beta travel time, we can see that in Model 1 and Model 2, the respondents immensely enjoy the UAM after Car, but Model 3 indicates that UAM is just somewhat more enjoyable than ODT.

In model 1, the travel time of PT in SP and bus were not significant. However, travel time of BRT and train were negative and significant. People preferred to use BRT rather than a train because the train was less accessible. The number of BRT stops and lines are higher than train stations and lines: 325 BRT lines and 22 commuter lines. Moreover, people still need to take other modes to go to the train station. The λ of income and distance for all models is negative and significant, as shown in Vrtic *et al.* (2010).

With regard to BIC and rho-square, Model 3 (MXL) outperforms the other models, and Model 2 is better than Model 1. Those models provide a better fit than previous studies of Greater Jakarta (Belgiawan *et al.*, 2019b; Ilahi *et al.*, 2019c).

Table 4.1: A Pooled SP and RP result

Variable	Model 1		Model 2		Model 3	
Baseline: MC	Par.	t-test	Par.	t-test	Par.	t-test
ASC Walk	-2.51	-23.08^{***}	-2.57	-23.27^{***}	-16.38	-14.23^{***}
ASC Bike	-4.22	-13.57^{***}	-4.24	-13.65^{***}	-39.48	-18.52^{***}
ASC PT	-3.50	-25.14^{***}	-3.75	-30.5^{***}	-21.02	-26.28^{***}
ASC Bus	-5.05	-13.47^{***}	-	-	-	-
ASC BRT	-4.74	-20.3^{***}	-	-	-	-
ASC Train	-0.29	-0.9	-	-	-	-
ASC Car	-1.20	-10.64^{***}	-1.09	-9.05^{***}	-14.29	-2.61^{***}
ASC Taxi	-3.94	-23.59^{***}	-3.82	-21.21^{***}	-25.72	-9.28^{***}
ASC ODT	-1.43	-9.32^{***}	-1.23	-7.64^{***}	-10.83	-7.02^{***}
ASC UAM	-3.54	-6.95^{***}	-3.23	-5.7^{***}	-22.03	-5.89^{***}
β Tcost	-1.42	-12.08^{***}	-2.08	-15.76^{***}	-6.46	-17.15^{***}
β Con. charg	-4.13	-8.46^{***}	-	-	-	-
λ Income, cost	-0.09	-3.06^{***}	-0.06	-2.83^{***}	-0.13	-3.80^{***}
λ Dist, cost	-0.83	-19.16^{***}	-0.75	-19.13^{***}	-0.84	-29.59^{***}
λ Dist, time_{UAM}	-14.73	-5.23^{***}	-12.3	-9.46^{***}	-12.01	-5.18^{***}
σ Sc. Par MC_{SP}	0.77	32.79^{***}	0.67	31.31^{***}	0.19	11.28^{***}
η Walk	-	-	-	-	-10.37	-22.82^{***}
η Bike	-	-	-	-	17.04	20.73^{***}

(To be continued)

Variable	Model 1		Model 2		Model 3	
Baseline: MC	Par.	t-test	Par.	t-test	Par.	t-test
η PT	-	-	-	-	-18.59	-15.97^{***}
η Car	-	-	-	-	17.96	2.20^{**}
η Taxi	-	-	-	-	-13.06	-5.97^{***}
η ODT	-	-	-	-	-14.70	-9.49^{***}
η UAM	-	-	-	-	13.75	9.03^{***}
β Ttime Walk	-0.36	-6.82^{***}	-0.52	-9.11^{***}	-2.04	-3.65^{***}
β Ttime Bike	-8.61	-4.55^{**}	-9.05	-4.79^{***}	-20.00	-3.73^{***}
β Ttime PT	-0.28	-1.22	-1.49	-7.17^{***}	-8.52	-5.29^{***}
β Ttime Bus	-1.18	-1.4	-	-	-	-
β Ttime BRT	-1.07	-2.36^{**}	-	-	-	-
β Ttime Train	-2.72	-6.39^{***}	-	-	-	-
β Ttime Car	-0.60	-3.64^{***}	-1.24	-6.33^{***}	-5.15	-3.01^{***}
β Ttime MC	-2.34	-10.25^{***}	-3.32	-12.68^{***}	-10.13	-5.88^{***}
β Ttime Taxi	-3.49	-8.03^{***}	-4.79	-9.23^{***}	-9.63	-1.73^{*}
β Ttime ODT	-5.10	-15.27^{**}	-6.26	-16.82^{***}	-15.67	-8.48^{**}
β Ttime UAM	-1.36	-31.83^{***}	-2.65	-3.12^{***}	-12.41	-2.37^{**}
β Actime UAM	-3.54	-1.82^{*}	-4.54	-2.02^{**}	-15.51	-1.38
β Male ODT	-0.38	-4.87^{***}	-0.42	-4.95^{***}	-3.13	-1.41
β Male UAM	-0.04	-0.2	1.08	4.17^{***}	0.29	0.19
β Age Walk	0.99	3.69^{***}	1.03	4.02^{***}	10.41	1.05
β Age Train	-0.15	-2.01^{**}	-	-	-	-
β Age MC	-0.93	-3.96^{***}	-0.83	-3.45^{***}	-2.72	-0.46
β Age ODT	-1.36	-4.05^{***}	-1.32	-3.75^{***}	-11.13	-0.94
β Age UAM	-3.44	-3.48^{***}	-3.76	-3.35^{***}	-22.91	-2.78^{***}

(To be continued)

Variable	Model 1		Model 2		Model 3	
Baseline: MC	Par.	t-test	Par.	t-test	Par.	t-test
β Uni. ODT	0.25	2.86***	0.27	2.97***	3.45	7.34***
β Uni. UAM	0.93	4.09***	1.08	4.17***	6.88	4.10***
β Agglo. UAM	0.34	1.47	0.15	0.58	0.19	0.13
Observations	52731		52731		52731	
Draws	-		-		1000	
Final-LL	-57153		-59103		-33948	
Rho-square	0.44		0.42		0.67	
AIC	114381		118267		67970	
BIC	114709		118533		68299	

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

4.2.2 *Value of travel time savings*

We measured the VTTS of a person in U.S dollars (USD). VTTS measures a person's willingness to pay in return for a reduction of travel time. As our scenario was conducted using Indonesia Rupiah (IDR), we converted IDR into USD, and calculated the VTTS using the following formula:

$$\mathrm{VTTS_{i,n}} = \frac{\mathrm{ffiV_{i,n}/ffiT_{i,n}}}{\mathrm{ffiV_{i,n}/ffiC_{i,n}}} = \frac{60,000}{14,000} * \frac{\mathrm{fiT}}{\mathrm{fiC}} \tag{4.3}$$

where $V_{i,n}$ represents systematic utility for an alternative i for person n, $T_{i,n}$ represents travel time for the person n choosing alternative i, and $C_{i,n}$ represents the cost for the person n choosing an alternative i. The parameters of time and cost are represented by βT and βC respectively.

Table 11 shows the results of the VTTS at the sample mean. It presents the willingness to reduce travel time by one unit. Public Transport in Table 11 and 12 refers to general public transport, as in the SP dataset the mode is an aggregate of all PT services. The specific modes of transport, like

Bus, BRT, and Train, are from the RP dataset based on the assumptions in Table 9.

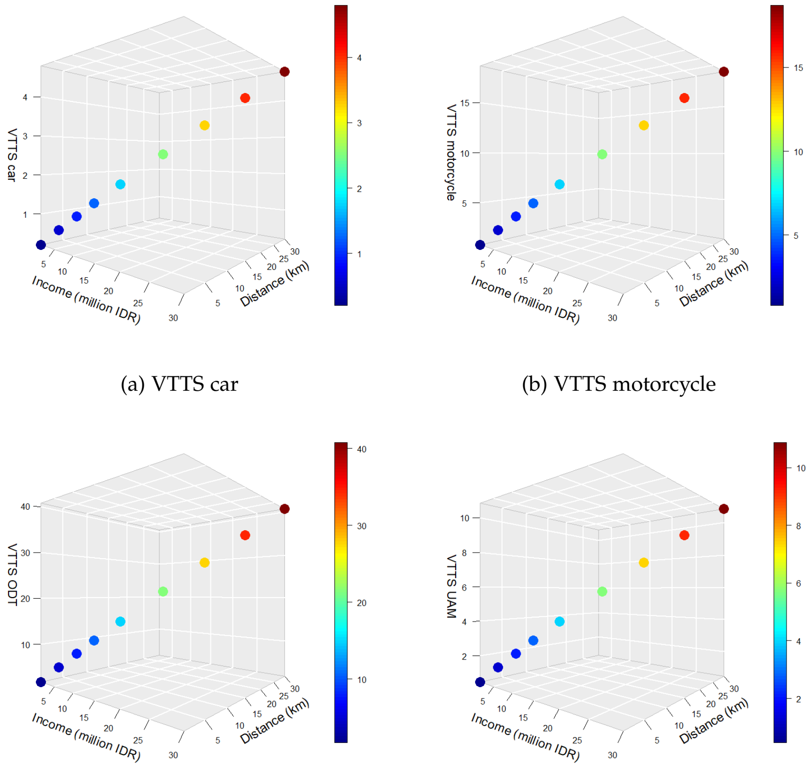
We found that the VTTS of the PT (in the SP data set) was the lowest, but the travel time parameter of PT was not significant. The VTTS of cars was lower than that of the other modes of transport. The result is the same as that of Ilahi *et al.* (2019c); Belgiawan *et al.* (2019b); however, it differs from that of Schmid *et al.* (2019); Shires and de Jong (2009); Wardman (2004) that shows the car VTTS was higher than public transport. The VTTS of ODT was the highest, followed by taxi, train, and MC. Model 2 gave a similar result, in which ODT was the highest, followed by taxi, MC, UAM, PT, and car. In Model 3, the highest VTTS is ODT, followed by UAM, taxi, and MC. Furthermore, the interaction between income and distance to VTTS is shown in Figure 7 for cars, MC, ODT and UAM.

In terms of VTTS vis-a-vis the Congestion Charge in Model 1, the ODT was the highest, followed by taxi, MC, and car, showing that the users of ODT and taxis were more sensitive to the charge. The income per capita in 2018 in Jakarta was 17,438 USD per year or 8.7 USD per hour for an assumed 2,000 hours per year. However, we must consider that Indonesia has a high Gini ratio. A small number of people have very high incomes per capita. The hourly income per capita is lower than 8.7 USD if we exclude the highest-income people. We found that the VTTS vis-a-vis travel cost of ODT, taxi, and train (Model 1), ODT and taxi (Model 2), and ODT and UAM (Model 3) were higher than income per capita/hour. The VTTS vis-a-vis of access cost was higher than income per capita/hour in all models.

As has been discussed in several studies (Schmid *et al.*, 2019; Jara-Díaz *et al.*, 2008; Hössinger *et al.*, 2020), the VTTS is the sum of the value of leisure (VOL) and VTAT. VOL represents the willingness to pay to reduce travel time to gain more utility, and VTAT is the indirect value disutility of travel time assigned to travel. VOL is calculated based on the proportion of hourly income and each country has a different value. Assuming that the VOL of Indonesians is similar to Chileans, approximately 66% of wages (Jara-Díaz *et al.*, 2008) the resulting VOL is 5.74 USD. As shown in Table 11, VTAT is positive for public transport, bus, BRT, car, and UAM (Model 1), and public transport, car, and UAM (Model 2). However, only the car and public transport in Model 3 have a positive VTAT. The higher the VTAT, the lower the VTTS. In the meantime, the lower the VTAT, the higher the VTTS. The positive value of VTAT shows disutility when travel time is reduced; it therefore reveals more about comfort, safety, and security during travel, than the travel time itself.

Model	Mode	Fuel/Ticket cost	Congestion cost	Access cost	VTAT
Model 1	PT	0.86	-	-	4.88
	Bus	3.56	-	-	2.18
	BRT	3.23	-	-	2.51
	Train	8.21	-	-	-2.47
	Car	1.80	0.62	-	3.94
	MC	7.06	2.43	-	-1.32
	Taxi	10.52	3.62	-	-4.78
	ODT	15.38	5.29	-	-9.68
	UAM	4.98	-	10.7	0.76
Model 2	PT	3.07	-	-	2.67
	Car	2.55	-	-	3.19
	MC	6.85	-	-	-1.11
	Taxi	9.88	-	-	-4.14
	ODT	12.92	-	-	-7.18
	UAM	5.47	-	9.35	0.27
Model 3	PT	5.65	-	-	0.09
	Car	3.42	-	-	2.32
	MC	6.72	-	-	-0.94
	Taxi	6.39	-	-	-0.65
	ODT	10.39	-	-	-4.65
	UAM	8.23	-	10.29	-2.49

Table 4.2: Value of time of mode of transport pool SP and RP (USD/hour)



(a) VTTS car

(b) VTTS motorcycle

(c) VTTS ODT

(d) VTTS UAM

FIGURE 4.1: VTTS related to income and distance

4.2.3 *Point elasticities of travel time*

In this section, we measured the direct point elasticities for all modes of travel. The method used were the same as those presented in Atasoy *et al.* (2013); Belgiawan *et al.* (2019a), as shown in Eq.3:

$$E^{w_i}_{iqX_{kiq}} = \sum_{q=1}^{Q_s} E_{iqX_{kiq}} \frac{w_q P_{iq}}{\sum_{q=1}^{Q_s} w_q P_{iq}} \tag{4.4}$$

where w_q represents the sample weight for individual q from sample Q_s from population Q and $E_{iq}X_{kiq}$ is the disaggregate elasticity on the demand of individual q for variations in attribute X_{kiq}. We weighted each observation in our data sets according to the representation of its age and gender category in the Greater Jakarta population.

The results are shown in Table 12. The sign of all the time and cost elasticities measurements were as expected, which means that a percentage increase in all travel time and travel cost would, on average, reduce the probability of choosing an alternative. We found that travel time for all modes was inelastic for Model 1, and ODT was elastic for Model 2. However, for Model 3, only walking and MC were inelastic. The bike had the highest reduction, in which a 1% increase in travel time would reduce the probability of choosing a bike by 4.26%, followed by ODT with a reduction of around 3.39%, and public transport of 2.83%.

Furthermore, for travel costs, only UAM was elastic for Model 1 and Model 2. However, travel cost was elastic for car, taxi, ODT, and UAM for Model 3. It showed that a 1% increase in travel cost for the UAM would reduce the probability of choosing UAM by 2.07% in Model 1, 2.96% in Model 2, and 9.55% in Model 3. In Model 3, we found that a 1% increase in travel cost would reduce the probability of choosing a car by 2.15%, a taxi by 2.23%, and ODT by 2.60%. The congestion cost was inelastic for Model 1. The access time was inelastic for Model 1 and Model 2, but was elastic for Model 3.

Model	Mode	Travel time	Travel cost	Congestion cost	Access time
Model 1	Walk	-0.17	-	-	-
	Bike	-0.94	-	-	-
	Bus	-0.46	-0.33	-	-
	BRT	-0.42	-0.05	-	-
	Train	-0.87	-0.05	-	-
	Car	-0.13	-0.36	-0.13	-
	MC	-0.24	-0.14	-0.05	-
	Taxi	-0.57	-0.44	-0.07	-
	ODT	-0.92	-0.43	-0.01	-
	PT	-0.06	-0.63	-	-
	UAM	-0.15	-2.07	-	-0.33
Model 2	Walk	-0.24	-	-	-
	Bike	-0.99	-	-	-
	Car	-0.26	-0.58	-	-
	MC	-0.34	-0.24	-	-
	PT	-0.54	-0.24	-	-
	Taxi	-0.75	-0.64	-	-
	ODT	-1.11	-0.68	-	-
	UAM	-0.29	-2.96	-	-0.39
Model 3	Walk	-0.95	-	-	-
	Bike	-4.26	-	-	-
	Car	-1.29	-2.15	-	-
	MC	-0.85	-0.72	-	-
	PT	-2.83	-0.67	-	-
	Taxi	-1.76	-2.23	-	-
	ODT	-3.39	-2.60	-	-
	UAM	-1.32	-9.55	-	-1.42

TABLE 4.3: Point elasticities of variables

5

AGENT-BASED MODEL OF GREATER JAKARTA

5.1 INTRODUCTION

Transport activities are not based on a single entity but consist of several complex interactions (Kitamura, 1988; Axhausen and Gärling, 1992). As social creatures, our trip activities are influenced by other people. It can be in intra-household interaction (Bradley and Vovsha, 2005). For instance, our activities in the morning can be influenced by our children. We need to drop off our child to school first before we go to the office. Each activity type also has different time allocations as stated by (Borgers *et al.*, 2001; Gliebe and Koppelman, 2005; Simma and Axhausen, 2001) and there are difference shares for each trip purpose (Axhausen *et al.*, 2002). As classified by (Arentze and Timmermans, 2004), there are several types of constraints in activities, such as household constraint, spatial constraints, time constraints, and spatial-temporal constraints. For example, in the shopping places, we have time constraint when it opens or closes, and how far it is from our home location. Nevertheless, trip activities can be influenced by bigger group interactions, in which people are also influenced by their needs, occupation, group, ethnicity, nationality, or even belief and interest. However, family has more influence factors as it is where we spend our time and have interaction the most.

To simulate above complex interactions, we used an agent-based model which has been used by many researchers. There are several tools using agent-based model, such as; ORIENT/RV (Axhausen, 1989), TRANSIMS (Smith *et al.*, 1995), SimMobility (Adnan *et al.*, 2016), SimTRAVEL (Pendyala *et al.*, 2012), Multi-Agent Transport Simulation MATSim (MATSim) (Balmer *et al.*, 2006; Horni *et al.*, 2016), and GEMSim (Saprykin *et al.*, 2019). However, in this research, we employed MATSim, which is considered suitable in the recent past to model large-scale cities, such as Singapore (Erath *et al.*, 2012). Besides, it is able to include micro-buses in simulation (Neumann *et al.*, 2015), and utilizes joint activities between household members as can be seen in (Dubernet and Axhausen, 2015). Besides, it can also simulate the impacts of emerging transportation options and policies, such as the impacts of car-sharing (Balac *et al.*, 2019), urban air mobility (Balac *et al.*, 2018), bike-sharing, congestion pricing, automated vehicles or equity effects, which are hard to investigate on a suitable level using the more traditional modeling techniques (Horni *et al.*, 2016).

The following reasons motivate the use of agent-based model. First, to the authors' knowledge, this research will be the first to make use of an agent-based model that incorporates microbus or "angkot" and simulates the daily behavior of people performing their activities in Greater Jakarta. There are several studies conducted in Jakarta. Yagi and Mohammadian (2010) simulated mode and destination choice based on discrete choice modelling, and Dharmowijoyo *et al.* (2016) measured variability of travel patterns in greater Jakarta. However, those studies do not take into account greater Jakarta as a whole object of study using agent-based modelling.

Second, our model integrates novel approach to the mode-choice model in simulation. As shown in Hörl *et al.* (2018, 2019), the approach can give faster convergence speed of simulation than scoring based (Balmer *et al.*, 2006; Horni *et al.*, 2016). Third, this research also adds to the growing literature on modelling large-scale cities especially when the data are scarce or not easily obtainable. Fourth, it is different from conventional four step model. Each individual is simulated as an agent, in which each of them has its attributes, such as socio-demographic, activity location, and mode of transportation. The attributes of each agent are used as an input plan, and then the model uses iterative processing to find the best plan that maximizes an agent utility.

The remainder of this chapter is structured as follows. The following section describes case studies of Greater Jakarta, and the third section explains the MATSim framework. Mode-choice in MATSim is presented in the forth section. Then, the fifth section presents the first results obtained for the commuting population of the Greater Jakarta. Finally, the conclusions, limitations and further recommendations are presented in the last section.

5.2 MATSIM FRAMEWORK

This study utilized a Multi-Agent Transport Simulation (MATSim), which performs a microscopic simulation of daily schedules of synthetic persons performing activities in the study area. The persons in MATSim are represented by agents, and each agent has its plan that represents its daily schedule of activities connected by trips. The plans are simulated using the mobility simulation (mobsim) for a number of iterations. Before each simulation starts, some of the agents are allowed to change a part of their plan in the re-planning phase. This simulation cycle can be seen in Figure 5.1.

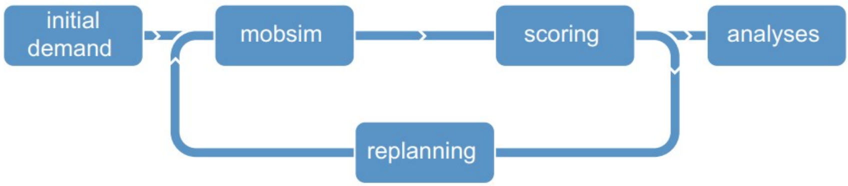


FIGURE 5.1: The MATSim loop (Source: MATSim book (Horni *et al.*, 2016))

In this work, however, we use a slightly different iterative approach proposed by Hörl *et al.* (2018, 2019) that integrates a mode-choice model with the micro-simulation in MATSim. In this approach, agents are allowed to change their modes of travel based on a discrete-mode choice model implemented. Since the mode-choice model in the re-planning phase uses estimates on the travel-times, costs, waiting times, etc. from the previous iteration, the scoring phase is no longer needed. This approach is then utilized to investigate mode choices for the commuting population of the Greater Jakarta population.

5.2.1 *Traffic flow model*

Traffic simulation model in MATSim uses a queue based approach, which has two attributes, storage and flow capacity. Storage capacity defines how many cars can be stored at a time on a road link, and flow capacity defines the outflow capacity of a link.

5.2.2 *Public transport network and counting stations*

In Jakarta and in Indonesia in general, there are different modes of transportation options available, some of them are formal ones, like car, motorcycle, commuter rail, Bus Rapid Transit (BRT), big buses and medium buses, and some informal ones are micro-buses (called "angkot", which has an informal service without a fixed schedule). Micro-bus, which is of a small size, has an ability to become a door to door services in Greater Jakarta. Simulation of micro-buses in an agent-based model has been previously used in the South-African context (Neumann *et al.*, 2015).

To create public transport network, we need open street map (OSM) and general transit feed specification (GTFS) data. However, there is no publicly available GTFS data for Greater Jakarta. Therefore, we have manually constructed the public transport schedules using the data from a company called Trafi (https://www.trafi.com/id/jakarta). The data scrapped from Trafi website are formatted to GTFS structure data (Google, 2019). There are several important files that must be available, which can be seen in Table 5.1.

File	Description
agency.txt	Public transport (PT) company that operates the pt lines. In our cases, it consists of BRT, Rail, Angkot, Bus. BRT is all PT that operate in busyway/dedicated line (Transjakarta), Bus is all big/medium bus operating in mixed traffic. Rail consists of commuter line operated by public company railway (PT. KAI), and angkot is all microbus lines.
stops.txt	Station/shelter location, which consists of shelter station coordinate, the name of the shelter station, name of shelter/station.
routes.txt	The name of public transport routes/lines. A route is a group of trips that are displayed to riders as a single service.
trips.txt	Trips for each route. A trip is a sequence of two or more stops that occur at specific time.
stop_times.txt	Times when a vehicle arrives at and departs from individual stops for each trip.

TABLE 5.1: Overview of GTFS files

Furthermore, OSM network and manually constructed GTFS schedules are converted to MATSim format using the pt2matsim extension (Poletti, 2016). BRT lines are categorized as dedicated lanes. In the end, the transit schedule is mapped to the MATSim network using the same extension. Finally, we obtain public transport lines within the network. There are 1,756 public transport lines in total. There are 325 BRT lines (including other bus companies that using BRT lines), 421 Bus lines, 22 commuter rail lines, and 988 micro-bus lines. There are also 20 counting stations that count the number of vehicles in 15min bins as can be seen in Figure 5.2.

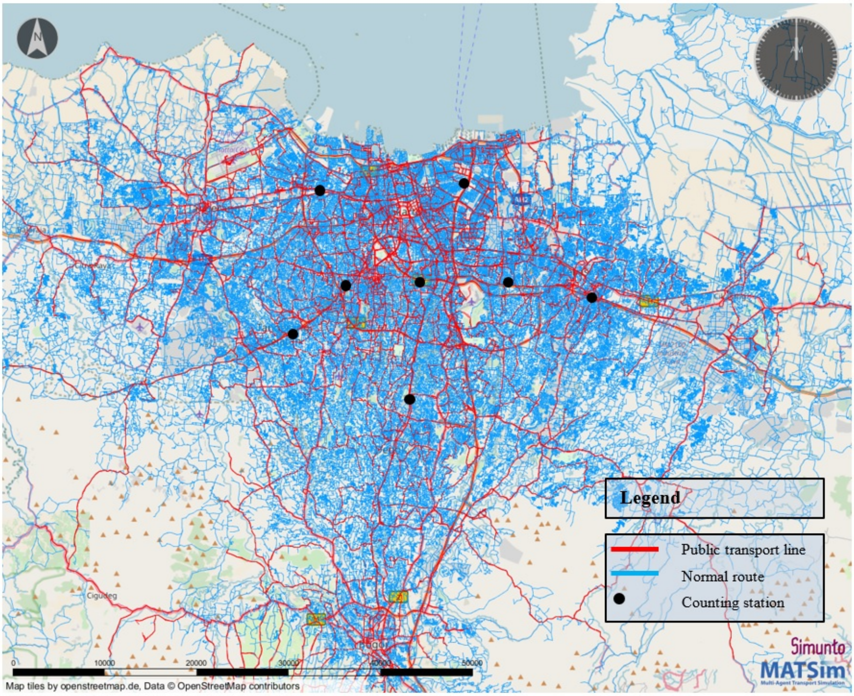


FIGURE 5.2: MATSim network, public transport lines, counting stations

5.2.3 *Private and public transport vehicles*

As the vehicles have different sizes and capacities in our simulation, we classify private and public transport vehicles as in Table 5.2. Car and motorcycle are classified as private vehicles, while BRT, bus, commuter, and micro-bus are classified as public transport. The pce (passenger car equivalent) of motorcycle used is 0.25.

Type	Mode	Symbol	Length/Width [m]	Capacity Seats/Standing	Number of lines
Private	Car	car	4.3/1.6	7/0	-
	Motorcycle	mc	1.7/1.0	2/0	-
Public	BRT	pt	2.5/50	50/30	325
	Bus	bus	2.5/35	35/15	421
	Commuter	rail	240/2.8	2.8/1000	22
	Micro-bus	angkot	4.2/1.6	1.6/8	988

TABLE 5.2: Public transports and private vehicles parameters

5.3 CALIBRATION

5.3.1 *Mode shares*

The model are calibrated using 1% of population using parameter estimated from MNL model based on chapter 4. The calibrated parameter estimated used for this model can be seen in Table 5.3. The average income used in this model is 5.327 Million IDR, and the average distance is 7.67 km.

Parameters	Estimation
ASC Walk	-7.00
ASC PT	-6.00
ASC Car	-0.75
ASC Car ODT	-6.50
ASC Motorcycle ODT	-4.00
β Travel cost	-2.08
λ Income, cost	-0.06
λ Distance, cost	-0.75
β Travel time Walk	-0.70
β Travel time PT	-2.00
β Travel time Car	-4.00
β Travel time Motorcycle	-6.10
β Travel time Car ODT	-8.26
β Travel time Motorcycle ODT	-7.00
β Waiting time PT	-0.05
β Access egress time PT	-0.05
β Male Car ODT	-0.42
β Male Motorcycle ODT	-1.15
β Age Walking	1.03
β Age Motorcycle	-0.83
β Age Car ODT	-1.32
β Age Motorcycle ODT	-1.32

Table 5.3: Calibrated parameters of mode choice model

We compare the mode shares and the mode shares for several distance band with the Household Travel Survey (HTS) of Mobility Jakarta Survey. After several calibration, we can see that the mode shares of the model fits the HTS, as shown in Figure 5.3, and Figure 5.4. It is not perfectly match, since we need to make sure that the mode shares match not only in aggregate view but also by distance band view.

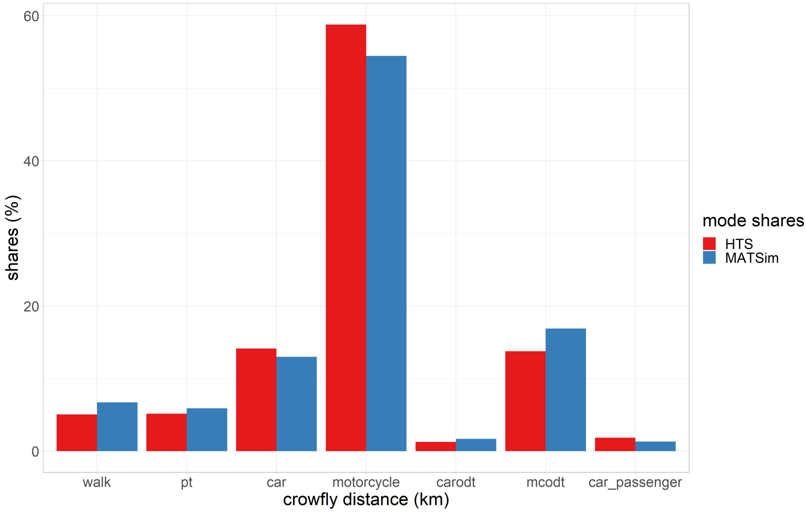


FIGURE 5.3: Mode shares

5.4 DISCUSSION AND CONCLUSIONS

In this chapter, we use an agent-based modelling framework to simulate the commuting population of Greater Jakarta. The methodology presented also utilizes a novel approach that integrates mode choice with a micro-simulation in MATSim. The results show that differences between the mode shares from MATSim and JICA mode shares is very low. It also shows that the mode shares for different distance band matches between MATSim and JICA.

This research simulate behavior of people living in Greater Jakarta including primary and secondary activities. Nevertheless, this research provides the backbone on which further research can be built. However, it takes a considerable amount of money and time for construction as the size of the city, and the population is huge.

FIGURE 5.4: Mode shares by distance band

The parameters of the mode choice model are based on the Greater Jakarta model. However, we adjusted with calibration value and factor of five mode transports (public transport, car odt, mc odt, car, motorcycle), as resented in Table 5.3. We use 1 (%) of the population in our simulations for calibration. The calibrated model further used for simulating impact of congestion pricing in Greater Jakarta.

As for now, however, the model in this thesis did not consider the joint trips of members in the same household, but it considers car passenger trips. The model also did not simulate public transport in mixed traffic to speed up the simulation time.

6

POLICY SCENARIO

Jakarta government had implemented several policy to reduce the number of vehicles for several arterial road in Greater Jakarta. Previously, a "3 in 1" throughout high-occupancy vehicle (HOV) policy had been implemented on several arterial roads in Jakarta's CBD for cars (Governor of DKI Jakarta Province, 2012).

The a 3 in 1 policy means that private cars with less than three persons are not allowed to travel on arterial roads during morning and evening peak hours. However, this policy has not been successful in improving urban transport conditions due to lack of control and some further implementation issues. For example, many people pay brokers to fulfill the minimum number of passengers (Anya and Wardhani, 2016). This led to the revocation of this policy (Governor of DKI Jakarta Province, 2012). Due to the failures of this policy, there is a movement to replace the 3 in 1 scheme with a more comprehensive approach such as Electronic Road Pricing (ERP).

It is argued that for the case of Jakarta ERP will produce financial resources for other projects (CMEA and JICA , 2012). ERP has already been successfully implemented in several countries (Agarwal and Koo, 2016; Eliasson and Mattsson, 2006; Rotaris *et al.*, 2010; Santos, 2005). For example, Singapore has succeeded to shift private car users to use public transport by 10%–20% (Agarwal and Koo, 2016). The ALS (Area Licensing Scheme) in London has significantly reduced the number of private vehicles (Santos, 2005). Other examples include Milan (Rotaris *et al.*, 2010) and Stockholm (Eliasson and Mattsson, 2006).

Several studies have already been conducted on the feasibility of ERP in Jakarta. Prayudyanto *et al.* (2013) evaluated several approaches that can support the implementation of congestion charging. Sugiarto *et al.* (2015, 2017) have explored the psychological factors that influence the public acceptance of ERP schemes. Their results emphasize that clear introduction and explanation of the benefits of the ERP policy will increase public acceptance of its implementation.

Furthermore, Belgiawan *et al.* (2019b); Ilahi *et al.* (2019c) estimated the impact of pricing in Greater Jakarta using Random Regret Minimization (RRM), and Mixed Logit Model (MXL). Belgiawan *et al.* (2019b) found that the road pricing of car and motorcycle are nearly inelastic. However, the car is substantially higher than the motorcycle, in which the probability of using a car might be lower when the congestion charging increases. Besides, Ilahi *et al.* (2019c) found that the motorcycle tends to be more willing to pay compares to the car for gaining more benefits.

6.0.1 *Scenario*

In this chapter, we discuss the road pricing scenario in Greater Jakarta using MATSim. There are three different road pricing schemes namely cordon toll pricing, distance toll pricing, and area toll pricing (Nagel, 2016).

Modelling of road pricing has been implemented by many researchers, as can be seen in (de Freitas *et al.*, 2017; de Palma and Lindsey, 2006; Kaddoura and Kickhöfer, 2014). However, those pricing schemes are simulated only for the car without considering the income of the agent. As Jakarta has more motorcycle than a car, we implemented the road pricing also for motorcycle with distance-based pricing. Besides, we also considered the income of the users to make the scenario more realistic, which, in reality, the higher the income, the users tend to be willing to pay the road pricing. We simulated 10% of population, which takes 10 day for computation time.

The congestion charging operates 7.00 a.m - 10.00 am and 4.00 pm to 7 p.m. The location of the scenario is at eight main roads in Jakarta as shown in Figure 6.1, which has a 24 km length in total, or 3 km length on average. There are three different scenarios as can be seen in Table 6.1. We increased the price for each scenario by 33% for both car and motorcycle.

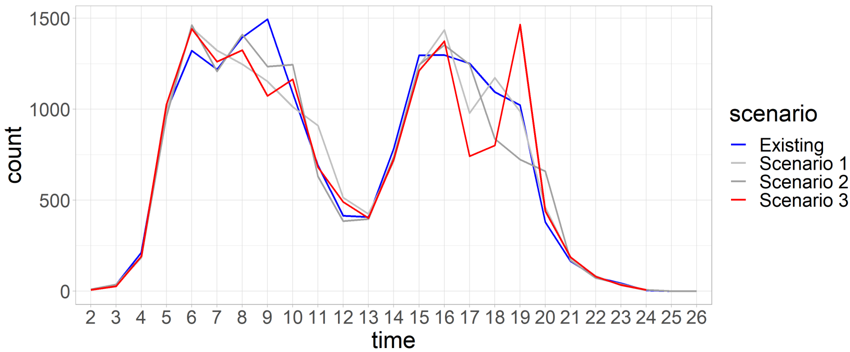
FIGURE 6.1: Mode shares

Type	Scenario 1	Scenario 2	Scenario 3
Car	3	4	5
Motorcycle	1.5	2.0	2.5

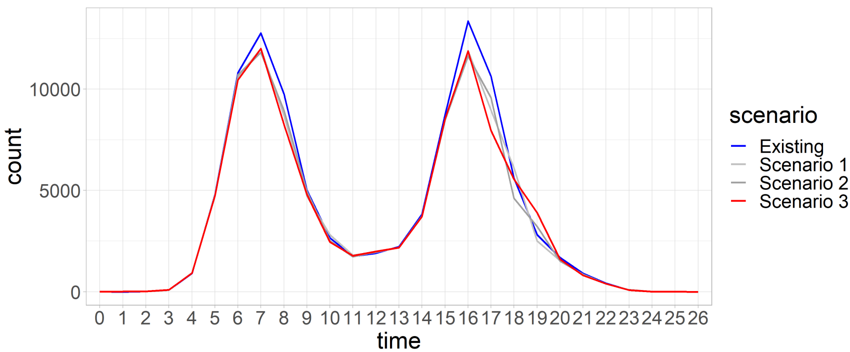
TABLE 6.1: Cost per km for car and motorcycle in road pricing in IDR Thousands

6.0.2 *Results*

The results in Figure 6.2 and Figure 6.3 show that the implementation of road pricing scenarios for both cars and MC simultaneously could reduce the traffic on the main road in greater Jakarta. It affects the other mode, such as car ODT and MC ODT, which is not charged. However, the number of cars and MC that decreased was not significant.

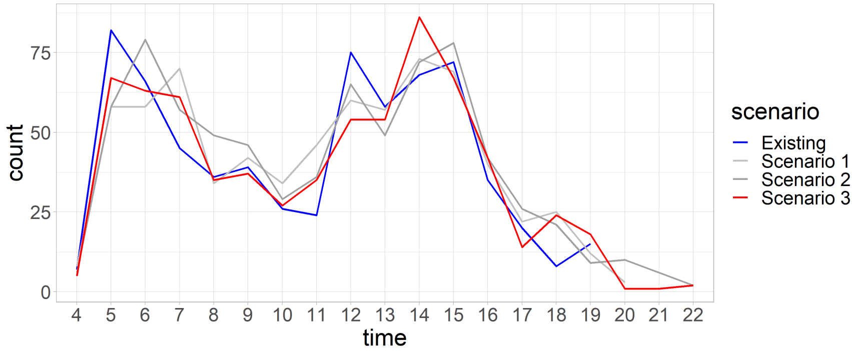


(a) Car

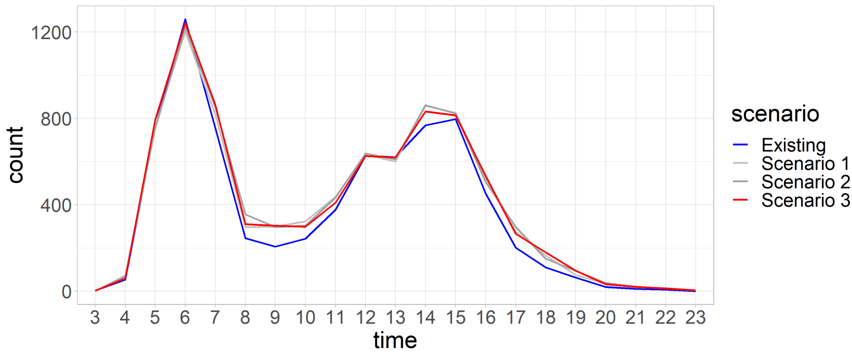


(b) MC

FIGURE 6.2: Traffic count on road pricing scenario for car and MC



(a) Car ODT



(b) MC ODT

FIGURE 6.3: Traffic count on road pricing scenario for car ODT and MC ODT

The traffic of cars and MC in morning peak decreases by increasing the price of congestion charging. In the morning peak from Table 6.2, at scenario 1, the car's number decreases by 9.30%, 6.21% at scenario 2, and 10.93% at scenario 3. For the motorcycle, the number of MC decreases by 7.77% at scenario 1, 6.69% at scenario 2, and 9.08% at scenario 3. However, in the evening peak from Table 6.3, the number of cars reduces by 1.56 % at scenario 1, 5.71 % at scenario 2, and 19.95% st scenario 3. Besides, it decreases for MC by 8.86% at scenario 1, 12.75% at scenario 2, and 14.21% at scenario 3.

Scenario	Mode	Count	% diff
Existing	Car	4,108	-
	Motorcycle	27,517	-
	Car ODT	120	-
	MC ODT	1,206	-
Scenario 1	Car	3,726	-9.30
	Motorcycle	25,380	-7.77
	Car ODT	146	21.67
	MC ODT	1,416	17.41
Scenario 2	Car	3,853	-6.21
	Motorcycle	25,677	-6.69
	Car ODT	152	26.67
	MC ODT	1,504	24.71
Scenario 3	Car	3,659	-10.93
	Motorcycle	25,018	-9.08
	Car ODT	133	10.83
	MC ODT	1,477	22.47

TABLE 6.2: Traffic count for each road pricing scenario on morning peak

The number of car ODT and MC ODT increases for all scenarios and peak times (morning and evening). In general, on both morning and evening peak times, the number of vehicles decreases when increasing the road pricing by 7.07% at scenario 1, 8.25%, and 11.00 % at scenario 3.

PT in the model was not simulated on mixed traffic to reduce the computing time. Therefore, we could not see the impact on PT directly regarding how much it increases the share of PT.

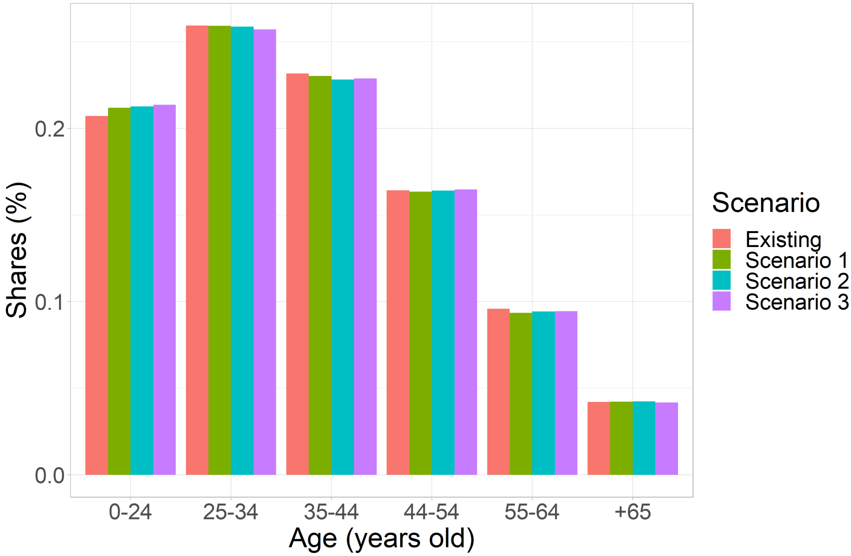
Scenario	Mode	Count	% diff
Existing	Car	3,644	-
	Motorcycle	29,578	-
	Car ODT	63	-
	MC ODT	766	-
Scenario 1	Car	3,587	-1.56
	Motorcycle	26,958	-8.86
	Car ODT	87	38.10
	MC ODT	966	26.11
Scenario 2	Car	3,436	-5.71
	Motorcycle	25,808	-12.75
	Car ODT	89	41.27
	MC ODT	955	24.67
Scenario 3	Car	2,917	-19.95
	Motorcycle	25,375	-14.21
	Car ODT	80	26.98
	MC ODT	979	27.81

Table 6.3: Traffic count for each road pricing scenario on evening peak

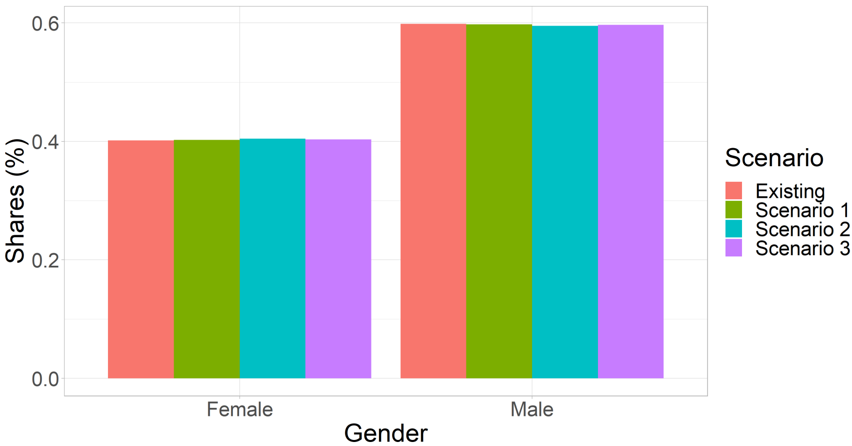
Overall, the reduction in the evening peak is higher than the morning peak. Several reasons that might not significantly decrease the traffic. Firstly, the price of road pricing is not high enough, making the agent willing to pay road pricing. Secondly, the agents' income that uses those roads is high, making the road pricing relatively low to income. Furthermore, As can be seen in Figure 6.4 and 6.5, we found that the distribution of age, gender, employment, and regional distributions of the agents are similar for existing and the future scenarios. The share of males, employment, and age between 25-34 years old is the largest distributions, and most of them are from south of Jakarta.

Times	Scenario	Count	% diff
Morning peak	Existing	32,951	-
	Scenario 1	30,688	-6.93
	Scenario 2	31,186	-5.36
	Scenario 3	30,287	-8.08
Evening peak	Existing	34,051	-
	Scenario 1	31,598	-7.20
	Scenario 2	30,288	-11.05
	Scenario 3	29,351	13.80
Both peak	Existing	67,002	-
	Scenario 1	62,226	-7.07
	Scenario 2	61,474	-8.25
	Scenario 3	59,638	11.00

Table 6.4: Traffic count for each road pricing scenario on different peak times

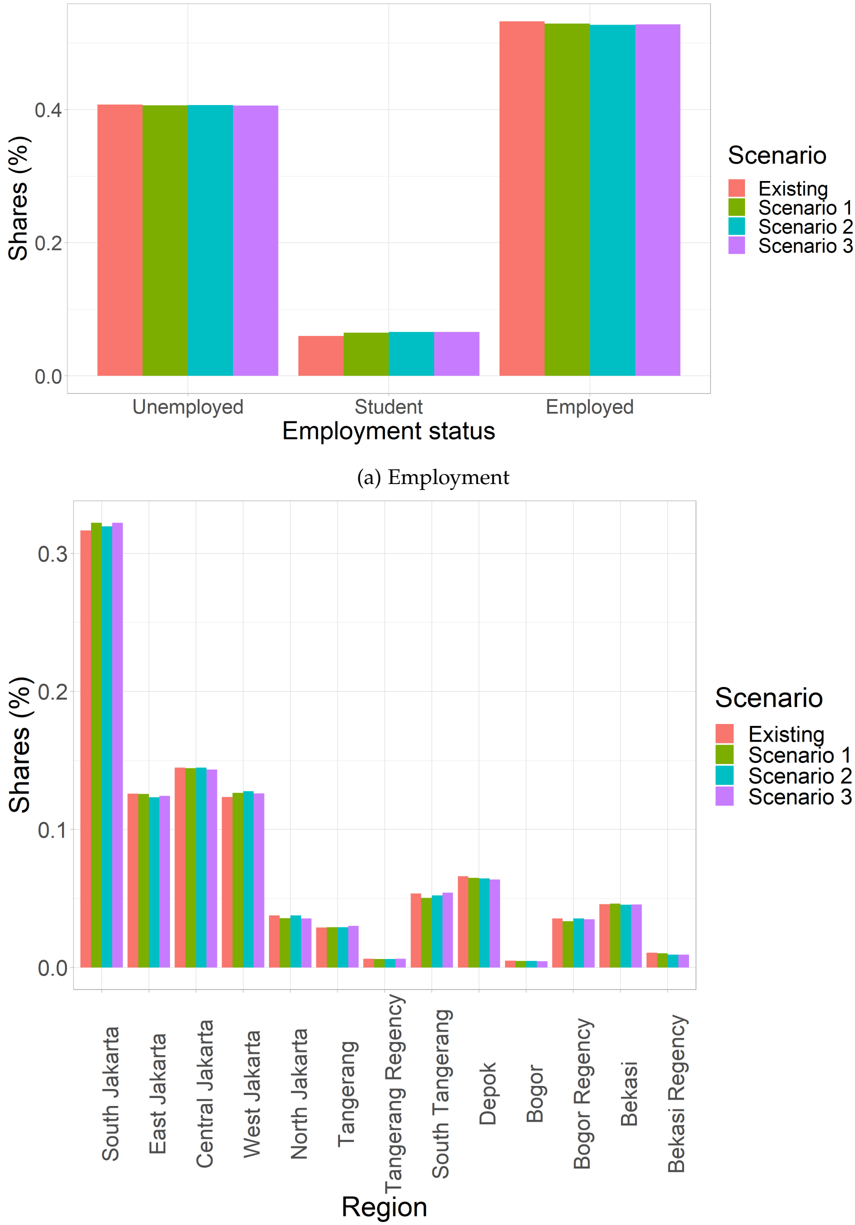


(a) Age



(b) Gender

FIGURE 6.4: Age and gender distribution on road pricing scenario



(a) Employment

(b) Region

FIGURE 6.5: Employment and region distribution on road pricing

7

CONCLUSIONS AND RECOMMENDATIONS

Summary of Findings The overall objective of this thesis, i.e., developing an agent-based model in Greater jakarta, has been broken down into five tasks. These are: 1) To Generate population synthesis. 2) To analyze travel behavior on each mode in Greater Jakarta. 3) To explore model estimations based on stated preference and revealed preference data set, which measure willingness to pay (WTP) for specific mode, i.e., Value of Travel Time Savings (VTTS), elasticities, Value of Travel Time Assigned to Travel (VTAT). 4) To combine the results of choice model, travel behaviour, and population synthesis into Agent-based model. 5) To give policy recommendations based on model and simulation.

To accomplish the first objective, we constructed population synthesis based on the data from JICA in 2012. However, the data are only limited to the profile of respondents and mandatory activities. Therefore, we did a Revealed Preference (RP) survey in Greater Jakarta, with total of 5,000 respondents, so that we could model the secondary activities. The second objectives has been accomplished by the same Revealed Preference (RP) data set. We analyzed the behavior for each mode transport. The third objectives has been carried out by creating mode choice experiments through stated preference (SP) survey. There are 5,000 respondents in this experiments. The model was estimated using Multinomial Logit (MNL) and Mixed Logit (MXL) model. The fourth objectives has been done by applying agent-based model. We simulated people behavior regarding their activities, interaction, and facility constraints. We used parameter estimated from discrete choice model and population synthesis. For the fifth objective, we simulated policy scenario in Agent-based model. In this cases, we created road pricing scenario for main arterial road in Jakarta. We implemented road pricing for both car and motorcycle, and we also considered the income of agents.

This thesis presents a travel diary survey and its outcomes for the Greater Jakarta region. It provides the most comprehensive sample to date of the mobility behavior of people living in Greater Jakarta. This paper discusses an early effort to understand On-Demand Transport services and Urban Air Mobility (UAM) and their impacts on mobility in Greater Jakarta. We identify the patterns of trip purposes for each mode of transport and distinguish the mode choice by its socio-demographic attributes. The results indicate who made the trips, as well as when and why. We also describe the response rate and response burden. This research further enriches the scarce literature on mobility patterns in developing countries, especially considering the multitude of mode choices, some of which are informal.

Our objectives were to capture travel behavior for each mode of transportation comprehensively. The WTP for each mode was also investigated, which includes the VTTS, the VTAT, and elasticity, using pooled SP and RP data sets. The attributes of time and cost were negative and significant in all models, as expected, except for PT and the bus in Model 1. We found that the VTTS of ODT was the highest. The low VTTS of the car in our result shows that car users enjoy riding in a car. Implementing road pricing might not significantly reduce the share of cars and MC, as it was inelastic in Model 1. However, increasing PT frequency or creating a special bike lane and bus priority lane to reduce travel time could increase the shares of those modes. As can be seen in all models, we found that cars and MC are inelastic for all variables, except for cars in Model 3.

UAM has the potential to develop, although it might be suitable only for high-income residents and long-distance travel. As seen in our results, the VTTS of UAM was relatively low. As the urban agglomeration cities in greater Jakarta are relatively far from Jakarta and the airport, UAM has an advantage compared to other choice alternatives. Moreover, severe congestion in Jakarta also provides an advantage for UAM over other alternatives. The infrastructure requirements of UAM that are needed to ensure that UAM services are adequate will take some time to cover all of Greater Jakarta. Greater Jakarta differs from cities in the U.S. such as Los Angeles and New York City, where UAM is currently available. Society still needs to be educated to become familiar with this system. Safety and security are also required to ensure that society accepts the development of this mode of transport.

In our results, ODT had the highest VTTS, suggesting that this system provide great benefits for people in Greater Jakarta. Both high and low-income people can use this transportation alternative. If a person has the app, s/he can comfortably ride the ODT service. The principal problems that still arise in the context of using ODT are related to regulation, as this system is not subject to the same regulations as other modes of travel, such as PT. MC ODT changed the behavior of conventional transaction; people do not need to negotiate the price upfront, and drivers do not need to wait for the customer at a particular place daily (called pangkalan). Regulations could be beneficial for both passengers and providers. ODT can also support the PT infrastructure. The positive value of VTAT in public transport shows that public transport's comfort, safety, and security should be considered for further improvement.

The results from the agent-based model show that the different road pricing is not significantly reduce the number of cars and motorcycles. It shows that motorcycling is more sensitive to road pricing. In contrast, other policies may reduce the number of cars and MC, such as by creating a quota for each household for owning a private vehicle, by making it difficult for people to afford those modes, and by implementing an odd and even license plate policy to control the number of cars in circulation.

A

UTILITY EQUATION

In this appendix, we present the utiliy fromulation for Model (A.1 -A.20), Model 2 (A.21-A.40), Model 3 (A.41-A.60). In Model 2 and Model 3, we combine all public transport mode (bus, BRT, and train) as single public transport (PT), and travel cost (TC) and congestion charging (CC) as a single travel cost (TC).

$$U_{walk,n,t} = 1 * (\alpha_{walk} + \beta_{TT_{walk}} * TT_{walk,n,t} + \beta_{Age} * Age_n) + \epsilon_{walk,n,t} \tag{A.1}$$

$$U_{bike,n,t} = 1 * (\alpha_{bike} + \beta_{TT_{bike}} * TT_{bike,n,t}) + \epsilon_{bike,n,t} \tag{A.2}$$

$$\begin{aligned} U_{bus,n,t} = &1 * (\alpha_{bus} + \beta_{TT_{bus}} * TT_{bus,n,t} + \beta_{TC} * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} \\ &* (\frac{Income}{AverageIncome})^{\lambda_{Income}} * TC_{bus,n,t}) + \epsilon_{bus,n,t} \end{aligned} \tag{A.3}$$

$$\begin{aligned} U_{BRT,n,t} = &1 * (\alpha_{BRT} + \beta_{TT_{BRT}} * TT_{BRT,n,t} + \beta_{TC} * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} \\ &* (\frac{Income}{AverageIncome})^{\lambda_{Income}} * TC_{BRT,n,t}) + \epsilon_{BRT,n,t} \end{aligned} \tag{A.4}$$

$$
\begin{aligned}
U_{train,n,t} =& 1 * (\alpha_{train} + \beta_{TT_{train}} * TT_{train,n,t} + \beta_{TC} * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} \\
& * (\frac{Income}{AverageIncome})^{\lambda_{Income}} * TC_{train,n,t} + \beta_{Age_{train}} * Age_n) + \epsilon_{train,n,t}
\end{aligned}
\tag{A.5}
$$

$$
\begin{aligned}
U_{car,n,t} =& 1 * (\alpha_{car} + \beta_{TT_{car}} * TT_{car,n,t} + \beta_{TC} * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} \\
& * (\frac{Income}{AverageIncome})^{\lambda_{Income}} * TC_{car,n,t}) + \epsilon_{car,n,t}
\end{aligned}
\tag{A.6}
$$

$$
\begin{aligned}
U_{MC,n,t} =& 1 * (\beta_{TT_{MC}} * TT_{MC,n,t} + \beta_{TC} * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} \\
& * (\frac{Income}{AverageIncome})^{\lambda_{Income}} * TC_{MC,n,t} + \beta_{Age_{MC}} * Age_n) + \epsilon_{MC,n,t}
\end{aligned}
\tag{A.7}
$$

$$
\begin{aligned}
U_{taxi,n,t} =& 1 * (\alpha_{taxi} + \beta_{TT_{taxi}} * TT_{Cartaxi,n,t} + \beta_{TC} * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} \\
& * (\frac{Income}{AverageIncome})^{\lambda_{Income}} * TC_{Cartaxi,n,t}) + \epsilon_{taxi,n,t}
\end{aligned}
\tag{A.8}
$$

$$
\begin{aligned}
U_{ODT,n,t} =& 1 * (\alpha_{ODT} + \beta_{TT_{ODT}} * TT_{CarODT,n,t} + \beta_{TC} \\
& * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} * (\frac{Income}{AverageIncome})^{\lambda_{Income}} * TC_{CarODT,n,t} \\
& + \beta_{Male_{ODT}} * Gender_{Male,n} + \beta_{Education_{ODT}} * Education_{University,n} \\
& + \beta_{Age_{ODT}} * Age_n) + \epsilon_{ODT,n,t}
\end{aligned}
\tag{A.9}
$$

$$U_{taxi,n,t} = 1 * (\alpha_{taxi} + \beta_{TT_{taxi}} * TT_{MCtaxi,n,t} + \beta_{TC} * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} * (\frac{Income}{AverageIncome})^{\lambda_{Income}} * TC_{MCtaxi,n,t}) + \epsilon_{taxi,n,t} \tag{A.10}$$

$$U_{ODT,n,t} = 1 * (\alpha_{ODT} + \beta_{TT_{ODT}} * TT_{MCODT,n,t} + \beta_{TC} * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} * (\frac{Income}{AverageIncome})^{\lambda_{Income}} * TC_{MCODT,n,t} + \beta_{Male_{ODT}} * Gender_{Male,n} + \beta_{Education_{ODT}} * Education_{University,n} + \beta_{Age_{ODT}} * Age_n) + \epsilon_{ODT,n,t} \tag{A.11}$$

$$U_{walk_{SP},n,t} = \sigma_{SP} * (\alpha_{walk} + \beta_{TT_{walk}} * TT_{Walk_{SP,n,t}} + \beta_{Age_{walk}} * Age_n) + \epsilon_{walk_{SP},n,t} \tag{A.12}$$

$$U_{PT_{SP},n,t} = \sigma_{SP} * (\alpha_{PT} + \beta_{TT_{PT}} * TT_{PT_{SP,n,t}} + \beta_{TC} * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} * (\frac{Income}{AverageIncome})^{\lambda_{Income}} * TC_{PT_{SP,n,t}}) + \epsilon_{PT_{SP},n,t} \tag{A.13}$$

$$U_{car_{SP},n,t} = \sigma_{SP} * (\alpha_{car} + \beta_{TT_{car}} * TT_{car_{SP,n,t}} + \beta_{TC} * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} * (\frac{Income}{AverageIncome})^{\lambda_{Income}} * TC_{car_{SP,n,t}} + \beta_{CC} * CC_{car_{SP,n,t}}) + \epsilon_{car_{SP},n,t} \tag{A.14}$$

$$
\begin{aligned}
U_{MC_{SP},n,t} =& \sigma_{SP} * (\beta_{TT_{MC}} * TT_{MC_{SP},n,t} + \beta_{TC} \\
& * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} * (\frac{Income}{AverageIncome})^{\lambda_{Income}} \\
& * TC_{MC_{SP},n,t} + \beta_{Age_{MC}} * Age_{MC,n} + \beta_{CC} * CC_{MC_{SP},n,t}) + \epsilon_{MC_{SP},n,t}
\end{aligned}
\tag{A.15}
$$

$$
\begin{aligned}
U_{Taxi_{SP},n,t} =& \sigma_{SP} * (\alpha_{Taxi} + \beta_{TT_{Taxi}} * TT_{CarTaxi_{SP},n,t} \\
& + \beta_{TC} * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} * (\frac{Income}{AverageIncome})^{\lambda_{Income}} \\
& * TC_{Cartaxi_{SP},n,t} + \beta_{CC} * CC_{Cartaxi_{SP},n,t}) + \epsilon_{Taxi_{SP},n,t}
\end{aligned}
\tag{A.16}
$$

$$
\begin{aligned}
U_{ODT_{SP},n,t} =& \sigma_{SP} * (\alpha_{ODT} + \beta_{TT_{ODT}} * TT_{CarODT_{SP},n,t} + \beta_{TC} \\
& * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} * (\frac{Income}{AverageIncome})^{\lambda_{Income}} * TC_{CarODT_{SP},n,t} \\
& + \beta_{CC} * CC_{CarODT_{SP},n,t} + \beta_{Male_{ODT}} * Gender_{Male,n} + \beta_{Education_{ODT}} \\
& * Education_{University,n} + \beta_{Age_{ODT}} * Age_{n}) + \epsilon_{ODt_{SP},n,t}
\end{aligned}
\tag{A.17}
$$

$$
\begin{aligned}
U_{Taxi_{SP},n,t} =& \sigma_{SP} * (\alpha_{Taxi} + \beta_{TT_{Taxi}} * TT_{MCTaxi_{SP},n,t} \\
& + \beta_{TC} * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} * (\frac{Income}{AverageIncome})^{\lambda_{Income}} \\
& * TC_{MCTaxi_{SP},n,t} + \beta_{CC} * CC_{MCtaxi_{SP},n,t}) + \epsilon_{Taxi_{SP},n,t}
\end{aligned}
\tag{A.18}
$$

$$\begin{aligned} U_{ODT_{SP},n,t} =& \sigma_{SP} * (\alpha_{ODT} + \beta_{TT_{ODT}} * TT_{MCODT_{SP},n,t} + \beta_{TC} \\ & * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} * (\frac{Income}{AverageIncome})^{\lambda_{Income}} * TC_{MCODT_{SP},n,t} \\ & + \beta_{CC} * CC_{MCODT_{SP},n,t} + \beta_{Male_{ODT}} * Gender_{Male,n} + \beta_{Education_{ODT}} \\ & * Education_{University,n} + \beta_{Age_{ODT}} * Age_n) + \epsilon_{ODT_{SP},n,t} \end{aligned} \tag{A.19}$$

$$\begin{aligned} U_{UAM_{SP},n,t} =& \sigma_{SP} * (\alpha_{UAM} + \beta_{TT_{UAM}} * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} TT_{UAM_{SP},n,t} \\ & + \beta_{AT_{UAM}} * AT_{UAM_{SP},n,t} + \beta_{TC} * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} \\ & * (\frac{Income}{AverageIncome})^{\lambda_{Income}} * TC_{UAM_{SP},n,t} + \beta_{Male_{UAM}} * Gender_{Male,n} \\ & + \beta_{Location_{UAM}} * Location_{Agglomeration,n} + \beta_{Education_{UAM}} \\ & * Education_{University,n}) + \epsilon_{MC_{UAM},n,t} \end{aligned} \tag{A.20}$$

$$U_{walk,n,t} = 1 * (\alpha_{walk} + \beta_{TT_{walk}} * TT_{walk,n,t} + \beta_{Age} * Age_n) + \epsilon_{walk,n,t} \tag{A.21}$$

$$U_{bike,n,t} = 1 * (\alpha_{bike} + \beta_{TT_{bike}} * TT_{bike,n,t}) + \epsilon_{bike,n,t} \tag{A.22}$$

$$\begin{aligned} U_{PT,n,t} =& 1 * (\alpha_{PT} + \beta_{TT_{bus}} * TT_{bus,n,t} + \beta_{TC} * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} \\ & * (\frac{Income}{AverageIncome})^{\lambda_{Income}} * TC_{PT,n,t}) + \epsilon_{PT,n,t} \end{aligned} \tag{A.23}$$

$$U_{PT,n,t} = 1 * (\alpha_{PT} + \beta_{TT_{PT}} * TT_{BRT,n,t} + \beta_{TC} * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} * (\frac{Income}{AverageIncome})^{\lambda_{Income}} * TC_{BRT,n,t}) + \epsilon_{PT,n,t} \tag{A.24}$$

$$U_{PT,n,t} = 1 * (\alpha_{PT} + \beta_{TT_{PT}} * TT_{train,n,t} + \beta_{TC} * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} * (\frac{Income}{AverageIncome})^{\lambda_{Income}} * TC_{train,n,t} + \beta_{Age_{train}} * Age_{n}) + \epsilon_{PT,n,t} \tag{A.25}$$

$$U_{car,n,t} = 1 * (\alpha_{car} + \beta_{TT_{car}} * TT_{car,n,t} + \beta_{Car} * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} * (\frac{Income}{AverageIncome})^{\lambda_{Income}} * TC_{car,n,t}) + \epsilon_{car,n,t} \tag{A.26}$$

$$U_{MC,n,t} = 1 * (\beta_{TT_{MC}} * TT_{MC,n,t} + \beta_{TC} * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} * (\frac{Income}{AverageIncome})^{\lambda_{Income}} * TC_{MC,n,t} + \beta_{Age_{MC}} * Age_{n}) + \epsilon_{MC,n,t} \tag{A.27}$$

$$U_{taxi,n,t} = 1 * (\alpha_{taxi} + \beta_{TT_{taxi}} * TT_{Cartaxi,n,t} + \beta_{TC} * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} * (\frac{Income}{AverageIncome})^{\lambda_{Income}} * TC_{Cartaxi,n,t}) + \epsilon_{taxi,n,t} \tag{A.28}$$

$$\begin{aligned} U_{ODT,n,t} =& 1 * (\alpha_{ODT} + \beta_{TT_{ODT}} * TT_{CarODT,n,t} + \beta_{TC} \\ & * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} * (\frac{Income}{AverageIncome})^{\lambda_{Income}} * TC_{CarODT,n,t} \\ & + \beta_{Male_{ODT}} * Gender_{Male,n} + \beta_{Education_{ODT}} * Education_{University,n} \\ & + \beta_{Age_{ODT}} * Age_n) + \epsilon_{ODT,n,t} \end{aligned} \tag{A.29}$$

$$\begin{aligned} U_{taxi,n,t} =& 1 * (\alpha_{taxi} + \beta_{TT_{taxi}} * TT_{MCtaxi,n,t} + \beta_{TC} * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} \\ & * (\frac{Income}{AverageIncome})^{\lambda_{Income}} * TC_{MCtaxi,n,t}) + \epsilon_{taxi,n,t} \end{aligned} \tag{A.30}$$

$$\begin{aligned} U_{ODT,n,t} =& 1 * (\alpha_{ODT} + \beta_{TT_{ODT}} * TT_{MCODT,n,t} + \beta_{TC} \\ & * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} * (\frac{Income}{AverageIncome})^{\lambda_{Income}} * TC_{MCODT,n,t} \\ & + \beta_{Male_{ODT}} * Gender_{Male,n} + \beta_{Education_{ODT}} * Education_{University,n} \\ & + \beta_{Age_{ODT}} * Age_n) + \epsilon_{ODT,n,t} \end{aligned} \tag{A.31}$$

$$U_{walk_{SP},n,t} = \sigma_{SP} * (\alpha_{walk} + \beta_{TT_{walk}} * TT_{Walk_{SP,n,t}} + \beta_{Age_{walk}} * Age_n) + \epsilon_{walk_{SP},n,t} \tag{A.32}$$

$$\begin{aligned} U_{PT_{SP},n,t} =& \sigma_{SP} * (\alpha_{PT} + \beta_{TT_{PT}} * TT_{PT_{SP,n,t}} \\ & + \beta_{TC} * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} * (\frac{Income}{AverageIncome})^{\lambda_{Income}} \\ & * TC_{PT_{SP,n,t}}) + \epsilon_{PT_{SP},n,t} \end{aligned} \tag{A.33}$$

$$\begin{aligned} U_{car_{SP},n,t} =& \sigma_{SP} * (\alpha_{car} + \beta_{TT_{car}} * TT_{car_{SP,n,t}} \\ & + \beta_{TC} * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} * (\frac{Income}{AverageIncome})^{\lambda_{Income}} \\ & * TC_{car_{SP,n,t}} + \beta_{CC} * CC_{car_{SP,n,t}}) + \epsilon_{car_{SP},n,t} \end{aligned} \tag{A.34}$$

$$\begin{aligned} U_{MC_{SP},n,t} =& \sigma_{SP} * (\beta_{TT_{MC}} * TT_{MC_{SP,n,t}} + \beta_{TC} \\ & * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} * (\frac{Income}{AverageIncome})^{\lambda_{Income}} \\ & * TC_{MC_{SP,n,t}} + \beta_{Age_{MC}} * Age_{MC,n} + \beta_{CC} * CC_{MC_{SP,n,t}}) + \epsilon_{MC_{SP},n,t} \end{aligned} \tag{A.35}$$

$$\begin{aligned} U_{Taxi_{SP},n,t} =& \sigma_{SP} * (\alpha_{Taxi} + \beta_{TT_{Taxi}} * TT_{CarTaxi_{SP,n,t}} \\ & + \beta_{TC} * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} * (\frac{Income}{AverageIncome})^{\lambda_{Income}} \\ & * (TC_{Cartaxi_{SP,n,t}} + CC_{Cartaxi_{SP,n,t}})) + \epsilon_{Taxi_{SP},n,t} \end{aligned} \tag{A.36}$$

$$\begin{aligned} U_{ODT_{SP},n,t} =& \sigma_{SP} * (\alpha_{ODT} + \beta_{TT_{ODT}} * TT_{CarODT_{SP,n,t}} + \beta_{TC} \\ & * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} * (\frac{Income}{AverageIncome})^{\lambda_{Income}} \\ & * (TC_{CarODT_{SP,n,t}} + CC_{CarODT_{SP,n,t}}) + \beta_{Male_{ODT}} * Gender_{Male,n} \\ & + \beta_{Education_{ODT}} * Education_{University,n} + \beta_{Age_{ODT}} * Age_{n}) + \epsilon_{ODt_{SP},n,t} \end{aligned} \tag{A.37}$$

$$\begin{aligned} U_{Taxi_{SP},n,t} =& \sigma_{SP} * (\alpha_{Taxi} + \beta_{TT_{Taxi}} * TT_{MCTaxi_{SP,n,t}} \\ & + \beta_{TC} * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} * (\frac{Income}{AverageIncome})^{\lambda_{Income}} \\ & * (TC_{MCTaxi_{SP,n,t}} + CC_{MCtaxi_{SP,n,t}})) + \epsilon_{Taxi_{SP},n,t} \end{aligned} \tag{A.38}$$

$$\begin{aligned} U_{ODT_{SP},n,t} =& \sigma_{SP} * (\alpha_{ODT} + \beta_{TT_{ODT}} * TT_{MCODT_{SP},n,t} + \beta_{TC} \\ & * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} * (\frac{Income}{AverageIncome})^{\lambda_{Income}} \\ & * (TC_{MCODT_{SP},n,t} + CC_{MCODT_{SP},n,t}) + \beta_{Male_{ODT}} * Gender_{Male,n} \\ & + \beta_{Education_{ODT}} * Education_{University,n} + \beta_{Age_{ODT}} * Age_n) + \epsilon_{ODT_{SP},n,t} \end{aligned} \tag{A.39}$$

$$\begin{aligned} U_{UAM_{SP},n,t} =& \sigma_{SP} * (\alpha_{UAM} + \beta_{TT_{UAM}} * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} TT_{UAM_{SP},n,t} \\ & + \beta_{AT_{UAM}} * AT_{UAM_{SP},n,t} + \beta_{TC} * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} \\ & * (\frac{Income}{AverageIncome})^{\lambda_{Income}} * TC_{UAM_{SP},n,t} + \beta_{Male_{UAM}} * Gender_{Male,n} \\ & + \beta_{Location_{UAM}} * Location_{Agglomeration,n} + \beta_{Education_{UAM}} \\ & * Education_{University,n}) + \epsilon_{MC_{UAM},n,t} \end{aligned} \tag{A.40}$$

$$U_{walk,n,t} = 1 * (\alpha_{walkRND} + \beta_{TT_{walk}} * TT_{walk,n,t} + \beta_{Age} * Age_n) + \epsilon_{walk,n,t} \tag{A.41}$$

$$U_{bike,n,t} = 1 * (\alpha_{bikeRND} + \beta_{TT_{bike}} * TT_{bike,n,t}) + \epsilon_{bike,n,t} \tag{A.42}$$

$$\begin{aligned} U_{PT,n,t} =& 1 * (\alpha_{PTRND} + \beta_{TT_{bus}} * TT_{bus,n,t} + \beta_{TC} * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} \\ & * (\frac{Income}{AverageIncome})^{\lambda_{Income}} * TC_{PT,n,t}) + \epsilon_{PT,n,t} \end{aligned} \tag{A.43}$$

$$U_{PT,n,t} = 1 * (\alpha_{PT RND} + \beta_{TT_{PT}} * TT_{BRT,n,t} + \beta_{TC} * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} * (\frac{Income}{AverageIncome})^{\lambda_{Income}} * TC_{BRT,n,t}) + \epsilon_{PT,n,t} \tag{A.44}$$

$$U_{PT,n,t} = 1 * (\alpha_{PT RND} + \beta_{TT_{PT}} * TT_{train,n,t} + \beta_{TC} * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} * (\frac{Income}{AverageIncome})^{\lambda_{Income}} * TC_{train,n,t} + \beta_{Age_{train}} * Age_n) + \epsilon_{PT,n,t} \tag{A.45}$$

$$U_{car,n,t} = 1 * (\alpha_{car RND} + \beta_{TT_{car}} * TT_{car,n,t} + \beta_{Car} * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} * (\frac{Income}{AverageIncome})^{\lambda_{Income}} * TC_{car,n,t}) + \epsilon_{car,n,t} \tag{A.46}$$

$$U_{MC,n,t} = 1 * (\beta_{TT_{MC}} * TT_{MC,n,t} + \beta_{TC} * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} * (\frac{Income}{AverageIncome})^{\lambda_{Income}} * TC_{MC,n,t} + \beta_{Age_{MC}} * Age_n) + \epsilon_{MC,n,t} \tag{A.47}$$

$$U_{taxi,n,t} = 1 * (\alpha_{taxi RND} + \beta_{TT_{taxi}} * TT_{Cartaxi,n,t} + \beta_{TC} * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} * (\frac{Income}{AverageIncome})^{\lambda_{Income}} * TC_{Cartaxi,n,t}) + \epsilon_{taxi,n,t} \tag{A.48}$$

$$\begin{aligned}U_{ODT,n,t} =& 1 * (\alpha_{ODTRND} + \beta_{TT_{ODT}} * TT_{CarODT,n,t} + \beta_{TC} \\ & * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} * (\frac{Income}{AverageIncome})^{\lambda_{Income}} * TC_{CarODT,n,t} \\ & + \beta_{Male_{ODT}} * Gender_{Male,n} + \beta_{Education_{ODT}} * Education_{University,n} \\ & + \beta_{Age_{ODT}} * Age_n) + \epsilon_{ODT,n,t}\end{aligned} \tag{A.49}$$

$$\begin{aligned}U_{taxi,n,t} =& 1 * (\alpha_{taxiRND} + \beta_{TT_{taxi}} * TT_{MCtaxi,n,t} + \beta_{TC} * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} \\ & * (\frac{Income}{AverageIncome})^{\lambda_{Income}} * TC_{MCtaxi,n,t}) + \epsilon_{taxi,n,t}\end{aligned} \tag{A.50}$$

$$\begin{aligned}U_{ODT,n,t} =& 1 * (\alpha_{ODTRND} + \beta_{TT_{ODT}} * TT_{MCODT,n,t} + \beta_{TC} \\ & * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} * (\frac{Income}{AverageIncome})^{\lambda_{Income}} * TC_{MCODT,n,t} \\ & + \beta_{Male_{ODT}} * Gender_{Male,n} + \beta_{Education_{ODT}} * Education_{University,n} \\ & + \beta_{Age_{ODT}} * Age_n) + \epsilon_{ODT,n,t}\end{aligned} \tag{A.51}$$

$$U_{walk_{SP},n,t} = \sigma_{SP} * (\alpha_{walkRND} + \beta_{TT_{walk}} * TT_{Walk_{SP},n,t} + \beta_{Age_{walk}} * Age_n) + \epsilon_{walk_{SP},n,t} \tag{A.52}$$

$$\begin{aligned}U_{PT_{SP},n,t} =& \sigma_{SP} * (\alpha_{PTRND} + \beta_{TT_{PT}} * TT_{PT_{SP,n,t}} \\ & + \beta_{TC} * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} * (\frac{Income}{AverageIncome})^{\lambda_{Income}} \\ & * TC_{PT_{SP,n,t}}) + \epsilon_{PT_{SP},n,t}\end{aligned} \tag{A.53}$$

$$
\begin{aligned}
U_{car_{SP},n,t} =& \sigma_{SP} * (\alpha_{carRND} + \beta_{TT_{car}} * TT_{car_{SP,n,t}} \\
& + \beta_{TC} * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} * (\frac{Income}{AverageIncome})^{\lambda_{Income}} \\
& * TC_{car_{SP,n,t}} + \beta_{CC} * CC_{car_{SP,n,t}}) + \epsilon_{car_{SP},n,t}
\end{aligned} \tag{A.54}
$$

$$
\begin{aligned}
U_{MC_{SP},n,t} =& \sigma_{SP} * (\beta_{TT_{MC}} * TT_{MC_{SP,n,t}} + \beta_{TC} \\
& * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} * (\frac{Income}{AverageIncome})^{\lambda_{Income}} \\
& * TC_{MC_{SP,n,t}} + \beta_{Age_{MC}} * Age_{MC,n} + \beta_{CC} * CC_{MC_{SP,n,t}}) + \epsilon_{MC_{SP},n,t}
\end{aligned} \tag{A.55}
$$

$$
\begin{aligned}
U_{Taxi_{SP},n,t} =& \sigma_{SP} * (\alpha_{TaxiRND} + \beta_{TT_{Taxi}} * TT_{CarTaxi_{SP,n,t}} \\
& + \beta_{TC} * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} * (\frac{Income}{AverageIncome})^{\lambda_{Income}} \\
& * (TC_{Cartaxi_{SP,n,t}} + CC_{Cartaxi_{SP,n,t}})) + \epsilon_{Taxi_{SP},n,t}
\end{aligned} \tag{A.56}
$$

$$
\begin{aligned}
U_{ODT_{SP},n,t} =& \sigma_{SP} * (\alpha_{ODTRND} + \beta_{TT_{ODT}} * TT_{CarODT_{SP,n,t}} + \beta_{TC} \\
& * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} * (\frac{Income}{AverageIncome})^{\lambda_{Income}} \\
& * (TC_{CarODT_{SP,n,t}} + CC_{CarODT_{SP,n,t}}) + \beta_{Male_{ODT}} * Gender_{Male,n} \\
& + \beta_{Education_{ODT}} * Education_{University,n} + \beta_{Age_{ODT}} * Age_{n}) + \epsilon_{ODt_{SP},n,t}
\end{aligned} \tag{A.57}
$$

$$
\begin{aligned}
U_{Taxi_{SP},n,t} =& \sigma_{SP} * (\alpha_{TaxiRND} + \beta_{TT_{Taxi}} * TT_{MCTaxi_{SP,n,t}} \\
& + \beta_{TC} * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} * (\frac{Income}{AverageIncome})^{\lambda_{Income}} \\
& * (TC_{MCTaxi_{SP,n,t}} + CC_{MCtaxi_{SP,n,t}})) + \epsilon_{Taxi_{SP},n,t}
\end{aligned} \tag{A.58}
$$

$$
\begin{aligned}
U_{ODT_{SP},n,t} =& \sigma_{SP} * (\alpha_{ODTRND} + \beta_{TT_{ODT}} * TT_{MCODT_{SP},n,t} + \beta_{TC} \\
& * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} * (\frac{Income}{AverageIncome})^{\lambda_{Income}} \\
& * (TC_{MCODT_{SP},n,t} + CC_{MCODT_{SP},n,t}) + \beta_{Male_{ODT}} * Gender_{Male,n} \\
& + \beta_{Education_{ODT}} * Education_{University,n} + \beta_{Age_{ODT}} * Age_{n}) + \epsilon_{ODT_{SP},n,t}
\end{aligned}
\tag{A.59}
$$

$$
\begin{aligned}
U_{UAM_{SP},n,t} =& \sigma_{SP} * (\alpha_{UAMRND} + \beta_{TT_{UAM}} * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} TT_{UAM_{SP},n,t} \\
& + \beta_{AT_{UAM}} * AT_{UAM_{SP},n,t} + \beta_{TC} * (\frac{Distance}{AverageDistance})^{\lambda_{Distance}} \\
& * (\frac{Income}{AverageIncome})^{\lambda_{Income}} * TC_{UAM_{SP},n,t} + \beta_{Male_{UAM}} * Gender_{Male,n} \\
& + \beta_{Location_{UAM}} * Location_{Agglomeration,n} + \beta_{Education_{UAM}} \\
& * Education_{University,n}) + \epsilon_{MC_{UAM},n,t}
\end{aligned}
\tag{A.60}
$$

BIBLIOGRAPHY

Adnan, M., F. Pereira, C. Lima Azevedo, K. Basak, M. Lovric, S. Raveau, Y. Zhu, J. Ferreira, C. Zegras and M. Ben-Akiva (2016) Simmobility: A multi-scale integrated agent-based simulation platform, 01 2016.

Agarwal, S. and K. M. Koo (2016) Impact of electronic road pricing (erp) changes on transport modal choice, *Regional Science and Urban Economics*, **60**, 1 – 11.

Ahmed, S. S., K. F. Hulme, G. Fountas, U. Eker, I. V. Benedyk, S. E. Still and P. C. Anastasopoulos (2020) The flying car—challenges and strategies toward future adoption, *Frontiers in Built Environment*, **6**, 106.

Akaike, H. (1974) A new look at the statistical model identification, in *Selected Papers of Hirotugu Akaike*, 215–222, Springer, Berlin.

Al Haddad, C., E. Chaniotakis, A. Straubinger, K. Plötner and C. Antoniou (2020) Factors affecting the adoption and use of urban air mobility, *Transportation Research Part A: Policy and Practice*, **132**, 696 – 712.

AngryWorkersWorld (2019) Gojek: Delivery workers struggle in indonesia, `http://libcom.org/blog/gojek-delivery-workers-struggle-indonesia-28062019`. Accessed: 2019-07-23.

Anya, A. and D. A. Wardhani (2016) No 3-in-1? get ready to get stuck in traffic, `https://www.thejakartapost.com/news/2016/03/30/no-3-1-get-ready-get-stuck-traffic.html`. Accessed: 2017-07-11.

Arentze, T. A. and H. J. Timmermans (2004) A learning-based transportation oriented simulation system, *Transportation Research Part B: Methodological*, **38** (7) 613 – 633.

Atasoy, B., A. Glerum and M. Bierlaire (2013) Attitudes towards mode choice in switzerland, *disP - The Planning Review*, **49** (2) 101–117.

Axhausen, K. (1989) Eine ereignirsorientierte simulation von aktivitätenketten zur parkstandswahl /, 01 1989.

Axhausen, K. W. (1995) Travel diaries: An annotated catalog, vol. 2, Institut fur Strassenbau und Verkehrsplanung.

Axhausen, K. W. (2008) Social networks, mobility biographies, and travel: Survey challenges, *Environment and Planning B: Planning and Design*, **35** (6) 981–996.

Axhausen, K. W. and T. Gärling (1992) Activity-based approaches to travel analysis: conceptual frameworks, models, and research problems, *Transport Reviews*, **12** (4) 323–341.

Axhausen, K. W. and C. Weis (2010) Predicting response rate: A natural experiment, *Survey Practice*, **3** (2).

Axhausen, K. W., A. Zimmermann, S. Schönfelder, G. Rindsfüser and T. Haupt (2002) Observing the rhythms of daily life: A six-week travel diary, *Transportation*, **29** (2) 95–124, May 2002.

Balac, M., H. Becker, F. Ciari and K. W. Axhausen (2019) Modeling competing free-floating carsharing operators – a case study for zurich, switzerland, *Transportation Research Part C: Emerging Technologies*, **98**, 101 – 117.

Balac, M., R. L. Rothfeld and S. Hörl (2019a) The prospects of on-demand urban air mobility in zurich, switzerland, paper presented at the *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, 906–913, Oct 2019.

Balac, M., A. R. Vetrella, R. Rothfeld and B. Schmid (2018) Demand estimation for aerial vehicles in urban settings, *IEEE Intelligent Transportation Systems Magazine*.

Balac, M., A. R. Vetrella, R. Rothfeld and B. Schmid (2019b) Demand estimation for aerial vehicles in urban settings, *IEEE Intelligent Transportation Systems Magazine*, **11** (3) 105–116.

Balmer, M., K. W. Axhausen and K. Nagel (2006) Agent-based demand-modeling framework for large-scale microsimulations, *Transportation Research Record*, **1985**, 125–134.

Barthelemy, J. and P. L. Toint (2013) Synthetic population generation without a sample, *Transportation Science*, **47** (2) 266–279.

Beckman, R. J., K. A. Baggerly and M. D. McKay (1996) Creating synthetic baseline populations, *Transportation Research Part a-Policy and Practice*, **30** (6) 415–429.

Belgiawan, P. F., I. Dubernet, B. Schmid and K. Axhausen (2019a) Context-dependent models (crrm, murrm, prrm, ram) versus a context-free model (mnl) in transportation studies: a comprehensive comparisons for swiss and german sp and rp data sets, *Transportmetrica A: Transport Science*, **15** (2) 1487–1521.

Belgiawan, P. F., A. Ilahi and K. W. Axhausen (2019b) Influence of pricing on mode choice decision in jakarta: A random regret minimization model, *Case Studies on Transport Policy*, **7** (1) 87 – 95.

Benkler, Y. (2002) Coase's penguin, or, linux and "the nature of the firm", *The Yale Law Journal*, **112** (3) 369–446.

Borgers, A., F. Hofman and H. Timmermans (2001) Conditional choice modelling of time allocation among spouses in transport settings, *European Journal of Transport and Infrastructure Research*, **2** (1).

Borowiak, C. and M. Ji (2019) Taxi co-ops versus uber: Struggles for workplace democracy in the sharing economy, *Journal of Labor and Society*, **22** (1) 165–185.

BPS-Statistics (2016a) *Banten in Figures*, BPS-Statistics of Banten Province.

BPS-Statistics (2016b) *Jakarta in Figures*, BPS-Statistics of DKI Jakarta Province.

BPS-Statistics (2016c) *West Java in figures*, BPS-Statistics of West Java Province.

Bradley, M. and P. Vovsha (2005) A model for joint choice of daily activity pattern typesof household members, *Transportation*, **32** (5) 545–571, Sep 2005.

Casati, D., K. Muller, P. J. Fourie, A. Erath and K. W. Axhausen (2015) Synthetic population generation by combining a hierarchical, simulation-based approach with reweighting by generalized raking, *Transportation Research Record*, **2493**, 107–116.

Cervero, R. (1991) Paratransit in southeast asia: A market response to poor roads?*, *Review of Urban & Regional Development Studies*, **3** (1) 3–27.

Cherchi, E. and J. d. D. Ortúzar (2011) On the use of mixed rp/sp models in prediction: Accounting for systematic and random taste heterogeneity, *Transportation Science*, **45** (1) 98–108.

ChoiceMetrics (2014) *Ngene 1.1.2 user manual: The Cutting Edge in Experimental Design, Choice Metrics*.

City of Copenhagen (2013) Bicycle statistics, `https://web.archive.org/web/20131212093813/http://subsite.kk.dk/sitecore/content/Subsites/CityOfCopenhagen/SubsiteFrontpage/LivingInCopenhagen/CityAndTraffic/CityOfCyclists/CycleStatistics.aspx`. Accessed: 2019-07-15.

CMEA and JICA (2012) *Jabodetabek Urban Transportation Policy Integration Project in the Republic of Indonesia*, Final Report, Jakarta.

Cohen, A., J. Guan, M. Beamer, R. Dittoe and S. Mokhtarimousavi (2020) *Reimagining the Future of Transportation with Personal Flight: Preparing and Planning for Urban Air Mobility*, UC Berkeley: Transportation Sustainability Research Center.

Contreras, S. D. and A. Paz (2018) The effects of ride-hailing companies on the taxicab industry in las vegas, nevada, *Transportation Research Part A: Policy and Practice*, **115**, 63 – 70. Smart urban mobility.

Cowell, R. G., P. Dawid, S. L. Lauritzen and D. J. Spiegelhalter (2006) *Probabilistic networks and expert systems: Exact computational methods for Bayesian networks*, Springer Science & Business Media, Berlin.

de Freitas, L. M., O. Schuemperlin, M. Balac and F. Ciari (2017) Equity effects of congestion charges: An exploratory analysis with matsim, *Transportation Research Record*, **2670** (1) 75–82.

de Palma, A. and R. Lindsey (2006) Modelling and evaluation of road pricing in paris, *Transport Policy*, **13** (2) 115 – 126. Modelling of Urban Road Pricing and its Implementation.

Deming, W. E. and F. F. Stephan (1940) On a least squares adjustment of a sampled frequency table when the expected marginal totals are known, *Annals of Mathematical Statistics*, **11**, 427–444.

Dharmowijoyo, D. B., Y. O. Susilo, A. Karlström and L. S. Adiredja (2015) Collecting a multi-dimensional three-weeks household time-use and activity diary in the bandung metropolitan area, indonesia, *Transportation Research Part A: Policy and Practice*, **80**, 231 – 246.

Dharmowijoyo, D. B. E., Y. O. Susilo and A. Karlström (2016) Day-to-day variability in travellers' activity-travel patterns in the jakarta metropolitan area, *Transportation*, **43** (4) 601–621, Jul 2016.

Dias, F. F., P. S. Lavieri, V. M. Garikapati, S. Astroza, R. M. Pendyala and C. R. Bhat (2017) A behavioral choice model of the use of car-sharing and ride-sourcing services, *Transportation*, **44** (6) 1307–1323, Nov 2017.

Dickey, M. R. (2020) Here's how much uber's flying taxi service will cost, `https://techcrunch.com/2018/05/08/heres-how-much-ubers-flying-taxi-service-will-cost/`. Accessed: 2020-10-25.

Downing, S. (2019) 7 urban air mobility companies to watch, `https://www.greenbiz.com/article/7-urban-air-mobility-companies-watch`. Accessed: 2019-12-23.

Dubernet, T. and K. W. Axhausen (2015) Implementing a household joint activity-travel multi- agent simulation tool: first results, *Transportation*, **42** (5) 753–769, Sep 2015.

Eker, U., S. S. Ahmed, G. Fountas and P. C. Anastasopoulos (2019) An exploratory investigation of public perceptions towards safety and security from the future use of flying cars in the united states, *Analytic Methods in Accident Research*, **23**, 100103.

Eker, U., G. Fountas and P. C. Anastasopoulos (2020a) An exploratory empirical analysis of willingness to pay for and use flying cars, *Aerospace Science and Technology*, **104**, 105993.

Eker, U., G. Fountas, P. C. Anastasopoulos and S. E. Still (2020b) An exploratory investigation of public perceptions towards key benefits and concerns from the future use of flying cars, *Travel Behaviour and Society*, **19**, 54 – 66.

Eliasson, J. and L.-G. Mattsson (2006) Equity effects of congestion pricing: Quantitative methodology and a case study for stockholm, *Transportation Research Part A: Policy and Practice*, **40** (7) 602 – 620.

Erath, A., P. J. Fourie, M. A. van Eggermond, S. A. Ordonez Medina, A. Chakirov and K. W. Axhausen (2012) Large-scale agent-based transport demand model for singapore, *Arbeitsberichte Verkehrs-und Raumplanung*, **790**.

ETH Zürich (2016) Getting started with clusters, `http://scicomp.ethz.ch/wiki/Getting_started_with_clusters`. Accessed: 2019-04-17.

Farooq, B., M. Bierlaire, R. Hurtubia and G. Flotterod (2013) Simulation based population synthesis, *Transportation Research Part B-Methodological*, **58**, 243–263.

Fu, M., R. Rothfeld and C. Antoniou (2019) Exploring preferences for transportation modes in an urban air mobility environment: Munich case study, *Transportation Research Record*, **2673** (10) 427–442.

Garrow, L. A., M. Ilbeigi and Z. Chen (2017) Forecasting demand for on demand mobility, paper presented at the *17th AIAA Aviation Technology, Integration, and Operations Conference*.

Gliebe, J. P. and F. S. Koppelman (2005) Modeling household activity–travel interactions as parallel constrained choices, *Transportation*, **32** (5) 449–471, Sep 2005.

Glover, F. and E. Taillard (1993) A user's guide to tabu search, *Annals of Operations Research*, **41** (1) 1–28.

Google (2019) Directions API, `https://developers.google.com/maps/documentation/directions/intro`. Accessed: 2019-07-10.

Google (2019) General transit feed specification reference, `https://developers.google.com/transit/gtfs/reference/`. Accessed: 2019-01-13.

Google (2019) Geocoding API. Accessed: 2019-07-10.

Governor of DKI Jakarta Province (2012) *Governor Regulation on the Revocation of Governor Regulation No 110/2012 on Traffic Control Area*, Governor Regulation 114/2016, Jakarta.

Guo, J. Y. and C. R. Bhat (2007) Population synthesis for microsimulating travel behavior, *Transportation Research Record*, **2014**, 92–101.

Habib, K. N. (2019) Mode choice modelling for hailable rides: An investigation of the competition of uber with other modes by using an integrated non-compensatory choice model with probabilistic choice set formation, *Transportation Research Part A: Policy and Practice*, **129**, 205 – 216.

Hadian, E. (2018) A demand and capacity analysis on bus semirapid transit network (case: Jabodetabek public transport network), paper presented at the *MATEC Web of Conferences*, vol. 181, 10001.

Hafezi, M. H. and M. A. Habib (2014) Synthesizing population for microsimulation-based integrated transport models using Atlantic Canada micro-data, *Procedia Computer Science*, **37**, 410–415.

Harsono, N. (2019) Grab unlocks rp 46t in additional income for drivers, merchants: Survey, `https://www.thejakartapost.com/news/2019/04/11/grab-unlocks-rp-46t-in-additional-income-for-drivers.html`. Accessed: 2019-10-21.

Henao, A. (2017) *Impacts of Ridesourcing-Lyft and Uber-on Transportation Including VMT, Mode Replacement, Parking, and Travel Behavior*, University of Colorado at Denver.

Hörl, S. and M. Balac (2020) Open data travel demand synthesis for agent-based transport simulation: A case study of paris and ile-de-france, *Arbeitsberichte Verkehrsund Raumplanung*, **1499**.

Horni, A., K. Nagel and K. W. Axhausen (2016) *The Multi-Agent Transport Simulation MATSim*, Ubiquity Press, London.

Hornỳ, M. (2014) *Bayesian networks*, Boston University School of Public Health, Boston.

Hössinger, R., F. Aschauer, S. Jara-Díaz, S. Jokubauskaite, B. Schmid, S. Peer, K. W. Axhausen and R. Gerike (2020) A joint time-assignment and expenditure-allocation model: value of leisure and value of time assigned to travel for specific population segments, *Transportation*, **47** (3) 1439–1475.

Huynh, N., M. Namazi-Rad, P. Perez, M. J. Berryman and Q. Chen (2013) Generating a synthetic population in support of agent-based modeling of transportation in sydney, paper presented at the *20th International Congress on Modelling and Simulation (MODSIM 2013)*, 1357–1363.

Hörl, S., M. Balac and K. W. Axhausen (2018) A first look at bridging discrete choice modeling and agent-based microsimulation in matsim, paper presented at the *The 7th International Workshop on Agent-based Mobility, Traffic and Transportation Models,Methodologies and Applications (ABMTrans)*.

Hörl, S., M. Balac and K. W. Axhausen (2019) Pairing discrete mode choice models and agent-based transport simulation with matsim, paper presented at the *98th Annual Meeting of the Transportation Research Board (TRB)*.

Ilahi, A. and K. W. Axhausen (2017) Measuring accessibility using an activity based model approach in jabodetabek, in *17th Swiss Transport Research Conference (STRC 2017)*, Ascona.

Ilahi, A. and K. W. Axhausen (2019) Integrating bayesian network and generalized raking for population synthesis in greater jakarta, *Regional Studies, Regional Science*, **6** (1) 623–636.

Ilahi, A., M. Balac, A. Li and K. W. Axhausen (2019a) The first agent-based model of greater Jakarta integrated with a mode-choice model, *Procedia Computer Science*, **151**, 272 – 278.

Ilahi, A., M. Balać and K. W. Axhausen (2019b) Existing urban transportation in greater jakarta. results of agent-based modelling.

Ilahi, A., P. F. Belgiawan and K. W. Axhausen (2018) Influence of pricing on mode choice decision integrated with latent variable, paper presented at the *15th International Conference on Travel Behavior Research (IATBR 2018)*.

Ilahi, A., P. F. Belgiawan and K. W. Axhausen (2019c) Chapter 8 - influence of pricing on mode choice decision integrated with latent variable: The case of jakarta greater area, in K. G. Goulias and A. W. Davis (eds.) *Mapping the Travel Behavior Genome*, 125 – 143, Elsevier.

Ilahi, A., P. F. Belgiawan, M. Balać and K. W. Axhausen (2019d) Understanding travel and mode choice with emerging modes. a pooled sp and rp model in greater jakarta.

Ilahi, A., A. I. Waro and P. Sumarsono (2015) Public transport reform in indonesian cities, paper presented at the *Proceedings of the Eastern Asia Society for Transportation Studies*, vol. 10.

Irawan, M. Z., P. F. Belgiawan, A. K. M. Tarigan and F. Wijanarko (2019) To compete or not compete: Exploring the relationships between motorcycle-based ride-sourcing, motorcycle taxis, and public transport in the jakarta metropolitan area, *Transportation*, Jun 2019.

Jara-Díaz, S. R., M. A. Munizaga, P. Greeven, R. Guerra and K. Axhausen (2008) Estimating the value of leisure from a time allocation model, *Transportation Research Part B: Methodological*, **42** (10) 946 – 957.

JICA (2009) Traffic data collected under "the Jabodetabek urban transport policy integration", JICA, Tokyo.

Kaddoura, I. and B. Kickhöfer (2014) Optimal road pricing: Towards an agent-based marginal social cost approach, in *VSP working paper 14-01, TU Berlin, transport systems planning and transport telematics*.

Kitamura, R. (1988) An evaluation of activity-based travel analysis, *Transportation*, **15** (1) 9–34, Mar 1988.

Lam, C. T. and M. Liu (2017) Demand and consumer surplus in the on-demand economy: the case of ride sharing, *Social Science Electronic Publishing*, **17** (8) 376–388.

Lovelace, R. and D. Ballas (2013) 'truncate, replicate, sample': A method for creating integer weights for spatial microsimulation, *Computers, Environment and Urban Systems*, **41**, 1–11.

Mackie, P., M. Wardman, A. Fowkes, G. Whelan, J. Nellthorp and J. Bates (2003) Values of travel time savings uk, Institute of transport studies, University of Leeds.

McFadden, D. (1973) *Conditional Logit Analysis of Qualitative Choice Behavior*, BART impact studies final report series: Traveler behavior studies, Institute of Urban and Regional Development, University of California.

Medeiros, R. M., F. Duarte, F. Achmad and A. Jalali (2018) Merging ict and informal transport in jakarta's ojek system, *Transportation Planning and Technology*, **41** (3) 336–352.

Moeckel, R., K. Spiekermann and M. Wegener (2003) Creating a synthetic population, in *8th International Conference on Computers in Urban Planning and Urban Management (CUPUM)*, Vienna.

Molloy, J., B. Schmid, F. Becker and K. W. Axhausen (2019) mixl: An open-source r package for estimating complex choice models on large datasets, vol. 1408, Institute for Transport Planning and Systems (IVT), ETH Zurich, Zurich.

Müller, K. (2016) Accelerating weighted random sampling without replacement, in *Arbeitsberichte Verkehrs-und Raumplanung*, vol. 1141, IVT, ETH Zürich, Zürich.

Müller, K. (2017) A generalized approach to population synthesis, Ph.D. Thesis, ETH Zürich, Zürich.

Müller, K. and K. W. Axhausen (2011) Population synthesis for microsimulation: State of the art, in *90th Annual Meeting Transportation Research Board*, Washington.

Nagel, K. (2016) *The Multi-Agent Transport Simulation MATSim*, chap. Road Pricing, Ubiquity Press London, London.

Neumann, A., D. Röder and J. W. Joubert (2015) Towards a simulation of minibuses in south africa, *Journal of Transport and Land Use*, **8** (1) 137.

OECD (2019) Conversion rates purchasing power parities (ppp), `https://data.oecd.org/conversion/purchasing-power-parities-ppp.htm`. Accessed: 2020-10-25.

Pendyala, R., K. Konduri, Y.-C. Chiu, M. Hickman, H. Noh, P. Waddell, L. Wang, D. You and B. Gardner (2012) An integrated land use–transport model system with dynamic time-dependent activity-travel microsimulation, *91st Annual Meeting of the Transportation Research Board, Washington, DC*, **2303**, 12 2012.

Pepić, L. (2018) The sharing economy: Uber and its effect on taxi companies, *Acta Economica*, **16** (28) 123–136.

Peticca-Harris, A., N. deGama and M. N. Ravishankar (2018) Postcapitalist precarious work and those in the 'drivers' seat: Exploring the motivations and lived experiences of uber drivers in canada, *Organization*.

Poletti, F. (2016) Public transit mapping on multi-modal networks in matsim, Thesis.

Prayudyanto, N., Z. Tamin, R. Driejang and D. Umami (2013) Will jakarta road pricing reduce fuel consumption and emission, paper presented at the *Proceedings of the Eastern Asia Society for Transportation Studies*, vol. 9, 67.

Pritchard, D. R. and E. J. Miller (2012) Advances in population synthesis: fitting many attributes per agent and fitting to household and person margins simultaneously, *Transportation*, **39** (3) 685–704.

R Core Team and others (2013) R: A language and environment for statistical computing.

Rayle, L., D. Dai, N. Chan, R. Cervero and S. Shaheen (2016) Just a better taxi? a survey-based comparison of taxis, transit, and ridesourcing services in san francisco, *Transport Policy*, **45**, 168 – 178.

Reiche, C., R. Goyal, A. Cohen, J. Serrao, S. Kimmel, C. Fernando and S. Shaheen (2018) *Urban Air Mobility Market Study*, UC Berkeley: Transportation Sustainability Research Center.

Rissanen, J. (1978) Modeling by shortest data description, *Automatica*, **14** (5) 465–471.

Rogers, B. (2018) The sharing economy: Uber and its effect on taxi companies, *University of Chicago Law Review Dialogue*, **82**, 123–136.

Rotaris, L., R. Danielis, E. Marcucci and J. Massiani (2010) The urban road pricing scheme to curb pollution in milan, italy: Description, impacts and preliminary cost–benefit analysis assessment, *Transportation Research Part A: Policy and Practice*, **44** (5) 359 – 375.

Russell, J. (2018) Go-jek officially announces southeast asia expansion to fill void left by uber's exit, `https://techcrunch.com/2018/05/23/go-jek-officially-announces-southeast-asia-expansion/`. Accessed: 2019-10-21.

Saadi, I., B. Farooq, A. Mustafa, J. Teller and M. Cools (2018) An efficient hierarchical model for multi-source information fusion, *Expert Systems with Applications*, **110**, 352–362.

Saadi, I., A. Mustafa, J. Teller and M. Cools (2016a) Forecasting travel behavior using Markov chain-based approaches, *Transportation Research Part C: Emerging Technologies*, **69**, 402 – 417.

Saadi, I., A. Mustafa, J. Teller, B. Farooq and M. Cools (2016b) Hidden Markov model-based population synthesis, *Transportation Research Part B: Methodological*, **90**, 1 – 21.

Santos, G. (2005) Urban congestion charging: A comparison between london and singapore, *Transport Reviews*, **25** (5) 511–534.

Saprykin, A., N. Chokani and R. S. Abhari (2019) Gemsim: A gpu-accelerated multi-modal mobility simulator for large-scale scenarios, *Simulation Modelling Practice and Theory*, **94**, 199 – 214.

Schlich, R., S. Schönfelder, S. Hanson and K. W. Axhausen (2004) Structures of leisure travel: Temporal and spatial variability, *Transport Reviews*, **24** (2) 219–237.

Schmid, B. and K. W. Axhausen (2019) Predicting response rates further updated, vol. 1412, IVT, ETH Zürich, Zürich.

Schmid, B., S. Jokubauskaite, F. Aschauer, S. Peer, R. Hössinger, R. Gerike, S. R. Jara-Diaz and K. W. Axhausen (2019) A pooled rp/sp mode, route and destination choice model to investigate mode and user-type effects in the value of travel time savings, *Transportation Research Part A: Policy and Practice*, **124**, 262 – 294.

Schwarz, G. *et al.* (1978) Estimating the dimension of a model, *The Annals of Statistics*, **6** (2) 461–464.

Scutari, M. (2010) Learning Bayesian networks with the bnlearn r package, *J STAT SOFTW*, **35** (3) 1–22.

Shaheen, S., A. Cohen and E. Farrar (2018) The potential societal barriers of urban air mobility (uam), UC Berkeley: Transportation Sustainability Research Center.

Shires, J. and G. de Jong (2009) An international meta-analysis of values of travel time savings, *Evaluation and Program Planning*, **32** (4) 315 – 325. Evaluating the Impact of Transport Projects: Lessons for Other Disciplines.

Simma, A. and K. W. Axhausen (2001) Within-household allocation of travel: Case of upper austria, *Transportation Research Record*, **1752** (1) 69–75.

Smith, L., R. Beckman and K. Baggerly (1995) Transims: Transportation analysis and simulation system, 07 1995.

Sothy, T. C. (2019) Uber has already made billions from its exits in china, russia and southeast asia, `https://techcrunch.com/2019/04/11/uber-global-exits-billions/`. Accessed: 2019-10-21.

Sugiarto, S., T. Miwa, H. Sato and T. Morikawa (2015) Use of latent variables representing psychological motivation to explore citizens' intentions with respect to congestion charging reform in jakarta, *Urban, Planning and Transport Research*, **3** (1) 46–67.

Sugiarto, S., T. Miwa, H. Sato and T. Morikawa (2017) Explaining differences in acceptance determinants toward congestion charging policies in indonesia and japan, *Journal of Urban Planning and Development*, **143** (2) 04016033.

Sun, L. and A. Erath (2015) A Bayesian network approach for population synthesis, *Transportation Research Part C*, **61**, 49–62.

Sun, L., A. Erath and M. Cai (2018) A hierarchical mixture modeling framework for population synthesis, *Transportation Research Part B: Methodological*, **114**, 199 – 212.

Taubenböck, H., N. Goseberg, G. Lämmel, N. Setiadi, T. Schlurmann, K. Nagel, F. Siegert, J. Birkmann, K.-P. Traub, S. Dech *et al.* (2013) Risk reduction at the "last-mile": an attempt to turn science into action by the example of padang, indonesia, *Natural hazards*, **65** (1) 915–945.

Templ, M., B. Meindl, A. Kowarik and O. Dupriez (2017) Simulation of synthetic complex data: The r package simpop, *Journal of Statistical Software, Articles*, **79** (10) 1–38.

TheJapanTimes (2019) Flying taxis in singapore to test cleaner, quieter sky ride, `https://www.`

`japantimes.co.jp/news/2019/10/08/asia-pacific/flying-taxis-singapore-test-cleaner-quieter-sky-ride/`. Accessed: 2020-10-25.

Train, K. (2003) *Discrete Choice Methods with Simulation*, Cambridge University Press.

United Nations (2016) The world's cities in 2016, UN, New York.

Voas, D. and P. Williamson (2001) Evaluating goodness-of-fit measures for synthetic microdata, *Geographical and Environmental Modelling*, **5** (2) 177–200.

Vrtic, M., N. Schuessler, A. Erath and K. W. Axhausen (2010) The impacts of road pricing on route and mode choice behaviour, *Journal of Choice Modelling*, **3** (1) 109 – 126.

Walton, D. and J. Buchanan (2012) Motorcycle and scooter speeds approaching urban intersections, *Accident Analysis Prevention*, **48**, 335 – 340.

Wardman, M. (2004) Public transport values of time, *Transport Policy*, **11** (4) 363 – 377.

Woodcock, J., A. Abbas, A. Ullrich, M. Tainio, R. Lovelace, T. H. Sá, K. Westgate and A. Goodman (2018) Development of the impacts of cycling tool (ict): A modelling study and web tool for evaluating health and environmental impacts of cycling uptake, *PLOS Medicine*, **15** (7) 1–22, 07 2018.

Yagi, S. and A. K. Mohammadian (2010) An activity-based microsimulation model of travel demand in the jakarta metropolitan area, *Journal of Choice Modelling*, **3** (1) 32 – 57.

Ye, X., K. Konduri, R. M. Pendyala, B. Sana and P. Waddell (2009) A methodology to match distributions of both household and person attributes in the generation of synthetic populations, paper presented at the *88th Annual Meeting of the Transportation Research Board, Washington, DC*.

Young, M. and S. Farber (2019) The who, why, and when of uber and other ride-hailing trips: An examination of a large sample household travel survey, *Transportation Research Part A: Policy and Practice*, **119**, 383 – 392.

Yudis (2019) Rasio apartemen di jakarta belum sampai 2 persen, potensi pasar masih sangat besar, `http://housingestate.id/read/2019/01/11/rasio-apartemen-di-jakarta-belum-sampai-2-persen-potensi/`. Accessed: 2019-10-21.

Zhang, D., J. Cao, S. Feygin, T. Dounan and Pozdnoukhov (2017) *Connected Population Synthesis for Urban Simulation*, UC Berkeley, Berkeley.

Zhu, Y. and J. Ferreira (2014) Synthetic population generation at disaggregated spatial scales for land use and transportation microsimulation, *Transportation Research Record*, **2429**, 168–177.

CURRICULUM VITAE

PERSONAL DATA

Name	Anugrah Ilahi
Date of Birth	September 23, 1989
Place of Birth	Bengkulu, Indonesia
Citizen of	Indonesia

EDUCATION

2011 – 2013	Universitas Gadjah Mada, Yogyakarta, Indonesia *Final degree:* Master of Science
2007 – 2011	Universitas Gadjah Mada, Yogyakarta, Indonesia *Final degree:* Bachelor of Engineering
2004– 2007	SMA N 2 Bengkulu (high school) Bengkulu, Indonesia

EMPLOYMENT

June 2013 – June 2016	Junior Advisor *Sustainable Urban Transportation Improvement Project*, Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) Jakarta, Indonesia

PUBLICATIONS

ARTICLES IN PEER-REVIEWED JOURNALS

Ilahi, A. and K. W. Axhausen (2019) Integrating Bayesian network and generalized raking for population synthesis in Greater Jakarta, *Regional Studies, Regional Science*, **6**, 623-636.

Ilahi, A., M. Balac, A. Li and K. W. Axhausen (2019) The first agent-based model of greater Jakarta integrated with a mode-choice model, *Procedia Computer Science*, **151**, 272-278.

Belgiawan, P. F., A. Ilahi and K. W. Axhausen (2019) Influence of pricing on mode choice decision in Jakarta: A random regret minimization model, *Case Studies on Transport Policy*, **7**, 87-95.

CONFERENCE CONTRIBUTIONS

Ilahi, A. and K. W. Axhausen (2017) Measuring accessibility using an activity based model approach in Jabodetabek, paper presented at the *17th Swiss Transport Research Conference*, Ascona, May 17-19, 2017.

Belgiawan, P. F., A. Ilahi and K. W. Axhausen (2017) Bali Trans Sarbagita: Comparison between Utility maximization and Regret Minimization, paper presented at the *12th Eastern Asia Society for Transportation Studies*, Ho Chi Minh City, September 18-21, 2017.

Ilahi, A., P. F. Belgiawan and K. W. Axhausen (2017) Mode choice of transit system in Denpasar greater area (Sarbagita), paper presented at the *12th Eastern Asia Society for Transportation Studies*, Ho Chi Minh City, September 18-21, 2017.

Ilahi, A. and K. W. Axhausen (2017) Population synthesis of Jakarta, paper presented at the *18th Swiss Transport Research Conference*, Ascona, May 16-18, 2018.

Ilahi, A., P. F. Belgiawan and K. W. Axhausen (2018) Influence of pricing on mode choice decision integrated with latent variable, paper presented at the *15th International Conference on Travel Behavior Research*, Santa Barbara, July 15-20, 2018.

Ilahi, A., M. Balac, A. Li and K. W. Axhausen (2019) The first agent-based model of greater Jakarta integrated with a mode-choice model, paper presented at the *15th International Conference on Travel Behavior Research*, Santa Barbara, April 29- May 2, 2019.

WORKING PAPERS AND BOOK CHAPTERS

Ilahi, A., P. F. Belgiawan and K. W. Axhausen (2018) Influence of pricing on mode choice decision integrated with latent variable, in Konstadinos, G. and A. W. Davis (ed.) Mapping the Travel Behavior Genome, Elsevier, 125-143.

Ilahi, A., P. F. Belgiawan, Balac. M and K. W. Axhausen (2020) Understanding Travel and Mode Choice with Emerging Modes. A Pooled SP and RP Model in Greater Jakarta, Arbeitsberichte Verkehrs- und Raumplanung, ETH Zurich.

Ilahi, A., M. Balac and K. W. Axhausen (2020) Existing urban transportation in Greater Jakarta. Results of agent-based modellin, Arbeitsberichte Verkehrs- und Raumplanung, ETH Zurich.