Improved Activity-Based Model for Mega-Events

by

Alaa Sindi

A thesis submitted to the Faculty of Graduate and Post Doctoral Affairs in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Civil and Environmental Engineering

Carleton University

Ottawa, Ontario

Abstract

The multiple discrete-continuous extreme value (MDCEV) model is an advanced model used to estimate activity duration. It contains three parameters: baseline utility, translating satiation, and pure satiation. The translating satiation parameter is expected to capture the constant marginal utility effect, but it does not. Therefore, a modified model was developed that adds a fourth parameter (the power parameter) to the translating satiation parameter to capture the constant marginal utility effect. In addition, this research applies the power parameter to either and both the translating satiation parameter and the pure satiation parameter to examine effects of the parameters' interaction to further improve the MDCEV model's accuracy. The proposed model was applied to data sets from two countries, Saudi Arabia and Germany, to test the applicability of the modified model to any data set. This research found that adding the power parameter exclusively to the translating satiation parameter was the best model structure to maximize the accuracy of the MDCEV model for both data sets. Because an activity duration model is part of an activity-based travel demand framework, this improvement will lead to better predictability of activity-based travel demand. As a result, transportation planners can make appropriate decisions regarding future transportation infrastructure projects, which in turn will lead to a reduction in costs associated with these projects and decreased delays for transportation system users.

The main step in estimating travel demand is establishing the modeling framework. Conceptual travel demand modeling frameworks for mega-events were established based on a literature review of frameworks and the studies that form the skeleton of these frameworks. Studies performed on mega-events demonstrated the

importance of modeling mega-events separately from regular daily activities. Studies conducted on market segmentation shed light on the importance of modeling mega-events participants separately from nonparticipants and have resulted in improved guidelines for mega-event host cities aiming to reduce road network congestion. Criteria were established to select statistical software suitable for project inputs (e.g., project size). Finally, the estimated modified MDCEV model was generalized for use in transportation planning around the globe where limited planning models are available.

Acknowledgements

I would like to thank professors Said Easa and A.O. Abd El Halim for their supervision, advice, continuous support, encouragement, patience, and guidance throughout the process of this research. It was an honour to have both as my supervisors who pushed me to explore my area of research in more depth, but also in other areas of transportation engineering. I appreciate all they have done for me and this is immeasurable.

I would also extend my thanks to Mr. Matthew Strathearn for his support in R software. Also, I would like to thank professors Amir Shalaby, and Khandker Habib for the discussion about the Makkah case study, and the support of professor Habib with the GAUSS coding and data. Special thanks go to Professor Chandra Bhat, Professor Yasser Hassan, and Dr. Ott Toomet for their inputs about the MDCEV model and the Maxlik package in R software.

This research would not have been possible without the kind support and help of King Abdulaziz University in Saudi Arabia, the Saudi Cultural Bureau in Canada, and the municipality of Makkah, Saudi Arabia. I would like to extend my sincere thanks to all of them. In addition, I would like to thank Carleton University and NSERC for their financial support during my study. Thanks also to all my colleagues, friends, proofreaders, and GAUSS software technical support staff for their support.

Finally, the completion of this thesis would not have been possible without the motivation, love, and support of my parents Rashad and Nadia, my wife Noorah, and my children Janna, Dan, and Adam, and my brother and sisters.

Table of Contents

List of Tables

List of Figures

List of Symbols

U_{it} : Total utility for alternative i and for person t.

V_{it} : Systematic utility for alternative i and for person t.

ε_{it} : Random utility for alternative i and for person t.

βio : Alternative specific constant for alternative i.

Xijt : Socioeconomic variable j or service attribute j.

βij : Coefficient of the variable $Xijt$.

V_{mi} : Systematic utility for mode of transportation *m* and for specific alternative mode *i*.

V_{di} : Systematic utility for destination *d* and for specific destination location *i*.

X_z : In-home activity variables for pure satiation parameter.

β_z : Coefficient of variable X_z.

X_a : Out-of-home activity variables for translating satiation parameter.

β_a : Coefficients of variable X_a.

X_v : Out-of-home activity variables for the baseline utility parameter.

β_v : Coefficients of variable X_v.

T_a : Total time spent on activity type *a*.

AVG_a : Average time spent on activity type *a*.

Frq_a : Frequency of activity type *a*.

ε : Error term (Type I extreme value distribution).

ε_z : Error term for total in-home activities.

z : Total time spent on all in-home activities.

M : Activities frequency when conducted more than once.

L : Total number of in –home and out-of-home activities.

σ : Scale parameter.

E : Total time budget.

U : Sub-utility for a specific activity.

U_{Total} : Total utility for all activities.

1 Chapter: Introduction

The safe and efficient movement of people and goods within a city is an important economic indicator for its functioning and growth. Cities' planning departments use travel demand modelling and forecasting to make effective decisions regarding transportation infrastructure. Beginning in the 1950s, the Urban Transportation Model System (UTMS) or the four-stage sequential model became the basis for the analysis of travel demands [Meyer and Miller (2001), and Habib (2007)]. The UTMS is a macroscopic transportation planning tool designed for large urban and regional transportation studies (Ortuzar and Willumsen, 1994). This approach consists of four stages: trip generation and attraction, trip distribution, mode choice, and trip assignment. Trip generation and attraction divides the study area into zones and collects population and land use information in order to estimate trip generation and trip attraction. Trip distribution then uses a gravity model to generate an origin–destination trip matrix. Next, travel mode shares are determined using a mode split model, which is based on a discrete choice model. The discrete choice model, based on the Random Utility Maximization (RUM) theory states a person attempts to maximize his/her utility when choosing an alternative mode of transportation. Also taken into consideration is a person's socioeconomic characteristics and alternative attributes (e.g., age and travel cost).

In the fourth and final stage of the UTMS, trip assignment, trips are allocated using different modes of transportation on transportation network routes. The output of this stage is the volume of automobiles and transit on each road link (Meyer and Miller 2001). This method is used to forecast travel demand which is used in transportation planning (Figure 1). Examples of the usage of the UTMS four-stage model include estimating the need for additional road lanes to accommodate the increase in travel demand, and assessing the need for building new roads by

quantifying the advantages and disadvantages (e.g., estimating the construction cost and the expected reduction in vehicle delay).

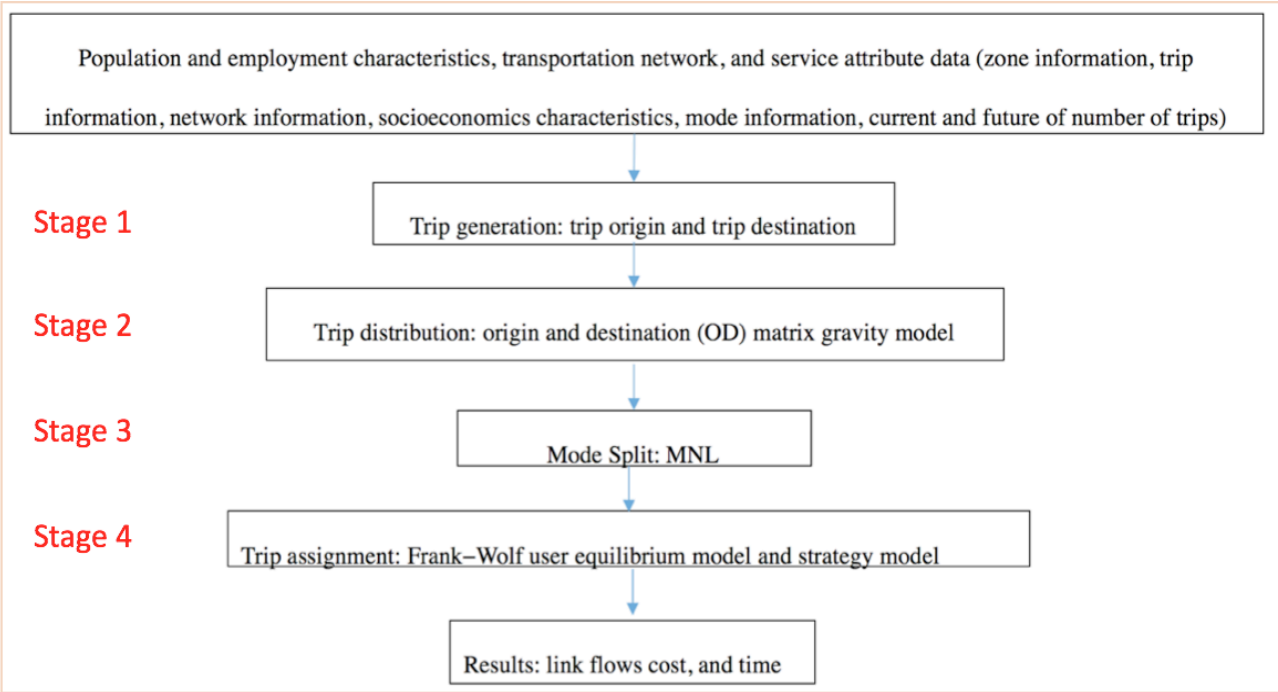


Figure 1: Four-stage sequential modelling.

In the 1970s, sociologist Chapin (1974), environmentalists Cullen and Godson (1975), and planner Hägerstrand (1970) formulated an activity-based travel demand model in an attempt to provide a better understanding of human travel behaviour and to improve policy analysis and project future travel demand. The fundamental assumption of the model is that people travel not for the sake of travel, but rather to conduct activities in- and out-of-home. It is thus important to understand the motivators of travel behaviour rather than considering the unit of a trip as the demand basis [Jones et al., (1983), and Habib, 2007]. Jones et al. (1983) proposed that travel behaviours are explained by the following:

1. The necessity to conduct activities;
2. A continuous sequence of activities connected by two key elements: time and space (location);
3. Constraints, such as:
    a) Locations of activities;
    b) Activities' time of the day;
    c) Availability of transportation facilities and their cost;
4. Household information, e.g., household income; and
5. Interactions between the members of a household, between individuals and urban development, and between individuals and transportation systems.

In addition, individuals' travel choices can be long-term or short-term. Long-term choices include mode of transportation choice (e.g., automobile ownership) and lifestyle choice (e.g., housing) which are conducted mostly once in a lifetime. Short-term activity choices include the chain of individuals' daily activities, i.e., those that change on a day-to-day basis.

Activity-based models are divided into two categories: (1) activity generation models and (2) activity scheduling models (Habib, 2007). Activity generation models capture individuals' desires to participate in different activities and include components such as start time, activity duration, activity location, and mode of transportation. Activity scheduling models organize a list of activities to be conducted during a specific time period, such as a day. The final output of this approach is the individuals' agendas during said time period. These individuals' agendas are inputs to simulation software, which contain a city map and the spatial characteristics of the city, in order to allocate travel demand on a city network. The first three steps of the four stage

sequential model are similar to activity generation and activity scheduling in that both represent a travel agenda. The four-stage sequential model does not contain activity start time and duration components, unlike the activity-based model. The final stage, trip assignment, is similar to the simulation component of the activity-based model. The predicted travel demand that resulted from these models aids in predicting future congestion within cities and inform improvements, such as changing signal timing, increasing the capacity of streets, or supporting the addition of overpasses. This research focuses on one type of activity generation model, the activity duration model, which uses the multiple discrete-continuous extreme value model (MDCEV) proposed by Bhat (2005).

While in the 1970s the UTMS demand model was the basis for forecasting travel demand and testing policies, nowadays many cities with large-scale travel demands, such as San Francisco, New York, Atlanta, San Diego, Seattle, Los Angeles, Boston, Chicago, Toronto, Calgary, Edmonton, and Jakarta, are using or planning to use the activity-based method to capture the complexity of travel behaviour in their cities [Roorda et al. (2008), Yagi and Mohammadian (2010), Shalaby et al. (2010), Vovsha et al. (2011), and Habib (2011)]. Activity-based travel demand modelling is well suited for modelling the travel demands of cities that host large-scale regional and special events, such as mega-events. Mega-events are usually, but not exclusively, sports-related events with worldwide recognition, such as international competitions (e.g., World Olympics, FIFA World Cup, Grand Prix racing), and may also include large-scale religious events, such as those that happen every year in the city/municipality of Makkah, Saudi Arabia. Makkah serves as the case study for this present research because it is a frequent host of mega-events.

Figure 2: Location of Makkah and the Al-Haram on maps

[Evacation.org (2016), and Maphill, (2017)].

To provide context, Makkah is the most sacred city for Muslims around the world. Its focal point is the Holy Mosque (Al-Haram), which has Al-Qaba (cubic building) at its centre (Figure 2). Al-Haram is open twenty-four hours a day, all year round for to up to 2 million worshippers (SUSRIS, 2015). Its impressive capacity is to account for the millions of Muslims who visit Makkah during the four religious events of the year –Hajj, Ramadan, Friday, and

Umrah– each of which is defined by the type of religious activity conducted during that time period (See Appendix A for more details about the city of Makkah mega events). During these mega-events, the city's transportation network becomes congested, causing serious traffic and safety problems. In addition, the city has future land use plans, such as the expansion of the Holy Mosque, which will attract more visitors and result in more traffic. This research recognizes that there is a need for a travel demand modelling framework that is capable of estimating future demand accurately to: better coordinate high transportation demand during Makkah's mega-events; facilitate planning for the city and its mega-events; and support decision making for transportation infrastructure projects. The motivation of this research is to model travel demand using activity-based modelling, which is an advanced technique. However, estimating all activity-based modelling components could take years of work. Therefore, this research focuses on one major modelling component: activity duration modelling.

1.1 Objectives

- To evaluate and improve the accuracy of an advanced modelling technique for the modelling of activity duration, the MDCEV model. This technique is based on a microeconomics framework that uses RUM theory (i.e., that a person attempts to maximize their utility when spending time on a specific type of activity).
- To improve the predictability of travel demand estimates and better plan for mega-events, such as those in Makkah. As a result, this will reduce the overall travelers' delay on the transportation network, and save the city losses from investing in unnecessary transportation infrastructure.

1.2 Research Plan

In order to achieve the objectives of this research study, six tasks were identified. These tasks are presented in Figure 3, below, and expanded in the sections that follow.

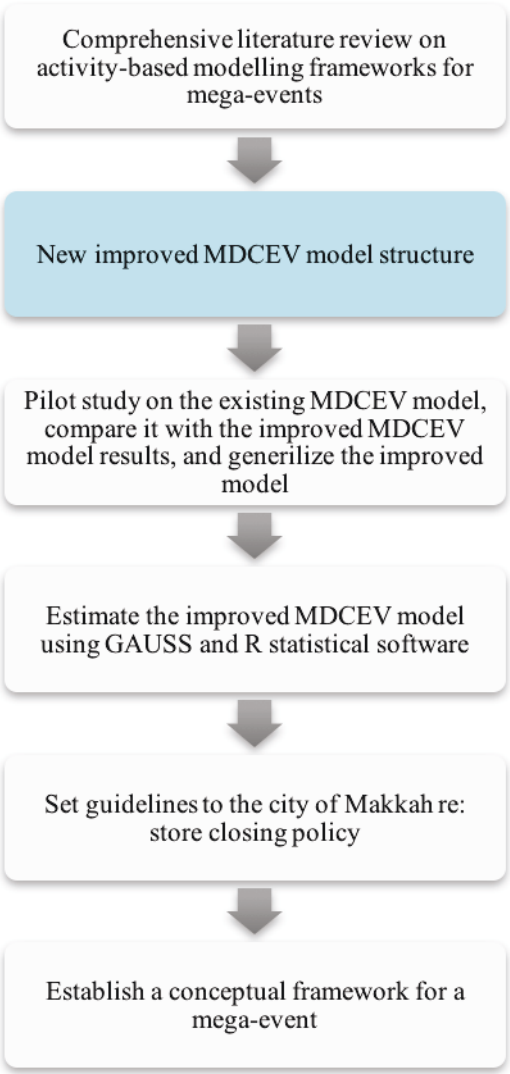


Figure 3: Research map.

(1) *Conduct a thorough literature review of activity-based modelling frameworks for mega-events.* Literature on activity-based modelling structures will be comprehensively reviewed, as well as details of the modelling components and how they are integrated

within the framework. The limitations of existing frameworks will be presented and improvements proposed. Finally, the limitations of other methods that estimate activity duration will be discussed alongside previous efforts in improving the MDCEV model. This portion of the research will help in establishing the general framework for mega-events, as there is limited research in the area.

(2) *Develop an improved MDCEV model.* Based on the results of the literature review, a new and improved MDCEV model will be proposed in order to obtain a more accurate estimate of activity duration. Different structures, such as the addition of new parameters to the existing model and the interaction between them, and the implementation of different utility structures will be explored.

(3) *Conduct a pilot study to evaluate the existing MDCEV model, and compare the base MDCEV model results to the improved MDCEV model to evaluate the improved MDCEV model.* These phases will help to better understand the model and ensure it is working as expected. It will also aid in identifying key factors in estimating activity duration; evaluate the improved model performance; and helping to design a more detailed modelling framework for Makkah. In addition, two datasets from different countries, Saudi Arabia and Germany are applied to the proposed model, and to ensure its applicability to different datasets.

(4) *Apply the improved MDCEV model using GAUSS and R statistical software.* The majority of researchers use GAUSS to estimate the MDCEV model [e.g., Habib et al., (2007), Bhat (2008), and Eluru et al. (2010)]; however, since software packages operate on different procedures, other options must be explored. The improved MDCEV model will therefore be estimated using both GAUSS and R statistical soft wares, after which

point one of the two will be recommended for future use. In addition, software selection criteria will be developed to aid in selecting the appropriate software for a specific project.

(5) *Apply the improved MDCEV model using the recommended software to set guidelines for a major policy for the city of Makkah.* Store closing during prayer is a main policy in Makkah which defined by frequency of closure and the length of time of closure that affects residents and visitors' daytime activities. This part of the research will study the difference in time expenditure behaviour between the residents of Makkah who do not visit Al-Haram and those who visit Al-Haram at least once a day. The investigation of time expenditure behaviour will be used to set guidelines for the city's store closing policy.

(6) *Establish a conceptual travel demand modelling framework for the city of Makkah's mega-event that fits the improved MDCEV model within it.* This stage of the research will require the following steps:

a) Presenting the history of Makkah, its transportation problems and policies, and its seasons, which include mega-events;
b) Using the literature review to establish the skeleton of the conceptual travel demand modelling framework for the city and its mega-events;
c) Using the pilot study and the city policy guidelines to better define the time frame of the conceptual framework; and
d) Using the improved model and recommending one of the two statistical software packages to be the tool used in the framework.

This study is the first step in building an advanced comprehensive travel demand modelling system (i.e. Activity-based travel demand modelling system) that will be used to accurately forecast travel demand, and inform effective decision-making regarding future transportation projects.

2 Chapter: Literature Review

2.1 Introduction

This chapter introduces mega-events and presents (1) a comprehensive review of the literature on travel demand activity-based modelling frameworks and models for mega-events, (2) a review of the travel demand modelling framework for the city of Makkah, and (3) an in-depth review of travel demand activity-based models, which are the most advanced models to date. This latter part of the literature review will overview the classifications of travel demand activity-based models and then focus on one classification in particular, the econometric model, which includes both activity scheduling and activity generation. One model used for activity generation, activity duration modelling, will then be explored in more depth, in particular as it uses the MDCEV model (see Figure 4 for a flow chart). Overall, this chapter will help form the basis of the conceptual modelling framework for mega-events, specifically those in the city of Makkah.

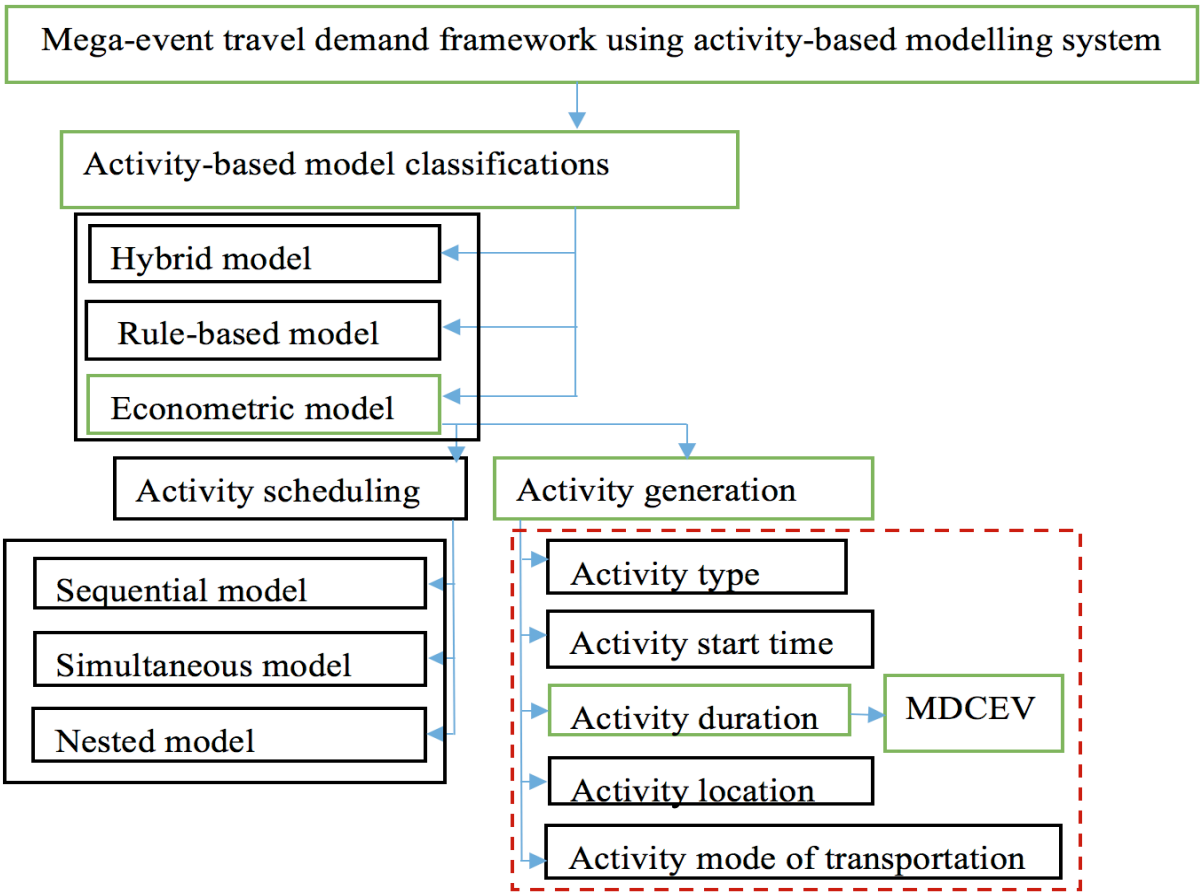


Figure 4: Map of literature review.

2.2 Mega-Events

Mega-events are "large-scale cultural (including commercial and sporting) events, which have a dramatic character, mass popular appeal and international significance" (Roche, 2000). These events, unlike regular events, attract a large number of visitors from outside the city or the country. Mega-events are described and categorized by their internal and external characteristics. Internal characteristics include the duration and scale of the event. External characteristics, on the other hand, refer to media coverage, tourist attraction, and the impact of the event on the host city (Malfas et al., 2004).

According to Roche (2000), media coverage (national or international) is one of the primary factors that determine the size of an event, but there is evidence that this is not always

the case. Some events with little media coverage succeeded in attracting a significant number of tourists and having a notable impact on the local economy. For example, the 1991 World Student Games in Sheffield, England, and the 1998 Spring Racing Carnival in Melbourne, Australia, resulted in a significant economic impact to the locale in which the events took place. Another example is the case study examined in the present work, Makkah, where yearly religious mega-events are the second major source of income for the country, bringing in roughly $12 billion, and yet media coverage is minimal (RT Question more, 2016).

Sola (1998) argues that the size of an event is also impacted by the significance of the economic returns for the city hosting the event, including the volume of visitors, visitor expenditures, and new facilities projects (Malfas et al., 2004). In some cases, countries compete to host mega-events that attract international investments (Dunn and McGuirk, 1999), where the expectation is that the event will make the city a better place to live post-event. For example, employment opportunities may increase, a spotlight on cultural and social life might improve tourism, and the built environment can improve with additions and upgrades to infrastructure for the event.

Benefits to the host community occur as a result of positive returns, which can be socioeconomic, sociocultural, and physical. Socioeconomic impacts include costs specifically related to the event, such as ticket sales, television rights, and sponsorship deals. They also include broader benefits to the city and the community at large, including job creation, investment opportunities (e.g., presenting the city as a tourist destination), and the potential for future investment and commercial activity in the region. Sociocultural impacts include motivating citizens to participate in activities related to the event and encouraging them to present themselves favourably to each other and to visitors (i.e., “show their best side”). The

residents of the host city may also gain more pride about their heritage and traditions as a result of the event (Essex and Chalkley, 1998). Lastly, positive physical impacts include the construction of new facilities or expansion of existing sporting facilities (and potentially the removal of old ones), as well as the improvement of the overall physical environment of the host city (e.g., new roads or improved public transportation services). However, there is also evidence of mega-events creating infrastructure that becomes problematic because of high operating and maintenance costs. This is true of many Olympic venues: in the case of the recent Olympics in London UK, most of the infrastructure was temporary in an attempt to avoid the cost of the operation and maintenance subsequent to the games (Manfred, 2012). The city of Makkah is a regular host for mega event; therefore, expansion projects are not problematic in terms of maintenance and operation costs.

To illustrate the scope and potential impacts of mega-events, two examples of mega-events are introduced and compared: FIFA World Cup and Hajj (one of the religious seasons in Makkah). A different country hosts the World Cup every four years, with events taking place in different cities within the host country. The last World Cup was held in Brazil in 2014 over a one-month period, with games taking place in twelve cities each with an average stadium capacity of 53,000 seats. An impressive 3.5 million tickets were on sale (Bovy, 2013). Figure 5 shows more information about previous and future World Cups and points to the large scale of the event.

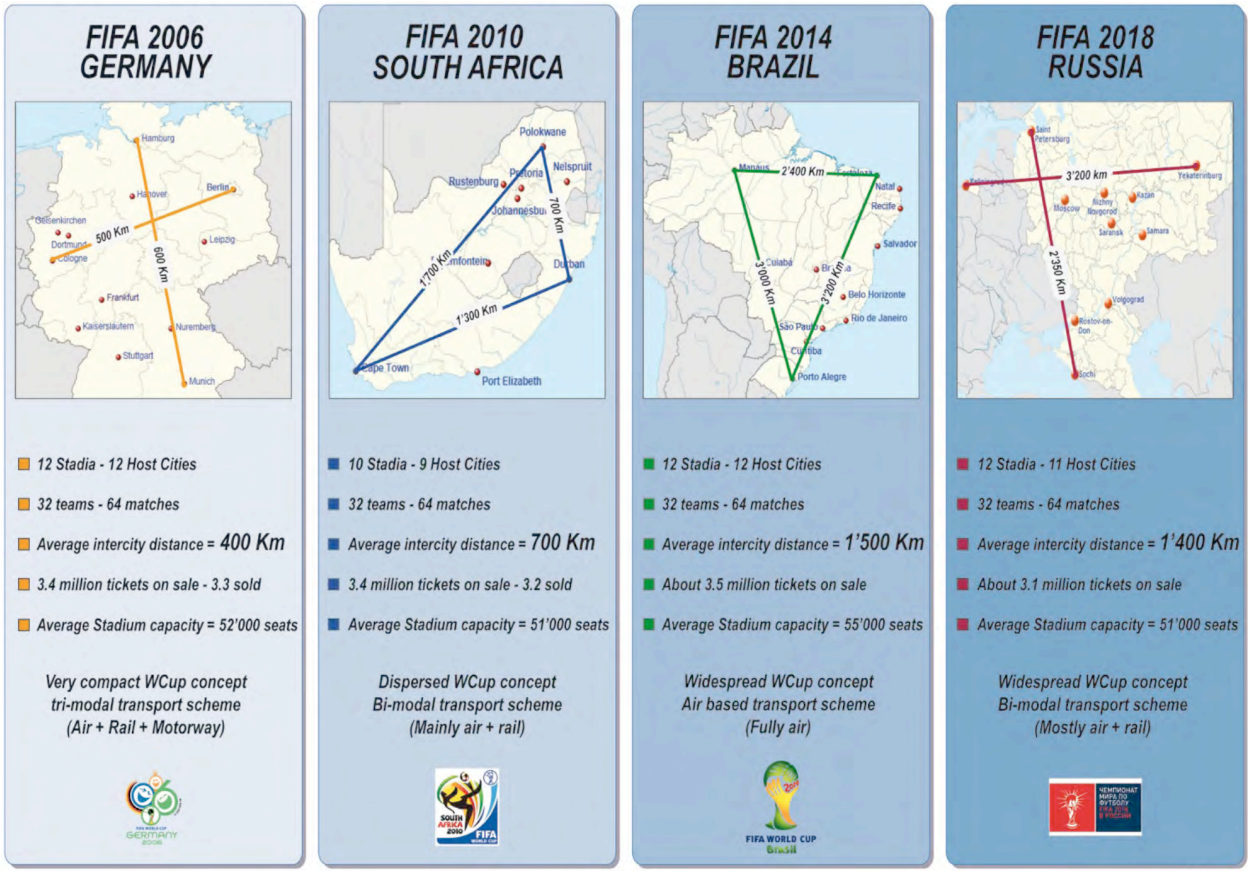


Figure 5: FIFA World Cup hosting countries from 2006 to 2018 (Bovy, 2013).

Though the size of the World Cup events is undoubtedly considerable, Hajj in Makkah is an even larger scale event, in part due to the 2-million person capacity of the Holy Haram. Hajj is held every year between the 8th and the 14th of the twelfth month of the lunar calendar and takes place at different locations in the city. If they are physically and financially able, Muslims around the globe are obliged to visit Makkah at least once in their lifetime to perform Hajj. In addition to religious activities, people visit Makkah for work, shopping, family visits, and studies, making Makkah one of the most diverse cities in the world [Fattah (2005), Seyyed, (2005), Khan (2009), Al-Laithy (2009), The Economist (2010), Taylor (2011), and Statistics Saudi Arabia (2014)]. Figure 6 shows the diversity and the volume of visitors in Makkah during Hajj season. Hajj is the classic example of Makkah mega-events because of its size. Umrah, a

smaller season than Hajj, is held any time during period of the second to the tenth month of the lunar calendar, and visitors are allowed to stay for a maximum of one month. Table 1 shows people distribution from the world that come to Makkah to perform Umrah in Al-Haram.

Table 1: Percentage of Makkah residents during Umrah season, by nationality (data provided by the municipality of Makkah, Saudi Arabia in 2010).

Nationality	Saudi	Arab	Asian	European	Other
Percentages	71.297	9.752	13.709	0.120	5.122

Figure 6: Number of male and female pilgrims during Hajj season in Makkah (MMM group and Moriyama and Teshima architects and planners, 2009).

2.2.1 Mega-Event Travel Demand Framework

To begin, Easa et al. (2002) describes the big umbrella of transportation planning which could be applied to mega-events as consisting of three dimensions (3D): system planning, project planning, and environmental analysis. System planning, which include data collection and modelling, is based on travel demand forecasting. Project planning considers transportation system operation such as transit operation and airport traffic circulation. Environmental analysis focuses on environmental issues, e.g., air quality. The focus of this research is on system planning such that the collected data will be used to develop the models.

As mentioned earlier, during Makkah's mega-events road transportation networks become overly congested due to the influx of visitors, causing serious traffic and safety issues. Bovy (2013) has stated that there is a limited amount of research related to mega-event travel behaviour. The existing research is limited to baseline data and quantitative studies of residents and visitors. In response to this gap, one of the objectives of the present work is to add to previous research by constructing a travel demand modelling framework designed specifically for mega-events. This section overviews existing research that will contribute to the development of such a model.

Yan et al. (2010) developed a travel demand modelling framework for the 2008 Olympics in Beijing. Their research aimed to establish travel demand and develop a forecasting traffic operation conditions analysis tool that considered both Olympic-related transportation demands and the daily travel demands of local residents. The framework included a system of models for both base demand and Olympic demand, as described below.

Base demand model involved modelling usual resident activities. First, the residents were divided into groups such as employee, non-employee, student, and retired. A model was

developed based on the activity chain, destination, and mode choice models, which are simulated using VISEM software (PTV, 2017). The activity chain included the order of daily activities (e.g., home-work-home) and the frequency or the number of people of the group (e.g. student) in the chain. Therefore, the number of people doing an activity chain of each group was the multiplication of the probability of the activity chain by the population of each group.

The framework included a total of nine types of activities: home (H), work (W), primary grade school (R), secondary grade school (E), university (U), shopping (S), leisure (L), private business (P), and employment or business related (B). Some examples of the activity chain are HBH, HSH, HBBH, and HBUH. Each activity chain was split into an origin and a destination to model the destination choice. For example, HWSH was split into three groups: HW, WS, and SH. The destination was estimated using the destination choice model (i.e., gravity model). The mode choice model was then estimated using the multinomial logit (MNL) model (Meyer and Miller, 2001). Seven modes of transportation were included in the model: car driving, car passenger, taxi, coach, bicycle, pedestrian, and public transport. These modes were divided into nontransferable (e.g., car) and transferable (e.g., walk) modes, where the model assumed that if a person chose a nontransferable mode for the first trip in the chain, he or she will use the same mode for the rest of the trips in the chain.

For the Olympic demand modelling, the arrival and departure times for each event were assumed to be few hours before and after the start time of each event. Based on a survey that Yan et al. (2010) conducted as part of the study, 94% of the Olympic spectator population was expected to use public transit. Trip distribution forecasting was modelled using an activity-chain based approach. More specifically, it was categorized based on the AM (morning), PM (afternoon), and evening periods and modelled using the single constrained gravity model.

The combination of both demands (i.e., base and Olympic) gave the total travel demand for the duration of the event. As a final step, the researchers evaluated several transportation demand management measures, including limiting vehicles entering the city centre, vehicle restriction by license number, and people leaving for vacation during the event.

While Yan et al.'s modelling framework is useful in testing the previously mentioned scenarios; it is limited in several ways. Five key limitations are presented below along with possible improvements. Firstly, in the base model, the framework used the adjusting factor to model tourist travel behaviour. The framework should have accounted for tourists who are not interested in the Olympic Games, as tourists may have different travel behaviours than spectators and residents.

Secondly, in the base model, the framework used activity chain data (e.g., HBH, HEH); this could be improved by using a more detailed modelling system, such as activity start time and duration for each purpose, which would result in more accurate outcomes to better serve policy analysis and lead to a better understanding of human behaviour.

Thirdly, the gravity model was used to model destination choice and the outcome was the zone number, but the framework could have used the MNL model instead. The MNL model is a statistical method that generalizes a regression model so that the probability of different alternatives can be predicted given their attributes. This would result in more detailed outcomes such as detailed locations; e.g., postal code (Eluru et al., 2010).

Fourthly, the available modes of transportation were divided into transferable methods (e.g., on foot, public transport, and taxi) and nontransferable methods (e.g., car and bike), where the model assumed that if a person chose a nontransferable mode for the first trip in the chain, he or she will use the same mode for the rest of the trips in the activity chain. This assumption may

not always hold true. For example, bicycles and cars could be considered transferable; cyclists might use the bus or train for part of the trip, while drivers might use park-and-ride options. Thus, a proposed improvement to this framework is to consider cars and bikes in the transferable mode category.

Finally, the Olympics demand model assumed that the spectator arrival and departure time profile would be similar to those of the Sydney and Athens Olympic operations (i.e., spectators arrive a few hours before the sessions). A more accurate method would be to model the start time and duration of the activity for each purpose so that previous data for the same event could then be used to estimate the initial parameter values. More accurate start and travel times result in better planning and better use of transportation facilities. Moreover, the modelling framework should use a more flexible time frame, not simply a 24-hour time frame (12am one day until 12am the following day). For example, the evening peak is expected to last until 1am the day following an event (Karlaftis et al., 2006); therefore, the 24-hour time frame should start from the time the person leaves their home.

2.2.2 Travel Demand Models for the City of Makkah

Although efforts have been made to develop travel demand models and plans for the city of Makkah since the 1970s, no study has yet established a comprehensive activity-based travel demand modelling framework. Bushnak (1977) developed a seven-step planning model in order to investigate different modes of travel that could be used in Makkah during a special event transportation system (SETS) –in Bushnak's case, Hajj– as well as analyze their performance. The seven planning steps were: (1) identify the system needs, regulations, and limitations; (2) describe the constraints and variables and estimate the difference between the supply and

demand; (3) develop performance measures that help to evaluate the scenarios based on the study goals, objectives, and constraints; (4) transform the analysis into plans and modify them to get the required performance; (5) evaluate alternatives using the multi-attribute utility function; (6) select the best alternative; and (7) implement the best choice. Due to the complexity of the transportation system in Makkah, there is a need for more advanced modelling techniques not only to explore the future transportation modes, but also to better test transportation policies and forecast travel demand.

A decade after Bushnak (1977) developed his model, Al-Harbi (1986) summarized travel factors that affect the movement of pilgrims between holy places in Makkah during Hajj, including the location of Al-Haram, routes and scheduled times to transport pilgrims, pedestrian–vehicle interactions, and the incomplete road network between the holy places and Al-Haram. Similarly, both Al-Yafi (1987) and Al-Turki (1989) highlighted the unique nature of Hajj activities; in Hajj, large groups of pilgrims move between certain locations at the same time, causing terrible traffic jams. They recommended that the flexibility in the departure time of different pilgrim groups could reduce this problem. Al-Rahman and Abdulaal (1989) suggested the use of different modes of transportation, such as trains and monorails, to transport pilgrims and goods. Other studies recommended the use of large buses to transport crowds of pilgrims from one place to another place [Al-Aqad (1988), Al-Jefry and Al-Hazmi (1989), Al-Shreef (1989), and MoMRA (1985b)]. However, the existing advanced travel demand modelling framework such as activity based-model has not been used to predict travel demand and test these suggestions (Shalaby et al., 2010).

Badabaan (2001) used the MNL model to study the mode choices of different pilgrim groups to travel between the holy locations in Makkah during Hajj. In addition, using a log-linear

formulation, Badabaan studied the interdependencies of these holy places. In 2011, similar travel demand modelling efforts were conducted using simulation software based on the four-stage sequential model to study traffic movement in the city of Makkah (Fayez, 2013). Fayez's study was useful to quantify the travel demand during the peak hours, but not to issue policies such as collecting a toll on some roads at specific times of the day.

2.3 Activity-Based Model Classifications

This section begins with an overview of approaches to travel demand activity-based modelling, with a focus on one in particular, the econometric approach. The two main models used in the econometric approach are then outlined, activity scheduling and activity generation, leading to the main model of the present work, the duration model. The MDCEV model and related studies will then be described and reviewed as it is used within duration modelling.

Activity-based models can be classified as (1) econometric, (2) microsimulation or computational, or (3) hybrid. The discrete choice model (the main model to the microeconomics modelling framework) assumes that the decision maker chooses the alternative from a choice set that maximize his or her utility, based on the aforementioned RUM theory (Castiglione et al., 2015).

In the econometric approach, utility refers to the measure of satisfaction or benefit. The total utility function is divided into systematic and random utility. Systematic utility is usually linear in its parameters and contains choice characteristics or service attributes, as well as socioeconomic characteristics (e.g., travel cost, and gender). Random utility is the random error term used to capture the uncertainty in choosing an alternative and in unobserved variables (Equations 1 and 2). The random utility contains the error term, where the general assumption

for the error term is Gumbel distribution, which results in the MNL model that is used most often to represent mode choice. The MNL model assumes that the error term is independently and identically distributed (IID) and consequently independent irrelevant alternatives (IIA) [Ben-Akiva et al., (1998), Ortuzar and Willumsen (1994), and Castiglione et al., (2015)]. The total utility function is obtained using the following formulas:

$$U_{it} = V_{it} + \varepsilon_{it} \quad [1]$$

$$V_{it} = \beta io + \sum \beta ij\ Xijt \quad [2]$$

The second approach to activity-based modelling is the computational approach. It is based on a set of rules that model the decision making process. Jones et al. (1983) developed CARLA, an algorithm which generates various activity patterns and uses a set of rules to reduce the number of patterns. Another example, ALBATROSS, a simulation system, considers the short, medium, and long term decision-making process in one comprehensive framework. This latter approach uses a decision tree algorithm to map decision steps, starting with the skeleton activities to fill part of the daily schedule. It then adds more activities to the remaining time using different sets of rules (Figure 7) (Arentze and Timmermans, 2000, 2004a, 2004b).

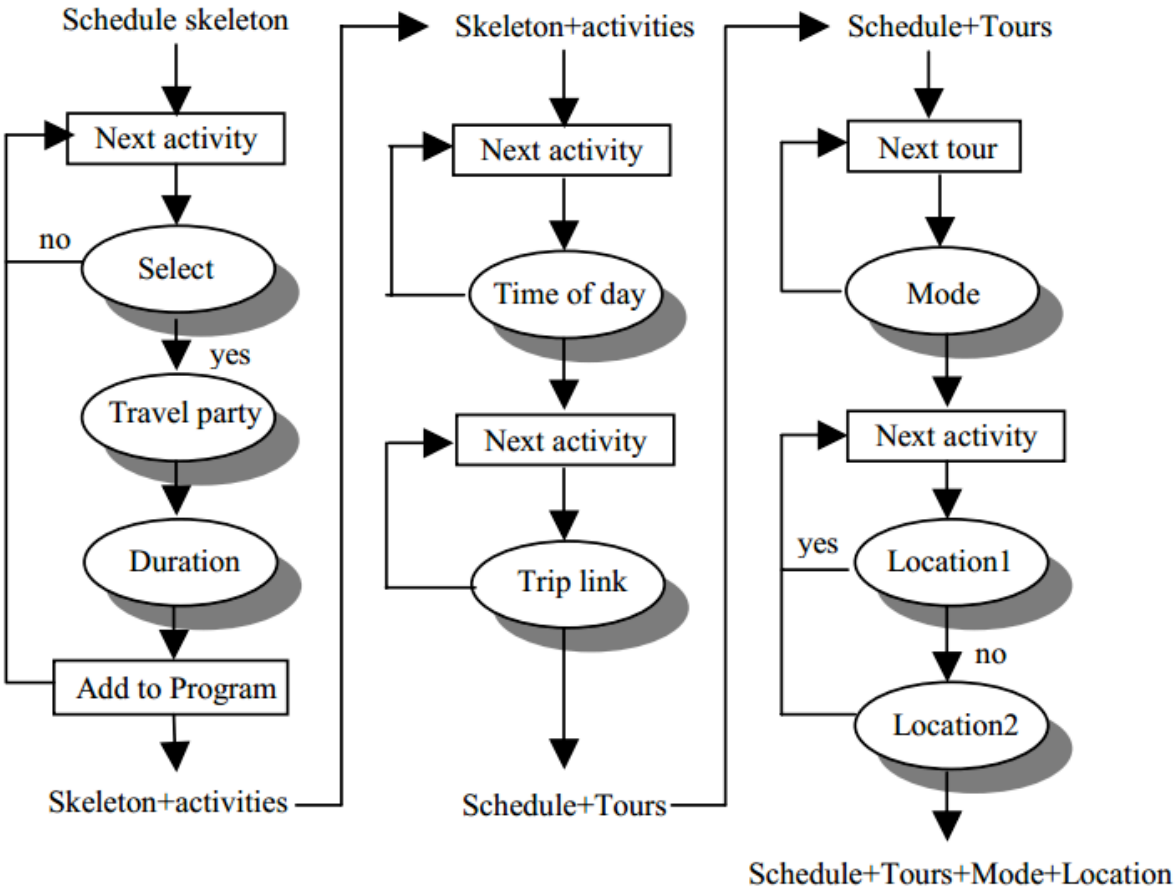


Figure 7: ALBATROSS scheduling process model (Arentze and Timmermans, 2004a).

The third approach, the hybrid approach, combines the utility maximization and computational approaches. For example, TASHA is a hybrid activity-based approach developed by Miller and Roorda (2003) that generates activities by selecting activity type, activity start time, and activity duration based on their distributions. It then models activity location using the entropy model that is based on employment population, activity density, and distance variables. Next, it models activity scheduling using rule-based models, after which it models a household tour mode choice using the MNL model. Finally, it assigns the trip using the simulation model EMME (Figure 8).

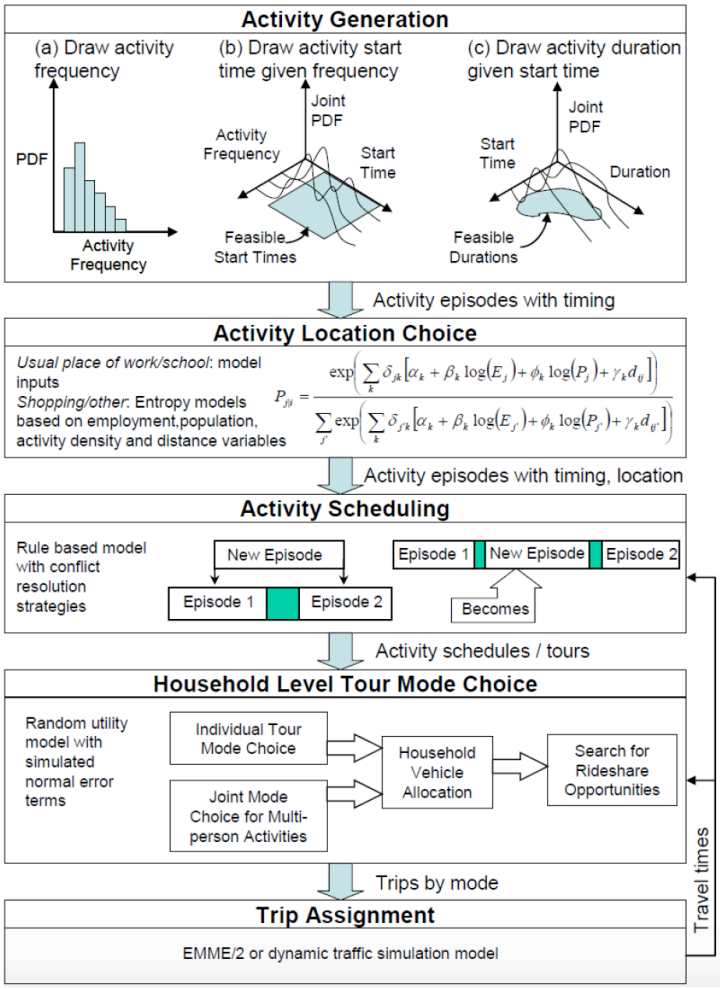


Figure 8: TASHA model framework (Hao, 2009).

The econometric approach is mainly based on the RUM theory which is the main and most widely used method when modelling activity and travel choices. It was found that RUM is a robust method when used over a wide range of choice making context. In addition, the method contains specific assumptions which when implemented in the model enable the users to account for measurement error and the random heterogeneity in the population, including missing variables and unobserved preferences. However, the RUM assumption stating that the decision maker is fully aware of all alternatives is unrealistic. For example, an individual might not be aware of a bus station close to his location so he/she may use a taxi (Castiglione et al., 2015).

Therefore, this research will use the econometric approach because it is the most widely used and robust method.

2.3.1 Activity Generation and Scheduling

According to Habib 2007, an econometric approach to activity-based modelling includes two main categories: activity generation and activity scheduling. Activity generation models capture people's desire to participate in different activities and contain different components, such as activity start time, activity duration, activity location, and mode of transportation. Socioeconomic characteristics and service attributes are also taken into account for each activity generation component during modelling. In addition, each of these components can act as input for another component, and several modelling frameworks link these components together. Following is the specification and discussion of each of the activity generation model components based on ARC 2012.

Many researchers such as Kawakami and Isobe (1982, 1988, 1989) and Habib (2011) considered activity type to be the first step when planning for the day because all other activity generation components (e.g., activity duration, and activity location) are related to type of activity. A person usually first decides on the activity to do during the day –such as work, school, shopping, recreation, and religious service– and then considers the other activity components, such as start time, duration, location, and mode of transportation. Activity types can be categorized in ways that help achieve more accurate model estimation. For example, school activity can be categorized into secondary, elementary, high school, college, and university. However, this categorization is also a factor of the total number of data points available within each category. In order to have a representative sample a sufficient number of data points should

be available in each category when estimating a model's parameters. If the number of data points is not sufficient, certain categories can be grouped together to solve the problem. New categories, such as before high school and after high school, could also be utilized.

The definition of time in both activity generation and scheduling is important for two reasons. Firstly, it defines the accuracy of the output result (i.e., activity start time), and secondly, it ensures the compatibility of the start time specification with the simulation time step. When dealing with time in the modelling process, the time of day could be divided into peak and off-peak periods and intervals of 15-minutes or less. The selection of time intervals is a trade-off between the accuracy of the model input and output, data collection efforts, and model estimation time. The more detailed the data, the higher the accuracy of the model. However, the collection and computational times increase, as well. The accuracy of reporting time (i.e., 15-minute intervals) is not high. Researchers overcome this problem by using GPS trackers in cell phones to collect accurate travel information (i.e., time and location) (Abdulazim et al., 2013).

Activity duration is collected in two ways: (1) report the start and end time of the trips or the activity, and (2) report the activity duration in minutes or seconds. When using the second collection method, activities with short durations are not captured in some cases because they last less than fifteen minutes (e.g., pick-up and drop off). Pick-up and drop off activities indicate that an individual was in a specific part of the road network or made an extra tour that was not captured. A detailed activity duration can capture these short activities.

Activity location can be collected using a postal code, full address, or zone number. Rich data should include detailed information such as the full address and postal code, or x-y coordinates. When only zone numbers are collected, researchers randomly assign an activity location within a zone because the exact location was not recorded. This will result in

aggregation bias. For example, if we assume that the randomly assigned location is on one side of a zone, and the actual location is on the other side of the zone, the error in this scenario will be significant. In addition, the error tends to be more significant in large zones because the travel distance and travel time between the two ends of a zone are sizable. Since this problem is mitigated in smaller zones, in cases of large zones researchers tend to divide the zone into smaller subzones based on the measurement of accessibility to transit so they can better allocate people or activity locations.

Zone characteristics or attractive measures such as area of the development, parking number, and type of zone are important factors when modelling activity location as these are the factors that attract the individuals to visit a specific location. A MNL model could be used to model activity location. The choice set could be zone numbers or postal codes. If two locations of the same activity type exist in the same postal code, one location could be chosen randomly; however, if the capacity constraint of the facility is included in the modelling framework; both locations should be included [Eluru et al. (2010), and Segovia (2016)].

Modes of transportation are modeled using the MNL model, where the required information to model a specific mode of transportation includes travel cost, travel time, and reliability. Reliability is quantified as the difference between the actual and the scheduled arrival time of a mode of transportation, such as a bus. If the bus arrives on time 95% of the time, we can say that this bus is highly reliable (Casello et al., 2009). Proposed or future modes of transportation shares could be forecasted using the MNL model once adequate information is collected. The current practice for collecting information about a new mode of transportation is through a stated preference survey in which a participant is given several scenarios and they choose the preferred mode of transportation (Yang et al., 2009). The mode choice model could

consider multimodal trips as in the case when people may use more than one mode of transportation during their entire trip, e.g., park-and-ride or kiss-and-ride options.

Researchers have realized the importance of modelling passengers joining a driver on a trip, specifically because it shows the social interaction among household members when choosing activities [Roorda et al. (2006), and Habib et al. (2008)]. Based on that, there is person-type segmentation and household-type segmentation. Every household member is a person segment that could be categorized into full-time or part-time worker, student, non-worker, preschooler, and so on. The household-type segmentation could be categorized by income into low, medium, and high-income groups. The purpose of segmentation is to reduce the variability of the full data and better estimate a model's parameters. Such segmentation is usually based on a research hypothesis.

In terms of modelling the above, activity type, location, and mode of transportation could be modeled using a MNL model. Activity start time and duration could be modeled using the Cox proportional hazards model or the MDCEV model [Bhat (2005), Habib (2007), Zhong et al., (2008), Yang et al. (2009), and Eluru et al. (2010)].

In summary, an activity generation model combines advanced modelling techniques for each component of an activity to accurately model human behaviour, travel patterns, and highly sensitive policies. Disaggregation (i.e., more detailed information) is important for model sensitivity. In addition, segmentation reduces the complexity of the model as well as model estimation time. Activity generation model components such as activity type, start time, duration, location, and mode of transportation could be added together in one framework to model travel demand, but there are several frameworks or modelling decision structures for use.

The above paragraphs described in depth the types of travel demand activity-based modelling and activity generation. A second category, activity scheduling, will be described below. Activity scheduling schedules people's activities during a specific time period such as a day or a week. It is a set of connected activities that comprise the person's travel agenda referred to as a tour. There are two major scheduling approaches to schedule tours. The first is the hierarchal approach, which starts by scheduling skeleton activities such as work and school. Then it schedules maintenance activities such as shopping. Finally, it schedules discretionary activities, such as social tours. Intra-household interaction that is used to model joint travel made by individuals of the same household is included after the mandatory activity tours, but before maintenance activity tours. This hierarchical scheduling method can be approached in different ways, such as starting by defining the time an individual leaves the home and returns back, and then adding traveler details such as activity start time, duration, and stops. To date, there is no unified approach used in scheduling time of the day, destination, and mode of transportation. For illustration, the mode of transportation could be an input to the time of the day model or the opposite structure could be used. A second method is building the schedule chronologically through the day (Castiglione et al., 2015).

According to Ben-Akiva and Lerman (1985), Ophem and Schram (1997), and Habib (2012) there are three types of choice models: sequential, simultaneous, and joint. These can be used to link activity generation components together or schedule people's activities. In a sequential model, the sequence of the model is chosen arbitrarily, but the model does not indicate which sequence is correct (current or opposite); consequently, the model will produce an incorrect forecast if the sequence is not correct. The sequential modelling framework ignores the relationship or interdependency of the different components of the activity-based model. The

sequential modelling decision is obtained using each of the previously mentioned activity generation and activity scheduling models independently as shown in Figure 9 and Equations 3 to 7. The output of one model is used as an input in the other model (similar to the four-stage sequential model).

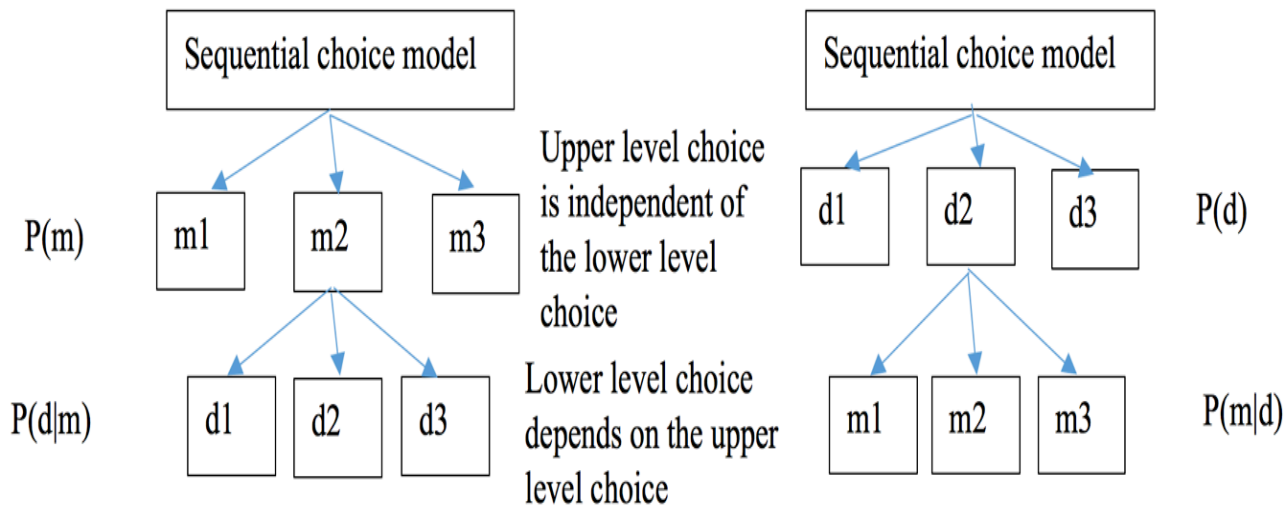


Figure 9: Sequential and nested choice model structure.

$$P(mi) = \frac{e^{Vmi}}{\sum_{i=1}^{i=n} e^{Vmi}} \quad [3]$$

$$P(\mathrm{dj}|\mathrm{mi}) = \frac{e^{V\mathrm{dj}|\mathrm{mi}}}{\sum_{j=1}^{j=k} e^{V\mathrm{dj}|\mathrm{mi}}} \text{ for mi} \quad [4]$$

$$P(dj) = \frac{e^{Vdj}}{\sum_{j=1}^{j=k} e^{Vdj}} \quad [5]$$

$$P(\mathrm{mi}|\mathrm{dj}) = \frac{e^{V\mathrm{mi}|\mathrm{dj}}}{\sum_{i=1}^{i=k} e^{V\mathrm{mi}|\mathrm{dj}}} \text{ for dj} \quad [6]$$

Joint probability

$$\mathrm{P(m2.d2)} = \frac{e^{Vm2}}{\sum_{i=1}^{i=n} e^{Vmi}} \cdot \frac{e^{V\mathrm{d2|m2}}}{\sum_{j=1}^{j=k} e^{V\mathrm{dj|m2}}} = \frac{e^{Vd2}}{\sum_{j=1}^{j=k} e^{Vdj}} \cdot \frac{e^{V\mathrm{m2|d2}}}{\sum_{i=1}^{i=k} e^{V\mathrm{mi|d2}}} \quad [7]$$

A simultaneous choice model could be used in the formulation of the travel mode and location choice for out-of-home activities while considering fixed choice sets. In other words, people choose the location of the activity and the mode of transportation at the same time. In this case, the utility of the mode is added to the utility of the destination, and this would be an MNL model structure as shown in Figure 10 and Equation 8.

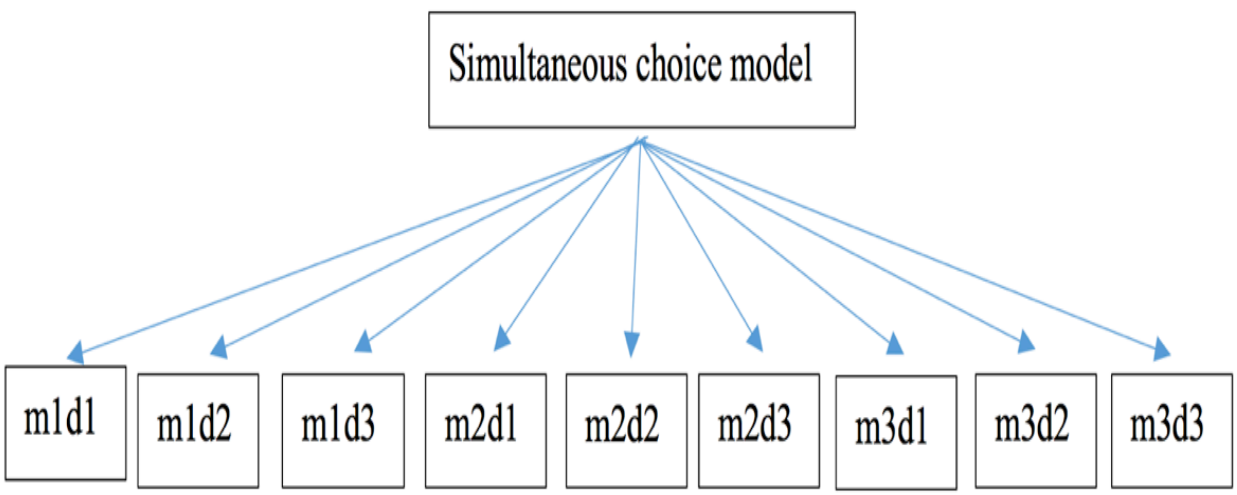


Figure 10: Simultaneous choice model structure.

$$P(midj) = \frac{e^{(Vmi+V\mathrm{dj})}}{\sum_{j=1}^{j=k} \sum_{i=1}^{i=n} e^{(Vmi+V\mathrm{dj})}} \quad [8]$$

A joint modelling framework assumes a correlation between the components of the activity-based model. The joint decision is modelled using a nested logic model (i.e., using a joint probability distribution of the decision). All three decision-making models are intended to reflect how people make decisions about their trips; but in a nested choice model, the choice of the location (lower level) is based on the mode of transportation (upper level) or the available choices of the mode of transportation (lower level) will depend on the location (upper level) (see Figure 9). The systematic utility and probability of destination choice are shown in Equations 9 to 13:

$\widetilde{Vd} = Vd +$ Expected maximum utility of choice of mode m for destination d [9]

$\widetilde{Vd} = Vd + Vm|d$ [10]

For $P(\mathrm{mi}|\mathrm{dj}) = \frac{e^{V\mathrm{mi}|\mathrm{dj}}}{\sum_{i=1}^{i=k} e^{V\mathrm{mi}|\mathrm{dj}}}$ for dj [11]

The maximum utility choice of mode choice for destination d is $ln \sum_{i=1}^{i=k} e^{V\mathrm{mi}|\mathrm{dj}}$ which is known by inclusive value.

$\widetilde{Vd} = Vd + ln \sum_{i=1}^{i=k} e^{Vmi|dj}$ [12]

$P(d1) = \frac{e^{\widetilde{Vd1}}}{\sum_{j=1}^{j=k} e^{\widetilde{Vdj}}}$ [13]

One limitation of the four-stage sequential model is its inability to consider the interrelationship between the various components of an activity. An appropriate solution is to model the components jointly using travel demand activity-based model. Some examples of

travel demand activity-based modelling components that were modelled jointly include Pendyala and Bhat (2004), who studied the interdependency of activity start time and duration. They found that for non-commuters activity duration precedes activity start time. However, the relationship between the activity start time and duration for commuters is loosely related. Habib et al. (2008) jointly modeled travel with whom (i.e., relatives or friends), activity start time, and activity duration. Similarly, Roorda et al. (2006) modeled interactions among household members to make a mode choice.

Several frameworks have been proposed by researchers in order to model travel demand using the econometric approach of the travel demand activity-based model. However, some of these frameworks do not consider the complete travel pattern as a unit of choice. Instead, they build the patterns step by step using a random utility model. For example, Kawakami and Isobe (1982, 1988, 1989) built the activity patterns for workers in two periods, before and after work. This framework adopts a hierarchical-choice structure. It begins with activity type, followed by activity-travel pattern, and ends with activity location. One criticism of this framework is that it does not take into account the interdependence of activity-travel behaviour during the day (Ben-Akiva and Bowman, 1995, 1996). In response, Ben-Akiva and Bowman modelled the activity scheduling hierarchically into primary and secondary home-based tours for both workers and non-workers. For each of the tours, a random utility-based nested logit model was used to model the choice of activity pattern; primary tour; mode and destination of the primary tour; start time of the secondary tour; and mode and location of the secondary tour. Taking a slightly different approach, Wen and Koppelman (1999) added household decision making to the model, which assigns a number of stops and vehicles to the spouses in a given household for out-of-home activities. Then, the model assigns tours to each individual and his or her choice of stops (Habib,

2007). Conversely, Habib (2011) proposed a framework for a dynamic activity scheduling method, which uses RUM. Habib's model is a series of nested models that choose the activity type, activity location, and activity duration. The model then selects the activity type again and all activity components for the rest of the day. The model is constrained by time, budget, and space (Figure 11).

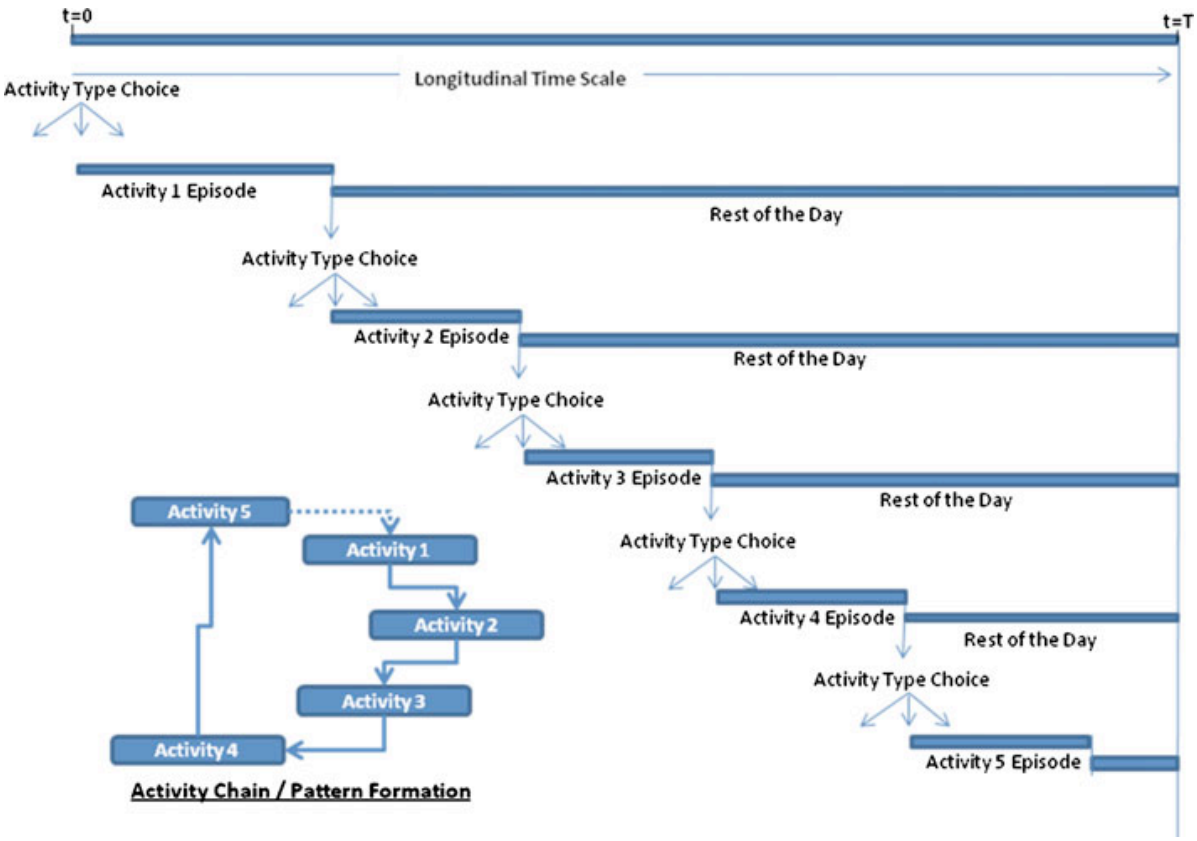


Figure 11: Activity pattern formation using dynamic discrete-continuous approach (Habib, 2011).

Bhat et al. (2003) proposed a comprehensive activity-based model system that considered both workers and non-workers in a household (Figure 12). The study modelled travel demand for metropolitan areas in Texas using an activity-based approach. It considered land-use, socio-demographic attributes, activity systems, and transportation level-of-service. The output of the

model is a complete activity-travel pattern during the day for each individual within a household. The framework contains short, medium length, and long-term household decision making processes and is divided into three models. Firstly, a generation-allocation model system models for all of the various sets of activity an adult participates in during the day. Activity start and end times are modelled for out-of-home activities and in-home activities. The model starts by modelling a person's wake-up and sleeping time, and then work and school start and end times. Pleasure or maintenance activities such as social activities or shopping are modelled last. The model then determines the number of adults within the household and distributes the household responsibilities between them. Finally, each activity is categorized as an in-home or out-home activity.

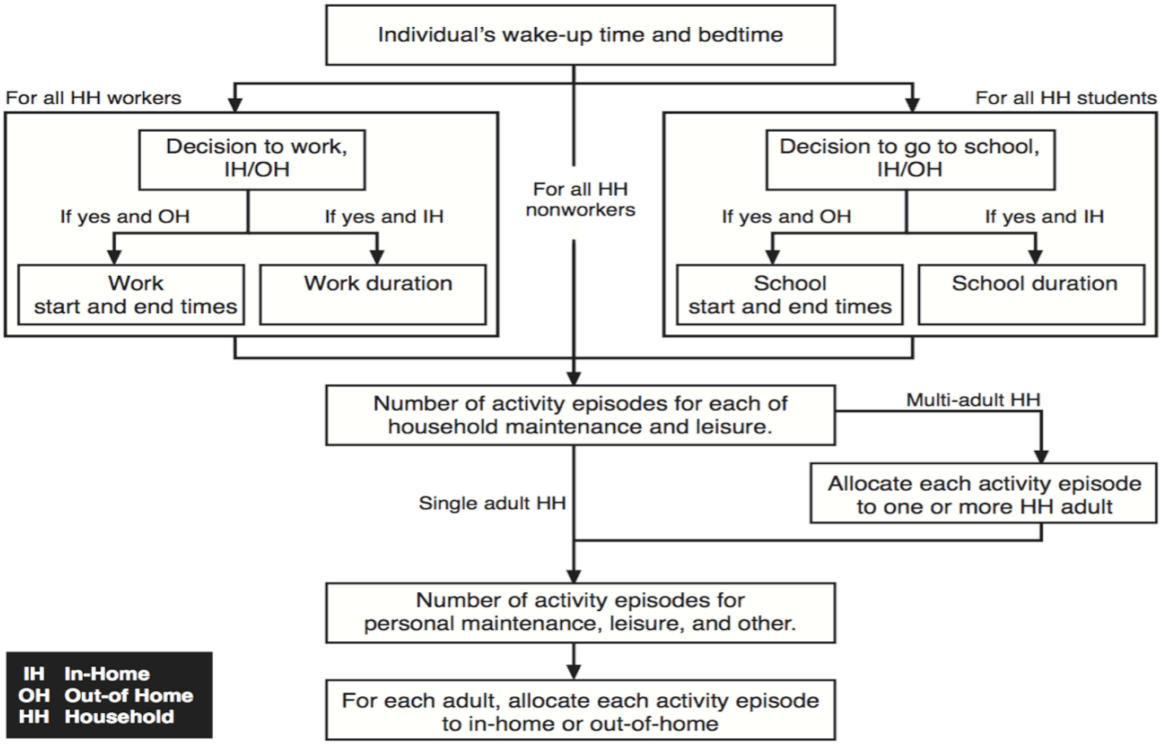


Figure 12: Generation- allocation modelling framework (Bhat et al., 2002).

Secondly, a pattern-level model system is meant to schedule all of the generated activities into time periods. The day is divided into five time periods: before work, commute time to work, working hours, commute time to home, and after working hours or the rest of the day. For people who participate in activities other than work, the day is divided into the number of tours undertaken before the work period, during work, and after the work period. For non-workers, the total number of tours is determined. Work and school activities are scheduled sequentially, i.e., first the mode is modelled, then number of stops, and finally duration. Then, the joint tours made by household members are identified. Lastly, the in-home activities are scheduled (Figure 13).

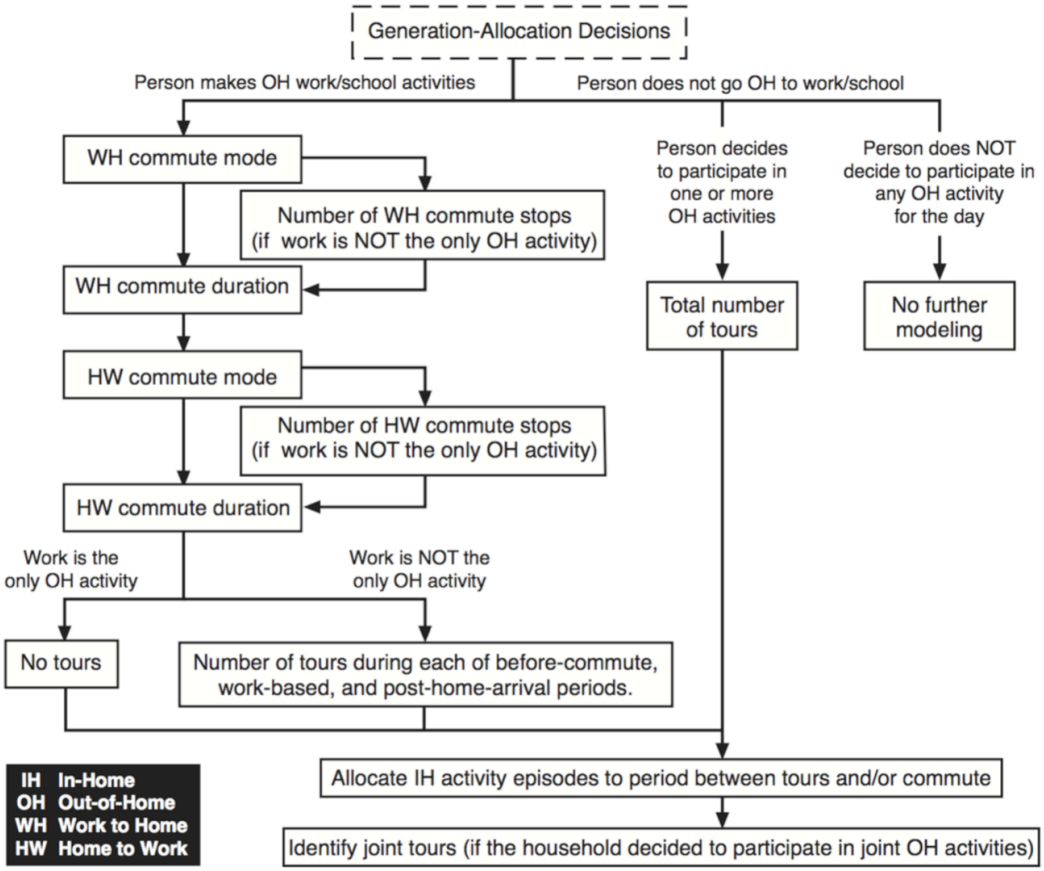


Figure 13: Pattern-level modelling framework (Bhat et al., 2002).

Thirdly is a tour- and stop-level model system, which involves more detailed scheduling decisions. The tour is classified into a joint or solo tour and the attributes of each stop are modelled sequentially in order to obtain the complete characteristics of a person's daily travel pattern (Figure 14).

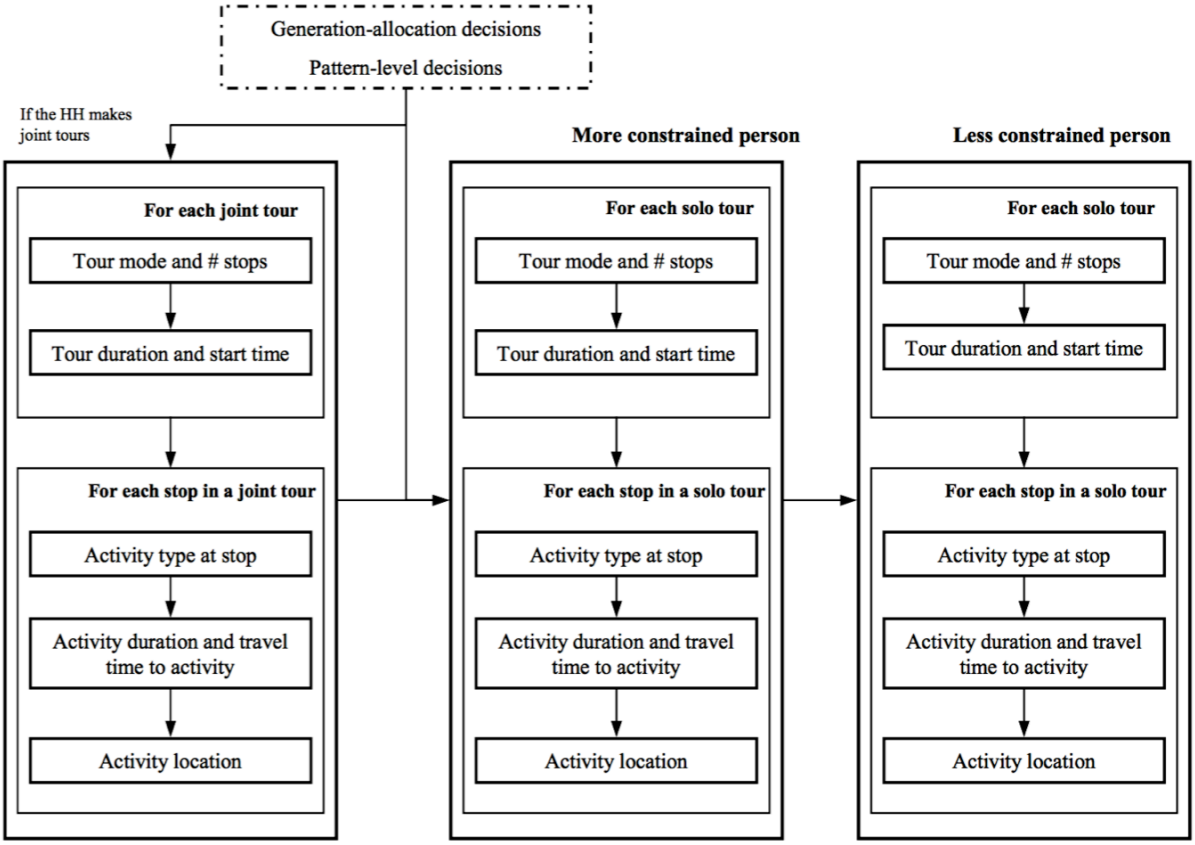


Figure 14: Tour-and stop modelling framework (Bhat et al., 2002).

Overall, an activity-based model is successful when it has an understanding of activity patterns, develops model specifications (such as modelling decision), and includes household interactions [Griesenbeck and Garry (2007), and Shalaby et al. (2010)]. According to Goulias (2007), no single activity-based model could be used for all cities because cities have different

characteristics and travel patterns. This research project builds on the above previous studies to create a conceptual travel demand modelling framework for mega-events in Makkah.

2.3.2 Activity Duration Models

Having established that the econometric approach is based on activity scheduling and activity generation models, this section of the literature review expands on one model within activity generation, the activity duration model. This final section of the literature review presents previous research and limitations of activity duration models. In addition, it focuses on MDCEV models, and suggests ways of improving them to make them better suited to the present work regarding travel demand modelling of mega-events in Makkah and other cities.

Researchers have shed light on the importance of building a travel demand modeling framework emphasizing activity duration models with the rationale that people spend time during weekdays differently from weekends. Zhong et al. (2008) observed that literature in this area acknowledges the differences in activity participation rates, and the start time, duration, and the location using statistics, log-rank and Wilcoxon, and the best-fit duration model during the weekdays and the weekends. Zhong and co-authors suggested differentiating between weekdays and weekends when modelling travel demand. Rachel and Bhat (2007) similarly conducted a descriptive study on school-aged children engaging in activities during the weekdays differently from the weekends. The study showed that during the weekends the number of trips to church is the highest with an average of 2.5 hours of activity duration, while trips to retail stores are also a relatively frequent activity duration from noon to 4pm. Moreover, children spend more time during weekends in sport practices, kids' clubs, and other meetings compared to weekdays. A key finding of the study was that a key factor in specifying the type of activities is age.

TASHA uses probability distribution to model activity duration, which does not reflect how an individual behaviourally allocates time to specific activities (see Figure 8). To illustrate, the method draws activity duration based on its distribution and the estimated start time, and it does not include any information about an individual, such as age. It is important in an activity-based modelling framework to relate activity duration to human characteristics rather than relying on activity duration distribution. For example, how a family of eight spends their time shopping compared to a family of two? It is expected that large families will spend more time shopping compared to small families. This enables the transportation engineers to understand how people spend their time out-of-home and to propose solutions and policies based on that (e.g. large families should live next to shopping malls), and thus solve transportation problems such as congestion. This will be demonstrated more thoroughly in section 2.4 (the following research sheds light on the role of the base MDCEV model parameters) and in Chapter 5. The following researches shed the light on how an individual behaviourally spends time on activities, and present limitations within some behaviuoral approaches (i.e., structural equation model and hazard-based duration model).

Ashish (2004) used a structural equation model (SEM) to model activity duration during weekdays and the weekends. The research used socio-demographic characteristics of the individuals and households to estimate the models and comments on the differences in people's behaviours in spending time during the weekdays and weekends. For example, Ashish results showed that people with an income over $55K spend more time at work during weekdays compared to weekends. However, according to Bhat et al. (2012) and Werner and Engel (2009), SEM becomes hard to deal with as dimensionality of the problem increases. Moreover, SEM faces identification problems as the number of dependent variables increases. In an attempt to

mitigate these problems, discrete-continuous models such as the MDCEV model were established. Xia et al. (2009) compared two discrete-continuous models for multiple time choices. The comparison was between the models proposed by Kim et al. (2002) and Bhat (2005) which is MDCEV model. The main difference between the two was the error term: Kim et al.'s model used log-normal distribution and Bhat used extreme value distribution. Xia et al. concluded that both methods produced accurate results; however, the accuracy of the results depended on the distribution of the data and the number of data points. Kim et al.'s model worked well with few data points, but Bhat's model performed well with a large number of data points (i.e., more than 1000 points). Kim et al. used the same baseline utility for all individuals, whereas Bhat allowed it to vary per individual, which provided a more detailed model. In addition, Bhat argued that the utility structure proposed by Kim et al. results in an indirect interpretation of the model parameters. Ultimately, the Bhat model is more suited for the model in this research because of the high number of data points and the need for more detailed travel behaviour that support policy analysis.

Prior to Kim et al. and Bhat's studies, early activity based models used regression models and hazard-based duration models to estimate activity duration. It was found in practice that the discrete choice model is easier to integrate with other model structures, easier to calibrate, and allows for testing of the sensitivity measure of travel condition for different times of the day (Castiglione et al., 2015). Hazard-based duration models include individuals' information to model behaviourally how an individual engages in a specific one out-of-home activity. However, Hazard-based duration models cannot consider all activity types within the framework, and so a more comprehensive model was necessary. The MDCEV model, which includes all types of activities conducted during the day, thus became popular in estimating activity duration.

2.4 MDCEV Model Formulation

As previously mentioned, the MDCEV is an advanced modelling technique for the modelling of activity duration. It was proposed by Bhat (2005) who based it on a microeconomics framework using RUM theory. In other words, a person tries to maximize his or her utility in spending time out-of-home while taking into consideration the time budget constraint. Basically, the model contains three parameters to capture behaviouarlly the trade off in time consumed between in-home and out-of-home activities: pure satiation parameter, translating satiation parameter, and baseline utility parameter [see also Bhat (2005, 2008) for the model formulation, derivations, and identification issue details]. These parameters include individual information, trip information, and household information obtained from a real dataset. The output of this model is the parameters' coefficients that maximize the likelihood function of the MDCEV modelling framework. The parameters' coefficients and their statistical significance are used to comment on the people behaviour in spending time out-of-home. The likelihood value is used to comment on the accuracy of the model or the goodness of fit of the estimated model. The sections below present the role of each parameter, the model specifications, and derivation.

2.4.1 Pure Satiation and Translating Satiation Parameters

The pure satiation parameter (μ) represents the satiation effect in out-of-home activity duration expenditure. The pure satiation parameter aims to represent behaviour in the context of the decrease in a person's preference in spending time in out-of-home activities with an increase in time consumption. For example, a person's willingness to spend time shopping decreases with the increase in time he or she spent shopping. In other words, because of a person constraints

(e.g. time budget) and person physical energy, a person cannot spend more than two hours shopping. In addition, the pure satiation effect for shopping for a family of eight people is lower than the family of two, such that the family of eight needs more time to spend on shopping. The model captures how different household structures or communities allocate time to specific activities during time budget (e.g., one day). It is obtained using the following equation:

$$\mu = 1 - exp\ (-\beta_z X_z) \qquad [14]$$

This parameter is a function of the variables of an in-home activity and shows the effect of the increase or decrease in in-home utility on the duration of out-of-home activities. The increase in the coefficient β_z increases μ, and increase the activity time spent out-of-home as shown in Equation 14.

The translating satiation parameter (Ω) represents the trade-off between out-of-home activities such that how long an individual is willing to trade time to spend on a specific activity compared with another. In addition, it ensures the corner solution (such as a person spending all day on one type of activity such as work). To illustrate, the corner solution is where work activity has a constant marginal utility and the rest of the activities have low or zero translating satiation parameter values and is obtained using the following equation:

$$\Omega = \exp\ (\beta_a X_a) \qquad [15]$$

As noted, the higher the coefficient β_a the higher the Ω value and the higher the out-of-home activity duration. The Ω value also includes the satiation effect of out-of-home activities, taking into consideration in-home activities (Kim et al., 2002).

2.4.2 Baseline Utility Parameter

The baseline utility parameter of specific activity a (B_a) is the marginal utility at the zero time expenditure point. The baseline utility parameter illustrates the increase or decrease in a person's preference in spending time out-of-home on a specific activity. For example, if the baseline utility parameter for work is higher than shopping, we can assess in quantitative terms that the person will choose work activities to spend more time on. Moreover, if a person did not conduct an activity during the day, B_a for that specific activity equals zero. This can also be contextualized with the knowledge that people will spend more time at work on a daily basis rather than shopping and is obtained using the following equation:

$$B_a = \exp\left((\beta_v X_v)_a + \varepsilon\right) \quad [16]$$

The higher the coefficient β_v is the higher the B_a value and the higher the out-of-home activity duration.

The total duration of specific activity a is the average time spent on the activity multiplied by the frequency of the activity, which is obtained using the following equation:

$$T_a = AVG_a \, Frq_a \quad [17]$$

2.4.3 MDCEV Utility Structure

The total individual utility for spending time in-home and out-of-home is the summation of the sub-utilities of each individual activity as follows:

$$Total\ Utility = \sum_{a=1}^{No.of\ Activities} (\Omega_a/\mu)\ \exp((\beta_v X_v)_a + \varepsilon)((\frac{T_a}{\Omega_a} + 1)^{\mu}\text{-}1) + \frac{1}{\mu}\exp(\varepsilon_z)(z)^{\mu} \qquad [18]$$

The total individual utility contains the pure satiation parameter, the translating satiation parameter, and the baseline utility parameter. The total utility function is used to estimate the model parameters and drive the likelihood function.

2.4.4 Derivation of the Likelihood Function

The Kuhn-Tucker (KT) optimality conditions are applied to the utility function (Equation 18) to transform it to an indirect utility part of specific activity (V_L) and an indirect utility part of the total in-home activities (V_z) considering a time budget constraint (Equation 19) as follows:

$$\sum \mathrm{T_a} + z = E \qquad [19]$$

$$\mathrm{V_L} = \mathrm{B_a} + (\mu - 1)\ln(\mathrm{T_a}/\Omega\mathrm{a} + 1) - \ln(\mathrm{AVGa}) \qquad [20]$$

$$\mathrm{V_z} = (\mu - 1)\ln(\mathrm{z}) \qquad [21]$$

Then, the transformation variable theorem and the error term are used to drive the likelihood function. The final shape of the likelihood function used to estimate the coefficients activity duration model is given by:

$$p(T_1, T_2, T_3 \ldots, 0,0,0, T_a) =$$

$$Ln\ (1/\sigma^{M\text{-}1})(\prod_{a=1}^{M}(M\text{-}1)\left(\prod_{a=1}^{M}\frac{1-\mu}{T_a+\Omega_a}\right)\left(\sum_{a=1}^{M}\frac{T_a+\Omega_a}{1-\mu}\ AVG\ FRQ\right)\left(\frac{\prod_{a=1}^{M} exp(V_L/\sigma)}{\sum_{L=1}^{L} exp\left(V_{L/\sigma}\right)^{M}}\right) \quad [22]$$

In summary, this section presented the MDCEV model formulation and derivation. It started by presenting the three model parameters (i.e., translating satiation parameter, pure satiation parameter, and the baseline utility parameter), their role, and how these parameters explain activity duration. Next, it introduced the utility structure that included the three parameters. Finally, it presented the derivation of the likelihood function which uses the KT optimality conditions and the time frame constraint. This likelihood function will be used to estimate the MDCEV model parameters in the Chapter 5.

2.4.5 MDCEV Model Studies

The MDCEV model has been diversely applied to activity duration applications, with researchers changing its structure and using different scale parameters to evaluate and improve the model. Paleti et al. (2010) used the MDCEV model to study children's allocation of time in various activity duration and location during the after-school period. Understanding child travel patterns and their linkage with parents' schedules is an important component for an accurate estimate of an activity based travel demand modelling system. The four major findings of Paleti et al. are as follows: (1) 55% of the children conduct one after-school activity during the day; (2) the children's characteristics, environmental characteristics, and activity travel pattern variables are important factors that affect children engagement in after-school activity pattern; (3) middle school children spend more time at home after returning from school compared to younger and

older school children; and (4) children that have internet access at home prefer to spend more time at home after returning from school compared to children with no internet access.

Considering a different context, Habib et al. (2007) used the MDCEV model to study the rhythms of activity time expenditure behaviour for the residents in German cities during a time budget of one week. Six weeks are considered the time frame for this study. Seven models were examined (i.e., a model for each individual week and a seventh model for all weeks combined). The research concluded that when modelling the rhythm of travel behaviour, one week is an adequate modelling time frame. However, the highest goodness of fit of the seven models is 0.237.

Several studies have aimed to improve on the MDCEV model. To improve the mean log-likelihood of the MDCEV model, Pinjari and Bhat (2010) proposed a multiple discrete-continuous nested extreme value (MDCNEV) model structure. Considering several time intervals of the day they modeled the activity duration of different activity types. This modelling structure benefit the activities scheduling process and the MDCNEV results in better goodness of fit compared to the MDCEV model. Eluru et al. (2010), used the MDCEV model to model activity type, duration, time of the day, and mode choice, and then used multinomial logit (MNL) model to model destination choice. They then connected these two models using nested logit (NL) model. This modelling structure was used to test the effect of travel cost and time variables, which results in changes to out-of-home participation in activities. The proposed model performs better in terms of goodness of fit compared to the sequential modelling structure. Taking a different approach, Sikder and Pinjari (2014) proposed to use a different scale parameter for each one of the choice alternatives to investigate the benefits of implementing this modelling structure

when compared with the MDCEV model. They found that the goodness of fit improved after using their proposed structure.

Easa (2014), proposed an improvement to the Muskingum model which is used in hydraulics for flood routing and contains three parameters. He added a power parameter which added more flexibility to the model and improved the model accuracy to up to 80%. The present research proposes an improvement to the existing MDCEV model using a different approach from the abovementioned studies but close to Easa's work. Specifically, it focuses on a particular limitation in the MDCEV modelling structure: one of its parameters, the translating satiation parameter, does not capture the constant marginal utility effect. The proposed model adds a power parameter to the current structure of the MDCEV model which will be discussed in more detail in Chapter 3. This research is intended to improve the structure of the MDCEV model and to produce more accurate activity duration estimates, which will in turn improve the predictability of travel demand forecasts for mega-events and regular events.

3 Chapter: Improved MDCEV Model

3.1 Introduction

This doctoral research proposes an improvement to the existing MDCEV model by focusing on the utility structure. To proceed with describing the rationale for proposing improvements to the existing MDCEV model, this chapter presents the base case MDCEV utility structures and selects the best structure for the existing data set. Then, it presents the limitation in the existing formulation by evaluating the MDCEV model parameters, followed by the remedy to this limitation. Finally, criteria are presented for building the proposed model. In addition, it also evaluates various utility structures and the interactions between parameters.

3.2 Utility Structures

According to Bhat (2005), several research studies have used the linear expenditure system (LES) utility function form in environmental economics [e.g., Phaneuf et al. (2000), Phaneuf and Herriges (2000), von Haefen (2003), Herriges et al. (2004), and von Haefen and Phaneuf (2005)]. The LES utility function form is obtained using the following equation:

$$U = \sum \mathrm{B_a}\, ln(\mathrm{T_a} + \Omega_a) \qquad [23]$$

Kim et al. (2002) used a translated constant elasticity of substitution (CES) direct utility form as follows:

$$U = \sum B_a (\mathrm{T_a} + \Omega_a)^{\mu} \qquad [24]$$

Based on Bhat (2005), the interpretation of the baseline marginal utility is not direct since it is a function of the translating parameter. Bhat used the following utility function which is

based on a generalized variant of the translated CES utility function (Equation 25) and used the LES utility function type first used by Hanemann (1978) (Equation 26):

$$U_{Total} = \sum_{a=1}^{No.of\ Specific\ Activities} (\Omega_a/\mu)\exp((\beta_v X_v)_a + \varepsilon)(\frac{T_a}{\Omega_a} + 1)^{\mu}\text{-}1) \quad [25]$$

$$U_{Total} = \sum_{a=1}^{No.of\ Specific\ Activities} (\Omega_a)\exp((\beta_v X_v)_a + \varepsilon) ln(\frac{T_a}{\Omega_a} + 1) - 1) \quad [26]$$

The difference between the previous work and the work conducted by Bhat is that the utility structure results used by Bhat are a direct interpretation of the baseline utility unlike the other forms which combine the baseline utility with the translating effect.

Based on the above discussion the Bhat model will be used. However, the following are the three forms of the indirect utility portion of the out-of-home activities and in-home activities proposed by Bhat (2005) as shown by the following equations:

First form- $V_L = B_a + (\mu - 1)\ ln\ (T_a\ + 1)\ -\ ln(AVGa)$; $V_z = (\mu - 1)\ ln(z)$ [27]

Second form- $V_L = B_a\text{-}\ ln(T_a/\Omega a\ + 1)\ -\ ln(AVGa)$; $V_z = (\mu - 1)\ ln(z)$ [28]

Third form- $V_L = B_a + (\mu - 1)\ ln(T_a/\Omega a\ + 1)\ -\ ln(AVGa)$; $V_z = (\mu - 1)\ ln(z)$ [29]

Bhat (2005) used three parameters in the third form: baseline parameter; translating satiation parameter; and pure satiation parameter to model the trade off in time expenditure between in- and out-of-home activities. The baseline parameter represents the marginal rate of utility at zero-time expenditure point. The translating satiation parameter changes the shape of the indifference curve and the utility curve and ensures the corner solution. The pure satiation parameter shapes the utility curve and represents the diminishing marginal utility. The satiation

effect is then represented in two different ways: indirectly using the translating parameter, and directly using the pure satiation parameter. The first form (Equation 27) uses the baseline and pure satiation parameters, and the second form (Equation 28) uses the baseline and translating satiation parameters. Bhat suggested trying the three forms and choosing the form that best fits the data based on statistical and intuitive considerations. Habib et al., 2007 used the third from in their study. In the present research, the best form that fits the Makkah data is the third form (Equation 29), which will be the basis for improving the MDCEV model and includes the three parameters. The rationale for using the third form over the other two forms is that the two forms result in lower goodness of fit as compared to the third form. Conceptually, the translating satiation parameter captures an individual's participation in one specific activity during the day, which is logical for the residents of the city of Makkah, where they provide assistance to the millions of visitors during the entire day for mega-events. Moreover, the pure satiation parameter captures the diminishing marginal utility with the consumption of a specific activity, which is a natural behaviour for activities such as shopping. Based on that, the third form is best suited for this data set. The third form of the total out-of-home utility structure used by Bhat is as follows:

$$Total\ Utility = U_{Total} = \sum_{a=1}^{No.of\ Activities} (\Omega_a/\mu)\ B_a\ ((\frac{X}{\Omega_a} + 1)^{\mu}\text{-}1) \qquad [30]$$

Consequently, the marginal utility for one specific out-of-home activity given the quantity of good x (or activity x) is given by the following equation:

$$\frac{dU}{dX} = \quad B_a\ (\frac{X}{\Omega_a} + 1)^{\mu-1} \qquad [31]$$

The marginal utility value decreases along with increases in good quantity. Increases in the pure satiation parameter (μ) value cause increases in the marginal utility value. The shape of the marginal utility as a result of the use of various pure satiation parameter values (μ) is shown in Figure 15. The marginal utility reaches a flat line (μ =1) which represents the constant marginal utility effect.

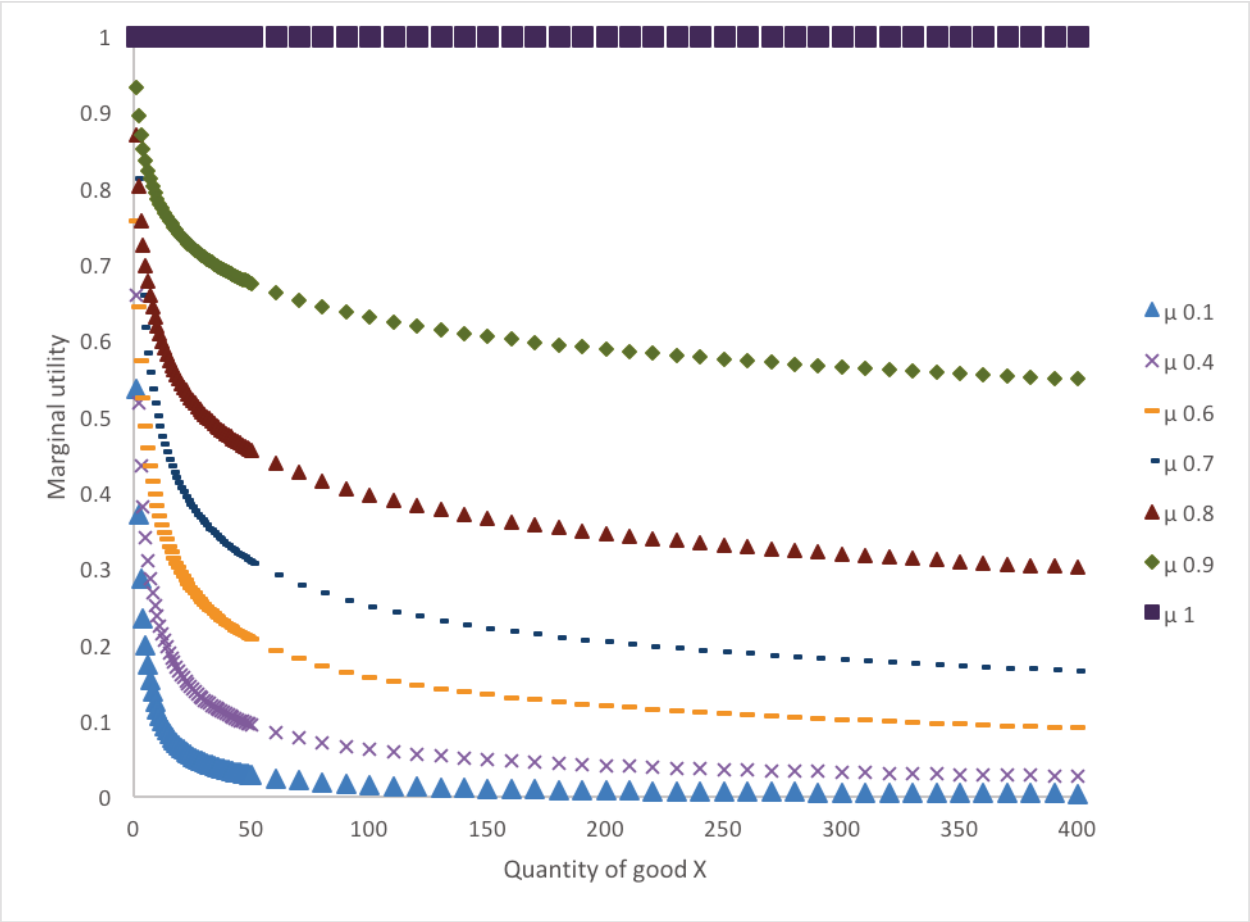


Figure 15: Effect of various pure satiation parameter values (μ) on the marginal utility.

As seen in Figure 16, the translating satiation parameter (Ω) appears to affect the marginal utility and, consequently, the total utility value. Increases in the Ω value cause increases in the marginal utility value. The marginal utility using Ω parameter does not reach a straight

line, as shown by the μ parameter. An apparent limitation to the model is that the translating satiation parameter value has to be very high in order for the marginal utility curve to reach a flat line.

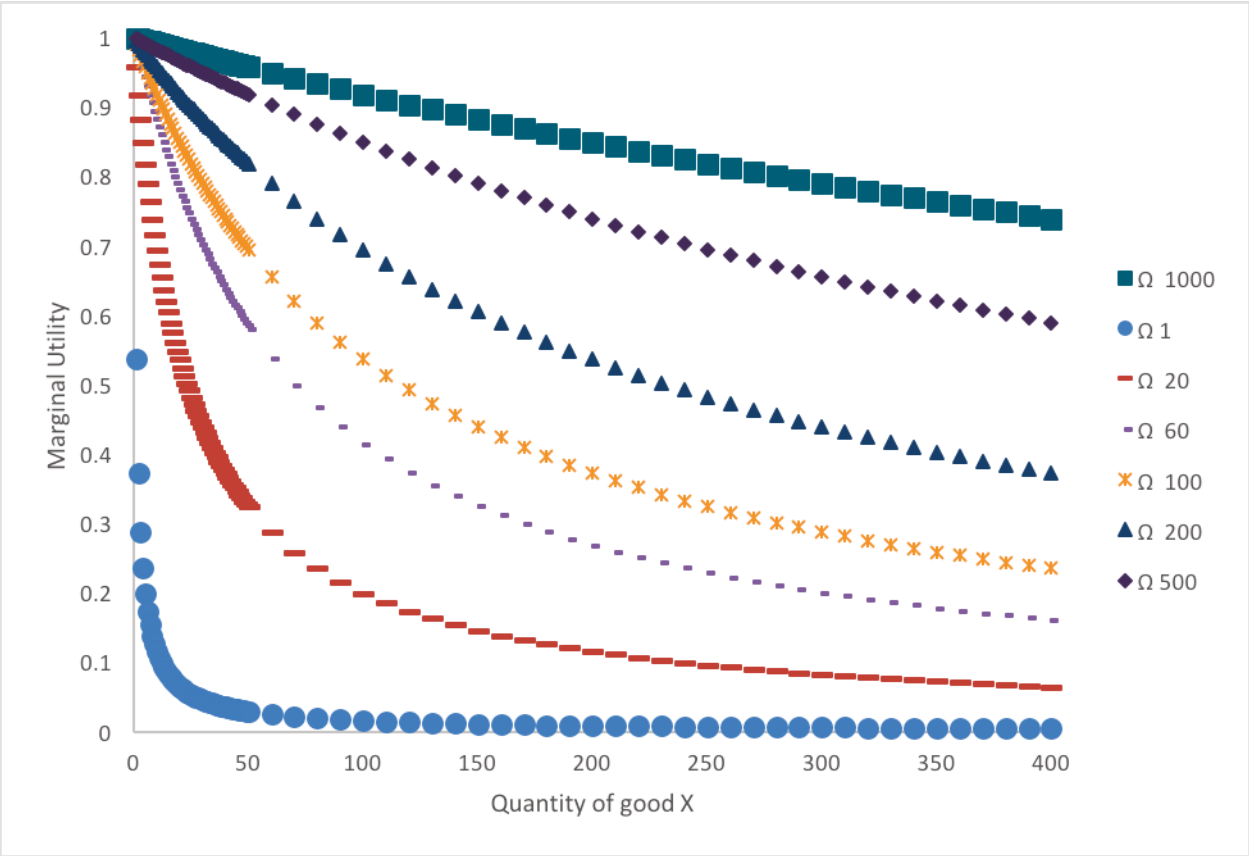


Figure 16: Effect of various translating satiation parameter values (Ω) on the marginal utility.

3.3 Improved MDCEV Modelling Structure

At this point in the research process, a new parameter was added to the Bhat utility function (Equation 30) in order to overcome the previously discussed limitation and consequently improve the estimated model. The new proposed structure adds a power, named Tpower, to the translating satiation parameter. Tpower determines the marginal utility function's rate of growth or deterioration, and specifies the function's shape. It allows the utility distribution

to range from a straight line to a diminishing curve with an increase in good consumption. The purpose of the parameter is to use different marginal utility distribution's shapes and ranges that better fit the data. In other words, it represents an individual behaviour in terms of diminishing marginal utility and consequently the tradeoffs between out-of-home activities. The proposed utility structure is as follows:

$$Total\ Utility = \sum_{a=1}^{No.of\ Activities} \left(\Omega_a{}^{Tpower}/\mu\right)\ B_a\ ((\frac{X}{\Omega_a{}^{Tpower}} + 1)^{\mu}\text{-}1)\ + \frac{1}{\mu} exp\ (\varepsilon_z)(z)^{\mu} \qquad [32]$$

Consequently, the marginal utility for one specific out-of-home activity given the quantity of good x (or activity x) is given by:

$$\frac{dU}{dX} = \ B_a \left(\frac{X}{\Omega_a{}^{Tpower}} + 1\right)^{\mu-1} \qquad [33]$$

When introducing the power to the translating satiation parameter, the following distribution was obtained:

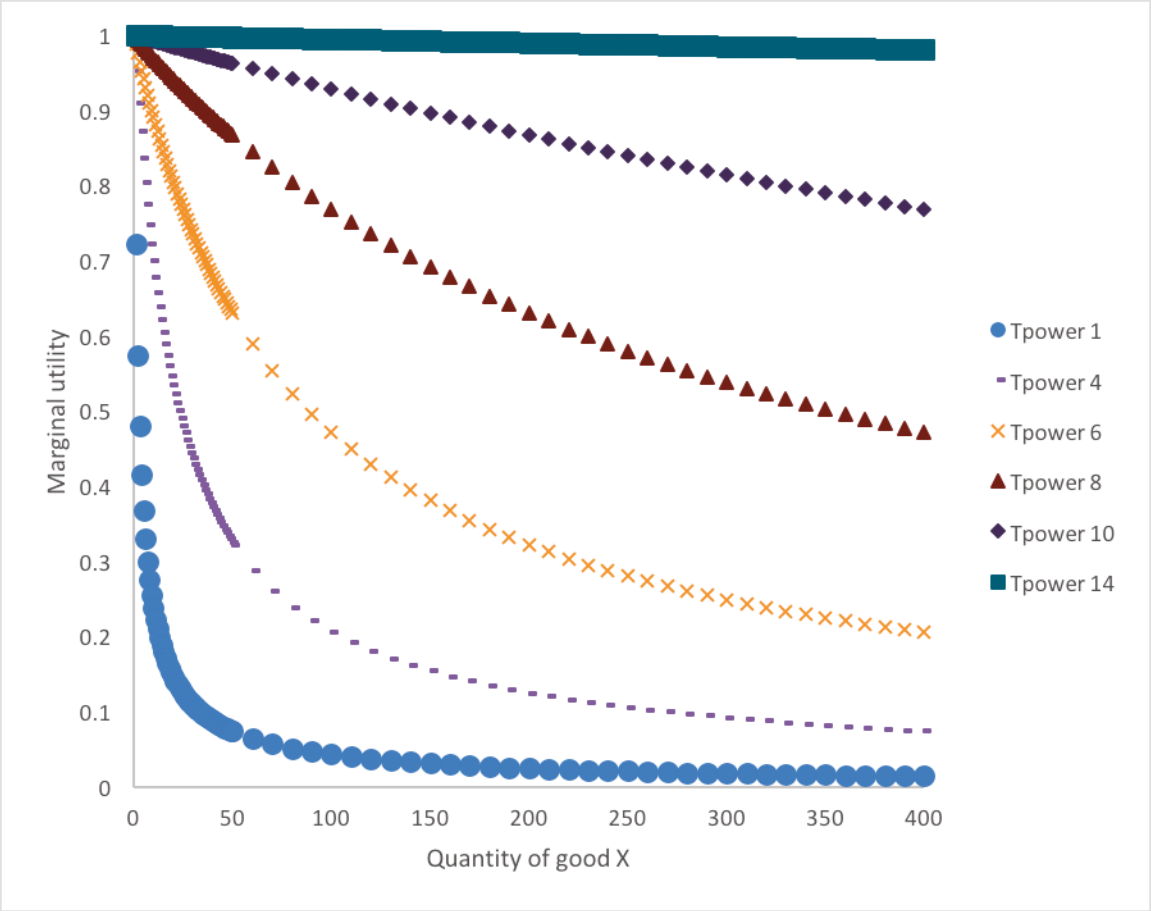
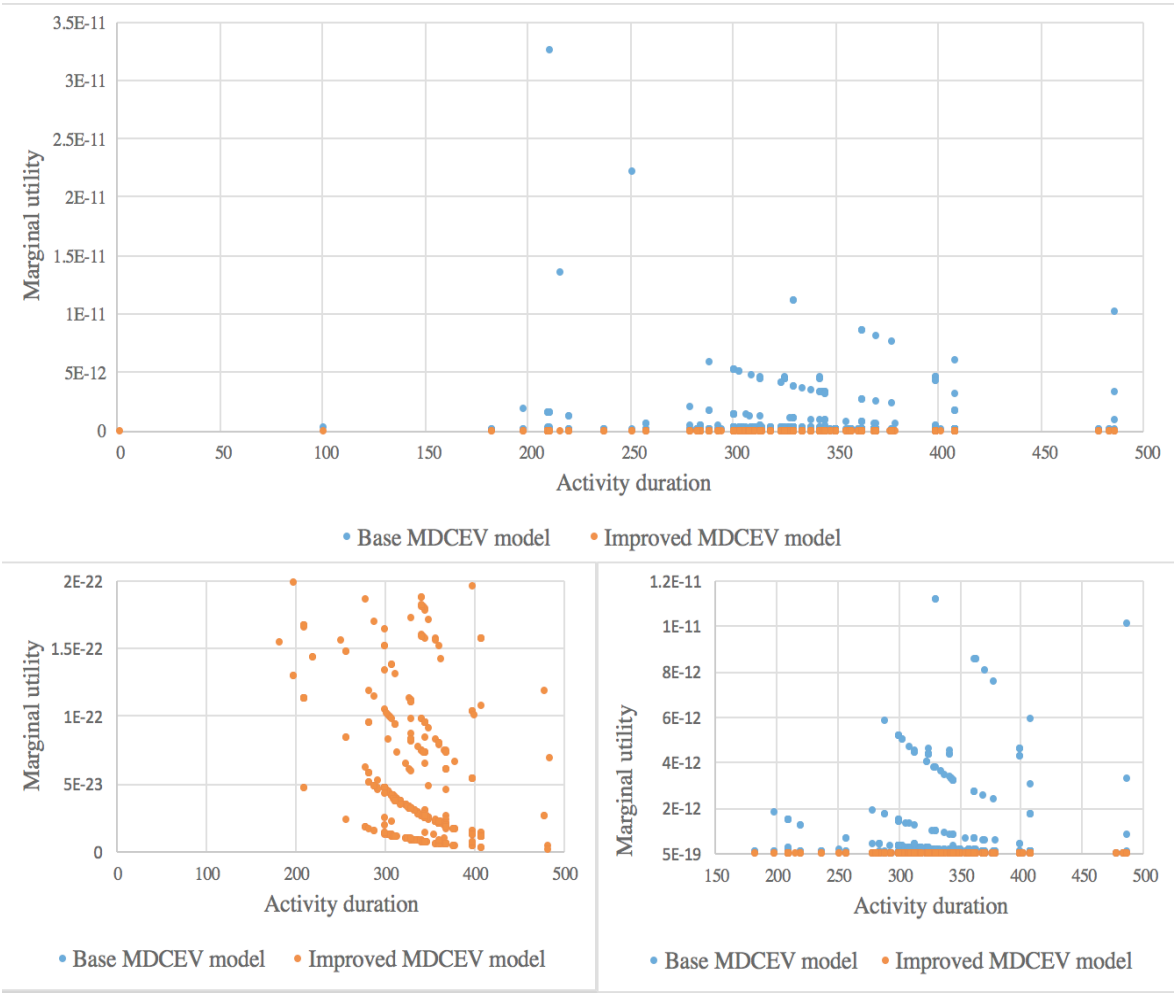


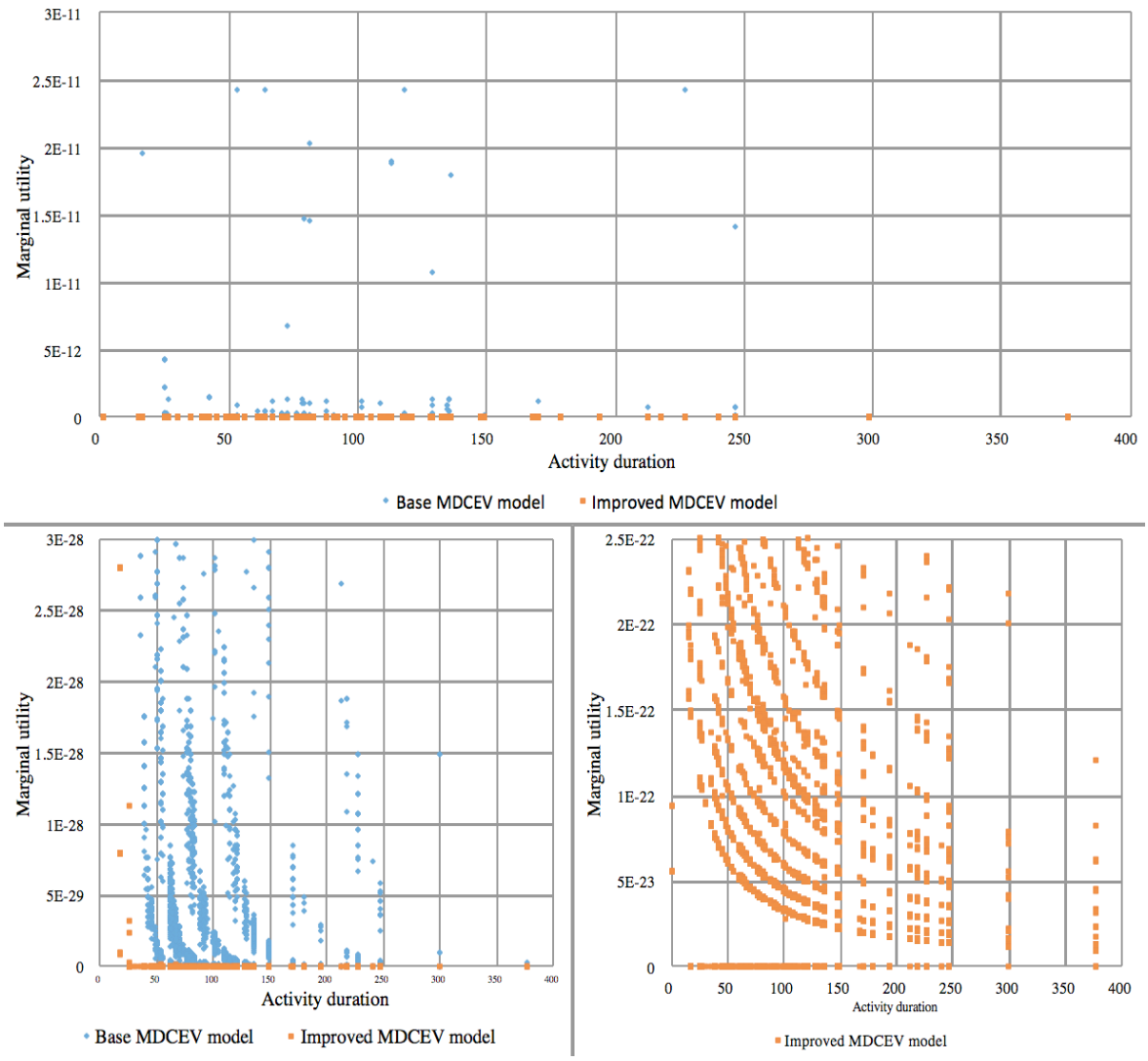
Figure 17: Effect of introducing an exponent to the translating satiation parameter on the marginal utility.

The diminishing rate of marginal utility using translating satiation parameter does reach a straight line after introducing Tpower 14 such that the Tpower value is 14 as shown in Figure 17. The constant marginal utility indicates that a person is willing to spend all his/her time on one specific out-of-home activity type such as work. The following is a practical implication of the base and the improved MDCEV model that shows the shortcoming of the base MDCEV model and the effect of introducing Tpower 14 and its function. By combining the effect of the three parameters (using the coefficients' values of the three parameters resulted from the estimated based and improved MDCEV Makkah model) the relationship between the marginal utility and

the activity duration time spent in a single work and shopping for both the base and the improved MDCEV model are shown in the following figures.



(a) Diminishing marginal utility with consumption of work activity for the base and improved MDCEV model.



(b) Diminishing marginal utility with consumption of shopping activity for the base and improved MDCEV model.

Figure 18: Relation between the marginal utility and the time spent at work and shopping activity for both the base and improved MDCEV model.

Tpower adds more flexibility to the MDCEV model by allowing the marginal utility curve to take different layouts and curvatures as shown in Figure 18. The base MDCEV model shows a high satiation effect compared to the improved MDCEV model for both single work and shopping activity. From Figure 18a, the base MDCEV model where Tpower is 1 is shown in blue

dots, reflecting the high diminishing rate of marginal utility, an example being where a person prefers to end his or her working day by spending a few hours in the vicinity of their workplace. However, the improved MDCEV model with Tpower 14 shows that a person is willing to spend more time at work compared to the base model. As work activities are a daily routine, the diminishing rate of marginal utility should not be sharp and should instead be flat, given that a person will spend their regular daily hours at work. For example, the lowest curve in the improved model shows that an individual spends around 500 minutes (around 8 hours) at work such that the marginal utility for work reaches zero in minute 500. On the other hand, the lowest curve in the base MDCEV model shows that an individual spends 300 minutes (roughly 5 hours) at work. Based on the ministry of labor kingdom of Saudi Arabia, the official working hours are 8 hours per day (Ministry of labor Kingdome of Saudi Arabia, 2017). However, these working hours differ from government to private sectors.

In terms of shopping activities, the base MDCEV model results in a faster diminishing of marginal utility compared to the improved MDCEV model, such that the improved MDCEV model indicates that an individual will spend more time shopping compared to the base MDCEV model. The base MDCEV model shows that the diminishing curve of marginal utility is an immediate satiation (perpendicular line on the activity duration axis), such that a person is not getting much benefit from shopping, which is not reasonable. The improved MDCEV model on the other hand, results in a lower satiation effect such that a person utility or benefit in spending time shopping reduces gradually. The coefficients' of each specific activity type together with the effect of Tpower shapes the diminishing marginal effect of a specific activity.

Tpower also reduces the variance of the translating satiation parameter, which results in a more accurate estimation of the model parameter and a better goodness of fit. The following

shows this effect statistically, the translating satiation parameter has an exponential distribution. One of the properties of the exponential distribution (see Equation 34) is the variance, which is given by $1/\mathrm{a}^2$. The variance shows that as the exponential function parameter value (a) increases, the variance decreases.

$$f(x) = \mathrm{a}\exp(-\mathrm{a}\,\mathrm{x}) \qquad [34]$$

By adding Tpower to the exponential distribution function we get the following function:

$$\exp(-\mathrm{a}\,\mathrm{x})^{Tpower} = \exp(\mathrm{Tpower}(-\mathrm{a}\,\mathrm{x})) \qquad [35]$$

The variance in this case is given by $1/(\mathrm{Tpower} * \mathrm{a})^2$. In other words, as the value of Tpower increases the variance decreases, hence a better goodness of fit as it is presented in Figure 19. Based on this discussion, the Tpower parameter allows for a wide range of diminishing marginal utility that best fit the data or peoples' behaviour, results in a more behaviourally rational consumption of time, and results in an improved goodness of fit.

3.3.1 Derivation of the Likelihood Function of the Improved MDCEV Model

The Kuhn-Tucker optimality conditions are applied to the utility function (Equation 32) while taking into consideration the time constraint. The indirect utility forms of a specific activity (V_L) and the indirect utility form of the total in-home activities (V_z) are as follows:

$$\mathrm{V_L} = \mathrm{B_a} + (\mu - 1)\ln(\mathrm{T_a}/\Omega_{\mathrm{a}}{}^{\mathrm{Tpower}} + 1) - \ln(\mathrm{AVGa}) \qquad [36]$$

$$V_z = (\mu - 1)\, ln(z) \qquad [37]$$

Then, applying the transformation variable theorem and the error term to drive the likelihood function used to estimate the coefficients of the activity duration model, which is given by:

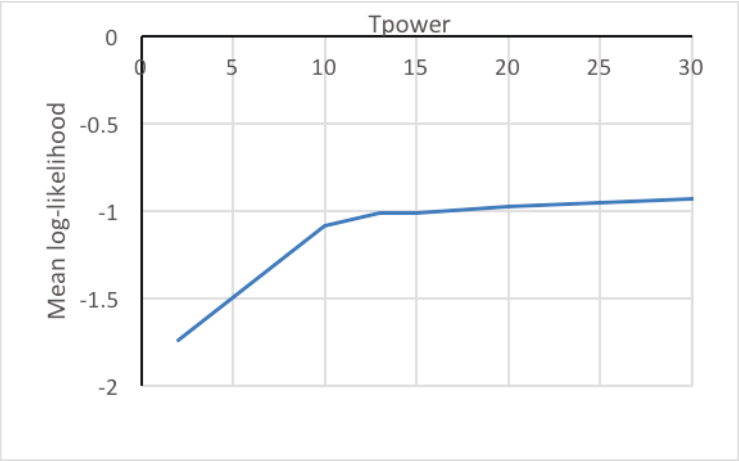
$$p(T_1, T_2, T_3 \ldots, 0,0, T_a) =$$

$$Ln\ (1/\sigma^{M\text{-}1})(\prod_{a=1}^{M}(M\text{-}1)\left(\prod_{a=1}^{M}\frac{1-\mu}{T_a+\Omega_a}\right)\left(\sum_{a=1}^{M}\frac{T_a+\Omega_a}{1-\mu}\ AVG\ FRQ\right)\left(\frac{\prod_{a=1}^{M}exp(V_L/\sigma)}{\sum_{L=1}^{L}exp\left(V_{L/\sigma}\right)^{M}}\right) \quad [38]$$

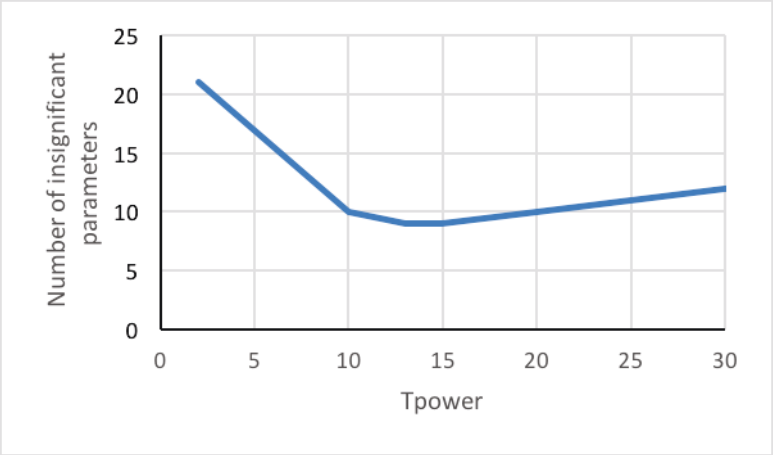
Based on Bhat (2005), the Jacobian term $\left(\prod_{a=1}^{M}\frac{1-\mu}{T_a+\Omega_a}\right)\left(\sum_{a=1}^{M}\frac{T_a+\Omega_a}{1-\mu}\right)$ did not consider the scale parameter (σ) when derived. Similarly, the Tpower parameter is not included in translating satiation parameter in the Jacobian term as shown in Equation 38 [see Bhat (2005) for Jacobian derivation and application details]. Bhat (2005) added a scale parameter (σ) to the utility function (see Equation 38). It has been reported that setting the scale value to 1 results in the best fitted model (Habib et al., 2007). Based on this scale selection approach, the following is the criteria for selecting the Tpower value. The pure satiation parameter already reached the constant marginal utility as shown in Figure 15. Therefore, adding the Tpower parameter to the pure satiation parameter will not improve the MDCEV model as tested in the parameter interaction section in this chapter.

3.3.2 Criteria for Selecting the Tpower Value

Several Tpower values were used to study the improvements in the mean log-likelihood using Makkah dataset (Figure 19a).



(a) Effect of various Tpower values on the mean log-likelihood.



(b) Effect of various Tpower values on the number of insignificant parameters.

Figure 19: Selecting the Tpower value.

As shown in Figure 19a, the improvement in the mean log-likelihood decreases along with the increase in the Tpower value. Figure 19b shows the number of insignificant parameters (based on 90% confidence interval) based on various values of Tpower. The minimum number of insignificant parameters is at Tpower 14. Bhat 2005 and Habib et al., 2007, advised to select the model form (Equation 27, 28, or 29) or the scale parameter that best fit the data as discussed previously. Based on that, Tpower 14 is selected because the improvement in the mean log-

likelihood is trivial and the highest number of significant parameters is at Tpower 14. It is difficult to parametrize Tpower, as the increase in the Tpower value will always lead to an improvement in the mean log-likelihood and in the goodness of fit. Therefore, Tpower is set to a specific value. Based on that, the model's users are advised to use their judgment in choosing the Tpower value that achieves the maximum number of significant parameters, and a high mean log-likelihood. The downside of having insignificant parameters is the reduction in the goodness of fit or the accuracy of the model. In addition, it results in irrational interpretation of the model parameters. Based on this, it is recommended that the model's users build their model again as a second step using the selected Tpower and include only the statistically significant parameters, which in return will further improve the accuracy of the model. These two modelling steps are shown in the Results Chapter where German model includes insignificant parameters (first step) and Makkah model include only statistical significant parameters (second step). In addition, a limitation of adding a high Tpower value is the decrease in the software estimation speed. However, the speed is acceptable for Tpower 14 as presented in the software comparison section (section 5.8) in Chapter 5.

3.3.3 Exponent Parameters' Interactions

The improved MDCEV model was built sequentially. To review, three utility structures proposed by Bhat (2005) were tested and the one with the highest goodness of fit was chosen for improvement. The best model based on the Makkah dataset was the model that included all three parameters—baseline utility, pure satiation and translating satiation. The other two models included only two parameters: either the baseline utility parameter and the pure satiation parameter, or the baseline utility parameter and the translating satiation parameter.

The effect of changing the scale parameter (σ) that affects the baseline utility parameter and the error term was tested. According to prior research, setting the scale value to 1 results in the best fitted model (see Habib et al., 2007). In the context of this study, as tested, setting the scale parameter to 1 does indeed result in the highest number of significant parameters (considering 90% confidence interval). Table 2 demonstrates this, and shows the results for Tpower values 10 and 14.

Table 2: The effect of the scale parameter on the mean log-likelihood and the number of significant parameters for different Tpower values.

Scale parameter	Tpower	Mean log-likelihood	Number of significant parameters
1	14	-8.058	47
2	14	-7.459	45
3	14	-7.347	45
4	14	-7.308	45
5	14	-7.291	44
6	14	-7.281	44
1	10	-8.327	48
2	10	-7.724	47
3	10	-7.611	47
4	10	-7.571	46
5	10	-7.553	45
6	10	-7.677	44
8	10	-7.642	44
9	10	-7.531	44

The baseline utility reflects the marginal utility at the zero-time expenditure point. The relationship between the baseline utility parameter and the marginal utility is linear. Thus, the baseline utility parameter is not related to the constant marginal utility effect unless the baseline value is fixed to a specific value, which is not feasible based on the formulation of the baseline utility parameter. Finally, the exponent Tpower parameter was added to the translating satiation

parameter as discussed in the previous section, which results in better goodness of fit and a higher number of statistically significant parameters compared to the base MDCEV model.

This section of the research examines various combinations of the exponent parameter. It is assumed that adding another exponent parameter and the interaction between these parameters could further improve the accuracy of the MDCEV model. As noted previously, the total utility of the base MDCEV model is given as follows:

$$Utility\ function = \left(\frac{\Omega_a}{\mu}\right) B_a \left(\left(\frac{X}{\Omega_a} + 1\right)^{\mu} - 1\right) + \frac{1}{\mu} exp\ (\varepsilon_z)(z)^{\mu} \qquad [39]$$

It contains three parameters: the baseline utility parameter (B_a), the translating satiation parameter (Ω_a), and the pure satiation parameter (μ). The scale parameter already affects the baseline utility parameter and does not account for the constant marginal utility effect, as noted above. Building on this, there is no need for further exploration of adding an exponent parameter to the baseline utility parameter. For the other parameters, various exponent combinations are applied to the translating satiation parameter and the pure satiation parameter with the aim of exploring if there are possible improvements that can be made to the accuracy of the MDCEV model. The following four forms are possible combinations of the exponent parameter applied to the translating satiation parameter and the pure satiation parameter, and the exponent is added to either of these or both, as follows:

Form 1:

$$\left(\frac{\Omega_a{}^{T}}{\mu^{S}}\right)^{A} = \left(\frac{\Omega_a{}^{TA}}{\mu^{SA}}\right) = \left(\frac{\Omega_a{}^{T^*}}{\mu^{S^*}}\right) \quad [40]$$

The utility function for form 1 is as follows:

$$Utility\ function = U = \left(\frac{\Omega_a{}^{T^*}}{\mu^{S^*}}\right) B_a\ ((\frac{X}{\Omega_a{}^{T^*}} + 1)^{\mu^{S^*}} \text{-}1)\ + \frac{1}{\mu^{S^*}} exp\ (\varepsilon_z)(z)^{\mu^{S^*}} \quad [41]$$

Form 2:

$$\left(\frac{\Omega_a{}^{T}}{\mu^{S}}\right) \quad [42]$$

The utility function for form 2 is as follows:

$$Utility\ function = U = \left(\frac{\Omega_a{}^{T}}{\mu^{S}}\right) B_a\ ((\frac{X}{\Omega_a{}^{T}} + 1)^{\mu^{S}} \text{-}1) + \frac{1}{\mu^{S}} exp\ (\varepsilon_z)(z)^{\mu^{S}} \quad [43]$$

Form 3:

$$\left(\frac{\Omega_a{}^{T}}{\mu}\right) \quad [44]$$

The utility function for form 3 is as follows:

$$Utility\ function = U = \left(\frac{\Omega_a{}^{T}}{\mu}\right) B_a\ ((\frac{X}{\Omega_a{}^{T}} + 1)^{\mu} \text{-}1)\ + \frac{1}{\mu} exp\ (\varepsilon_z)(z)^{\mu} \quad [45]$$

Form 4:

$$\left(\frac{\Omega_a}{\mu^{S}}\right) \quad [46]$$

The utility function for form 4 is as follows:

$$Utility\ function = U = \left(\frac{\Omega_a}{\mu^{S}}\right) B_a\ ((\frac{X}{\Omega_a} + 1)^{\mu^{S}} \text{-}1)\ + \frac{1}{\mu^{S}} exp\ (\varepsilon_z)(z)^{\mu^{S}} \quad [47]$$

An improvement to the model is attempted by adding an exponent (T or Tpower) to the translating satiation parameter (Ω_a), as in Form 3. The first and the second forms are identical such that $\left(\frac{\Omega_a{}^{T}}{\mu^{S}}\right) = \left(\frac{\Omega_a{}^{T^*}}{\mu^{S^*}}\right)$ and the value of T and S could differ or be the same in value. In total, there are three forms of the exponent parameters to be tested and compared (see Table 3 for possible combinations). First, these three forms are compared, considering the scale value of the baseline utility parameter being equal to 1. These combinations are applied to the Makkah and the German cities datasets which are from the obtained data and the previously referenced Habib et al., study of 2007.

Table 3: The four forms and possible T and S exponent parameters values.

Form number	T	S
Form 1, Form 2: base MDCEV model	1	1
Form 1, Form 2	2	2
Form 1, Form 2	2	3
Form 1, Form 2	3	2
Form 3: improved MDCEV model	2	1
Form 4	1	2

For the German dataset, the number of statistically significant parameters when the S parameter is 1 is always greater than the number of statistically significant parameters when the S parameter is 2 for the same T value. Additionally, in most cases where the S parameter is greater than 2, the model failed to converge. The low number of statistical parameters when the S parameter is equal to 2 reflects the instability of the model in this specific case (see Table 4). For the Makkah dataset, the only case when the model succeeds to converge is when the scale factor is set to 1.

Table 4: Mean log-likelihood and the number of statistical significant parameters based on German data set for scale parameter 1.

Scale parameter = 1	S = 1	S = 1	S = 2	S = 2
T	Mean log-likelihood	Number of significant parameters	Mean log-likelihood	Number of significant parameters
1	-17.759	37	-9.042	4
2	-12.765	53	-9.042	3
3	-10.811	52	failed to converge	
4	-9.873	52	failed to converge	
5	-9.335	52	-8.262	50
6	-8.988	52	failed to converge	
7	-8.747	52	failed to converge	
8	-8.570	51	failed to converge	
9	-8.434	51	failed to converge	
10	-8.327	50	failed to converge	
11	-8.241	50	failed to converge	
12	-8.1691	49	failed to converge	
13	-8.109	49	failed to converge	
14	-8.058	49	failed to converge	

In the second scenario, the scale parameter is set to 10. Different scale parameters are then tested in an attempt to reach a strong conclusion regarding the interaction of the exponent parameters. For the German dataset, results show that the model does not converge when the S parameter is more than 2 in value. Moreover, the mean log-likelihood is always smaller in value when the S parameter is equal to 1, compared to when the S parameter is equal to 2, which results in a better goodness of fit. In addition, the number of statistically significant parameters when the S parameter is equals 1 is more compared to the case where the S parameter is 2 (see Table 5). For the Makkah dataset, the model did not converge when the S parameter was greater than 1. It was found that in this context, adding an exponent to the pure satiation parameter complicates the model and makes it harder to estimate the MDCEV model. To illustrate, the model works well when the exponent of the pure satiation parameter is set to 1 in the context of the Makkah dataset.

Table 5: Mean log-likelihood and the number of statistical significant parameters based on German data set for scale parameter 10.

Scale parameter= 10	S = 1	S = 1	S = 2	S = 2
T	Mean log-likelihood	Number of significant parameters	Mean log-likelihood	Number of significant parameters
1	-16.768	33	-17.288	10
2	-11.850	53	-16.813	30
3	-9.941	53	-12.445	45
4	-9.115	50	-10.369	48
5	-8.622	50	-9.235	44

To sum, pure satiation parameter does not require many variables or an exponent parameter to increase the pure satiation parameter value and captures the constant marginal utility effect. Therefore, adding the exponent parameter to the pure satiation parameter will not improve the MDCEV model as tested. To conclude, the best-fitted model is the one that includes an exponent to the translating satiation effect only (Form 3, as noted above), and testing confirms that the exponent parameter interaction does not improve the accuracy of the MDCEV model.

4 Chapter: Data Collection and Preparation

4.1 Introduction

The previous chapters shed light on the MDCEV model and some of the variables used to estimate and analyze activity duration models. This chapter provides the research with the required data to model activity duration using the MDCEV model and to estimate activity models, which serve the purpose of completing the research tasks. It begins by presenting the organization that collected the data for the city of Makkah and their process of collecting it. Next, it shows the effort of reviewing and organizing this large data set. Then, it presents the variables included in the data, and discusses logical outputs. After that, it provides some background information about the cities of Karlsruhe and Halle and comparisons with the city of Makkah in terms of activities and city structure. Finally, it presents the variables available in the Karlsruhe and Halle data set used to model the cities' activity duration model. This data is the basis when examine the research tasks.

4.2 Makkah Data Set

This section describes the process of collecting the household travel survey for the city of Makkah, including the size of the data, the group of people included in the survey, and all implied assumptions when collecting this information. During the third and the fourth months of the lunar calendar in 2010 (i.e., Umrah and Friday seasons), the municipality of Makkah conducted a survey of 5,305 households for the residents of Makkah (1.5% of the total population). A household was defined as any number of people living together in the same house, where a single person living on their own also qualified as a household. Only family members who were eight years of age and older, and who stayed home at least three nights

during the weekdays were interviewed. As a preliminary step, the interviewer asked the householder in person for permission to conduct the survey, briefly explained the survey to the head of the household (i.e., the oldest male), then identified the household by address. The interviewer interviewed all eligible participants of the household using a fixed set of questions, which were asked in a particular order. The interviewer took measures to ensure that responses were complete and as clear as possible before leaving the household. Finally, each member was given an identification number for organizational and analytical purposes.

The survey questions covered three topics: household, members, and trips. Household-related questions asked about socioeconomic status and some general information (e.g., number of household members, number of cars owned, and the salaries of working members). Questions about household members requested information such as each person's age, occupation, place of work, and if individuals possessed a driver's license. Trip questions elicited information about the chain of trips made by each household member the day before the interview (e.g., start and end time of each trip in the chain, origin and destination, mode of transportation used, number of people in the car, and purpose of the trip). The origin of the second trip in the trip chain was defined as the destination of the first.

Data has to be in one file in order for both GAUSS and R software to be able to use it in the estimation process. Therefore, organizing the data involved matching each individual record from three different Excel spreadsheets (i.e., household information, individual information, and trip information) and merging it all into one. In addition, each individual record contained a maximum of eight activities during the day presented in long form (i.e., the individual sequence of trips are included in several rows) that had to be transformed into a wide form (i.e., all of the individual's and trip's information are included in a single row). As a result, each line of data

included an individual's record that contained a specific activity type, and all the individual's attributes. Individuals who spent more than 24 hours in out-of-home activities per day were removed from the data set. If a duplicate of an individual's ID existed, one of the IDs was chosen arbitrarily. Non responses were also removed from the data set. After organizing the data set and transforming it from three wide shape Excel sheets into one long shape Excel sheet, a total of 22,062 individual cases were retained for the model estimation. The eight out-of-home activities available in the data set are as follows:

1. Single stop work activity (ACT 1).
2. Single stop school activity (ACT 2).
3. Single stop shopping activity (ACT 3).
4. Single stop recreation activity (ACT 4).
5. Single stop other activity (e.g. hospital, clinic) (ACT 5).
6. Multiple recreation activity stops (e.g. trip from gym to restaurant) (ACT 6).
7. Multiple shopping activity stops (ACT 7).
8. Religious activities (ACT 8).

These activity types can be combined so that a tour may be made up of a single stop work and shopping activity so that an individual can make a shopping stop in a work tour. Or these activities could be used individually such that an individual work tour is independent of shopping tour. This combination could be done to better represent how an individual conducts his/her activities (Vuk, 2012). It is assumed in this research that an individual's engagement in a specific one out-of-home activity is independent from the others. The abovementioned diary was collected to model the travel demand of Makkah's residents, including all constraints and any

information needed to conduct activities and plan for the day. This is suitable for travel demand activity-based modelling and the MDCEV model. In addition, the diary contains complex and rich information that requires advanced models to capture it. Each resident made a plan based on some complex constraints (e.g., money, time, and space) as well as transportation network conditions. The diaries also contained heterogeneous travel plans for the population because each resident had a unique plan. These plans included obvious and unapparent scheduling attributes such as activity start time and location, mode of transportation, interactions among persons (i.e., family members, friends, or neighbours), dynamic route choice, and changes in plans during the day to reach travel destinations. These travel destinations are based on some basic rules and constraints, such as who will join the trip, the cost of the trip, and available time window. Also included are city constraints, such as work or school start and end times and their location and travel time based on road congestion levels. The diary also included optional trips, such as shopping, social visits, and religious activities. Irregular plan destinations resulting from social interactions, which were included in the diaries, and dynamic trips during the day were also included in the diaries (i.e., activity chain). However, these diaries do not contain trips that were intended to be made because a resource –time, vehicle, money, and so forth- was not available. This kind of data would help to examine how an individual changes his plan based on the previously mentioned resources. The following is the household travel information collected by the municipality of Makkah that reflects the previously discussed travel behaviour.

4.2.1 Data Description

The present research uses the MDCEV model to estimate activity duration for different purposes, using the survey-derived diary data described above. First, data variables were divided

into ratio/interval variables and categorical variables. Ratio/interval variables included the number of cars per household, number of parking spaces available per household, travel time, number of persons per household, occupancy (i.e. number of joining the trip), and age. The number of parking spaces was assumed to correlate with the number of cars; by this logic, the number of parking spaces available limits the number of cars and, consequently, the number of trips. It was also expected that as the travel distance increases, the number of trips will decrease and the duration of activities will increase. Regarding age, retired people tended to stay at home and have fewer out-of-home activities compared to younger people.

As mentioned in Chapter 2, categorization is expected to reduce the variability in each model and better predict the model. However, a minimum number of observations should be achieved in order to estimate the model parameters and obtain an unbiased model. Categorical variables in this diary included household income, gender, nationality, education level, work status, acquisition of a driver's license, and profession. The categories were: Household income [per month in Saudi Riyal (SR)] 1 – Less than 3,000 SR; 2 – 3,000–7,000 SR; 3 – 7,000–10,000 SR; 4 – 10,000–15,000 SR; 5 – 15,000 –20,000 SR; 6 – More than 20,000 SR; 7 – No response. The average family income in 2013 based on Saudi Statistics is 13,610 SR which is in the middle of the income category (General authority for statistics Saudi Arabia, 2013). Therefore, the gradation of income is reasonable. It was assumed that as income increases, the duration for work activities will also increase; because people who work more or work overtime gain more money compared with those who work less time. However, different types of jobs have different salaries therefore this assumption has its limitations.

There are other important factors to consider in relation to activity. For example, in Makkah only males eighteen years and older are allowed to drive; consequently, males were

expected to have more out-of-home activities than females. Gender was coded as 1 – male, 2– female. Moreover, different social groups have different needs, social characteristics, residents' status, and schools of thought which results in different travel behaviour. It was thus expected that each group will show different travel patterns. The identified groups included are shown in Table 6. The identification is given by the following example:

A Saudi who is undergraduate and works as businessman who owns private car can be coded as 1, 5, 1, 1, and 2.

Table 6: Identified groups from Makkah dataset.

Groups	Nationality	Education level	Work status	Profession	Mode of transportation
1	Saudi	Not educated	Work	Businessman	Private auto (chauffeur)
2	Arab	Preschool	Student	Specialist in science, art, and humanities	Private auto (owner)
3	Asian	Primary	Does not work	Technician in science, art, and humanities	Taxi (autos owned by a company)
4	European	Secondary	Retired	Writer	Private Taxi (auto owned by individuals)
5	Other	Under graduate	Other	Merchant	Minivans for goods
6	-	Master's/PhD	-	Services	Delivery truck (pickups)
7	-	-	-	Agricultural	Truck
8	-	-	-	Industrial	Private school bus
9	-	-	-	Engineering	Private company bus
10	-	-	-	Driver	Walk
11	-	-	-	Maid	Bicycle and motorcycle

4.3 Karlsruhe and Halle Data Sets

The German cities of Karlsruhe and Halle are approximately 500 kilometers apart from one another. The city of Karlsruhe is the second-largest city in Baden-Württemberg state, which is located in southwest Germany. The center of the city is the oldest part of the city and is home to the Schloss palace tower and radial streets go outwards from the tower (Koshizuka and Kurita

1986). Halle is located in the southern part of the Saxony-Anhalt region. It is an economic hub and an educational center for the central-eastern part of Germany. It has one of the largest and oldest universities in Germany and the city is considered as the heart of the central German area (Ioesn, 2017). The cities of Karlsruhe and Halle make a good comparison to the city of Makkah because they are similar in terms of their importance for their countries. Like the city of Makkah, Halle is an economic and an educational hub that attracts visitors and business from around the world. Meanwhile, Karlsruhe and Makkah have a similar transportation network: specifically, the network structure in both cities consists of radial roads originating from the city center (i.e., the oldest part of the cities) as shown in Figure 20. However, no two cities are typically alike. There is a difference for the mode of transportation choice between the two cities, which is one factor that affects travel behaviour. Karlsruhe's mode of transportation share is approximately 25% for each of walking, cycling, public transport, and private automobiles. The city of Karlsruhe contains an advanced public transportation system (tram train) that provides a rapid connection between the suburb with and the city center (Puhe and Schippl, 2014). The city of Makkah is planning to introduce metro lines and to improve the bus transportation system. Based on the obtained data from the city of Makkah, private auto users are around 50%, public mode is 34%, other modes of transportation (walking, cycling, and motorcycles) form 13%, and the rest share is for trucks. The mode shares and the city's structure are mainly related to the activity's travel time, which is one factor that could affect the activity's duration. However, each individual in both cities tries to maximize his/her own benefits in allocating time to specific out-of-home activities constrained by a time budget. Therefore, the individual's/family's or the characteristics along with the time budget constraints are the major factors in allocating activity's duration of the different families' structures.

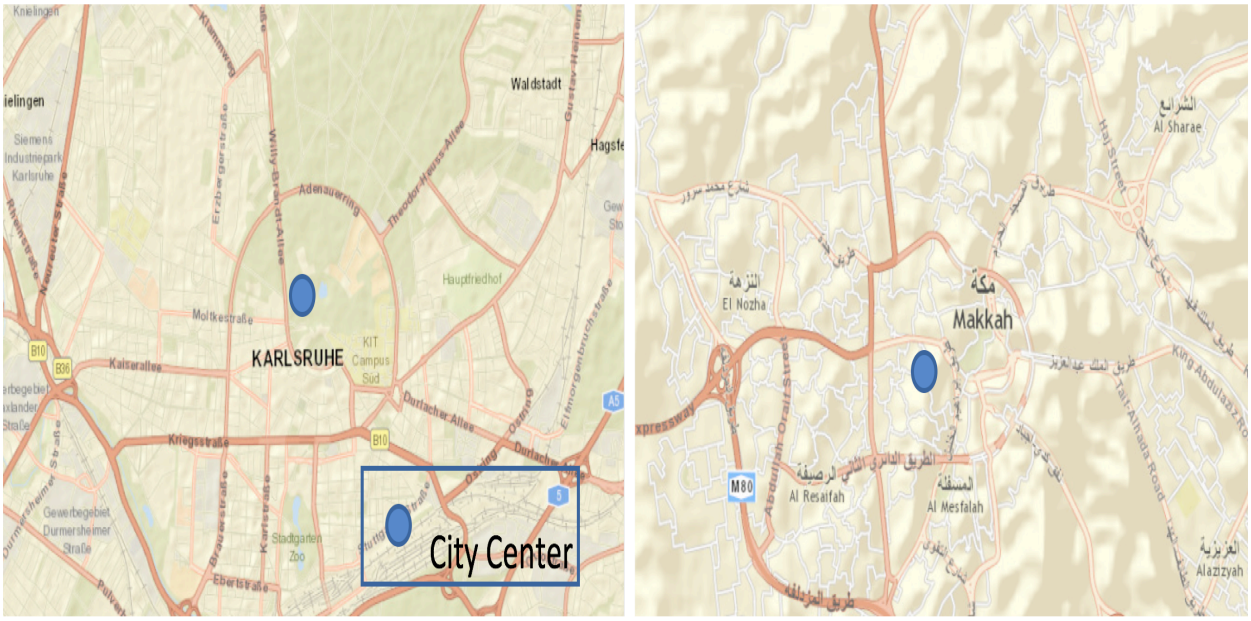


Figure 20: Radial roads starting from the city center of the city of Karlsruhe and the city of Makkah.

The travel diary MobiDrive was collected in 1999 and archived at the Institute of Transport Planning and Transportsysteme (IVT), Zürich, Switzerland. The goal of the data collection project was to capture the individual engagements in weekly life rhythms. However, the data did not show if there were a special events during the period of data collection. A week is considered as the time frame for their study (i.e., 168 hours). Individuals older than 6 years old completed a weekly survey form and 333 individuals were selected for modelling (Habib et al., 2008). The out-of-home activity types considered are:

1. Basic Needs (e.g., walk/stroll, lunch outside home, and household obligation activities such be present at kids event).
2. Work / School activities.
3. Pick up / drop off person.
4. Shopping activities which include window shopping.

5. Services (e.g., auto service, auto refueling, and gardening).

6. Recreation and entertainment activities (e.g., sports).

7. Social activities (e.g., restaurants, meeting with friends, and families).

8. Other activities.

This household's travel survey data contains un apparent information regarding how a person allocates time in various out-of-home activities during the week. The data, contains all constrains (e.g., money and time) during the week that affect the individual's time spent in out-of-home activities. In addition, it includes how an individual changes his/her plan during the week based on the person's constraints (e.g., time and money constraints) and the city's constraints (e.g., road closures). However, the planned activities are not collected in this travel survey. In addition, the survey includes the dynamic time allocation during the week such that how an individual allocates shopping activities across different weeks. For example, a person could be shopping once a month or weekly.

4.3.1 Data Description

This research uses the household travel survey data described above to model activities in the German cities using MDCEV model. The variables that are included in the dataset that are used in modelling are: city specific, gender, acquisition of a driver's license, existence of parents in home, home location, student status, employment status, marital status, existence of vehicles per household, and number of children per household. The city specific dummy variable, which consists of Karlshruhe and Halle. The gender variable includes male and female. Driver license variable which is either have or does not have a driver license. The parents in home variable

include either the existence of a parent in home or not. Home location includes if the home is located in the central business district (CBD) or not. Student variable include either a student or not, which is the same for employment and marriage. Number of vehicles per household is divided into household with no vehicles and with vehicles. Number of children per household is divided to either one child or more than one. Categorical variable includes age, which is divided into eight categories a follows: 1- (21-30), 2- (31-40), 3- (41-45), 4- (46-50), 5- (51-55), 6- (56-60), 7- (61-65), 8- (65+). The continuous variables include number of employed members of household, distance from home to bus stop, number of people per household.

The size of the city is an indicator of the variety of activities in the city. Moreover, larger cities have more verities in activities which reflect a higher demand; hence the residents of larger cities spend more time out-of-home than smaller cities. However, residents who live in the CBD usually have their needs met without traveling out of the CBD area. Therefore, they do not need to spend more time out-of-home. The increase in the number of vehicles per household reflects the need to conduct more out-of-home activities. Employees are expected to spend more time out-of-home compared to unemployed since they have to spend time at work. However, it is hard to comment on whether male spend more time on social and service activities compared to female. The increase in the number of children per household should increase the time spent on shopping such that there is a need to spend more time shopping for the larger family compared to a smaller family.

5 Chapter: Application Results

5.1 Introduction

The data in the previous chapter is used to estimate the tasks of this research that include pilot study, developing the improved MDCEV model, generalize the improved MDCEV model, implement the model in two statistical soft wares, and evaluating the store closing policy for the city of Makkah. Finally, all the above tasks will be used in building a conceptual travel demand modelling framework for a mega-event. The results of these tasks will be presented in this chapter.

5.2 Pilot Study - Evaluation of the MDCEV Model

This section presents the results of a pilot study that was conducted as part of the present dissertation work in order to examine the application of the MDCEV model to Makkah. The results comprise the baseline results to be compared with the improved MDCEV model of this dissertation. The study used the specifications of the model developed by Bhat (2005). The hypothesis of this study is that during a special event on Friday (i.e., Friday prayer) Makkah's residents' travel behaviour differs from the rest of the week. The hypothesis was tested to model the time expenditure behaviour for the residents of Makkah and to assist in constructing a travel demand framework for the city. The following summarizes the pilot study and its major findings.

In Makkah, any travel demand framework should recognize the differences in travel behaviour not only between weekdays and weekends, but also on Fridays, which holds particular significance for Muslim residents. Most residents spend time with their families and friends on Thursdays, which is considered the first day of the weekend, while only some go to work. Millions perform the Friday prayer in Al-Haram which takes place around noon on Fridays, i.e., the second day of the weekend, and most stores are closed until late afternoon. This research

study intends to fully capture the activity pattern of the residents of Makkah by presenting modelling of the activity duration for the residents on weekdays, Thursdays, and Fridays.

In 2010, the municipality of Makkah collected the aforementioned full-day travel diary for the residents of Makkah for all days of the week. The relationship between the variables and the expected activity duration outcome are presented in the Chapter 4. For this pilot study, the diary data were separated into the weekday and weekend categories. An MDCEV model was then used to estimate the activity duration for each of the three separate groups (i.e., weekdays, Thursdays, and Fridays), and a fourth model was used for all the days of the week. The four empirical models (i.e., weekdays, Thursdays, Fridays, and all-days) modelled the differences in time consumption for the residents of Makkah for each group. Each model comprised eight out-of-home activity durations with a time budget constraint of 16 hours for out-of-home activities. Individuals with more than 960 minutes (i.e., 16 hours) of out-of-home activities are excluded from the data because data was collected assuming residents spend at least eight hours a day at home (e.g., sleeping and doing other in-home activities).

The mean log-likelihood and goodness of fit varied between the four models. The goodness of fit for the Friday model was the highest, highlighting the importance of including Friday as a separate category from weekdays and Thursdays. Rho Square was the index used to measure the goodness of fit for this model as given by Equation 48. It ranged from between 0 and 1 (1 representing the best-fit), where values between 0.1 and 0.3 are considered reasonable for such a complex model (Habib et al., 2007). The Rho-square values in this study showed an extraordinarily high value of 0.69. Figure 21 and Table 7 show the details of the goodness of fit of each of the four models.

$$\text{Rho Square} = 1 - \frac{\text{Explained information}}{\text{Total information}} = 1 - \frac{\text{Mean log-likelihood of the full model}}{\text{Mean log-likelihood of the null model}} \quad [48]$$

The variability in travel plans was the lowest for the Friday model since it has the highest goodness of fit. This is a reasonable output because most people spend their mornings at home, as there are no working or shopping activities until the afternoon. In other words, Makkah's residents have only a half-day to conduct out-of-home activities and subsequently a limited number of out of-home activities to engage in.

Table 7: Summary of the number of parameters and number of cases for the pilot study.

Goodness of fit model components	All-days	Weekdays	Thursdays	Fridays
Total number of cases	18434	18434	2700	928
Number of parameters of the full model	59	65	54	61
Number of parameters of the null model	3	3	3	3

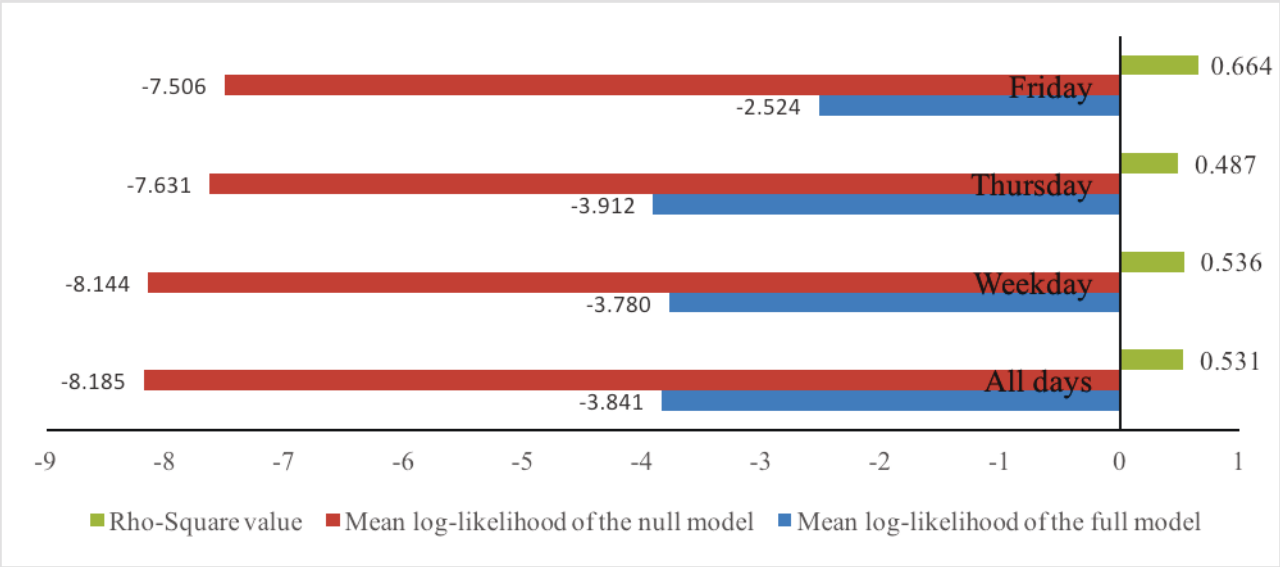


Figure 21: Goodness of fit for the pilot study.

The parameter coefficients varied from 54 to 65 based on the model (Table 7 and Figure 21). A 90% confidence interval was used to test the significance of the parameters and to decide whether to keep or omit the coefficient for each variable. Only statistically significant parameters are kept in the model. As previously discussed, the model has three parameters: translating satiation, pure satiation, and baseline utility. The pure satiation parameter (μ) contained a constant and the number of people per household (see Table 8). In the all-days and the weekday models, the coefficients for the number of people per household were negative. This indicates that the number of people per household affects activity duration during weekdays, but not weekends. In addition, as the number of residents per household increases, the time spent out of home decreases. This is not a logical explanation resulting from the base model. The increase in the number of people per household is expected to increase the time spent out of home, such that a family of eight needs more time for shopping compared to a family of two.

Table 8: Pure satiation parameters for the four models.

Pure satiation components	All-days	Weekdays	Thursdays	Fridays
Constant	-2.146	-2.158	-2.075	-2.474
Number of people/household	-0.003	-0.003	0.000	0.000

The combined number of people taking part in the activity and the mode of transportation are the variables that define the translating satiation parameter (Ω). The highest occupancy coefficient is for Thursday, which indicates that the residents prefer to spend time out-of-home when the travel group is large (Table 9).

Table 9: Constants and occupancy coefficients for the four models.

Translating satiation components	All-days	Weekdays	Thursdays	Fridays
Constant ACT1	5.496	0.630	2.738	3.631
Constant ACT2	5.496	0.630	2.738	3.631
Constant ACT3	-0.568	-1.44	-1.542	3.631
Constant ACT4	2.068	0.840	-1.980	3.631
Constant ACT5	0.748	-1.260	-6.069	3.631
Constant ACT6	2.146	0.985	2.738	3.631
Constant ACT7	1.681	0.841	-0.967	3.631
Constant ACT8	1.852	1.914	2.880	3.631
Occupancy	0.353	0.483	2.839	1.415

Based on the values obtained for the mode of transportation parameter, residents who use a private company bus spend less time in out-of-home activities such as work compared to residents who use their private car or motorcycle (i.e., the highest coefficient value) on weekdays. On Thursdays, company owned taxis are the least preferred mode of transportation when spending time outside of the home. On Fridays, both company owned and individually owned taxis are the least preferred mode of transportation when spending time outside of the home (Table 10).

Table 10: Mode of transportation coefficients for the four models.

Mode of transportation	Weekdays	Thursdays	Fridays
Private auto (chauffeur)	0.503	0.000	-0.948
Private Auto	2.559	1.810	-1.460
Taxi (owned by a company)	0.000	-1.156	-3.293
Taxi (owned by individuals)	1.095	0.000	-2.637
Private school bus	0.000	3.845	-1.407
Private company bus	-1.059	0.000	-1.013
Walk	0.438	0.000	0.000
Motorcycle	2.461	6.908	0.000

The baseline utility for out-of-home activities uses resident information such as constant, gender, and age for each model. Based on the values obtained for the constant parameter, residents prefer to spend more out-of-home time on Thursdays compared to the rest of the week (Table 11). Based on Zhong et al., (2007), during weekends, residents spend more time on shopping and recreational activities compared to weekdays. Hook (2012) found that some family members work overtime during the weekends, and spend less time with their families. In other words, the residents not only spend time on shopping and recreational activities during weekends, but they also work in shopping malls, recreational facilities or in private companies as well. In the city of Makkah, residents also work during the weekends to serve the high number of visitors visiting Al-Haram. Travel time and driver's license values were found to be significant only for the Friday model. The multiple recreational or social activity parameters contained the highest coefficient values, indicating that residents spent more time on these specific activities (Table 13). As discussed previously, only males 18 years and older are allowed to drive, and males show as preferring to drive to work, school, and single shopping destinations. Since the rest of the activities are mostly conducted by females, who are not allowed to drive, the coefficients were negative (Table 12). The coefficient values for parameters vary for different age groups as well (Figure 22), showing a difference in time expenditure behaviour between the groups.

Table 11: Baseline utility constants for the four models.

Baseline utility components	All-days	Weekdays	Thursdays	Fridays
Constant ACT1	-30.312	-30.638	-27.630	-41.346
Constant ACT2	-30.312	-30.638	-27.630	-34.057
Constant ACT3	-30.312	-30.638	-27.630	-39.790
Constant ACT4	-30.312	-30.638	-27.630	-38.107
Constant ACT5	-30.312	-30.638	-27.630	-37.929
Constant ACT6	-30.312	-30.638	-27.630	-44.329
Constant ACT7	-30.312	-30.638	-27.630	-41.346
Constant ACT8	-30.312	-30.638	-27.630	-41.782

Table 12: Gender specific dummy variable for the four models.

Gender specific dummy variable (male)	All-days	Weekdays	Thursdays	Fridays
ACT1	4.208	4.136	3.896	0.693
ACT2	6.306	6.297	6.155	0.693
ACT3	2.936	2.866	2.952	0.693
ACT4	3.053	3.069	3.670	0.693
ACT5	-0.468	-0.480	0.000	0.693
ACT6	-1.173	-1.115	-1.182	0.693
ACT7	-2.324	-2.472	-1.933	0.693
ACT8	-1.375	-1.393	-1.066	0.693

Table 13: Travel time and driver license coefficients for Friday model.

Baseline utility components	Travel time	Driver license
ACT1	0.155	1.344
ACT2	0.120	0.420
ACT3	0.159	1.789
ACT4	0.229	-0.009
ACT5	0.169	-0.009
ACT6	0.497	-0.009
ACT7	0.497	-0.009
ACT8	0.328	-0.009

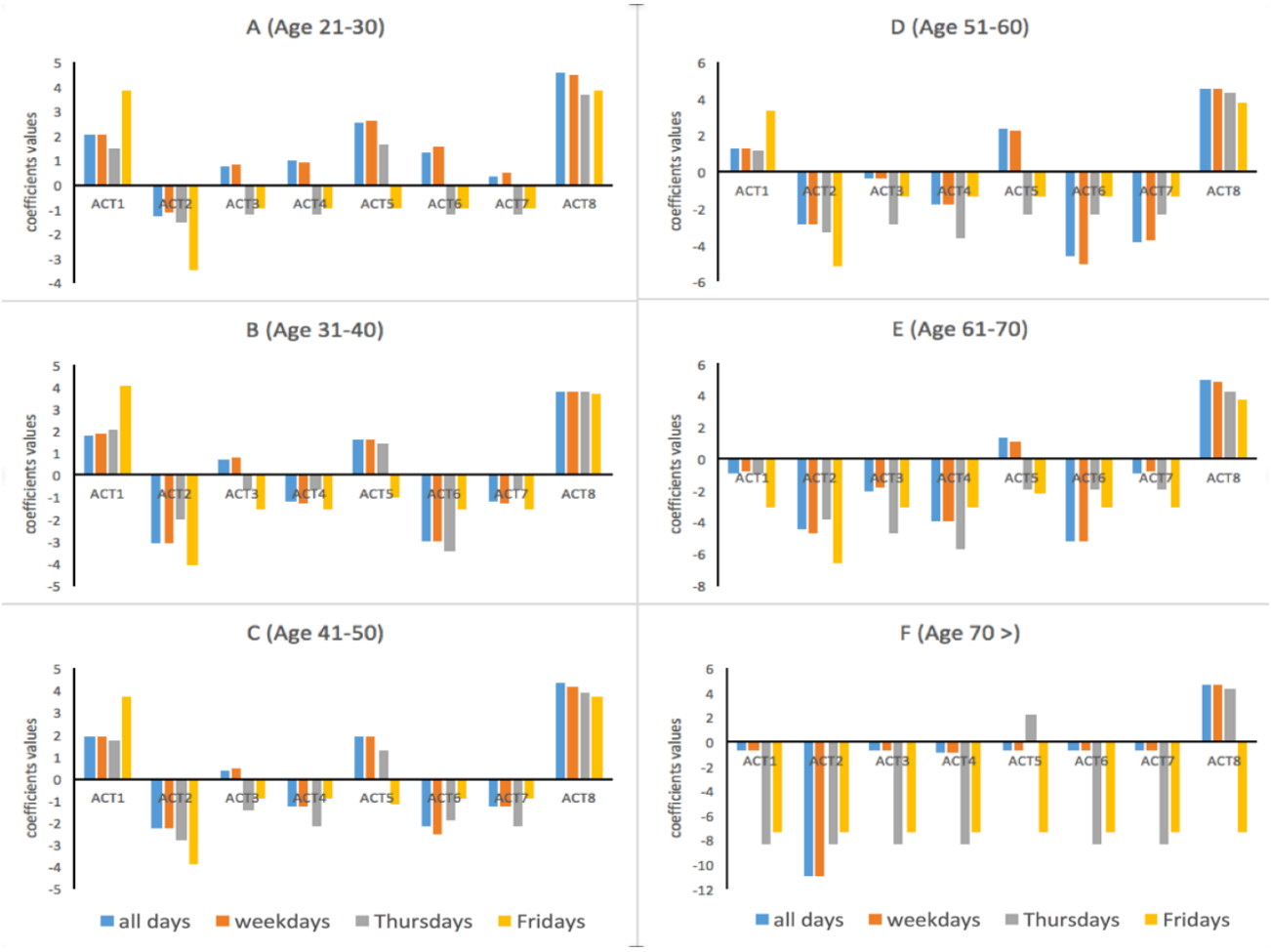


Figure 22: Age variable coefficients for different age categories for pilot study.

In summary, variations in the signs and values of the coefficients of the parameters were visible across all of the models, supporting the hypothesis that there are differences in time expenditure behaviour on weekdays, Thursdays, and Fridays. More specifically, the observed differences are related to the fact that the study took weekdays, weekends, as well as the specific event of Friday prayer into consideration as affecting all activities.

This pilot study was conducted to ensure that the MDCEV model was working as anticipated; to study the time expenditure behaviour for the residents of Makkah; to help in

building the travel demand framework for the city; and to obtain the base results to compare with the improved model. The models captured the trade-off in time expenditures between in- and out-of-home activity duration with a high goodness of fit (0.664), assuming that individuals spent a minimum of eight hours a day resting or sleeping.

5.2.1 Improved MDCEV Model Performance Using Makkah Data Set

In this section, results from the pilot study that used Bhat's model specification are compared with the improved modelling structure proposed in this thesis. More specifically, this section discusses the differences and similarities between the MDCEV model when estimated using the base and the improved model structures. Rho-Square index is used to measure the accuracy of the model where 1 is the best-fit model and 0 is the worst-fit model.

In the base model, highest goodness of fit (0.66) was found in the Friday model. In the new model, all four models have similar goodness of fit (0.85), an improvement from the base model. It can be said, therefore, that the new model explains the variability in the data better than the base model. Based on Rho-Square index values, the percentages of improvement for the four models are 26.9% for the Friday model, 74.3% for the Thursday model, 59.9% for the Weekdays model, and 60.6% for the all days' model. The number of cases within each of the four models is the same as the pilot study. A 90% confidence interval is considered to decide whether to keep or omit the variables' coefficients. All variables' coefficients in the results are statistically significant in both the base and the improved models.

The parameter variables contained in the new model in the baseline utility parameter are constant: gender, age, travel time, number of vehicles per household, and acquisition of driver's

license. The constants are negative in value and similar in all the four models, which means that Makkah's residents' preference for spending time out-of-home is the same (Table 14).

Table 14: Constant value in the baseline utility parameter for the four models of the improved model structure.

Baseline utility components	All days	Weekdays	Thursdays	Fridays
Constant ACT1	-49.249	-50.406	-46.846	-59.747
Constant ACT2	-49.461	-50.599	-45.912	-53.149
Constant ACT3	-50.375	-51.525	-50.061	-57.900
Constant ACT4	-51.180	-52.443	-50.134	-58.300
Constant ACT5	-51.175	-52.334	-49.311	-56.413
Constant ACT6	-55.919	-57.047	-54.713	-65.235
Constant ACT7	-56.351	-57.570	-54.080	-59.747
Constant ACT8	-52.135	-53.338	-50.299	-57.810

As mentioned earlier, in Makkah only males 18 years and older are allowed to drive. It is logical to say that the acquisition of a driver's license increases the number of out-of-home activities and, consequently, activity duration. In the base model, the acquisition of driver's licenses was found to be positive for ACT 1, 2, and 3, and negative for the other activities. In the new model, all of the variable coefficients' signs are positive for all out-of-home activities, which imply that acquisition of driver's licenses allows the residents to spend more time out-of-home. The number of vehicles per household variable is significant in all four improved models, but was not significant in the base model. An increase in the number of vehicles per household increased the time spent at work on Fridays compared to the rest of the week. This is a logical result of the residents offering services for the millions of prayers at Al-Haram on Fridays. On the other hand, an increase in the number of vehicles per household reduced the time spent in other out-of-home activities during the rest of the week. (For context, Makkah residents own a higher number of vehicles per household –five vehicles per household– than the residents of

other cities worldwide. The high number of vehicles per household is a cultural issue, in part a result of the low and consistent oil prices in the region.) In other words, owning more than one vehicle does not mean that the household members are using all of them simultaneously. Consequently, interpretation of this variable is different for the case of Makkah than for other cities, as will be presented in the next section. To illustrate, it will also be mentioned that an increase in the number of vehicles per household increases the activity duration out-of-home.

The base and the improved model also studied the difference in time consumption for each specific out-of-home activity by gender. The positive coefficient values indicate that males spend more time engaged in out-of-home activity as compared to females. The base model contains disparities in the coefficients' signs and values between the out-of-home activity types among the four models for the gender-specific variable. However, in the new model, gender-specific coefficient contains only one value for each one of the four models (Table **15**).

In the base model, for the all days model, males spend less time than females in multiple recreational activities, multiple shopping activities, other activities, and religious activities. In the new model, females spend less time in all out-of-home activities in all the four models compared to males, which is a more logical result compared to the base model because of the cultural restrictions on driving.

In the base model, travel time variable was found to be significant in the Friday model only. In the new model, for all of the three models except Friday, the travel time variable was significant. The Thursday model contained a specific constant for each out-of-home activity. However, in the all days and weekdays models a single coefficient was included for each model. All the values of the coefficients are positive, indicating that the increase in the travel time to an activity increases the residents' time spent out-of-home for a specific activity. Travel time is

added to this modelling framework based on the assumption that an individual makes a decision regarding activity duration based on the travel time. Based on a review of the literature, there is no unified modelling framework that states which modelling component is an input to which such as mode of transportation is an input to activity duration or the opposite.

Table 15: Driver's license, number of autos per household, gender, and travel time coefficients for the improved models.

Four models	Activity type	Driver's license	Number of autos per household	Gender	Travel time
Friday	ACT1	0.000	0.861	0.933	0.000
	ACT2	0.000	-0.368	0.933	0.000
	ACT3	0.000	-1.035	0.933	0.000
	ACT4	0.000	-0.753	0.933	0.000
	ACT5	0.000	-0.753	0.933	0.000
	ACT6	0.000	-0.753	0.933	0.000
	ACT7	0.000	-0.753	0.933	0.000
	ACT8	0.000	-0.204	0.933	0.000
Weekday	All activities	0.000	-0.017	1.231	0.127
All days	All activities	0.000	0.008	1.261	0.127
Thursday	ACT1	0.522	-0.106	1.031	0.118
	ACT2	0.522	-0.106	1.031	0.112
	ACT3	1.576	-0.106	1.031	0.117
	ACT4	2.244	-0.106	1.031	0.139
	ACT5	0.522	-0.106	1.031	0.204
	ACT6	3.009	-0.106	1.031	0.269
	ACT7	0.522	-0.106	1.031	0.180
	ACT8	0.762	-0.106	1.031	0.163

Seven age categories were used to account for the non-linearity of age effect on time expenditure behaviour. In both the base and the new model, the age coefficient decreases with the increase in age. In other words, as age increases, activity duration spent out-of-home decreases. This behaviour is better shown by the results of the new model compared to the base model (Figure 23).

In both the base and new model the residents of Makkah spend more time at work during Fridays compared to weekdays. Unlike the base model, the new model shows that the residents spend more time in Al-Haram on Friday as compared to weekdays. This is an intuitive result since millions of residents go to Friday prayer in Al-Haram and use many services offered before and after the prayer.

The new model shows that the residents spend more time on shopping and recreational activities during the weekend, as opposed to the base model that shows that residents spend more time on this activity during weekdays. This finding is consistent with many other researchers (e.g., Zhong et al., 2007) who indicate that people conduct leisure and shopping activities during weekends. However, residents of other cities might behave differently.

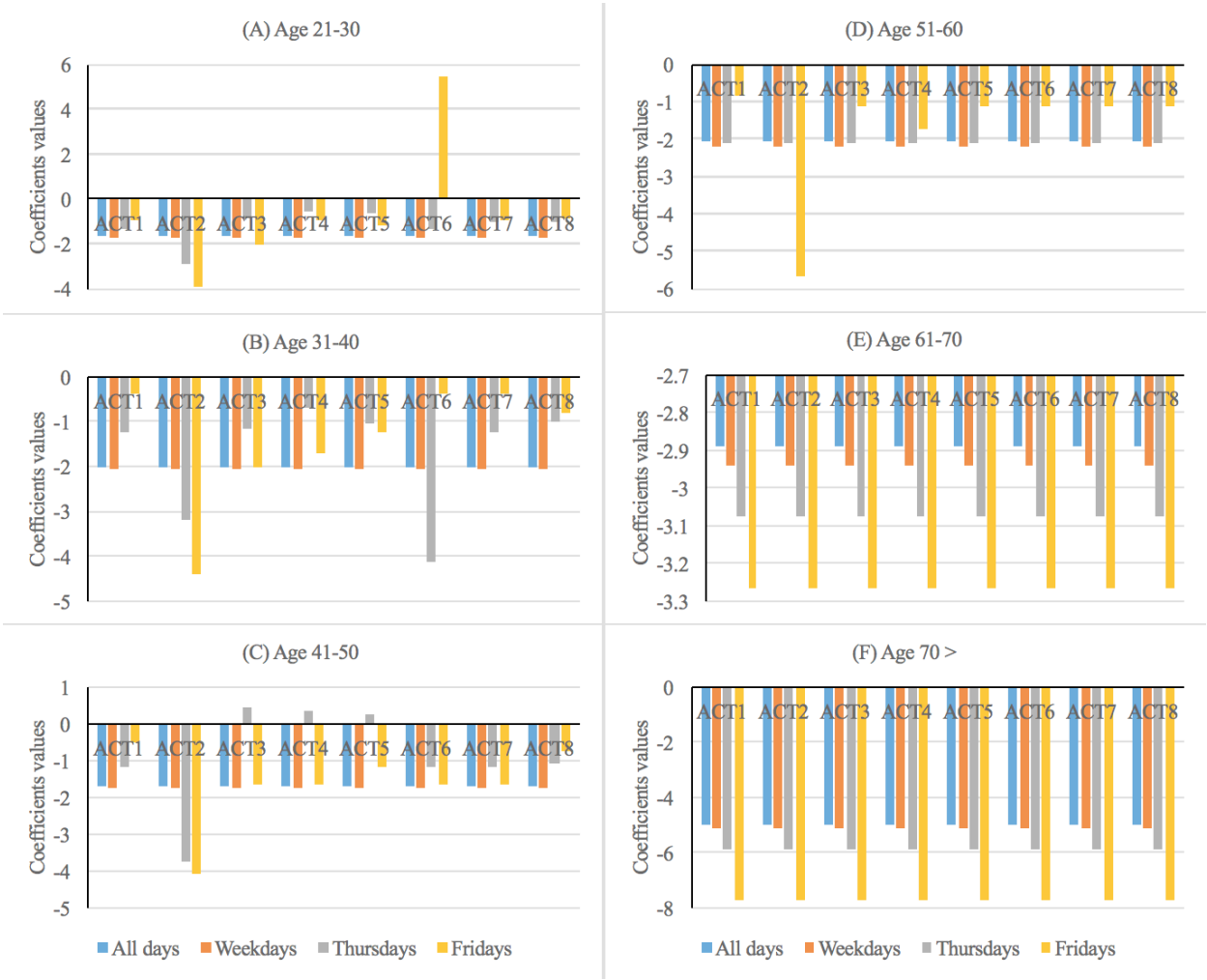


Figure 23: Age variable coefficients for different age categories of the improved four models.

For the new model only, constants are included in the pure satiation parameter, which are very similar in values for the four models and negative in sign (Table 16). Different variables were tested to be included in the pure satiation parameters, but found to be insignificant. However, the base model included the number of people per household variable in addition to the constant. Therefore, the translating satiation parameter apparently captures the satiation effect since the pure station parameter includes constants only.

The translating satiation parameter includes the constant, occupancy of the mode of transportation, and the type of mode of transportation. In the new model, based on the constant coefficient, Makkah residents spend more time on Friday at work compared to the rest of the week. In addition, when the travel group is large, residents spend more time out-of-home on Friday compared to the rest of the week. These outcomes are a result of the high travel demand on Friday due to Friday prayer; a result similar to the base model results. The same specification for the mode of transportation in the based model was tested in the new model; however, the new model resulted in different specification compared to the base model. The new model contains multiple coefficients for different out-of-home activity alternatives, whereas the base model contains one coefficient for all out-of-home activity types. The coefficient values during the weekends (i.e., Thursday and Friday) in the new model are very close in value, which shows that the residents have the same mode of transportation preferences during the weekends.

Table 16: Pure satiation and translating satiation coefficients for the four improved models.

Parameters	Parameters' components	Coefficients' structure	All-days	Weekdays	Thursdays	Fridays
Pure satiation	Constant	All activities	-2.793	-2.818	-2.744	-2.907
Translating satiation	Constant	ACT1	0.486	0.487	0.563	1.106
		ACT2	0.532	0.536	0.496	1.030
		ACT3	0.144	0.148	0.371	0.681
		ACT4	0.387	0.389	0.397	0.904
		ACT5	0.269	0.272	0.159	0.078
		ACT6	0.429	0.435	0.106	0.703
		ACT7	0.396	0.401	0.444	0.703
		ACT8	0.320	0.325	0.294	0.552
	Occupancy	All activities	0.023	0.021	0.083	0.170
	Mode of transportation	Private auto (chauffeur) for all other activities	0.000	0.000	-0.110	-0.102
		Private auto (chauffeur) for ACT3	0.000	0.000	0.000	-0.424
		Private auto (chauffeur) for ACT4	0.000	0.000	-0.191	-0.400
		Private auto (chauffeur) for ACT5	0.000	0.000	-0.151	-0.136
		Private auto (chauffeur) for ACT6	0.000	0.000	0.220	0.000
		Private auto	0.000	0.000	0.000	0.119
		Taxi (owned by company) for all other activities	0.000	0.000	-0.086	0.000
		Taxi (owned by company) for ACT2	0.000	0.000	-0.149	0.000
		Taxi (owned by company) for ACT5	0.000	0.000	-0.038	0.000
		Private school bus	0.000	0.000	0.000	-0.460

5.2.2 Improved MDCEV Model Performance Using German Data Set

In this thesis work, the improved MDCEV model is used to estimate one of Habib et al.'s (2007) estimated models which is Week 1 model (Six models are estimated for each week and

the seventh model is estimated for all weeks combined), using the same set of MobiDrive data. The Saudi Arabia and Germany data sets were used to test the applicability of the new model to any data set.

Habib et al. (2007) estimated activity duration for the residents of Karlsruhe and Halle in 1999 using the base MDCEV model. The goal of their research as mentioned in the literature review was to study the weekly rhythm of nested activity-travel decisions using six weeks' worth of travel diaries. They used the MDCEV model proposed by Bhat (2005) to estimate the models. The same model structure used by Habib et al. is estimated using the base MDCEV model structure and the improved MDCEV model structure in this study for comparison reasons. Tpower 14, which is used in the Makkah study, is used in this study to estimate the improved MDCEV model for comparison reasons.

The comparison showed that the new model performs better than the base model by 21.2 % based on the goodness of fit, where the base model goodness of fit is 0.2476 and the new model goodness of fit is 0.3. Habib et al.'s study kept multiple insignificant parameters in their model and argued that these parameters provide an insightful result. They also argued that if the number of observations increased these parameters might turn out to be significant. Habib and co-authors normalized the scale parameter to 1, as it gives the highest number of statistically significant parameters. The scale parameter of value 1 is used in the base and the new model. 90% confidence interval is used to test the significance of the parameters in both the base and the improved model. The new model resulted in 26 insignificant parameters and the base model results in 29 insignificant parameters. It is important to first set a hypothesis for the rational outcomes based in the model variables to comment on the rationality of the model results (see

Table 17 and Chapter 4). In other words, how do these variables explain human behaviour? Ultimately, the major findings of the base and the new model are explained later.

Table 17: Hypothesis of how the variables affect out-of-home activity time.

Variables	Rational explanation
Number of vehicles per household	The need for more vehicles reflects the need to engage in more out-of-home activities, which reflects the increase in time spent out-of-home
Residence location	The CBD is usually well developed, contains a variety of activities close to each other, and a high accessibility area, which results in a high demand on the activities in this area and consequently more time spent on activities in the CBD. Consequently, more time is needed to conduct the activities by visitors to the CBD due to the high demand. In addition, the residents of the CBD try to avoid congestion areas by spending less time in the out-of-home activities in the CBD area.
Work status	Employed people need to spend time at work. In addition, having a salary reflects the need to spend some of it on activities such as shopping and recreation, which reflects the need to spend more time in out-of-home activities.
City size	Bigger cities are expected to have more out-of-home activities and consequently increased out-of-home activity duration.

First, since the German city of Karlsruhe is the second largest city ($173.5km^2$) in Baden-Württemberg state, it is expected that the city produces longer durations of activities outside of the home compared to the city of Halle ($135km^2$). In the base model, it was found that in some activities the residents of Halle spent more time compared to the city of Karlsruhe, such that the sign and coefficients values vary between the weeks. In addition, it was difficult to get significant coefficients for work and school activities (Habib et al., 2007). In the new model, Karlsruhe residents spend more time in all out-of-home activities compared to the Halle residents with statistically significant coefficients for all activities.

Next, for the gender variable in the base model, most of the coefficients are statistically insignificant and males spend more time than females in most out-of-home activities. In the new model, males again spend more time in all out-of-home activities except for "other" activities, and all of the coefficients are significant. In the base model's week 3 model, males spend more time than females in work and school activities, which match the results in the new model. In the base model for shopping activity, females spend more time than males in both the week 5 and pooled models. In contrast, females spend less time in shopping activities than males based on the results from the week 3 and 4 models, and these results comply with the new model's results.

In the base model, in week 3 and the pooled model, males spend more time than females in service activities, which is the same result as in the new model. In both the new model and the base model in the week 3, 5, and pooled models, males spend more time than females in social activities with a statistically significant coefficient. In both the new model and the base model in the week 1, 2, and the pooled data models, females spend more time in "other" activities compared to males with a statistical significant parameter. Based on the gender variable results presented above, the signs of the coefficients from all of the significant coefficients in the base model match those in the new model. That indicates that the new model improves the accuracy of the base model in terms of the significance of the coefficient value and the sign of the coefficient.

In both the base and the new models, having more than one vehicle increases the time spent out-of-home. In the base model, the parents variable is insignificant, though it is significant in the new model, which indicates that parents spend more time in-home compared to out-of-home time. In both the base and the new models, for the living in the central business district (CBD) variable it was found that the coefficients are statistically significant and the residents

living away from the CBD spend more time out-of-home compared to the residents who live in the CBD. In the base model, the full time employment variable coefficients for the basic need, work, school, shopping, and social activity categories are statistically significant and positive. However, for the other activities there is considerable variation in coefficients' sign among the weeks or the models. In the new model, the employment variable coefficients are positive and statistically significant, indicating employed people are willing to spend more time out-of-home compared to the unemployed.

The number of children per household negatively affects the time expenditure on most out-of-home activities in both the base and the new models. However, it positively affects work and school activities with an insignificant coefficient, except in week 4 that shows a negative effect with a statistically significant coefficient. The new model contains a statistical significant and negative coefficient for work and school activities; in other words, an increase in the number of children prevents an adult from engaging in work and school activities. In both the base and the new models, the increase in age reduces the residents time to spend out-of-home which is comply with the results from Makkah studies in this dissertation (see Table 18 for results details). It can also be noted that in the translating satiation effect all of the coefficients' signs are the same in both the base and the new models. In contrast, the new model results in less significant parameters compared to the base model. The variables in the pure satiation parameters are significant. Based on the previous discussion about the results, the Tpower adds more significant parameters to the baseline utility and the pure satiation parameters compared to the base model, but less significant parameters in the translating satiation effect. It appears that the Tpower value helps to better explain the translating satiation parameter with fewer variables.

However, the three parameters' coefficients interact with each other and result in a final individual choice.

Table 18: Base and improved MDCEV model results for cities of Karlsruhe and Halle, Germany.

Baseline utility components	Activity type	Base model		New model	
		Parameter	t-value	Parameter	t-value
CONSTANT: GENERIC	All activities	-14.400	-17.280	-59.825	-88.619
City of Karshruhe	Basic need	0.750	3.380	0.9699	3.899
	Work / School	-0.190	-0.890	0.5715	2.673
	Drop off / Pick Up	-0.180	-0.840	0.4972	2.367
	Shopping	0.090	0.390	0.5795	2.711
	Services / Private business	-0.170	-0.770	0.5172	2.186
	Recreation	0.510	2.230	1.1253	4.924
	Social	0.770	3.350	1.1388	4.502
	Others	0.290	0.510	0.9826	1.565
Gender (male)	Basic need	0.410	1.740	0.8818	3.113
	Work / School	0.000	-0.020	0.7312	3.445
	Drop off / Pick Up	0.020	0.080	0.6871	3.021
	Shopping	-0.030	-0.120	0.6009	2.719
	Services / Private business	0.230	1.020	0.7417	3.343
	Recreation	-0.230	-0.90	0.4825	1.822
	Social	0.260	1.080	0.8495	3.157
	Others	-1.140	-2.00	-0.3031	-0.486
Household income	Basic need	0.070	1.340	-0.100	-1.736
	Work / School	-0.040	-0.630	-0.115	-2.333
	Drop off / Pick Up	-0.30	-5.090	-0.294	-6.039
	Shopping	-0.030	-0.630	-0.217	-4.107
	Services / Private business	0.000	-0.040	-0.152	-3.067
	Recreation	-0.180	-3.090	-0.216	-4.435
	Social	-0.080	-1.610	-0.169	-3.41
	Others	-0.920	-7.320	-0.843	-7.042
Driving license	All activities	0.350	1.020	0.113	0.547
Household vehicle (more than or equal to 1)	All activities	0.510	1.720	2.295	11.239
Employed household member	All activities	0.600	2.130	1.279	7.300
Parent in home	All activities	0.140	0.450	1.118	6.014
LN(Distance of Bus Stop From Home)	All activities	-0.030	-0.360	-0.038	-0.727
Home location (in CBD)	All activities	-1.480	-3.980	-2.881	-13.573
Student	All activities	-0.790	-2.260	-1.260	-5.599
Full time employee	Basic need	0.330	1.180	1.367	4.195
	Work / School	2.040	7.650	2.488	9.791
	Drop off / Pick Up	-0.440	-1.660	0.627	2.374
	Shopping	0.020	0.090	1.193	4.154
	Services / Private business	-0.410	-1.540	0.892	3.239
	Recreation	-0.110	-0.380	1.131	3.666
	Social	0.030	0.110	1.190	3.771
	Others	-1.490	-1.000	-0.375	-0.282

Baseline utility components	Activity type	Base model		New model	
		Parameter	t-value	Parameter	t-value
Number of children per household (More than or Equal to 1)	Basic need	-0.900	-1.810	-1.920	-4.151
	Work / School	0.320	0.630	-0.934	-2.463
	Drop off / Pick Up	-1.740	-3.470	-2.429	-6.585
	Shopping	-1.880	-3.540	-2.280	-5.409
	Services / Private business	-2.550	-5.350	-2.897	-8.200
	Recreation	-0.980	-2.080	-1.840	-4.933
	Social	-1.010	-2.190	-1.978	-5.201
	Others	-1.850	-2.010	-2.653	-2.938
Age	21-30	-1.990	-3.450	-3.111	-8.850
	31-40	-2.150	-3.460	-3.993	-10.343
	41-45	-2.910	-4.440	-5.320	-12.671
	46-50	-1.510	-2.410	-1.700	-4.252
	51-55	-3.220	-4.470	-7.810	-17.629
	56-60	-3.280	-4.760	-7.941	-18.157
	61-65	-3.210	-4.330	-9.318	-20.185
	65+	-3.690	-5.020	-9.017	-19.039
Translating satiation parameter of specific activities					
Constant	Basic need	2.570	12.980	0.496	1.459
	Work / School	2.860	5.920	0.548	0.780
	Drop off / Pick Up	3.90	15.580	0.613	2.181
	Shopping	2.130	9.170	0.469	2.792
	Services / Private business	1.270	5.640	0.363	1.808
	Recreation	3.680	3.650	0.642	0.455
	Social	3.190	6.00	0.621	0.949
	Others	4.840	1.280	0.754	0.308
Travel ratio	Basic need	-0.440	-4.360	-0.030	-0.153
	Work / School	-0.710	-2.930	-0.101	-0.403
	Drop off / Pick Up	-0.760	-14.890	-0.057	-1.450
	Shopping	-1.010	-7.990	-0.069	-1.222
	Services / Private business	-0.560	-7.280	-0.042	-0.990
	Recreation	-1.340	-1.660	-0.114	-0.098
	Social	-1.300	-3.370	-0.169	-0.333
	Others	-1.430	-0.420	-0.144	-0.157
Number of people accompanied in the trip	Basic need	0.140	7.180	0.025	0.311
	Work / School	0.460	6.480	0.058	0.262
	Drop off / Pick Up	0.110	1.770	0.017	0.172
	Shopping	0.200	7.110	0.022	0.555
	Services / Private business	0.390	8.310	0.053	0.812
	Recreation	0.190	1.970	0.014	0.081
	Social	0.200	4.880	0.029	0.318
	Others	-0.050	-0.080	-0.002	-0.002
Pure satiation parameter					
Constant	All activities	-1.320	-28.380	-2.643	-248.69
Married	All activities	0.000	-0.100	-0.007	-2.802
LN(Number of people per household)	All activities	0.030	1.890	0.057	20.113
Number of vehicle per household (>=2)	All activities	0.040	1.630	0.044	10.125

5.2.3 Remarks on the Studies

Using the municipality of Makkah's 2010 survey data about its residents' activities, four models were estimated using MDCEV proposed by Bhat (2005). These models formed the base case to be compared with the results from the improved model structure. The study investigated the differences in time expenditure among the residents of Makkah during weekdays, weekends, and Friday, which is a day of Friday prayer. It was found that Friday should be modeled differently from the weekends since the travel behaviour is different. Based on the base and the improved models' results it was found that there are some similarities and differences in the estimated models in terms of the sign, value, and statistical significance of the parameter coefficients. In general, the improved model provides more intuitive results compared to the base model.

This improved model structure was then applied to Habib et al.'s (2007) study that used the MDCEV model for travel activity in two German cities. This step was to ensure that the improved model would work with any data set, not only Makkah, even if the time budget is long, such as a week (i.e., high variability in travel plan). It was found that the improved modelling structure improved the accuracy of the base MDCEV model. As activity duration model is a part of activity-based travel demand model, the improvement in the estimation accuracy of the activity duration model (i.e. use the improved MDCEV model) will lead to improvement in the predictability in travel demand. In other words, the more accurate activity duration model will aid in better travel planning for cities and in testing policies which result in better transportation planning and more sustainable cities and infrastructures.

5.3 Generalization of the MDCEV Model

The data collection and planning process for any city's transportation needs is important such that it reduces traffic and individual delays; however, it is also a costly undertaking. There is therefore a need for transportation planning models that could be used anywhere around the world, particularly where limited transportation planning models and resources are available for a specific city. These models would save transportation planning agencies the time and cost involved in collecting the required data and estimating the models to plan for a transportation system for a specific city.

Every year, the city of Makkah attracts millions of people from around the globe; moreover, there is substantial diversity in the city's residents' travel patterns and behaviours (as previously discussed in the Introduction). The diversity of travel patterns in Makkah and the high number of data points makes Makkah an ideal city for the generalization of the activity duration model (i.e., the improved MDCEV model). In particular, the diversity of the population helps in capturing heterogeneity in time expenditure behaviour, while the high number of data points helps in providing a representative sample and thus estimating a reliable model (as discussed in the Literature Review). This research thus claims that generalization of the MDCEV model can be achieved through capturing the travel behaviours accounted for in the Makkah dataset.

Using all of the data points obtained from the city of Makkah's dataset, this section generalizes the improved MDCEV model to be used anywhere around the world. The model needs to be validated against other datasets: to accomplish this, the improved MDCEV model coefficients estimated for the city of Makkah using all data points (18,434 data point) are used to predict activity duration in the dataset for the German cities of Karlsruhe and Halle (henceforth

referred to as the German cities) which contain 333 data points (see Pinjari and Bhat, 2011 for prediction procedure details).

As noted in previous sections, there are similarities in the activity types and variables in both the Makkah and German cities datasets, though they differ in terms of time budget. More specifically, the variables gender, age, and acquisition of driver licences in the Makkah dataset are similar in the German cities dataset. Activity duration types in common include work, school, shopping, and recreational or social activities. The definition of social activity in the German dataset is the same as recreational activities in the Makkah dataset, both including meeting with friends in coffee shops or restaurants. Recreational activities in the German dataset include walks in nature and going to the movie theatre, however, activities not defined in the Makkah dataset. For illustration purposes in this section, the recreational activities in Makkah are considered the same as the social activities in the German dataset and both named social activities.

In terms of activity duration, the time budget is different in the dataset for Makkah than in the dataset for the German cities. In the Makkah dataset, activity duration has a maximum time budget of 24 hours, whereas the German dataset accounts for a time budget of one week, such that the total time budget is 24 hours * 7 days * 60 minutes = 10,080 minutes per week. To be able to predict the activity duration for German cities, the time budget for German cities is modified to be 24 hours—work and school activities are divided by six days (i.e., six business days per week), and the remaining shopping and social activities are divided by two days of the week. The result is a daily activity duration which is compatible with the prediction model outcome.

To evaluate the accuracy of the prediction model, the average activity duration from the predicted model is compared to the average activity duration from the observed dataset. The results are shown in Table 19 and summarized in the following paragraphs.

Table 19: Average activity duration for the residents of Makkah and of the German cities.

Activity types	Work	School	Shopping	Social
Frequency of the activity per week	6	6	2	2
Average predicted activity duration in hours per day	2.923	3.206	2.059	2.217
Average observed activity duration in hours per day	3.368	3.368	1.573	2.188
Prediction accuracy %	86.77	95.19	76.39	98.71

Work and school activities in the German dataset are considered one category that includes full- and part-time work and study, and non-worker/non-student. Work and school activities in the Makkah dataset are under separate categories, but also account for full- and part-time segments and non-worker nor student. The results show that the prediction model accurately predicts activity duration for work, school, shopping, and social activities. The average observed work and school activity duration is around 3.4 hours per day, while the average predicted work and school activity is 2.9 hours and 3.2 hours per day respectively. The prediction model underestimate work and school duration for German cities by 0.5 hour and 0.2 hour respectively. To justify, the prayer time policy in Makkah results in reducing work and school hours which results in underestimating work and school duration. However, the prediction accuracy of 86.77% for work duration is considered good.

Looking at other activities, the average activity duration in the prediction model (Makkah model) for shopping activity is 2 hours and 2.2 hours for social activities, while the average activity duration in the observed German dataset for shopping and social activities is around 1.6 hour and 2.2 hour, respectively. The predicted model therefore slightly overestimates the

observed shopping activity duration. This is logical results. Relative to the German cities, there are not as many types of recreational activity in Makkah (e.g., going to the movies or walking in nature). The result is that the residents of Makkah spend more time shopping compared to the residents of the German cities. Given the accuracy of the prediction model for shopping as 76.39%, the prediction model considered good.

To sum up, from the previous section, the size of dataset in German cities is small, which limits the ability to predict an accurate activity duration model. Therefore, this part of the research use the improved MDCEV model, which is based on the larger dataset to predict the activity duration for German cities and to generalize the model to be used in other cities. The performance of MDCEV model in predicting activity duration has some shortcomings. Based on Sikder and Pinjari (2014), this prediction over/under estimates activity duration compared to the average activity duration observed from the data set. Based on that, Sikder and Pinjari (2013) suggested more investigation on this issue. Therefore, the prediction of activity duration is not carried over the entire research. Despite this issue and the variables being limited in the model used to predict the German activity duration, the model accurately predicts work, school, shopping, and social activities. Therefore, a transportation planner from anywhere around the globe could use this model to predict activity duration for a city.

5.4 Application of the Improved MDCEV Model

This section presents several applications of the improved MDCEV model. First, two different statistical software and their optimization algorithms for MDCEV modelling, GAUSS and R, are presented in an attempt to better understand how the algorithms work. Comparing both software using the Makkah and German datasets sheds light on estimation differences between the two and results in GAUSS being recommended for future MDCEV modelling for

mega-events since GAUSS is more efficient in estimation cost compared to R. However, the study set criteria helps selecting the appropriate software to be used depending on the project needs. This section also sets guidelines for a major policy in Makkah –i.e., the store closing policy during the five prayer times including Friday prayer– based on the improved model structure and using GAUSS software. The section closes with a conceptual travel demand framework for the city of Makkah during Umrah season. Using GAUSS alongside the physical and cultural considerations of Makkah forms the benchmark for future development of the travel demand framework for the city.

5.4.1 Statistical Software Comparison

May and Easa in 1981, compared six traffic assignment software. They took into consideration the model input requirements, the approach used to represent driver behaviour, the software operation interactions, the software application history, and the possibility for improving the software. This research is similar to May and Easa framework in that it compares two software in terms of the model estimation methodology, the software framework interaction, application history of the software, and recommend improvements to both software.

There is a long list of statistical software and related packages that are used for maximum likelihood problems, including GAUSS, R, SAS, SPSS, and Stata. Each one of these has different underlying theoretical assumptions and estimation techniques and may result in different outcomes. GAUSS has been the choice software to estimate the MDCEV model such as in Habib, 2007; Bhat, 2008; and Eluru et al., 2010. GAUSS is a programming software that essentially manipulates matrices. It has tools to transform and manipulate matrices and programs to solve numerical optimization (Aptech Systems, 2002). R software, though used less frequently

than GAUSS for MDCEV modelling, was chosen for comparison with GAUSS in this research because it possesses a strong collection of advanced statistical optimization algorithms. R is a combination of software that facilitates data calculation. It is effective with data handling and storage, suitable for an array of calculations, and uses a simple programming language. The software includes approximately 25 standards and recommended packages (R manual, 2015). Furthermore, there are no costs associated with its download and use, making it a convenient and accessible choice.

Both GAUSS and R allow for estimating the MDCEV model utilizing a maximum likelihood approach, which is a unified approach. The maximum likelihood approach aims to maximize the likelihood function, which contains a set of observed variables and their parameters. The functional form of the likelihood is a joint distribution of a random variable that is the multiplications of the probability distribution of random variables. To illustrate, the maximum likelihood function estimate is the maximum value of the joint probability achieved through the estimated parameters. Moreover, solving for a higher number of parameters in a likelihood function requires matrix manipulation techniques and numerical methods. Likelihood functions can be represented as a matrix. The gradient vector of the likelihood (i.e., first derivative of the likelihood function) for any parameter value indicates the location of the solution. The second derivative of the likelihood function is the Hessian matrix, which indicates which way to move and the distance of the step to increase the likelihood function (Greene, 2002; and GAUSS, 2012).

As mentioned earlier, statistical software applications use different optimization algorithms to estimate likelihood parameters. These algorithms include: Newton-Ralphson, BHHH, DFP, and BFGS (Table 20). Of these, the Newton-Ralphson is considered one of the

most famous methods for solving a root-finding problem because of its simplicity and power. One of the limitations of this method is that the model does not guarantee that the likelihood function will increase in every iteration step (Miranda and Fackler, 2002). The Brendt-Hall-Hall-Hausman (BHHH) optimization algorithm is a scoring method that uses the cross-product of the likelihood function first derivative matrix to estimate the Hessian matrix. It estimates the gradient of the likelihood function using Jacobian and the summation of the cross-product of this Jacobian (Aptech Systems, 2004). BHHH ensures the increase in the likelihood function after each iteration [Berndt (1974), and Hurn (2009)]. However, it is noted in R that the simulation search algorithms (SANN, CG, and NM) which does not depend on the gradient and Hessian to estimate the MDCEV model parameters are very efficient when compared to BHHH.

Both the Davidon-Fletcher-Powell (DFP) and Broyden-Fletcher-Goldfarb-Shanno (BFGS) use an approximation of a Hessian matrix instead of an actual Hessian matrix during the first steps to reach the maximum likelihood more rapidly. When they are close to the maximum likelihood, an actual Hessian matrix is used. The BFGS is better than the DFP in terms of calculating speed and thus provides more efficient algorithms [Andrew (2008), and Ding et al. (2010)]. Better estimation of the parameters will always lead to less error and a better model, but based on the above discussion some optimization algorithm (e.g., NEWTON) might not work for the complex MDCEV optimization problem. Only the algorithms in both GAUSS and R software will be tested and compared in order to find the algorithm best suited for the MDCEV model.

Table 20: Optimization algorithms in the R and GAUSS statistical software (R manual 2015 and GAUSS manual 2012).

Optimization Algorithms	R	GAUSS
Broyden, Fletcher, Goldfarb, Shanno method (BFGS)	Yes	Yes
Davidon, Fletcher, Powell method (DFP)	No	Yes
Newton-Raphson method (NEWTON)	Yes	Yes
Berndt, Hall, Hall, Hausman method (BHHH)	Yes	Yes
Simulated ANNealing (SANN)	Yes	No
Conjugate Gradients (CG)	Yes	No
Nelder-Mead (NM)	Yes	No

Base on Theussl and Borchers (2016), the R Optimization Infrastructure (ROI) package framework solves several optimization problems, including linear, quadratic, and non-linear problems. Task View on Optimization (TVoO) and Mathematical Programming (Theussl and Borchers, 2016) lists the optimization function and its related package that aim to help in evaluating options (Nash, 2014). There are several optimization methods implemented in R statistical software such as optim, nlm, nlminb, and maxlik. For this research, the maxlik package was selected.

Maxlik has a single and unified interface for various optimization searching methods such as BHHH, and BFGS that is suitable for maximum likelihood (ML) optimization problems. In R, maxlik includes several sub-functions to get the final results. These functions are used to evaluate the gradient, fix some parameters values, extract the Hessian, return the log likelihood value, and offers other roles that serve to solve the optimization problem and present relevant results. The package contains an internal function (maxNRCompute function) based on the Newton-Raphson type optimization method and is applicable to multipurpose optimization methods. The Newton-Raphson method is used to approximate the function at a given location in a multidimensional parabola environment, and then use the estimated maximum to set the initial

values that will be used in the following iteration. This requires the information about the gradients and the Hessian. The Hessian is computationally intensive and hence there are several methods that aim to approximate it, such as BFGS and BHHH. As the Hessian is approximated based on individual information, it is calculated by taking the negative of the sum of the products of the gradients of individual observations.

The maxlik package contains two layers. The first layer deals with the maximization layer, which allows the users to switch between optimization routines or searching methods. The function is designed for various optimization problems related and unrelated to likelihood problems. The function provides information about the solution, such as the estimated coefficients, log likelihood value, and statistical significance. The second layer is the likelihood maximization layer, where the most important interior function is the maxlik. It deals with linear equality and inequality constraints. This function uses the input information in computing the variance-covariance matrix, which is the inverse of the negative Hessian. On the other hand, it does not correct the variance-covariance matrix and only a warring is printed [Henningsen and Toomet (2011), and Toomet et al. (2015)].

The maxlik package in GAUSS software is a set of procedures that leads to an estimation of the models parameters when the log-likelihood function is maximized using maximum likelihood method, and considering parameters constraints. The user provides the log-likelihood function and its parameters for an individual observation or set of observations. Based on that, the package returns the individual log-likelihood for each individual observation or, in case of a matrix of observations; it returns a vector of log-likelihoods. Searching for the parameter values is an iterative method that starts with the initial parameters values provided by the user and these values are updated every iteration. The parameter update procedure is based on direction vector

and step length. The step length is calculated by minimizing the likelihood function considering the parameter of a specific step. Consequently, it is required to calculate the Hessian and the gradient, which could be calculated numerically using the package if they are not provided by the user. In some cases, the Hessian is computationally intensive; poor starting values could cause convergence problems. For this reason, several searching methods that are computationally efficient such as BFGS, DFG and BHHH could be used to estimate the Hessian, instead of computing it directly at each iteration. The method reaches convergence when the relative gradient is less than the specified tolerance (i.e., the default is 1e-5) (GAUSS, 2012).

5.4.1.1 Results of GAUSS and R Software

Makkah, Saudi Arabia:

The improved multiple discreet-continuous extreme value model with a Tpower value of 14 is coded in GAUSS and R exactly the same way for the Friday model for the city of Makkah. In the Friday model the mode of transportation is excluded from the model framework, as it is assumed that activity duration is an input to the mode of transportation model (as discussed in the Literature Review). However, some commands in the software are different, but do the same job. It is noticed that some commands in GAUSS that take one line is transformed into R in two lines. In addition, GAUSS uses pure compiler which is much faster than R which uses both a compiler and interpreter when executing the code [GNU (1991), R manual (2015), and GAUSS customer support (2017)]. The mean log-likelihood of the improved models in GAUSS and R are -0.951 and -1.025, respectively, which are very close to each other and implies that the maxlik package in both software are the same. The small difference is due to the digits each software count on during calculations.

The coefficients were estimated in R and GAUSS using the BHHH searching algorithm since it is one of the most efficient algorithms. The BFGS method is the only one that converges when the starting values are set to ones; however, when the starting values are set close to the optimum values, BHHH performs the best among all other algorithms. The tolerance value used is 1e-5. In comparing the improved MDCEV coefficients outcomes of GAUSS and R, there are minor differences in coefficient values and in the standard error values. However, the signs are the same. All of the parameters are statistically significant based on the 90% confidence significance. Notably, the time it takes GAUSS to estimate the parameters is much lower than R considering the same searching algorithms, starting values, and the tolerance value. GAUSS estimated the model in 305.3 minutes, whereas R estimated the model in 11 days, 3 hours. The coefficient values were discussed in the previous sections (section 5.2.1). This section focuses on the differences in the estimation results, given by percentages [Difference % = coefficient result from (R - GAUSS) / GAUSS *100] (see Table 21, 22, and 23).

For the pure satiation parameter, GAUSS coefficients estimated value is more than the R estimate by 0.2%, such that, activity duration spent out-of-home estimated by GAUSS is more than the one estimated by R. For the translating satiation parameter, R coefficients estimated value are greater than GAUSS for constant and occupancy coefficients except for ACT 6 and 7 constant coefficients where GAUSS estimates are greater than R by 5.6%. For the baseline utility parameter, the difference in estimation results in travel time variable is the highest (3.4%) and the lowest is 0.007% for driver's license.

The highest differences are in estimating the constant of the translating satiation parameter with roughly 5% difference (Table 23). Based on the pervious discussion, it is not clear if R overestimates the coefficients values or GAUSS does. However, there is not much

difference in the estimated coefficients values, which is a reasonable output since both software use the unified maximum likelihood approach (see Figure 24). The significant difference between the two soft wares is in the estimation speed, which is explained by the type of compiler each software use.

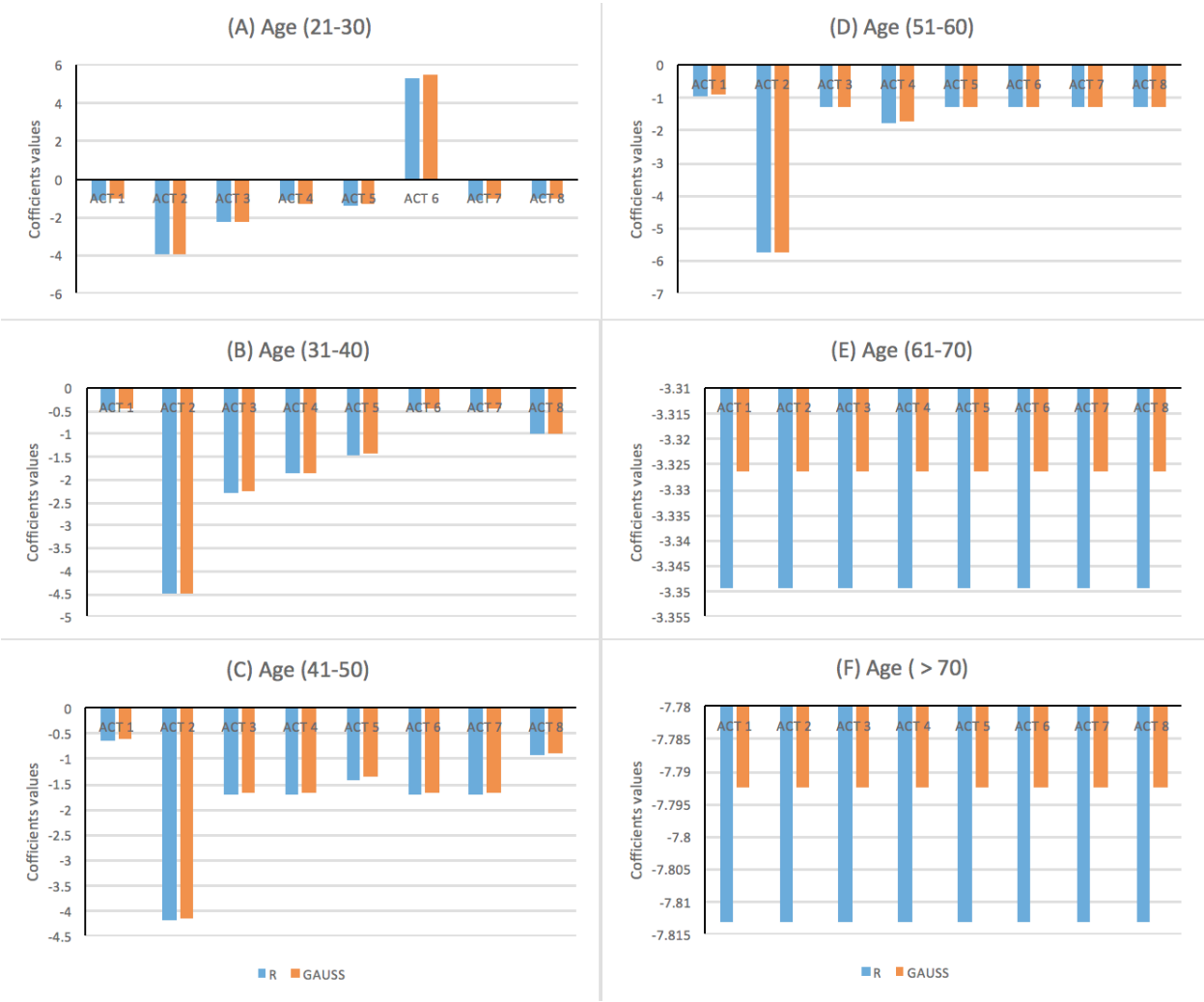


Figure 24: R and GAUSS results for age variable of the improved Friday model.

Table 21: Constant value in the baseline utility parameter for Friday model of the improved model structure.

Baseline utility components	R	GAUSS	Difference %
Constant ACT1	-61.792	-61.429	0.591
Constant ACT2	-55.216	-54.854	0.660
Constant ACT3	-59.769	-59.410	0.603
Constant ACT4	-60.320	-59.905	0.693
Constant ACT5	-58.394	-58.042	0.605
Constant ACT6	-67.120	-67.055	0.097
Constant ACT7	-61.792	-61.429	0.591
Constant ACT8	-59.854	-59.500	0.594

Table 22: Driver's license, number of autos per household, gender, and travel time coefficients for the improved model using GAUSS and R software for Friday model.

Four models	Activity type	Driver's license	# autos per household	Gender	Travel time
R	ACT1	2.646	0.857	0.950	0.164
	ACT2	0.187	-0.385	0.950	0.137
	ACT3	2.208	-1.095	0.950	0.120
	ACT4	3.326	-0.761	0.950	0.175
	ACT5	-1.852	-0.761	0.950	0.336
	ACT6	0.187	-0.761	0.950	0.492
	ACT7	0.187	-0.761	0.950	0.492
	ACT8	0.187	-0.199	0.950	0.115
GAUSS	ACT1	2.633	0.852	0.946	0.164
	ACT2	0.185	-0.385	0.946	0.136
	ACT3	2.208	-1.099	0.946	0.120
	ACT4	3.295	-0.773	0.946	0.174
	ACT5	-1.839	-0.773	0.946	0.335
	ACT6	0.185	-0.773	0.946	0.510
	ACT7	0.185	-0.773	0.946	0.510
	ACT8	0.185	-0.202	0.946	0.115
Difference %	ACT1	0.497	0.566	0.449	0.375
	ACT2	0.953	-0.030	0.449	0.495
	ACT3	0.007	-0.382	0.449	-0.123
	ACT4	0.954	-1.509	0.449	0.543
	ACT5	0.707	-1.509	0.449	0.483
	ACT6	0.953	-1.509	0.449	-3.425
	ACT7	0.953	-1.509	0.449	-3.425
	ACT8	0.953	-1.570	0.449	0.059

Table 23: Pure satiation and translating satiation coefficients for the improved model using GAUSS and R software for Friday model.

Parameters	Parameters' components	Coefficients' structure	R	GAUSS	Difference %
Pure satiation	Constant	All activities	-2.944	-2.938	0.217
Translating satiation	Constant	ACT1	0.507	0.508	-0.035
		ACT2	0.463	0.463	-0.002
		ACT3	0.311	0.311	0.214
		ACT4	0.387	0.385	0.376
		ACT5	0.093	0.093	-0.358
		ACT6	0.359	0.380	-5.681
		ACT7	0.359	0.380	-5.681
		ACT8	0.292	0.292	0.150
	Occupancy	All activities	0.095	0.094	0.538

Karlsruhe and Halle, Germany:

The improved multiple discreet-continuous extreme value model is estimated for the German cities as in section 5.2.2, but this time using both GAUSS and R statistical software to compare the performance of these software. The same code and the Tpower value used for the Makkah model in the above section are used for German model. This section focuses on the results of this comparison, specifically on the mean log-likelihood, the coefficients' value, t-value, and the time of the model estimation.

The results of the mean log-likelihood in both soft wares are the same, 8.0576. The coefficient sign results in both soft wares are the same, but there are small differences in the estimated coefficients' values. Further interpretation of the coefficients' values is provided in section 5.2.2. The translating satiation parameter variable result shows the highest difference in the estimated coefficients' values: the difference in the coefficients' values for the number of people sharing the mode of transportation in a trip variable is 6.7% (see Table 24). The difference in all variables' coefficient values for baseline utility parameters between the two soft

wares is less than 0.17%. The maximum difference between the variables' coefficients for the pure satiation parameter is 3.1%. This shows that both soft wares produce the same variables' coefficients values and both are reliable software to estimate the improved MDCEV model (see Table 24).

The major difference between the two soft wares is in the estimation speed. R estimated the full improved MDCEV model in 47 minutes, while GAUSS estimated the same model in only 0.085 minutes. It is notable that the estimation speed for German cities is much faster than the estimation speed for Makkah. This is a logical output due to the difference in the number of data points used in the estimation process—the German dataset uses 333 data points, whereas the Makkah dataset for Friday uses 928 data points. Based on that, R is considered acceptable tool to estimate the improved MDCEV model, but ideally for a small dataset. It is noted that the simulation searching algorithms in R are more efficient than BHHH. Based on that, it is recommended to support GAUSS with these algorithms to further improve the estimation speed.

Table 24: The estimated coefficients values of the MDECV model using GAUSS and R software for German data set.

Baseline utility component	Activity type	New model using GAUSS		New model using R		Parameter difference %
		Parameter	t-value	Parameter	t-value	
Constant	All Specific Activity	-59.825	-88.619	-59.8248	-88.619	-0.000
City of Karshruhe	Basic Need	0.970	3.899	0.970	3.898	0.002
	Work / School	0.572	2.673	0.572	2.673	0.000
	Drop off Pick Up	0.497	2.367	0.497	2.367	-0.001
	Shopping	0.580	2.711	0.580	2.712	-0.005
	Services / Private Business	0.517	2.186	0.517	2.186	0.016
	Recreation	1.125	4.924	1.125	4.924	0.001
	Social	1.139	4.502	1.139	4.502	-0.006
	Others	0.983	1.565	0.983	1.565	0.008

Baseline utility component	Activity type	New model using GAUSS		New model using R		Parameter difference %
		Parameter	t-value	Parameter	t-value	
		Parameter	t-value	Parameter	t-value	
Gender (male)	Work / School	0.731	3.445	0.731	3.445	0.003
	Drop off Pick Up	0.687	3.021	0.687	3.021	-0.002
	Shopping	0.601	2.719	0.601	2.719	-0.004
	Services / Private Business	0.742	3.343	0.742	3.343	-0.001
	Recreation	0.483	1.822	0.483	1.822	-0.003
	Social	0.850	3.157	0.850	3.156	0.000
	Others	-0.303	-0.486	-0.303	-0.486	0.000
Household income	Basic Need	-0.100	-1.736	-0.100	-1.736	0.034
	Work / School	-0.115	-2.333	-0.115	-2.333	0.150
	Drop off Pick Up	-0.294	-6.039	-0.294	-6.039	-0.057
	Shopping	-0.217	-4.107	-0.217	-4.109	0.174
	Services / Private Business	-0.152	-3.067	-0.152	-3.068	-0.260
	Recreation	-0.216	-4.435	-0.216	-4.435	-0.170
	Social	-0.169	-3.410	-0.169	-3.409	0.160
	Others	-0.843	-7.042	-0.846	-7.043	0.055
Driving license	All Specific Activity	0.113	0.547	0.112	0.547	-0.489
Household vehicle (more than or equal to 1)	All Specific Activity	2.295	11.239	2.295	11.239	0.006
Employed household member	All Specific Activity	1.279	7.300	1.279	7.300	-0.023
Parent in home	All Specific Activity	1.118	6.014	1.118	6.014	0.040
LN(Distance of Bus Stop From Home)	All Specific Activity	-0.038	-0.727	-0.038	-0.727	0.379
Home location (in CBD)	All Specific Activity	-2.881	-13.573	-2.881	-13.573	-0.015
Student	All Specific Activity	-1.26	-5.599	-1.260	-5.598	-0.023
Full time employee	Basic Need	1.367	4.195	1.367	4.195	0.026
	Work / School	2.488	9.791	2.488	9.791	0.011
	Drop off Pick Up	0.627	2.374	0.627	2.374	-0.030
	Shopping	1.193	4.154	1.193	4.154	0.007
	Services / Private Business	0.892	3.239	0.893	3.239	0.052
	Recreation	1.131	3.666	1.131	3.666	0.021
	Social	1.190	3.771	1.190	3.771	-0.023
	Others	-0.375	-0.282	-0.375	-0.283	-0.083
Number of children per household (More than or Equal to 1)	Basic Need	-1.920	-4.151	-1.919	-4.151	-0.056
	Work / School	-0.934	-2.463	-0.934	-2.463	0.003
	Drop off Pick Up	-2.429	-6.585	-2.429	-6.584	-0.011
	Shopping	-2.280	-5.409	-2.281	-5.409	0.022
	Services / Private Business	-2.897	-8.200	-2.897	-8.199	-0.013
	Recreation	-1.840	-4.933	-1.840	-4.933	-0.006
	Social	-1.978	-5.201	-1.978	-5.201	0.011
	Others	-2.653	-2.938	-2.653	-2.939	-0.008
Age	21-30	-3.111	-8.850	-3.111	-8.849	-0.011
	31-40	-3.993	-10.343	-3.993	-10.34	-0.009
	41-45	-5.320	-12.671	-5.320	-12.668	0.004
	46-50	-1.700	-4.252	-1.700	-4.251	-0.031
	51-55	-7.810	-17.629	-7.806	-17.627	-0.047
	56-60	-7.941	-18.157	-7.941	-18.155	-0.002
	61-65	-9.318	-20.185	-9.318	-20.183	-0.005
	65+	-9.017	-19.039	-9.016	-19.037	-0.006

Translating satiation parameter component	Activity type	New model using GAUSS		New model using R		Parameter difference %
		Parameter	t-value	Parameter	t-value	
Constant	Basic Need	0.496	1.459	0.497	1.468	0.100
	Work / School	0.548	0.780	0.548	0.780	0.035
	Drop off Pick Up	0.613	2.181	0.613	2.181	-0.037
	Shopping	0.469	2.792	0.469	2.792	-0.100
	Services / Private Business	0.363	1.808	0.363	1.809	-0.098
	Recreation	0.642	0.455	0.641	0.453	-0.160
	Social	0.621	0.949	0.620	0.943	-0.100
	Others	0.754	0.308	0.754	0.308	0.012
Travel ratio	Basic Need	-0.030	-0.153	-0.030	-0.156	0.4500
	Work / School	-0.101	-0.403	-0.101	-0.403	-0.1693
	Drop off Pick Up	-0.057	-1.450	-0.057	-1.451	-0.7965
	Shopping	-0.069	-1.222	-0.069	-1.222	-0.3493
	Services / Private Business	-0.042	-0.990	-0.042	-0.990	-0.9833
	Recreation	-0.114	-0.098	-0.114	-0.098	-0.0807
	Social	-0.169	-0.333	-0.169	-0.330	0.1154
	Others	-0.144	-0.157	-0.144	-0.157	-0.2840
Number of people accompanied in the trip	Basic Need	0.025	0.311	0.025	0.311	-0.0800
	Work / School	0.058	0.262	0.058	0.262	0.7328
	Drop off Pick Up	0.017	0.172	0.017	0.172	0.2647
	Shopping	0.022	0.555	0.022	0.555	1.5545
	Services / Private Business	0.053	0.812	0.053	0.812	-0.5434
	Recreation	0.014	0.081	0.014	0.081	2.9071
	Social	0.029	0.318	0.029	0.318	-1.5897
	Others	-0.002	-0.002	-0.002	-0.001	-6.7000
Pure satiation parameter						
Constant	All Activities	-2.643	-248.69	-2.643	-248.60	0.0074
Married	All Activities	-0.007	-2.802	-0.007	-2.802	3.1429
LN(Number of people per household)	All Activities	0.057	20.113	0.057	20.113	0.6333
Number of vehicle per household (>=2)	All Activities	0.044	10.125	0.044	10.125	-0.2091

In this chapter, two statistical soft wares, GAUSS and R, were used to estimate the improved MDCEV model considering a Tpower value of 14, which was based on the same data set used in the previous research (sections 5.2.1 and 5.2.2). The outputs of the two soft wares were fairly similar; minor differences in estimation results were due to the number of digits each software accounted for when estimating the parameters. The estimation speed was dramatically different: GAUSS was much faster than R. Based on that, it is recommended to improve the compiler used by R. Projects such as activity-based travel demand forecasting study have time constraints just like any other project. Therefore, the speed of the software is a critical factor in estimating the activity-based models such as the improved MDCEV model and in submitting the

project on time. However, R performs well when small data set was used. In addition, R is free to download and use which is suitable for small research projects. The simulation searching algorithm in R showed better performance in terms of estimation speed compared to BHHH. The simulated searching algorithm does not rely on the gradient and the Hessian, and consequently results in fast execution of the code as tested in R. It is additionally recommended for GAUSS to add simulation approaches, which can serve to further increase the speed of estimation time.

Based on this notable difference, this research recommends the use of GAUSS to estimate the improved MDCEV model for mega-events. To sum-up the discussion of the statistical software section, criteria are established to select the suitable software for a specific project based on the project inputs which is presented in Table 25.

Table 25: Criteria for selecting the software to estimate MDCEV model.

Software	Speed	Estimation	Searching algorithm			Data points		Cost
			BHHH	BFGS	Simulation	High (>300)	Low (<300)	
GAUSS	Fast	Minor difference in estimation results	Use when coefficients' values are known	Use when coefficients' values are not known	Not applicable	√	√	$
R	Slow				√	Slow	√	Free

5.4.2 City of Makkah Store Closing Policy

Having established that this research will use the improved model and GAUSS software for an application stage, this section will set guidelines for a major policy for the city of Makkah, store closing during prayer times. The city of Makkah is sacred, unique, and complex: the city operates with strict policies with respect to religious activities – among them; all workplaces are closed during the five prayer times of the day. Violating the policy can result in a fine. During prayer times, residents go to a central area, whether Al-Haram or the closest mosque, or they

may pray at home. This research project aims to study the difference in time expenditure between the residents who perform rituals in the central area of the city (i.e., Al-Haram) and those who do not, specifically during Umrah season. Given that daily prayers take place throughout one day, a day was used as the unit and time frame for this model. As discussed in this chapter, a full day travel diary for the residents of Makkah was used to model the duration of various activities. The improved MDCEV model with Tpower 14 value is used to estimate activity duration using the recommended GAUSS software.

Two models were created to incorporate the inclusion or exclusion of religious activities as an out-of-home activity. The inclusion model considers eight out-of-home activity types (including religious activities) and the data includes residents who make at least one trip to Al-Haram during the day or night. Prayer in Al-Haram results in spending more time in the process of praying than the praying time: people who pray in Al-Haram spend time traveling to it, finding a parking spot, or walking through the large area of Al-Haram arenas (Al-Haram is the largest mosque in the world at 400,000 m^2). In addition, some people conduct other rituals at Al-Haram, which also take time. Based on that, the business owners who pray in Al-Haram close their stores longer than business owners who do not. As a consequence, residents wait a longer time to shop in those stores than in stores where the owners return sooner.

The exclusion model, on the other hand, considers seven out-of-home activity types (excluding religious activities) and the data includes residents who do not make any trips to Al-Haram during the day or night, i.e., residents who pray at home, in a store (shoppers or business owners), or in a nearby mosque. For residents praying at home, prayer time takes 5 minutes on average. Therefore, for the residents who prayed at home and planned to travel during prayer time, travel–time will be less given that most people have stopped for prayer in the mosques,

leaving the streets empty. Some stores allow shoppers to stay in the store to continue shopping even if the store is closed for prayer, though cash registers are closed. By the time the prayer time is over shoppers are ready to buy their merchandise, therefore they are not affected by the store closing policy. Praying in the mosque nearby takes less time than praying in Al-Haram, as the walking distance to the praying area in Al-Haram is generally much longer compared to local mosques. In addition, the extra activities that take place in Al-Haram (e.g., attending a lecture or praying more) further extend prayer time duration. Based on this combination of factors it is assumed that this group of residents (i.e., residents who do not make any trips to Al-Haram) is not affected by the store closing policy that this research aims to inform.

5.4.2.1 Models Estimation Results

The modelling framework explained in this research project considers the above two models. The goodness of fit is measured using Rho-Square index. In this section, Rho-square values for the first and the second model are 0.82 and 0.72. The highest goodness of fit is for the first model, which indicates that the first model should be included in the modelling framework (i.e., because it provides a more accurate model than the second). In addition, the variability of travel plans in the second model is higher than the first since it has lower goodness of fit. This is a rational output since the residents who pray in Al- Haram schedule their activities in the time slots between prayer times and, consequently, their variability in travel plans is lower. Besides that, the first model has a much lower number of coefficients that explains residents' behaviour compared to the second model (i.e., 16 coefficients for the first model and 40 coefficients for the second model) which indicate that the prayer time organizes the residents time schedule during

the day and the model does not need more variables to explain the residents time consumption behaviour (Table 26).

Table 26: Summary of the number of parameters and number of cases.

Models information	Trips to Al-Haram	No trips to Al-Haram
Total number of cases	4000	13500
Number of parameters of the full model	16	40
Number of parameters of the null model	3	3

A 90% confidence level of interval is used to test the statistical significance of the parameters' coefficients. Only the statistically significant coefficients remained in the models. Various households' related variables were tested on the pure satiation parameters, but they were statistically insignificant. Based on that, only constants are included for this parameter, as shown in Table 27. The constants are negative in value in both models and highest in value in the second model, which shows the higher preference of the residents of the second group for spending more time out-of-home compared to the first group. This is logical result since making trips to Al-Harm reduces the person's time budget (24 hours a day) and results in spending less time in other activities compared to the group that does not pray in Al-Haram.

Table 27: Pure satiation parameters components for the two models.

Pure satiation components	Trips to Al-Haram (First model)	No trips to Al-Haram (Second model)
Constant	-3.348	-2.687

The translating satiation parameter includes the constant and number of people combined in the activity variable. Table 28 shows the specifications of the coefficients. Both the constant and the occupancy signs are positive, which shows that the residents are willing to spend more

time out-of-home except for the occupancy coefficient of the single stop shopping activity (i.e., ACT 3) which contains a negative sign. The constant value of the first model is lower than the second, which shows that the residents of the first group prefer to spend less out-of-home time compared to the second group which is consistent with the pure satiation parameter results. The occupancy coefficient value of the first model is greater than the occupancy coefficient value of the second model. This reflects the first group's residents' preference for spending more time out-of-home when the travel group is large compared to the second group. Traveling in high occupancy vehicles reduces the congestion to Al-Haram when the number of travelling groups is high. However, if the traveling groups are using small vehicles, more traffic will be on the road and therefore there will be more congestion.

Table 28: Translating satiation components for the two models.

Translating satiation components	Trips to Al-Haram (First model)	No trips to Al-Haram (Second model)
Constant		
ACT1	0.335	0.421
ACT2	0.472	0.571
ACT3	0.272	0.421
ACT4	0.438	0.421
ACT5	0.234	0.421
ACT6	0.438	0.455
ACT7	0.335	0.418
ACT8	0.335	0.000
Occupancy		
ACT1	0.071	0.091
ACT2	0.071	0.008
ACT3	0.071	-0.022
ACT4	0.071	0.008
ACT5	0.071	0.008
ACT6	0.071	0.008
ACT7	0.071	0.008
ACT8	0.071	0.000

The baseline utility contains constants, gender, and age variables for both models. The constant value of the second model is more than the constant value of the first model, which indicates that the residents prefer to spend more time out-of-home when no store closing policy is implemented. The gender variable shows that males spend more time compared to females out-of-home in the first model in all out-of-home activities. For the second model, males also spend more time in all out-of-home activities, except for ACT 5, other activities (Table 29).

Table 29: Baseline utility components for the two models.

Baseline utility components	Trips to Al-Haram	No trips to Al-Haram
Constant		
ACT1	-88.800	-42.951
ACT2	-88.800	-42.616
ACT3	-88.800	-44.835
ACT4	-88.800	-45.536
ACT5	-88.800	-42.616
ACT6	-88.800	-50.291
ACT7	-88.800	-53.346
ACT8	-88.800	0.000
Gender		
ACT1	1.102	0.606
ACT2	1.102	1.627
ACT3	1.102	0.458
ACT4	1.102	0.008
ACT5	1.102	-3.331
ACT6	1.102	0.321
ACT7	1.102	3.648
ACT8	1.102	0.000

For the age variable, we can generalize that residents of different age groups spend more time in out-of-home activities when the store closing policy is not applied. The specification of the coefficients for the first and the second models are shown in Figure 25. The first model indicates that the out-of-home activity duration is reduced with age, while the second model shows that the work duration for the residents decreases with an increase in the age. The second

model indicates that the youngest age group (ages 21-30) spends more time in school compared to other age groups. In addition, the residents whose ages range between 41 and 50 spend more time in activities 3, 4, 5, 6, and 7 compared to other groups.

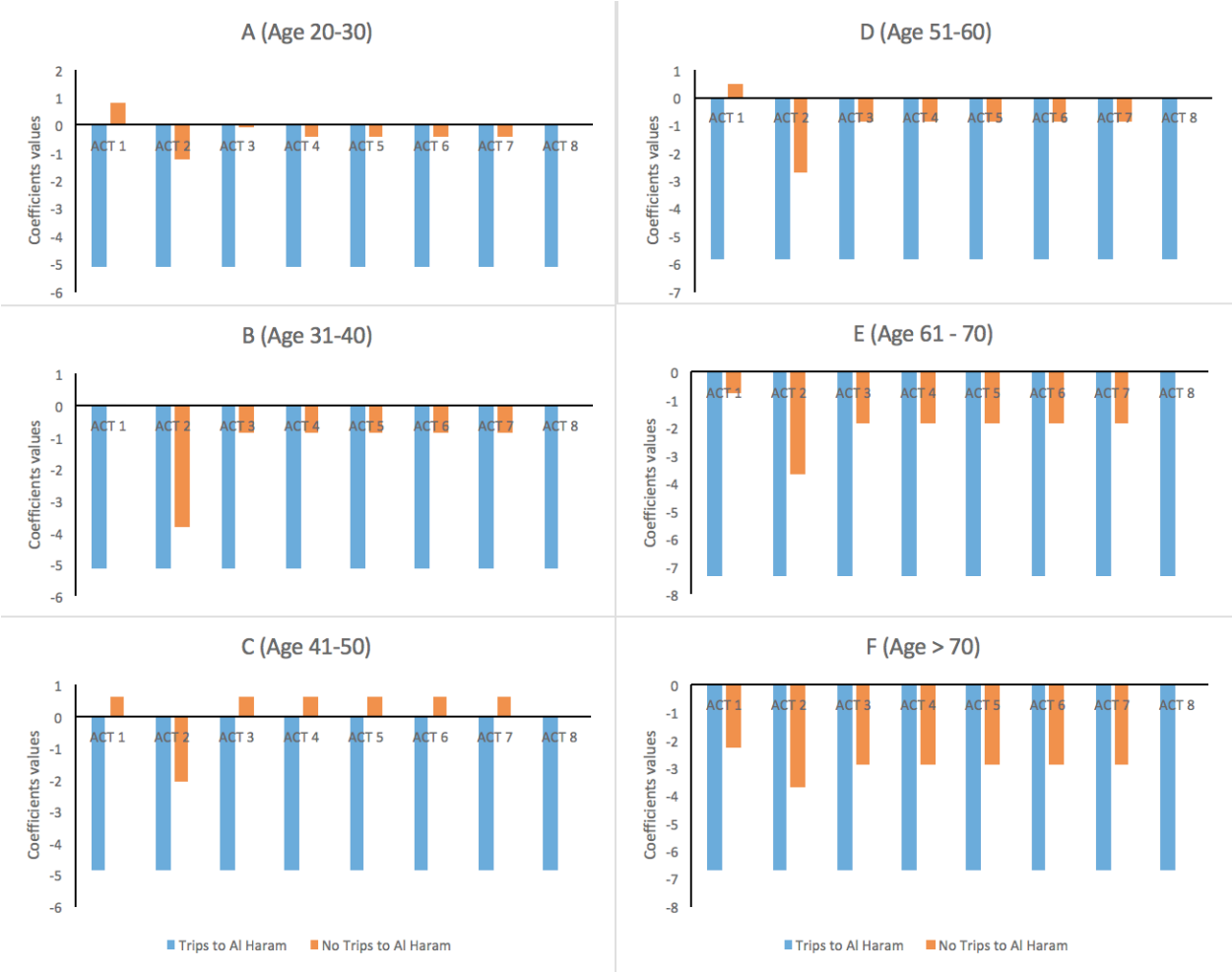


Figure 25: Age variable coefficient for different age categories.

Based on the previous discussion, it is clear that residents' time expenditure behaviour differs when the store closing policy is applied or not. In other words, residents' time expenditure behaviour is different when Al-Haram is included in their daily agenda and when it is not. In general, trips to Al-Haram take from the residents' time budget and physical energy and results in spending less time in other out-of-home activities.

5.4.2.2 Evaluating Makkah's Store Closing Policy

Taking into account the above findings, this research intends to discuss the store closing policy that is aimed at relieving the traffic pressure on the roads leading to Al-Haram and discusses the policy's effects on residents and businesses activities. Based on the results of the constant of the three parameters (i.e., the pure satiation, translating satiation, and the baseline utility parameters), along with age variable parameters, the residents spend more time out-of-home when the store closing policy is not applied. The rationale of this finding is that the time when the stores are closed and the residents are praying in Al-Haram is taken from the person's daily time schedule. Less time is left for other activities to be conducted. Conversely, the residents have more time to conduct out-of-home activities when the store closing policy is not in effect. Furthermore, keeping stores open during prayer time motivates residents to spend more time out-of-home. Some storeowners violate the store closing policy by keeping the stores open as much as they can, while others allow customers to stay on the premises during prayer times. It is expected that observing store closing time, increasing closing time, or better enforcing the store closing policy will shift the residents to pray at home or in a nearby mosque, which, as a result, would relieve the pressure on the roads leading to Al-Haram. Moreover, increasing the store closing time is expected to reduce the activities of the residents in the age group 41-50 (i.e., parents taking care of their dependents who need to spend more time in out of home activities such as shopping to shop for their dependents), thus, reducing congestion. To compensate for store closing time, one possibility is that business hours could be extended by a few hours at night. However, the store closing policy already implemented and business owners and customers have adapted to this policy as it is, therefore this solution may not positively affect business owners and customers.

Finally, traveling in large groups motivates residents to stay out-of-home, especially when Al-Haram trips are included in their agenda. In addition, traveling in large groups in a high occupancy vehicle helps reduce traffic in general, since traffic problems would occur if large groups travelled in individual vehicles. Based on that, it is advised to prevent small vehicles from entering the center area of Al-Haram. Also, it is recommended to support the city with Bus Rapid Transit (BRT) or metro lines leading to Al-Haram to accommodate the high demand.

In summary, this section of the research studied the time expenditure differences between the residents who visit Al-Haram and those who do not visit Al-Haram during the day. It evaluated the store closing policy for the city of Makkah, which dictates that when the mosques call for prayer all stores should be closed until the prayer is done. Results showed the importance of including residents who pray in Al-Haram in the modelling framework such that the goodness of fit of their group is higher than the other group (i.e., residents who do not pray in Al-Haram). The models also shed the light on the time expenditure differences between the groups. The research concluded that keeping the store closing policy, is expected to reduce the congestion of the streets. In other words, increasing the store closing time and better enforcing it is expected to reduce the traffic in the city, especially on the roadways leading to Al-Haram because residents who stop for praying will leave the street open for the others who prayed at home and need to travel faster during prayer time. The same way, if a city holding mega-event applied store closing policy to the stores on the major roads leading to the event (i.e., before and after the event), the travel demand on the major road are expected to reduce. For the city of Makkah, if the city canceled the store closing policy and allowed business activities during prayer times, the city could use the no trip to Al-Haram model to model activity duration, which reflects this scenario.

5.4.3 Conceptual Travel Demand Modelling System Framework for Makkah During Umrah Season

Makkah has frequent mega-events, special urban characteristics, and a distinct environment, making it an ideal case study for modelling travel demand (background information about the city of Makkah is provided in the appendix). This research project implements one of the most advanced modelling techniques; activity-based modelling, in the analysis of travel demand on Umrah season which include Friday season, which will serve as the basis for possible future work.

Hajj is a bigger event than Umrah that result almost all city links to be congested which is not the same condition as the world cup or the Olympics. In Umrah, the network tends to be less crowded than in the other seasons. Consequently, Makkah residents could add more activities to their plans. However, it is assumed that Makkah residents follow a similar routine each day, and any changes in daily plans are captured in the aforementioned diaries from residents in the city. However, visitors' routines may vary from day to day. Based on that, the framework will capture both residents and visitors movements, such that a model will be estimated for the residents and another for the visitors. As the residents of the city of Makkah are from around the globe and the number of data set used in modelling Makkah mega-events is high. In addition, some of the residents participate in the city mega events and others do not. Based on that, travel behaviour of the residents of Makkah during mega-events should be representative to other mega-events such as the Olympics. In addition, as the activity duration model (MDCEV model) is based on maximizing an individual utility or benefit in spending time on out-of-home activities constrained by a time budget, the major factors in estimating activity duration are the individual, household, and trip information. Therefore, the activity duration model for the residents of

Makkah estimated in this research could be used in other cities or events as explained in section 5.6. This activity duration model is an input to the activity-based travel demand framework. The activity-based modelling framework is not exclusive for the mega events. It could be used for other events by not including the activity duration of mega-event in the modelling framework.

Activities in Makkah can be ranked into three categories based on the city policy and frequency of the activities observed in the available data. Therefore, the sequential modelling framework is based on the activity type ranking. Based on the ranking, the activity generation modelling components of the first rank category is an input to the individual time budget (i.e., 24 hours) as first, second and so on. The most important activity, ranked first, is the mega-event or in Makkah case the religious activity. People visit Al-Haram to perform different ritual activities such as the five daily prayers, Friday prayer, Janaza prayer, and Umrah. Al-Haram visitors also study, work and do volunteer jobs at Al- Haram as well as break their fast there on Mondays and Thursdays. The city has a very strict prayer time policy in which all activities, including work and shopping, are halted for praying. Engaging in anything outside of prayer activities at the designated times can result in fines. As a result, Al-Haram visits are considered major activities because of the prayer time policy, variety of activities that take place there, and the high number of visits to it.

Work and school activities rank second because of their high daily frequency and their importance in modern life. Both work and school have different start times and durations that depend on the sector to which they relate, such as governmental or private sector work, or primary/secondary school or university. Minor activities are ranked third and include shopping and recreational activities.

Makkah has five prayer times and three peak hours (Figure 26 and Figure 27 shows the skeleton of the modelling framework). Prayer times are identified by the time of the day (i.e., the location of the sun during the day) and the city location – as a result, prayer time varies from day to day. The first prayer of the day is the Fajr (predawn) prayer. It takes place during the morning twilight, which starts when the sun is below the horizon and sunlight appears across the sky, and this occurs around 5:00 am. The second prayer is the Dhuhr (midday) prayer, and starts around 12:30 pm when the sun is perpendicular to the earth, when shadows are right underneath objects. The third prayer is the Asr (afternoon) prayer, and starts around 4:00 pm when the length of an object's shadow is the same length of the object itself. The fourth prayer is the Maghrib (sunset), and starts around 6:30 pm. It starts at sunset and lasts until the sunset's red light is gone. The fifth prayer is the Isha (night) prayer. It starts around 8:00 pm, after the red light disappears from the sky. The five prayer times could take place in Al-Haram or the city mosques or home. Prayers are ranked as a first category if they are conducted in Al-Haram or at the mosques. However, if they are conducted at home then they are considered as home activates (last category).

The three peak travel hours in Makkah take place in the morning, afternoon, and evening. The morning peak hour (at 7:00 am) and the afternoon peak hour (at 2:00 pm) result from work and school activities. The evening peak hour (at 6:30 pm) is the result of religious activities, shopping, and activities such as social visits. During the five prayer times, the number of trips decreases, as discussed in the store closing policy section. Figure 26 and Figure 27 show the three peak times, as well as the decrease in the number of trips during prayer times.

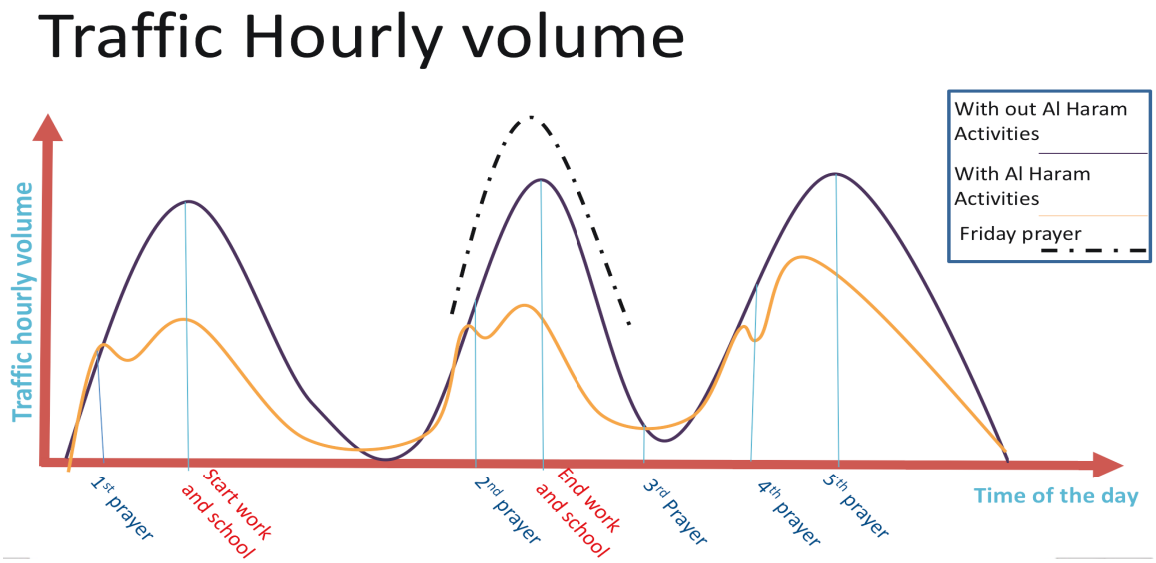


Figure 26: Conceptual traffic hourly volumes showing the three peak hours and the increase in demand at Friday prayer.

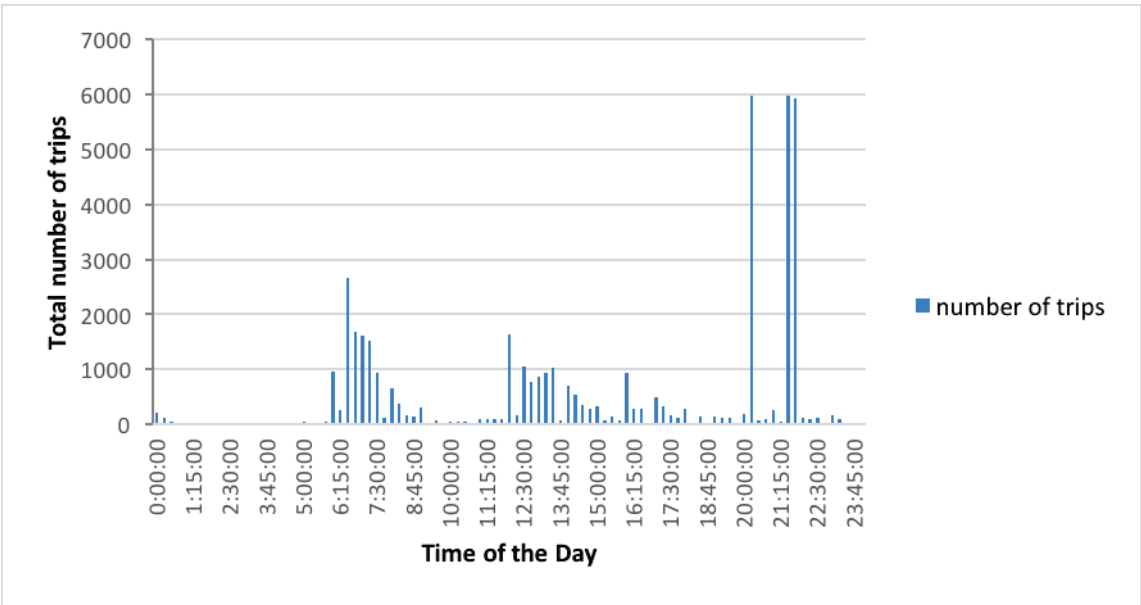


Figure 27: Total number of trips per 15 minutes based on the obtained data set showing the three peak travel periods (Makkah data set, 2010).

The conceptual travel demand framework should recognize market segmentation such as residents/visitors or gender (Shalaby et al., 2010). Male and female residents require different models, as females are dependent on males to make their trips in Makkah because they are not permitted to drive. Visitors are expected to make more trips to Al-Haram than residents, because visitors are visiting Makkah for a short period of time (e.g., Umrah visitors are only allowed to stay a maximum of one month). The estimated activity duration models from the Chapter 5 included residents only, as the data does not contain visitors' records. In addition, the estimated model included gender as a variable and showed the differences in time expenditure between the two genders.

The framework also should be a household-based and accounts for household interaction because females depend on males for their out-of-home activities, where, for example, a woman wanting to go to the shopping mall may need her brother to drive her there, pick her up and drive her back home. The framework uses a 24-hour time budget (16 hours for out-of-home activities), and different models are estimated for weekdays and weekends. This segmentation was done in the Chapter 5 to capture people's different activity patterns on different days. For instance, on Saturday through Wednesday, people mainly spend the day at work or school and the night at home. During the weekends (i.e., Thursday and Friday), most residents spend time with family and friends, although some residents work in places such as grocery stores and shops. On Fridays, a large number of people pray Friday's prayer at Al-Haram. The framework aims to recognize space constraints during weekdays, such as the greatest distance a person could travel and get back home in that time constraint. On weekends, families could also potentially travel outside of the city.

This study recommended GAUSS statistical software package over R to estimate the activity duration model (i.e., MDCEV model) for Umrah season. GAUSS software and both BHHH and BFGS searching algorithm demonstrated more efficient results than R software. As estimating all activity generation components would take years, this research focuses on an activity duration model to narrow down the focus. Future work will model activity type, location, and mode of transportation using a MNL model.

The final stage of building the conceptual modelling framework is to connect activity generation components and then schedule people's activities. Therefore, different modelling decisions are suggested in this thesis dissertation –sequential, joint, and hybrid econometric– that will be tested in future work, as well. Each of these decisions reflects how people conduct their trips. The following are modelling decisions suggested in this thesis dissertation.

Firstly, the framework will model activity-based model components (i.e., activity generation and activity scheduling) sequentially, with each activity-based model component independent of the others. Figure 28 presents a framework that uses a sequential modelling decision. It ranks activities by priority such as religious, work and school, and lastly, other. No other activities can substitute religious activity in the context of Makkah because of the city's policy discussed previously. Consequently, these activities are input in to every person's schedule by default. People who do not pray in Al-Haram pray in the mosques closest to their current location (the number of mosques in Makkah is considerable), but when mosque locations are not available or sufficiently convenient, the assumption is that people pray where they are (Figure 28 shows the prayer time in red line in a person schedule). For people who pray in Al-Haram, Al-Haram activity is considered a first priority, and then work and school, and finally other activities. For people who consider work as a first priority activity work duration and start

time are considered a blocked period during which no other activities can take place. However, this group of people could pray in their work place. Activities with low ranking could fill the periods between religious and work activities. Shopping and recreational activities are assumed to be the lowest-priority and they should be modelled last. Thus, the modelling decision in this case is a sequential modelling framework based on the ranking of activity types. Within each activity type, the other modelling components such as start time, duration (improved MDCEV model), location, and mode of transportation could be modelled sequentially, jointly, or simultaneously (see Figure 28 for sequential modelling decision of the activity-based components).

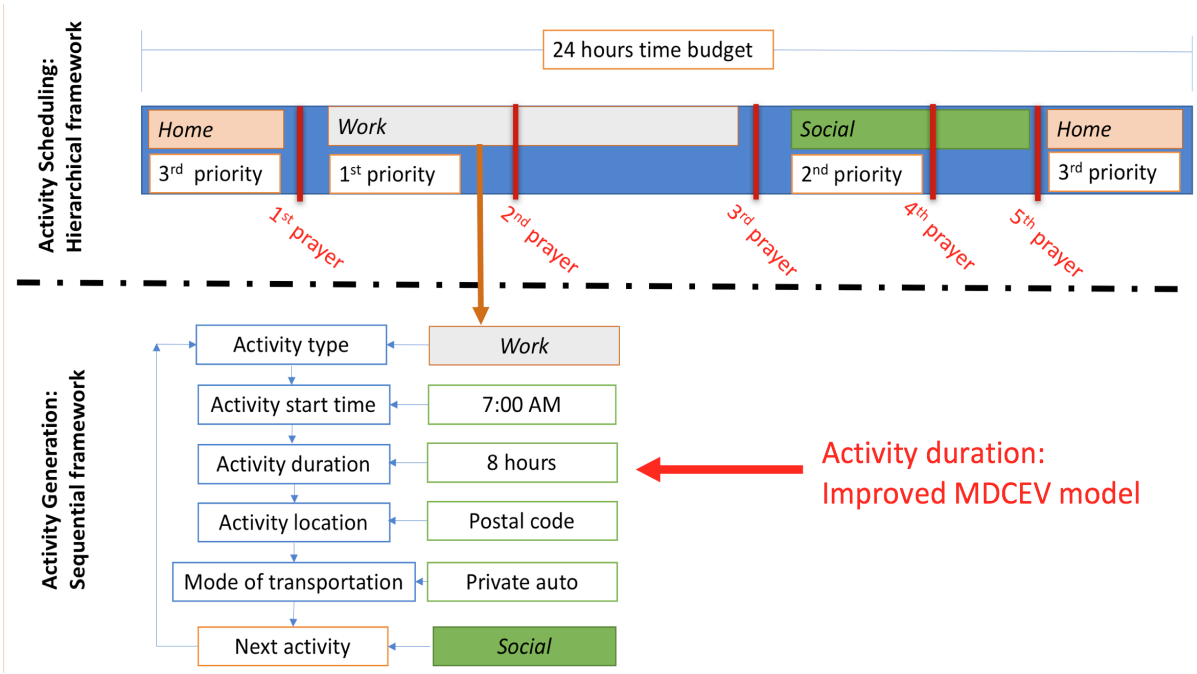


Figure 28: Sequential modelling decision.

Secondly, the framework will model activity-generation components in a hybrid manner. In this case, activity type, start time, duration (improved MDCEV model), and location will be

modelled jointly. Then, with whom, mode of transportation, and the route choice will be modelled jointly following the first model. This is referred to as a hybrid framework because it combines joint and sequential decisions (Figure 29). This method is a suggested framework will be tested in future work. The hypothesis is that only some of the activity-generation modelling components should be modelled jointly, but not all, as modelling all components jointly would be computationally cumbersome and could affect the significance of the parameters. Activity-generation model components could be further divided into activity and transportation segments. The activity segment includes activity type, start time, and duration (improved MDCEV model), whereas the transportation segment includes mode of transportation, with whom, and activity location. It reflects the stages of how people make their travel decisions. Activity scheduling still could be modeled sequentially.

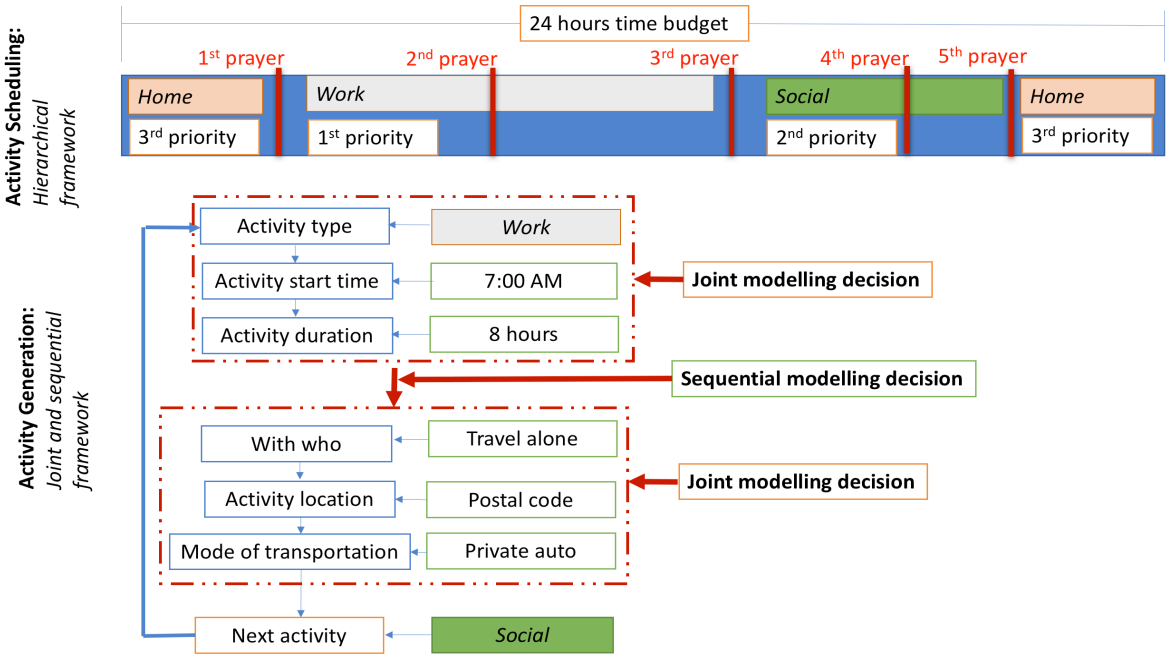


Figure 29: Joint and sequential modelling decision.

Thirdly is the joint modelling framework, whereby all of the activity-based model components (i.e., activity generation and activity scheduling) will be modelled jointly. This framework will consider all activity-generation components such as activity types, their start time, duration, location, mode of transportation, and route choice in a joint framework. Each one of these components is considered as a separate level of choice decision-making that depends on the other levels. In addition, for activity scheduling, the model will be dynamic and build on the activity patterns step by step. For each step, the model will recognize the consumption of time or the time and space constraints to build the final schedule and the consequent tours such as the one proposed by Habib, 2011. The final output of these activity-based modelling frameworks is the individual's agenda during the day. These agendas are executed in simulation software which contain the spatial characteristics of a city and which distribute travel demand on the city transportation network to present the predicted travel demand on each road link. Based on that, the travel demand result from these three proposed modelling frameworks can be compared with current travel demand to conclude on the best framework that capture individuals' travel behaviour. The simulation software not only helps in quantifying travel demand on a city network, but also a good visualizing tool. Based on the simulation results, transportation engineer researchers can specify the problematic section on the city network and test scenarios that aim to solve these problems. In addition, it can help in predict travel demand when some roads are closed. The simulation software, therefore, will distribute travel demand accordingly.

The above description of the modelling process demonstrates that there are many components to consider in improving upon a model for a framework for travel demand and to improve city planning. This project takes a slightly different approach in analyzing travel demand than others that usually consider work-related activities at the forefront. Most attention

paid to work-related activities in the existing literature [Bhat and Singh (2000), and Bhat (2001)]. This research has highlighted the uniqueness of mega-events in the city of Makkah and has emphasized how religious activities can impact an urban space and transportation issues, and that activities within certain time frames need to be considered when planning for the future. All the models estimated in this research are for the residents of Makkah during Umrah season. Therefore, all the models outputs in Chapter 5 which include considering Friday as a separate category from the weekends, the policy guidelines, and these studies models specification shapes part of the conceptual framework for the city of Makkah during Umrah season.

6 Chapter: Conclusion

A travel demand model is a valuable tool for the development and planning of any city and its mega-events. Decision makers can use travel demand modelling to decide how to implement major transportation infrastructure and land use projects. Transportation agencies have used the UTMS (four-stage) modelling approach developed in the 1950s and 60s, yet it is limited in its ability to present a bigger picture of travel demand. Activity-based modelling, a newly implemented and advanced technique, may be better suited to testing policies, studying human behaviours, and forecasting travel demand. The major assumption underlying activity-based modelling is that people travel not for the sake of travel, but to conduct activities. The model is divided into activity generation and activity scheduling. The present doctoral research focused on the former, more specifically, on the activity duration model proposed by Bhat (2005), which uses the multiple discrete-continuous extreme value (MDCEV) model. This research aimed to evaluate and improve the MDCEV model for mega-events, using the city of Makkah as a case study.

To achieve this goal, this dissertation first presented a review of literature concerning the travel demand frameworks for mega-events, including Makkah's travel demand framework, activity-based model components and frameworks, and activity duration modelling. One of the purposes of the literature review was to assist in establishing the major components of the modelling framework for the city of Makkah and its mega-events. Yan et al., (2010) conducted a relevant study of travel demand modelling for the Beijing Olympics using the advanced modelling approach (i.e., activity-based) to model the residents and visitors' travel demand during the mega-event. This research shed the light on the limitations of the travel demand modelling framework and informed improvements to it. Ultimately, a limited amount of research

has been conducted on this topic, despite the importance of the transportation, safety, and economical considerations of mega-events.

Following the broad literature review of travel demand and mega-events, this dissertation looked more closely at the city of Makkah's travel demand frameworks. It was found that much of the relevant research suggested improvements to the transportation system based on site visits and not based on transportation evaluation tools. Limited efforts have been made to estimate travel demand using the four stage model; regardless, this modelling method is not appropriate for the city of Makkah because the city is planning to implement new policies, modes of transportation, and land use developments (as described in the background information about Makkah in the Appendix A). These plans require an advance modelling approach to test future planning scenarios. In addition, previous research focused on one of Makkah's mega-events, Hajj, as it is biggest of four seasonal events in Makkah. It was assumed that if transportation planning could solve traffic problems in Hajj it would work for the rest of the events; however, based on the background information about the city of Makkah, travel demand during each of the events varies significantly. Therefore, there is a need to establish a travel demand framework for each individual event, which would require a household travel survey to be conducted for each season.

Activity-based models (activity generation and scheduling) and their modelling components were also reviewed to help build the conceptual travel demand modelling framework for the city of Makkah. Activity generation models contain several components, including activity type, duration, and mode of transportation. Activity scheduling decisions also have various frameworks, namely sequential, simultaneous, and nested. In theory, activity generation and scheduling could be combined in several ways to build travel demand framework

and consequently estimate travel demand, but establishing a full activity-based modelling framework with all modelling components and testing several scenarios would take years of work. Thus, testing several scenarios is kept for future work and this research project focused on only one of the two modelling components, activity generation models more specifically activity duration model, which is considered one of the first steps in establishing activity-based modelling framework.

This research further reviewed previous efforts in estimating activity duration, including data description and modelling techniques, and found that the best method to estimate activity duration is the MDCEV model proposed by Bhat (2005). The MDCEV model overcomes the limitations of other modelling approaches such as structural equation modelling and the hazard model. The MDCEV model structure was reviewed, and applications and proposed improvements to the MDCEV model presented in an attempt to further improve the accuracy of the MDCEV estimation.

Following the above described literature review, a new MDCEV modelling structure was proposed to improve the accuracy of the existing model. First, various utility structures were tested using the existing data and the appropriate utility structure that results with the high goodness of fit and maximum number of statistically significant parameters was chosen. Then, based on the literature review of the MDCEV model, the three model parameters (i.e., baseline utility, translating satiation, and pure satiation parameter) were evaluated. It was found that one of the model parameters, the translated satiation parameter, hardly captured the constant marginal utility effect. Following Easa's (2014) model in Hydraulics, it was found that adding an exponent parameter (Tpower) to the translating satiation parameter in the indirect utility function allowed the function to range from a straight line (i.e., constant marginal utility effect) to a

diminishing curve. This finding provided a methodology to select the most appropriate Tpower value that results with a high goodness of fit and maximum number of statistically significant coefficients. The baseline utility parameter of the MDCEV model already contained a scale parameter. It was found in this study and in Habib et al. (2007) that setting the scale value to one results with the best model. Moreover, the pure satiation effect of the MDCEV model already captures the constant utility effect. It was found that adding a power parameter to it does not improve the goodness of fit. The Tpower parameter adds more statistically significant coefficients to the baseline utility parameter and the pure satiation parameter, yet it reduces the number of statistically significant coefficients in the translating satiation parameter. This is considered a side effect of the Tpower parameter.

Once improvements to the MDCEV model had been proposed, Bhat's base MDCEV model was used in a pilot study to evaluate activity duration in the city of Makkah. The results were used as the basis for comparison with the results using the improved model structure. Specifically, the pilot study investigated the time expenditure behaviour during weekdays, weekends, and Fridays. The evaluation results showed the importance of including Friday as a separate category, as well as that this model framework which included Friday special event could be used for any other special event, such as a farmer's market during weekends.

To reinforce that the improve MDCEV could improve the accuracy of activity duration for any city, first, part of the MDCEV model (week 1 model) estimated by Habib et al. (2007) for two German cities was estimated using the base model. Then, the improved MDCEV model was tested on two data sets from two different cities (Saudi Arabia and German data sets) using different time budgets (i.e., a day and a week) and the results were compared. The results showed that the improved MDCEV model structure provides more accurate activity duration estimates

compared to the base MDCEV model, with up to a 74.3% improvement in accuracy and the interpretation of the results are more intuitive. Second, the improved MDCEV model estimated from Makkah dataset is used to predict activity duration for the German cities. It was found that the improved MDCEV model estimated based on Makkah dataset accurately predicts the activity duration for work, school, shopping, and social activity duration for the German cities. As a result, the improved MDCEV model estimated based on Makkah dataset can be used to help cities with limited transportation planning recourses to better plan their cities.

Next, the improved MDCEV model was applied using two statistical software packages, GAUSS and R, to examine differences in the estimated results. It is assumed that different software packages have different frameworks and assumptions, and therefore possibly different outputs. The same code, and specifications are used in both software. Results indicated that there were slight differences in the estimated results due to rounding (i.e., the maximum difference in estimated coefficients is 6.7%). The coding effort in R was expected to motivate further development of the MDCEV model because it is free and accessible software; however, the study ultimately recommends the use of GAUSS, as it was more efficient in terms of calculation speed. The study noted that the speed of R software could be improved significantly by developing a better compiler than the existing one that used maxlik package, which is the primary reason for the slow estimation time. The extra lines in R code compared to GAUSS do not contribute much to the speed of the software. Each of the software have their own capabilities that helps selecting the appropriate software for the needed project based on the study set criteria established. Faster software in estimating the required travel demand models will help in submitting travel demand project on time and avoid penalties associated with delay.

As an important next step, this research studied the time expenditure behaviour between Makkah's residents who visit Al-Haram and those who do not visit Al-Haram during the day and night. The study showed the effect of Al-Haram visits, specifically that residents spend more time in-home when they do pray at Al-Haram. Based on this finding, the effect of the city's store closing policy during prayer time was evaluated. It was assumed that residents who pray at Al-Haram do not shop during prayer time (i.e., the store closing policy is carried out) since all stores are closed during the five prayer times of the day. As a result, residents who pray at home have the option to travel around the city more easily, i.e., the store closing policy reduces congestion on the road transportation network. Therefore, it is advised for the city of Makkah to keep the store closing policy and to better enforce it. It was important to address this policy in the modelling framework because it gave insight into how the policy changes residents' travel behaviour.

To extend these results, this study could be implemented for the city of Madina, Saudi Arabia, since it is similar in city structure and policy to Makkah. Like Makkah, Madina also has a Holy mosque in the center of the city that attracts millions during Friday prayer. This modelling structure could also be applied to other cities hosting mega-events, such as the Olympic games. In these cases, people could be divided into two groups: the first group includes people who attend the games (or other event) and the second group is people who do not attend the games. Then, a study of the effect of each group on the transportation network could be conducted.

In closing, this dissertation work developed a conceptual activity-based travel demand framework for the city of Makkah, which included mega-events (Umrah and Friday), fit the improved MDCEV model in the framework, and recommend the use of GAUSS statistical

software. There is currently no unified activity-based modelling framework that states which modelling component depends on which (i.e., activity duration depends on mode of transportation or vice versa). Therefore, a literature review on mega-events and activity-based models informed a conceptual framework that set the major modelling components (e.g., mode of transportation and activity location) and presented various activity scheduling structures that will be tested in the future work. There is also a need for adequate household travel survey data for each season of the year in the city of Makkah. The currently available data suffices for the Umrah and Friday seasons, but future data collection should consider the two other seasons, Hajj and Ramadan. In addition, data should be collected from more than 5% of the population in order to better estimate the models. Still, the Friday model estimated in the pilot study used around 900 data points, while the MDCEV model estimated by Habib et al. (2007) used around 300 data points, which is nonetheless considered good. Though more data is desirable, collecting data for mega-events can be very expensive. It is advisable to use cellphone technology in collecting travel diaries to significantly reduce the cost of data collection. It is also worth considering including the data in one Excel sheet instead of three to reduce the effort in reviewing and organizing data. Future work will combine several improvements of the MDCEV model (the one developed in this research with others available in the literature) in a single framework for a possibility of further improvement in the estimation accuracy of the MDCEV model.

To conclude, this research evaluated and improved the MDCEV model and fit the model in an activity-based modelling framework for mega-events. The improved MDCEV model improved the predictability of the travel demand model and, consequently, will aid in better planning for cities and mega-events.

References

Abdulazim, T., Abdelgawad, H., Habib, K., and Abdulhai, B. (2013). Using Smartphones and Sensor Technologies to Automate the Collection of Travel Data. *in Annual Meeting of the Transportation Research Board. Washington, D.C.* 2013.

Al-Harbi, M. (1986). The Management of Pilgrims Transport Movement. *Proceedings from the Transportation Symposium: Makkah, Saudi Arabia: Ministry of Communication*, pp. 27-34. 1986.

Al-Aqad, S. (1988). The Importance of Public Transport in the Hajj Services. *Proceedings from the Transportation Symposium. In proceeding of the Third Hajj. Makkah, Saudi Arabia: Ministry of Communication*, pp. 125-141. 1988.

Al-Jefry, M., and Al-Hazmi, M. (1989). The Use of Two-floor Buses in the Hajj Transport. *Proceedings from the Transportation Symposium. In proceeding of the Fourth Hajj. Makkah, Saudi Arabia: Ministry of Communication*, pp. 337-341. 1989.

Al-Laithy, A. (2009). What Everyone Should Know About the Qur'an. *King Fahd National Library Cataloging-in-Publication Data.* 2009.

Al-Rahman, H., and Abdulaal, W. (1989). Towards General Strategy of Transport during the Hajj. *Proceedings from the Fourth Hajj Transportation Symposium: Makkah, Saudi Arabia: Ministry of Communication*, pp. 37-57. 1989.

Al-Shareef, F. (1989). The Use of Buses System and the Imagination of the Users. *in proceeding of the Fourth Hajj Transportation Symposium, Makkah, Saudi Arabia: Ministry of Communication, pp 25-33,* 1989.

Al-Turki, M. (1989). Evaluation and Organization of Traffic during the Hajj. Makkah, Saudi Arabia: Ministry of Communication, pp. 197-204. 1989.

Al-Yafi, A. (1987). The Management of Movement Regulations during the Hajj. *Proceedings from the Second Hajj Transportation Symposium: Makkah, Saudi Arabia: Ministry of Communication*, pp. 221-230. 1987.

Andrew, A. (2008). Overview of Quasi-Newton optimization methods. *Informal tutorial at Microsoft Research.* Retrieved from http://www.bibme.org/citation-guide/apa/lecture/. 2008.

Aptech Systems, Inc. (2002). Using the GAUSS Engine successfully in a real world application: MapInfo GAUSS Case Study. November 2002.

Aptech Systems, Inc. (2004). Maximum Likelihood Estimation for GAUSS Version 5.0. *Part Number: 001414, Aptech Systems, Inc.*, Maple Valley, WA. 2004.

ARC (2012). Activity-Based Travel Model Specifications: Coordinated Travel – Regional Activity Based" Atlanta Regional Commission, Modeling Platform (CT-RAMP) for the Atlanta Region. *Atlanta Regional Commission*. December, 2012.

Arentze, T., and Timmermans, H. (2000). ALBATROS: A Learning Based Transportation Oriented Simulation System. *European Institute of Retailing and Service Studies (EIRSS)*. 2000.

Arentze, T., and Timmermans, H. (2004a). ALBATROSS Version 2.0: A learning based transportation oriented simulation system. *European Institute of Retailing and Service Studies (EIRSS)*. 2004.

Arentze, T., and Timmermans, H. (2004b). A learning based transportation oriented simulation system. *Transportation Research Part B 38, pp 613-633*. 2004.

Ashish, A. (2004). A comparison of weekend and weekday travel behavior characteristics in urban areas. *MS thesis, University of South Florid.* 2004.

Badabaan, M. (2001). TRAVEL BEHAVIOUR OF PILGRIMS DURING THE HAJJ, SAUDI ARABIA. *Ph.D. dissertation, Department of City and Regional Planning, Faculty of Engineering and Environmental Design, Cardiff University, 2001.*

Ben-Akiva, M., and Bowman, J. (1998). Activity Based Travel Demand Model System. *P. Marcotte et al. (eds.), Equilibrium and advanced Transportation Modelling, Springer Science Business Media, New York*. 1998.

Ben-Akiva, M., Bowman, J., and Gopinath, D. (1996) Travel demand model system for the information era. *Transportation (23), pp 241-266*. 1996.

Ben-Akiva, M., and Bowman, J. (1995) Activity based disaggregate travel demand model system with daily activity schedules. *Massachusetts Institute of Technology.* 1995.

Ben-Akiva, M., and Lerman, S., (1985). Discrete choice analysis: theory and application to travel demand. *MIT Press series in transportation studies, Marvin L. Manheim, editor.* 1985.

Berndt, E., Hall, B., Hall, R., and Hausman, J. (1974). Estimation and inference in nonlinear structural models. *Annals of Economic and Social Measurement, 3, pp 653–665. 1974*.

Bhat, C., and Singh, S., (2000). A comprehensive daily activity-travel generation model system for workers. *Transportation Research Part A, volume 34, issue 1, pp 1-22. 2000.*

Bhat, C. (2001). Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model. *Transportation Research Part B: Methodological, volume 35, issue 7, pp 677-693.* August 2001.

Bhat, C., Srinivasan, S., and Guo, J. (2002). Activity-Based Travel-Demand Modelling for Metropolitan Areas in Texas: Data Sources, Sample Formation, and Estimation Results. *Research Project FHWA/TX-03/0-4080, Conducted for the Texas Department of Transportation in cooperation with the U.S. Department of Transportation Federal*

Highway Administration by the Center for Transportation Research The University of Texas at Austin. September 2002.

Bhat, C., Srinivasan, S., Guo, J., and Sivakumar, A., (2003). Activity-based travel demand modelling for metropolitan area in Texas: A micro-simulation framework for forecasting. *Research Report FHWA/TX-03/4080-4, Research Project 0-4080, Conducted for the Texas Department of Transportation in cooperation with the U.S. Department of Transportation Federal Highway Administration by the Center for Transportation Research The University of Texas at Austin. February 2003.*

Bhat, C. (2005). A multiple discrete-continuous extreme value model: Formulation and spplication to discretionary time-use decisions. *Transportation Research Part B: Methodological, volume 39, issue 8, pp 679-707. 2005.*

Bhat, C. (2008). The multiple discrete-continuous extreme value (MDCEV) model: Role of utility function parameters, identification considerations, and model extensions. *Transportation Research Part B: Methodological, volume 42, pp 274-303.* 2008.

Bhat, C,. Goulias, K., Pendyala, R., Paleti, R., Sidharthan, R., Schmitt, L., and Hu, H. (2012). Accommodating Intra-Household Interactions in Activity-Based Model Systems: How Much Does it Matter from a Forecasting and Policy Sensitivity Standpoint?. *Presented at the Innovations in Travel Modeling (ITM) conference, Tampa, FL.* 2012.

Bhat, C., Goulias, K., Pendyala, Paleti, R., R., Sidharthan, R., Schmitt, L., and Hu, H. (2013). A Household-Level Activity Pattern Generation Model with an Application for Southern California. *Spring Link, Transportation, volume 40, issue 5, pp 1063-1086.* September 2013.

Bovy, P., (2013). Transport planning for FIFA-2018 in Russia. *CCTM--Coordination Council for Traffic Management, Moscow, Training Seminar. Retrieved from http://www.mobility-bovy.ch/resources/Doc5-Russia-2018-Nov-Seminar-BY.pdf. October 2013.*

Bushnak, A. (1977). Planning and evaluation of special event transportation system with application to the Hajj. *Ph.D. dissertation University of Michigan, Michigan*. 1977.

Casello, J.M., Nour, A., Hellinga, B., 2009. *Quantifying impacts of transit reliability on user costs.* Transp. Res. Rec. 2112 (145), 136–141.

Castiglione, J., Bradley, M., and Gliebe, J. (2015). Activity-Based Travel Demand Models: A Primer. *TRB SHRP 2 Report S2-C46-RR-1.*

Chapin, F. (1974). Human activity patterns in the city: Things People Do in Time and in Space. *John Willy and Sons, Social indicators Research, volume 2, issue 2, pp.261-264.* 1974.

Cullen, I., and Godson, V. (1975). Urban network: the structure of city patterns. *In progress in planning, edited by D. Diamond and J. B. McLoughlin, volume 4, issue 1, New York: Pergamon Press, 1975.*

Ding, Y., Lushi, E., and Li, O., (2010). Investigation of Quasi-Newton methods for unconstrained optimization. Inter*national Journal of Computer Application, volume 29, pp 48-58, 2010.*

Dunn, K., and Mcguirk, P. Hallmark events. In Cashman R. and Hughes A. (eds) (1999). Staging the Olympics: The Event and its Impacts. *Centre for Olympic Studies, University of New South Wales Press (UNSW), pp 18-32, Sydney,* 1999.

Easa, S., Strauss, T., Hassan, Y., and Souleyrette, R., (2002). Three-dimensional transportation analysis: Planning and design. *Journal of Transportation Engineering, volume 128, issue 3.* May 2002.

Easa, S. (2014). New and improved four-parameter non-linear Muskingum model. *Proc. Inst. Civ. Eng. Water Manage., volume 167, issue 5, pp 288–298.* 2014.

Eluru, N., Pinjari, A., Pendyala, R., and Bhat, C., (2010). An econometric multi-dimensional choice model of activity-travel behaviour. *The International Journal of Transportation Research, Transportation Letters, volume 2, issue 4, pp 217-230.* 2010.

Essex, S., and Chalkley, B., (1998). Olympic Games—catalyst of urban change. *Leisure Studies, volume 17, issue 3, pp 187 – 206.* 1998.

Evacation.org (2016). Masjid Al-Haram – Sacred Place of Muslims. Retrieved from http://www.evacation.org/masjid-al-haram-sacred-place-of-muslims/. January 2016.

Fattah, H. (2005). Islamic Pilgrims Bring Cosmopolitan Air to Unlikely City. *The New York Times, Middle east, Mecca Journal,* January 20, 2005.

Fayez, Z., (2013), personal communication, unpublished data.

GAUSS (2012). Maximum Likelihood Estimation 5.0 for GAUSSTM Mathematical and Statistical System. Aptech Systems, Inc. Retrieved from http://www.aptech.com/wp-content/uploads/2012/11/ml5.0.pdf . last visited January 2017.

General authority for statistics Saudi Arabia (2013). *Household expenditure and income survey* (2013). Retrieved from "https://www.stats.gov.sa/sites/default/files/msH_nfq_wdkhl_lsr_2013m.pdf"

GNU GENERAL PUBLIC LICENSE (1991). Byte Compilation. Version 2, Free Software Foundation, Inc. Cambridge, MA 02139, USA. *Retrieved from* https://ftp.gnu.org/old-gnu/Manuals/elisp-manual-21-2.8/html_chapter/elisp_16.html. Last visited 10th April, 2017.

Greene, W. (2002). ECONOMETRIC ANALYSIS. FIFTH EDITION, *Prentice Hall, Upper Saddle River, New Jersey 07458, 2002.*

Griesenbeck, B., and Garry G., (2007). Comparison of activity-based tour model to four-step model as a tool for metropolitan transportation planning. *Proceedings of the National Transportation Planning Applications Conference, May 6-10, Daytona Beach, Florida,* 2007.

Goulias, K. (2007). Activity based travel demand model feasibility study. *Final report submitted to the Southern California Association of Governments. Contact number 07-046-C1. Work Element Number 07-070. SCGC08. June, Solvange, CA, Retrieved from http://www.scag.ca.gov/modeling/mtf/agendas/mtf052207_4_1.pdf*. 2007.

Griesenbeck, B. and G. Garry (2007). Comparison of activity-based tour model to four-step model as a tool for metropolitan transportation planning. *Proceedings of the National Transportation Planning Applications Conference, May 6-10, Daytona Beach, Florida.* 2007.

Habib, K., Miller, E., (2006). Modelling Skeletal Components of Workers' Daily Activity Schedules. *Transportation Research Record: Journal of the Transportation Research Board, No. 1985, pp 88–97.* 2006.

Habib, K. (2007). Modelling activity generation process*es*. *Ph.D. dissertation University of Toronto, Toronto.* 2007.

Habib, K., Miller, E., and Axhausen, K., (2007). Weekly Rhythm in Joint Time Expenditure to all At-Home and Out-of-Home Activities: Application of Kuhn-Tucker Demand System Model using a Multi- Week Travel Diary Data. *87th Annual Meeting of Transportation Research Board, Washington D.C., volume 2054. pp 64-73.* 2007.

Habib, K., Carrasco, J., and Miller, E., (2008). Social context of activity scheduling: Discrete-continuous model of relationship between "with whom" and episode start time and duration. *Journal of the Transportation Research Board, volume 2076, pp 81-87*. 2008.

Habib, K. (2011). A random utility maximizing (RUM) based dynamic activity scheduling model: application in weekend activity scheduling. *Spring Link, Transportation, volume 38, issue 1, pp 123-151.* January 2011.

Habib, K. (2012). Disaggregate Econometric Choice Theory. CIV 1538: Demand Analysis. University of Toronto, Winter 2012.

Hao, J., (2009). TASHA-MATSim Integration and its Application in Emission Modelling. *Master thesis, Department of Civil Engineering, university of Toronto.*

Hägerstrand, T. (1970). What about people in regional science?. *paper of Ninth European Congress of the Regional Science Association, volume 10, pp 61-77.* 1970.

Hanemann, W. (1978). A methodological and empirical study of the recreation benefits from water quality improvement. *Ph.D. dissertation, Department of Economics Harvard University, Massachusetts*. 1978.

Henningsen, A., and Toomet, O. (2011). maxLik: A package for maximum likelihood estimation in R. *Springer Link, Computational statistics. Volume 26, Issue 3, pp 443–458.* September 2011.

Herriges, J., Kling, C., and Phaneuf, D. (2004). What's the use? Welfare estimates from revealed preference models when weak complementarity does not hold. *Journal of Environmental Economics and Management, volume 47, issue 1, pp 55-70*. 2004.

Hook, J. (2012). *Working at the Weekend: Fathers' Time with Family in the United Kingdom.* Journal of Marriage and Family, volume 4, issue 74, pages 631-642.

Hurn, S. (2009). Likelihood methods in financial econometrics. Three-day course *Economic Research Southern Africa University of Stellenbosch. Retrieved from https://econrsa.org/wkshops/LikelihoodMethods.pdf* . April 2009.

Ioesn (2017). Halle (Saale). International exchange erasmus student network. Retrieved from http://halle.esn-germany.de/halle-saale. Last visited January 12, 2017.

Islamiclandmarks.com, 2017. The City of Makkah. Last visited in October 2016. http://www.islamiclandmarks.com/makkah-haram-sharief/city-of-makkah.

Jones, P., Dix, M., Clarke, M., and Heggie, I. (1983). Understanding travel behaviour. *Aldershot, UK: Gower Publishing.* 1983.

Karlaftis, M., Konstantinos, K., Antony, S., and Dimitriou, D. (2006). Public transportation during the Athens 2004 Olympics: From planning to performance evaluation. *Presented at the 85th annual meeting of the Transportation Research Board of the National Academies, Washington, D.C,* 2006.

Kawakami, S., and Isobe, T., (1982). Development of a one-day travel-activity scheduling model for workers. *In S. Carpenters and P. Jones (ed.) Recent Advances in Travel Demand Analysis. Gower, Aldershot, pp 314-344,* 1982.

Kamakami, S., and Isobe, T., (1988). Development of activity-travel scheduling model incorporating choice process of the number, location and time allocation of activities. *Paper presented at the Oxford Conference on Travel and Transport*. 1988.

Kawakami, S., and Isobe, T., (1989). Development of travel-activity scheduling model considering time-constraints and temporal transferability test of the model. *Selected Proceedings of the 5th World Conference on Transport Research, volume 4, pp 221-233*. 1989.

Kaysi, I., Shalaby, A., Mahdi, Y., Darwish, F. (2010). Back ground material toolkit. *Center of research excellence in Hajj and Umrah at Umm Al Qura university.* 2010.

Khan, A. (2009). Historical value of the Quran and Hadith. *The Book Published by Global Vision Pub, New Delhi, India.* January 2009.

Kim, J., Allenby, G., and Rossi, P., (2002). Modelling consumer demand for variety. *Marketing Science, volume 21, pp 229-250.* 2002.

Koshizuka, T. and Kurita, O. (1986). A theoretical study on the role of a circular part of radial-circular networks. Papers of the Annual Conference of the City Planning Institute of Japan, volume 21, pages 217-222.

Malfas, M., Theodoraki E., and Houlihan, B. (2004). Impacts of the Olympic Games as mega-events. *Proceedings of the Institution of Civil Engineers (ICE), volume 157, issue 3, pp 209-220.* September 2004.

Manfred, T. (2012). *This $65 Million, Recyclable Basketball Arena Might Change The Way Countries Build Olympic Venues.* Business insider sport. Retrieved from http://www.businessinsider.com/olympics-basketball-arena-temporary-pictures-2012-7.

Maphill (2017). Savanna Style Location Map of Makkah. http://www.maphill.com/saudi-arabia/makkah/location-maps/savanna-style-map/. Last visited 22 January 2017.

May A., and Easa S. (1981). Assessing traffic management strategies in residential neighborhoods. *Transportation Research Record Journal of the Transportation Research Board.* 1981.

McFadden, D., (1975a). Economic applications of psychological choice models. *W.P. 7519, Urban Travel Demand Forecasting Project Institute of Transportation Studies, University of California Berkeley, CA.*

McFadden, D., (1975b). On independent, structure, and simultaneity in transportation demand analysis. *Urban Travel Demand Forecasting Project Institute of Transportation Studies, University of California Berkeley, CA.*

McFadden, D., and Talvitie, A., and Associates (1977). Demand Model Estimation and Validation. *Urban Travel Demand Forecasting Project Phase 1 Final Report Series, Volume V, Special report UCB-ITS-SR-77-9, the institute of transportation studies university of California Berkeley and Irvine,* June 1977.

McFadden, D., (1978). Modelling the choice of residential location. In A. Karlqvist, L. Lundqvist, F. Snickars, and J. Weibull (ed s.), Spatial Interaction Theory and Planning Modles, 75-96, North Holland: Amsterdam, 1978, Reptinted in J. Quigley (ed.), The Economics of Housing, Vol. I, 531-552, Edward Elgar: London, 1997.

Meyer, M., & Miller, E. (2001). Urban transportation planning: A decision-oriented approach. *Boston: McGraw-Hill.* 2001.

Miller, E., and Roorda, M. (2003). Prototype Model of Household Activity- Travel Scheduling. *In Transportation Research Record: Journal of the Transportation Research Board, No. 1831, TRB, National Research Council, Washington, D.C., pages 114–121.* 2001.

Miranda, M., and Fackler, P. (2002). Applied Computational Economics and Finance. *The Ohio State University and North Carolina State University, MIT Press, Cambridge, Ebooks.*

MMM group and Moriyama & Teshima architects and planners (2009). Interim Transportation Report and Workshop Summary Comprehensive Plan for Makkah, Madinah and Mashaer. October 2009.

Ministry of labor Kingdome of Saudi Arabia (2017). Official working hours. Retrieved from https://4.bp.blogspot.com/-X4vDhxktm-

g/WVz4OOaIkoI/AAAAAAAARQk/auFe2e2g_AYml0mGdxx80lSWb-XC8sICgCLcBGAs/s1600/WORKING%2BHOURS%2BAS%2BPER%2BSAUDI%2BLABOR%2BLAW.jpg" in 2017.

MoMRA (1985). Makkah Region Comprehensive Development Plan, Existing Conditions. *Volume 2: Holy Environs. Riyadh: Ministry of Municipal and Rural Affairs, Deputy Ministry of Town Planning,* 1985.

Nash, J., (2014). On Best Practice Optimization Methods in R. *Journal of Statistical Software volume 60, Issue 2.* September 2014.

Nativepakistan.com, 2012. Photos of Mecca (Makkah)-An Extremely Rare Sketch Showing Mecca City in 1850. Last visited in Jannuary 2017. http://nativepakistan.com/photos-of-mecca-makkah/.

Ophem, H., and Schram, A. (1997). Sequential and multinomial logit: A nested model. *Empirical Economics, volume 22, pp 131-152.* 1997.

Ortuzar, J., and Willumsen, L. (1994). Modelling transport. *England: John Wiley and Sons.* 1994.

Paleti, R., Copperman, R. B., and Bhat, C. R. (2010). An empirical analysis of children's after school out-of home activity-location engagement patterns and time allocation. *Transportation, volume 38, issue 2, pp 273-304.* 2010.

Pendyala, R. M., and Bhat, C. R. (2004). An exploration of the relationship between timing and duration of maintenance activities. *Transportation, volume 31, issue 4, pp 429-456.* 2004.

Pendyala, R., and Bhat, C. (2004). An exploration of the relationship between timing and duration of maintenance activities. *Transportation, volume 31, issue 4, pp 429-456.* 2004.

Phaneuf, D., and Herriges, J., (2000). Choice set definition issues in a Kuhn-Tucker model of recreation demand. *Marine Resource Economics, volume 14, pp 343*-355. 2000.

Phaneuf, D., Kling, C., and Herriges, J., (2000). Estimation and welfare calculations in a generalized corner solution model with an application to recreation demand. *The Review of Economics and Statistics, volume 82, issue 1, pp 83-92.* 2000.

Pinjari, A., and Bhat, C. (2010). A multiple discrete-continuous nested extreme value (MDCNEV) model: Formulation and application to non-worker activity time-use and timing behaviour on weekdays. *Transportation Research Part B: Methodological*, *volume 44, issue 4, pp 562-583*. 2010.

Pinjari, A., and Bhat, C. (2011). Computationally Efficient Forecasting Procedures for Kuhn–Tucker Consumer Demand Model Systems: Application to Residential Energy Consumption Analysis. *Technical paper. Department of Civil and Environmental Engineering, University of South Florida, Tampa, 2011.*

Puhe, M. and Schippl, J. (2014). User Perceptions and Attitudes on Sustainable Urban Transport among Young Adults: Findings from Copenhagen, Budapest and Karlsruhe. Journal of Environmental Policy and Planning , Volume 16, Issue 3, pages 337-357.

PTV GROUP (2017). PTV Visum Functions. Retrieved from http://vision-traffic.ptvgroup.com/en-uk/products/ptv-visum/functions/.

R manual (2015). An Introduction to R. version 3.2.2 (2015-08-14). https://cran.r-project.org/doc/manuals/r-release/R-intro.html. 2015.

Rachel, C., and Bhat, C., (2007). An Exploratory Analysis of Children's Daily Time-Use and Activity Patterns Using the Child Development Supplement (CDS) to the US Panel Study

of Income Dynamics (PSID). *Transportation Research Record, volume 2021, pp 36-44.* 2007.

Roche, M., (2000). *Mega-events and modernity: Olympics and expos in the growth of global culture*. *London, UK: Routledge.* 2000.

Roorda, M., Miller, E., and Kruchten, N. (2006). Incorporating within-household interactions into a mode choice model using a genetic algorithm for parameter estimation. *In Transportation Research Record: Journal of the Transportation Research Board, No. 1985, 85th annual meeting of transportation research board TRB, National Research Council, Washington, D.C., pp 171-179.* 2006.

Roorda, M., Miller, E., and Habib, K. (2008). Validation of TASHA: A 24-hour activity scheduling microsimulation model. *Transportation Research Part A: Policy and Practice, volume 42, issue 2, pp 360-375.* 2008.

RT Question more (2016). Saudi Arabia to move from oil, earn more from Hajj. https://www.rt.com/business/337135-saudi-hajj-revenues-oil/. Published time: 25 Mar, 2016 07:08. Last visited 7th March 2016.

Saad (2001). Preparation study for internal public transport plan in the holy city of Makkah. *high supreme of Makkah development, Saudi company for mass transit.* 2001.

Segovia, C., (2016). Destination Choice Modeling Primer. Retrieved from http://www.fsutmsonline.net/images/uploads/fsutms_training/FSUTMS_Webinar_Series_Destination_Choice_A_Modeling_Primer_20150910.pdf. Last visited May 2016.

Seyyed, N. (2005). Mecca, The Blessed, Medina, The Radiant: The Holiest Cities of Islam. *Aperture.* 2005.

Shalaby, A., Habib, N., Al-Zaharani, A., Miller, E. (2010). Activity-based Travel Demand Model for Makkah: Conceptual Framework and Survey Design. *center of research excellence of Hajj and Omrah at Umm Al Qura university,* 2010.

Sikder, S., and Pinjari, A., (2013). Spatial transferability of person-level daily activity generation and time-use models: An empirical assessment. *In Transportation Research Record: Journal of the Transportation Rresearch Board, No. 2343, Transportation Research Board of the National Academies, Washington, D.C., pages 95-104, 2013.*

Sikder, S., and Pinjari, A. (2014). The Benefits of Allowing Heteroscedastic Stochastic Distributions in Multiple Discrete-Continuous Choice Models. *Journal of Choice Modelling, Volume 9, page 39-56*. 2014.

Sola, F. (1998). The impact of mega-events. *Annals of Tourism Research, volume 25, issue 1, pp 241-245.* 1998.

SUSRIS (2015). Third Expansion for the Grand Mosque in Mecca. *Saudi-US information service.* Retrieved from http://susris.com/2015/07/13/third-expansion-for-the-grand-mosque-in-mecca/

Statistics Saudi Arabia (2014). central department of statistics and information, Retrieved from http://www.cdsi.gov.sa/english/, last visited in December 2014.

Systra (2008). *Makkah Conceptual Master plan*. presentation slides.

Taylor, J. (2011). Mecca for the rich: Islam's holiest site 'turning into Vegas. *The Independent, London, Retrieved from http://www.independent.co.uk/news/world/middle-east/mecca-for-the-rich-islams-holiest-site-turning-into-vegas-2360114.html. Last visited March 20th 2017.*

The Economist, (2010). Saudi tower: Mecca versus Las Vegas: Taller, holier and even more popular than (almost) anywhere else. *The Economist, Cairo, retrieved from http://www.economist.com/node/16438996, last visited March 20th 2017.*

Theussl, S., and Borchers, H., (2016) .CRAN Task View: Optimization and Mathematical Programming. Last visited November 2016. http://finzi.psych.upenn.edu/views/Optimization.html.

Toomet, O., Henningsen, A., Graves S., and Croissant Y., (2012). maxLik: Maximum Likelihood Estimation. *R package 1.1–2.* February 20, 2015.

Toomet, O., Henningsen, A., Graves, S., and Croissant, Y. (2015). Package ‘maxLik. Retrieved from https://cran.r-project.org/web/packages/maxLik/maxLik.pdf. Last visited 26 November 2016.

Von Haefen, R., (2003). Incorporating observed choice into the construction of welfare measures from random utility models. *Journal of Environmental Economics and Management, volume 45, issue* 2, *pp 145-165.* 2003.

Von Haefen, R., and Phaneuf, D. (2005). Kuhn-Tucker demand system approaches to nonmarket valuation. *In Applications of Simulation Methods in Environmental and Resource Economics, R. Scarpa and A.A. Alberini (Eds.), Springer.* 2005.

Vovsha, P., Freedman, J., Livshits, V., and Sun, W. (2011). Design features of activity-based models in practice: coordinated travel-regional activity modeling platform. CT-RAMP experience. *Proceedings of 90th Annual Meeting of Transportation Research Board, Washington DC, No. 2254, pp 19–27.* 2011.

Vuk, G. (2012). *An innovative approach to activity based travel demand modelling*. Proceeding from the annul transportation conference at Aalborg University. ISSN 1603-9696.

Wen, C., and Koppelman, F., (1999). An integrated model system of stop generation and tour information for the analysis of activity and travel patterns. *Paper presented at the 78th Annual Meeting of the Transportation Research Board, Record 1676, Washington D.C, pp 136–144*. 1999.

Werner, C., Engel. K., (2009). Structural equation modeling: advantages, challenges, and problems. *Introduction to Structural Equation Modeling with LISREL, Goethe University, Frankfurt*. 2009.

Xia, F., Yuan, H., and Lo, H.P. (2009). A comparison of two models for the use of leisure time. *Transportation and Geography, volume 2*, *pp* 941-950. 2009.

Yagi, S., and Mohammadian, A. (2010). An activity-based microsimulation model of travel demand in Jakarta Metropolitan Area. *Journal of Choice Modelling, volume 3, issue 1, pp 32-57*. 2010.

Yan, L.; Yang, S.; and Fu, G. (2010). Travel demand modelling for Beijing 2008 Olympic Games. *ASCE Journal of Transportation Engineering, volume 136, pp 537-544*. 2010.

Yang, L., Choudhury, C., Ben-Akiva, M., Silva, J., and Carvalho, D., (2009). Stated Preference Survey for New Smart Transport Modes and Services: Design, Pilot Study and New Revision. MIT Portugal Program Transportation Systems Focus Area. Paper#: ITS-SCUSSE-09-02. Research Project: Smart Combination of Passenger Transport Modes and Services in Urban Areas for Maximum System Sustainability and Efficiency (SCUSSE).

Zhong, M., Hunt, D., and Lu. X. (2008). Studying Differences of Household Weekday and Weekend Activities: A Duration Perspective. *In Transportation Research Record:*

Journal of the Transportation Research Board, No. 2054, Transportation Research Board of the National Academies, Washington, D.C., pp 28–36. 2008.

Appendix A

Background of the City of Makkah

This research aims to describe a conceptual modelling framework established specifically for the city of Makkah. Since each city has unique characteristics and requirements and therefore needs a specific framework designed to address them, this section will detail the history of Makkah and the religious seasons of the year. This information is key to understanding the spatial and physical aspects of the city. There will be a specific focus here on transportation problems, future land use, and transportation projects.

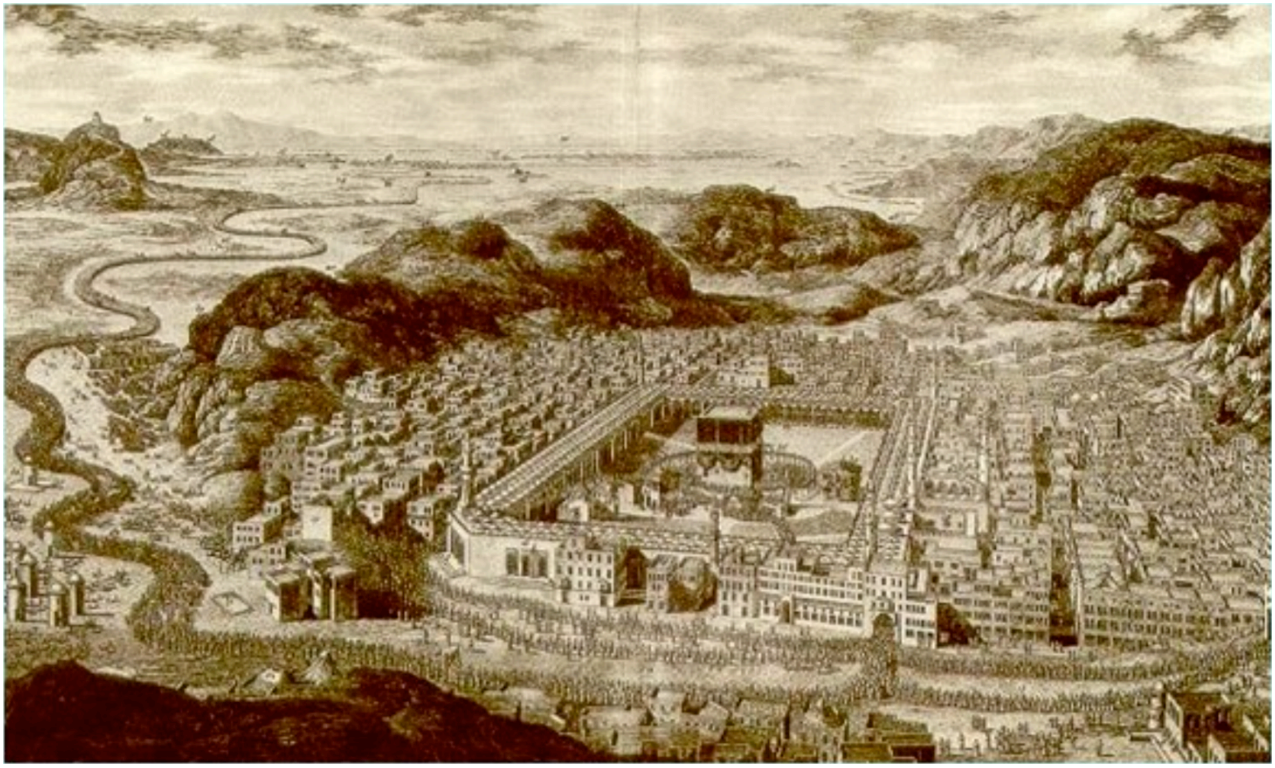


Figure 30: Old picture of the city of Makkah showing the houses surrounding Al-Haram (Nativepakistan.com, 2012).

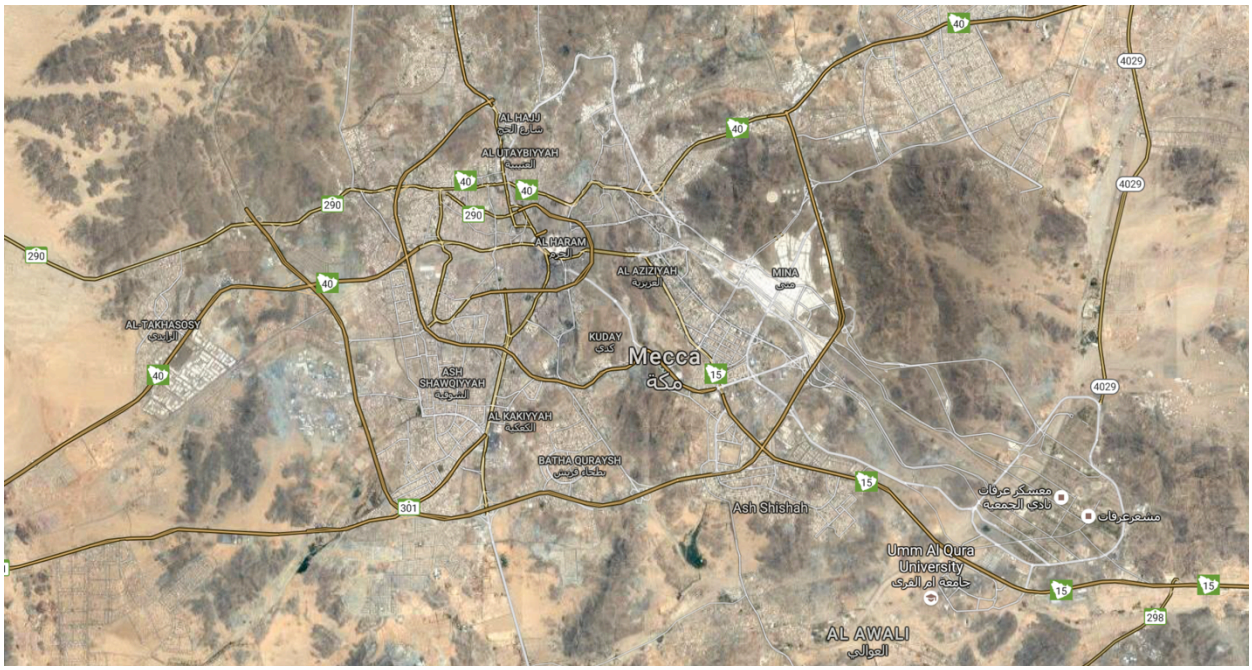


Figure 31: City of Makkah road network showing Al-Haram at the center of the city (google map October 5th, 2016).

Makkah is located in the Western part of the Kingdom of Saudi Arabia in the Hejaz region and has a mountainous topography. The first structure built in the city was Al-Qaba (Kaaba), which is a cubic building located at the lowest elevation of the city. Eventually, people started to come from all around the globe to be near Al-Qaba and build their houses around it in a circular pattern (Figure 30). As a result, the current road network in Makkah is circular (Figure 31). In 2003, Makkah's population was around 1.4 million and the area of the city was around 152 km^2; it is expected that the population will increase to 2.9 million by 2029 (Kaysi et al., 2010). The land use distribution of the city of Makkah is shown in Figure 32. The city of Makkah is changing in its topography, as the city is removing some of the mountains for an expansion of Al-Haram and for new residential areas.

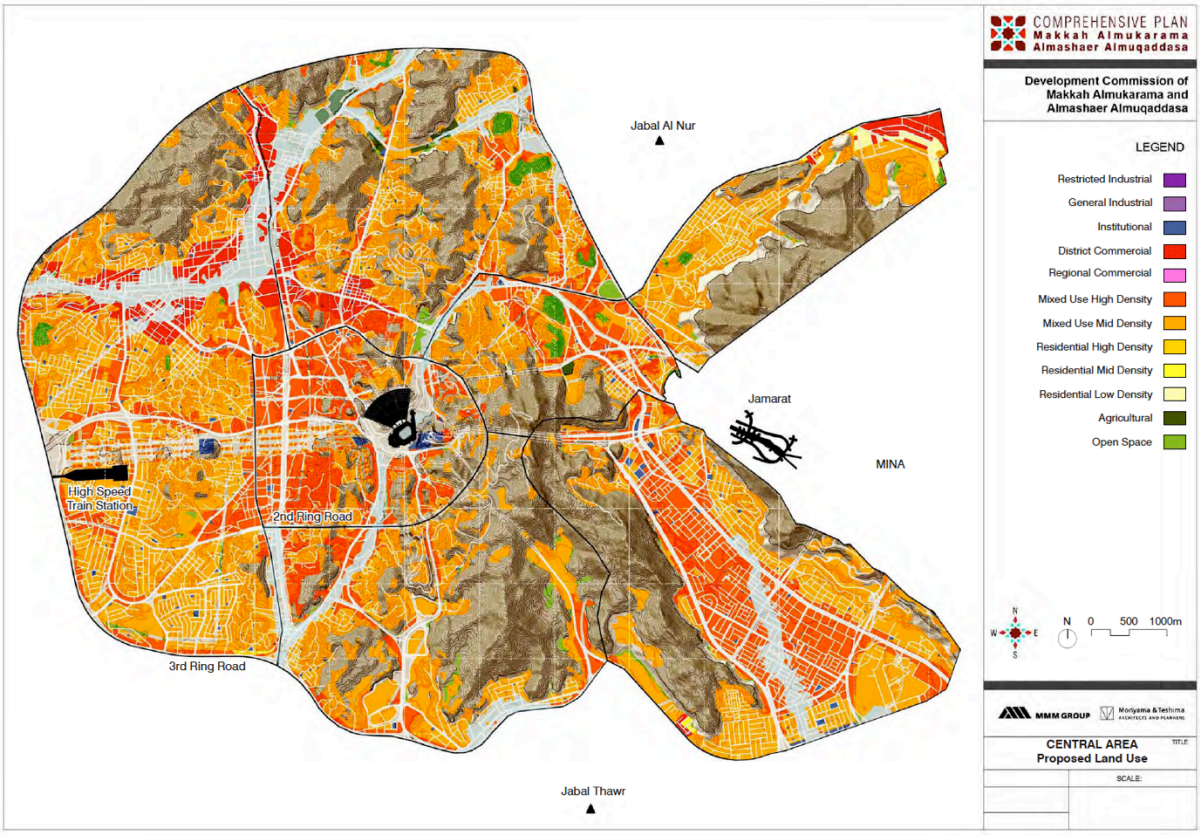


Figure 32: The city of Makkah land use and topography

(MMM group and Moriyama and Teshima architects and planners, 2009).

The municipality of Makkah defined four Islamic seasons during the year: Hajj, Ramadan, Friday, and Umrah. During these times, the travel pattern varies significantly. Hajj and Ramadan are considered to have the highest travel demand of the four seasons (Kaysi et al., 2010). Lower levels of congestion occur during other seasons such as Umrah which result in higher variability or flexibility in scheduling activities. Therefore, travel demand during Umrah can be more difficult to model. In 2001, the total daily travel demand for all modes of transportation was estimated as 10.1 million trips for Hajj, making it the largest mega-event in Makkah, 4.58 million for Umrah, 4.37 million for Friday, and 4.1 million for Ramadan. For context, although Ramadan is a much smaller mega-event than Hajj, on the twenty-sixth day of Ramadan the number of pedestrians going in and out of Al-Haram (per hour) reaches a peak of 366,000 (Kaysi et al., 2010).

Hajj includes pilgrims performing a sequence of religious activities that take place in several locations (the Holy Mosque, Arafat, Muzdalifa, and Mina) from the eighth to the twelfth of Dhu al-Hijjah, which is the twelfth month of the lunar calendar. Most pilgrims travel in groups by bus or metro, but some walk. Most of the pilgrims' movement between locations is well defined, scheduled, and organized by pilgrimage companies.

Umrah takes place in the Holy Mosque, and Muslims can perform it an unlimited number of times at any time of the day throughout almost the entire year (Table 30). It takes approximately one to two hours to perform Umrah since residents of Makkah must visit the Al-Taneem mosque, located outside Makkah, before performing Umrah. After performing Umrah, pilgrims should have their hair cut either at a barbershop or by a volunteer stationed outside Al-Haram. These key ritual components require individuals to move and travel within the city.

Ramadan is the month of fasting, where Muslims fast from sunrise to sunset for the entire month. It is the ninth month of the lunar calendar, and many people prefer to break their fasts in Al-Haram at sunset and perform Umrah during Ramadan. In addition, they pray the extra prayers in Al-Haram in the early evening and after midnight; because of this, the radial roads leading to Al-Haram are busy at almost all times during Ramadan. Throughout all of the seasons, Friday (Jema'ah) prayer is performed every Friday around noon which results in traffic congestion around Al-Haram.

Table 30: Monthly distribution of Umrah activities during the first nine months of the lunar calendar (MMM group and Moriyama and Teshima architects and planners, 2009).

1st	2nd	3rd	4th	5th	6th	7th	8th	9th	Total
187	12,007	83,395	123,303	194,580	212,921	256,236	458,067	814,165	2,154,861
0.01%	0.6%	3.9%	5.7%	9.0%	9.9%	11.9%	21.3%	37.7%	100%

People living in Makkah are classified as either residents or visitors. Both groups can be Saudi or non-Saudi with a mixed family structure. Each family member has his or her needs and activities, and for some, these activities vary based on the season. Taking paid or volunteer jobs increases the activities of some individuals. For instance, some residents use their private vehicles as taxis during the different seasons, while others work overtime to meet the increase in demand caused by visitors in almost all of the city's departments and workplaces (e.g., municipal organizations, travel agencies, and restaurants). Youths participate in some seasons, such as Hajj. They help lost pilgrims find their camp, sell food, or distribute free food. Conversely, the activities of some individuals decrease during these times, as they try to avoid congested roads.

As a result of all of these behavioural complexities, the streets, sidewalks, parking lots, and modes of transportation (i.e., cars, taxis, buses, SUVs, minivans, motorcycles) are often limited in capacity, causing serious traffic and safety problems during the high demand seasons. For example, a study by Saad (2001) observed that pedestrians share the right-of-way with cars because of limited sidewalk capacity (Figure 33). Saad also noted that private cars are preferable to buses for the residents for several reasons: bus service only exists on the main streets and never serves the local streets, the connectivity between bus lines is not good, and the headway is unstable and too long in some seasons. Making matters worse, there is no exclusive right-of-way for buses in Makkah. People sometimes park their cars at the bus stop, forcing the bus to stop further away from its station and therefore passengers to walk greater distances. (The weather in Makkah in the summer is quite hot (50-65 C^o), and so the city is a less-than-ideal walking environment). Traffic officers must change bus routes when the central area near Al-Haram is filled with people praying, and these changes prevent buses from reaching Al-Haram. Moreover, residents with private vehicles who drive people around force a competitive fare with regular bus

fares. Residents continue to provide this service despite receiving multiple traffic violations for providing an illegal service. Another key reason for individuals to drive is fuel prices; because the price is low and the fuel supply is constant, many people drive their cars and frequently, impacting traffic as well as the environment.

Figure 33: Pedestrians share the roads with vehicles during Umrah season (the picture is taken by Sindi, in January 2017).

As there are many illegal residents in Makkah, obtaining accurate records and identification can be close to impossible. It can be very difficult to stop illegal practices, such as unregulated taxi services via private cars. In addition, buses and SUVs are often loaded over capacity during certain seasons, which is an obvious safety hazard and difficult to regulate

during busy seasons. Another obstacle that overwhelms Makkah's traffic network performance are drivers who may not be familiar with the network or the traffic rules. (Of course, their familiarity will depend on the number of visits they have made to Makkah.) In addition, different traffic management plans are imposed each year by the General Department of Traffic in Makkah, such as opening and closing roads and changing their direction. Some people, especially visitors, are not aware of these plans. The busyness, violation of regulations, and unfamiliarity with the road system can make transportation quite chaotic in busy periods.

On the positive side, there are existing and ongoing projects aimed at reducing traffic congestion in the Makkah transportation network. For example, there are park-and-ride services located at various entry checkpoints, which serve as a way to reduce the quantity of traffic on Makkah streets. The Development Commission of Makkah and Masha'er, which is concerned with the development of Makkah, the Holy Mosque, Arafat, Muzdalifa, and Mina, is planning to introduce railway and metro lines, improve existing bus lines, and use Intelligent Transportation System (ITS) technology to reduce congestion on the streets (MMM group and Moriyama and Teshima architects and planners, 2009).

Many government and private organizations deal with Hajj and Umrah pilgrims. Their mission is to host pilgrims coming to perform Hajj or Umrah by providing suitable transportation, organizing their trips, offering appropriate accommodations, and providing meals. There are also groups of people or individuals with similar objectives as the larger organizations that personally organize trips during these seasons. This type of organization, large and small, can reduce confusion and streamline transportation processes by having trips and transit organized in advance.

Figure 34: The current expansion project of the Holy Haram (Islamiclandmarks.com, 2016).

Adding to the complexity of modelling Makkah's travel demands, Al-Haram has undergone many expansion projects. In 2010, the capacity of Al-Haram was 560,000 worshippers and is expected to reach 2 million worshippers in 2017 (SUSRIS, 2015) (Figure 34). The municipality has a structural plan for the coming years, as the city is working to make a comprehensive transportation and land use plan that focuses on establishing a low-density population area around Al-Haram and a high-density population are far from it. It will use a metro network to connect Al-Haram with the high-density population locations. It will implement bus rapid transit (BRT) with park-and-ride programs and ITS technology will guide drivers to available parking spots. The plan proposes that the management of traffic on the ring roads and the radial roads be done using ITS technology. In addition, a high-speed rail that will

connect with neighbouring cities (i.e., Medina and Jeddah) has recently been completed. It was designed to meet the capacity of Makkah visitors during Ramadan and Friday prayer [Kaysi et al., (2010), and MMM group and Moriyama and Teshima architects and planners (2009)]. A number of large-scale projects are also in the planning stage, but there are concerns about the development's impact on the current and future transportation system, which should be supported by a transportation system framework and models.