AMERICAN UNIVERSITY OF BEIRUT

DEVELOPING ECONOMETRIC MODELS TO FORECAST THE DEMAND FOR A SHARED-RIDE TAXI: AN APPLICATION TO AN ORGANIZATION-BASED CONTEXT

by
ZAHWA SAMI AL-AYYASH

A thesis
submitted in partial fulfillment of the requirements
for the degree of Master of Engineering
to the Department of Civil and Environmental Engineering
of the Faculty of Engineering and Architecture
at the American University of Beirut

Beirut, Lebanon

April 2015

AMERICAN UNIVERSITY OF BEIRUT

DEVELOPING ECONOMETRIC MODELS TO FORECAST THE DEMAND FOR A SHARED-RIDE TAXI: AN APPLICATION TO AN ORGANIZATION-BASED CONTEXT

by
ZAHWA SAMI AL-AYYASH

Approved by:

Dr. Maya Abou-Zeid, Assistant Professor Department of Civil and Environmental Engineering	Advisor
Dr. Isam Kaysi, Professor Department of Civil and Environmental Engineering	Member of Committee
Dr. Ibrahim Alameddine, Assistant Professor Department of Civil and Environmental Engineering	Member of Committee

Date of thesis defense: April 23, 2015

AMERICAN UNIVERSITY OF BEIRUT

THESIS, DISSERTATION, PROJECT RELEASE FORM

Student Name: Al-Ayyash Zahwa Sami
Last First Middle

⊗ Master's Thesis ○ Master's Project ○ Doctoral Dissertation

[X] I authorize the American University of Beirut to: (a) reproduce hard or electronic copies of my thesis, dissertation, or project; (b) include such copies in the archives and digital repositories of the University; and (c) make freely available such copies to third parties for research or educational purposes.

[ ] I authorize the American University of Beirut, **three years after the date of submitting my thesis, dissertation, or project,** to: (a) reproduce hard or electronic copies of it; (b) include such copies in the archives and digital repositories of the University; and (c) make freely available such copies to third parties for research or educational purposes.

Zahwa Sami Al Ayyash May 7, 2015
Signature Date

ACKNOWLEDGMENTS

First, I would like to express my deep sense of Gratitude to Professor Isam Kaysi who offered his generous assistance and encouragement throughout the course of this thesis. Foremost, I like to express my Gratitude to Professor Maya Abou-Zeid for providing me with the direction and the technical support, always. It was certainly very rewarding to me being a student for both of them as this experience enriched me on both levels- educationally and personally.

Special thanks also go to my thesis committee member Professor Ibrahim Alameddine for his insightful comments and generous feedback.

I should also acknowledge that this research would not have been possible without the Neighborhood Initiative Project. I give my sincere thanks for their contribution to this research.

Finally,

I would like to express my profound Gratitude to my mother Jihan for her unconditional love and support. Undoubtedly, she is the reason for who I am today. I also like to thank my loving brother and sister, Toufic and Alia who were always pushing me to do the best I can. I must not forget to thank my aunts Samia, Shadia, and Salwa for their endless love and support- of all kind.

I am very lucky.

AN ABSTRACT OF THE THESIS OF

Zahwa Al-Ayyash for Master of Engineering
Major: Civil and Environmental Engineering

Title: Developing Econometric Models to Forecast the Demand for a Shared-Ride Taxi: An Application to an Organization-Based Context

Traffic congestion has become a worldwide concern. One way to address this problem is to enhance the performance of the transport system by means of sound public transportation that is capable of appropriately addressing the demand of travelers, especially in highly urbanized areas of the world. The implementation of shared-ride transportation has been a viable transportation solution in many areas. Providing an efficient shared-ride transport system requires judicious planning, especially that pertaining to the study of users' travel demand and behavior. The purpose of this thesis is to evaluate the market demand potential of a Shared-Ride Taxi (SRT) service in an organization-based context. It presents a modeling framework that extends the transportation literature to include the estimation of both: discrete choice and count data models and using appropriate selection criteria, it concludes with the best fitting model, through which analysis is done. The modeling framework incorporates the level of service attributes of the Shared-Ride Taxi (e.g. fare, vehicle size, and internet availability), the socioeconomic characteristics of the users, as well as their attitudes towards ridesharing and technology. Following model estimation, one model (proving the best inference and fit) is applied to predict the SRT ridership (characterized by the percentage of students willing to use the SRT) and examine how the latter varies with variation in the values of the attributes of the new taxi service. The study involves extensive analysis of practical policy scenarios through which the impact of cost incentives (subsidies) and multiple Shared-Ride Taxi attributes on travelers' behavior is examined.

Using Stated Preference (SP) data, the evaluation sheds light on the case of the students at the American University of Beirut. Research results can be used as an initial step towards studying the market potential of a Shared-Ride Taxi in a university setting in Lebanon. Results reveal that 30% to 50% of the students are willing to utilize the Shared-Ride Taxi service under practical scenarios and that subsidies are likely to play a key role in increasing SRT ridership.

CONTENTS

Appendices

ILLUSTRATIONS

TABLES

Table Page

ABBREVIATIONS

HOV	High Occupancy Vehicle
HP	Hurdle Poisson
ICLV	Integrated Choice and Latent Variable
iid	independent and identically distributed
L.L.	Lebanese Lira
LL	Log-Likelihood
LL_m	Log-Likelihood of the fitted model
LL_0	Log-Likelihood of the null model
LRT	Likelihood Ratio Test
ML	Maximum Likelihood
MNL	Multinomial Logit
NI	Neighborhood Initiative
PC	Private Car
PT	Public Transport
R^2	R-squared
RE	Random Effect
RP	Revealed Preference
S.E.	Standard Error
SP	Stated Preference
SRT	Shared-Ride Taxi
UC	University of California
VOT	Value of Time
ZINB	Zero-Inflated Negative Binomial
ZIP	Zero-Inflated Poisson

To my father, Sami...

CHAPTER 1

INTRODUCTION

1.1. Study Motivation

Road traffic congestion is a serious growing problem. Large segments of the population in many urban areas of the world are in need for adequate urban mobility that would cater for the diversity of their activities. Since the beginning of the 21st century, auto ownership continues to increase dramatically, deteriorating the environment as well as the quality of urbanized life that it brings forth. A study by Dargay et al. (2007) on 45 countries which encompass 75% of the world's population shows that the total vehicle stock will increase from 800 million in 2002 to more than 2 billion in 2030. By that time, China will increase its vehicle stock from 16 to 269 per 1000 people (Dargay et al., 2007). In 2014, statistics in U.S., Germany, China, India, and Brazil proved that over 80% of Generation Y consumers (age 20 to 37 years) in these countries are interested in owning or leasing private cars in the coming 5 years (Deloitte, 2014) (See Figure 1 below). As such, it seems there is a general tendency for travelers to rely on private vehicles rather than public transportation and non-motorized modes because owning a car is no longer seen as a luxury, but a right and a need. The number of cars owned is rising as populations expand and national incomes increase. Buliung et al. (2009) assert that the use of cars for commuting along with the environmental costs of car ownership has reached exceptionally high levels in Canada and the United States. The increasing number of private cars can result in a wide range of environmental problems that is accompanied by smog and toxic emissions affecting public health. It also leads to drivers' frustration, aggressive behavior, and consequently increases accident rates. Increased number of private cars also leads to lost times on the roads and less productive work force.

To provide an efficient transport system, judicious management strategies are inevitable, especially those pertaining to the increase in the effective capacity of roadways or the reduction of travel demand on them. One way of improving the system's performance is by supply management which works on ameliorating the efficiency of existing infrastructure or adding more capacity to it. While providing a variety of mobility requisites to travelers, demand management strategies on the contrary, aim at controlling or eliminating vehicle trips on the system. According to the Federal Highway Administration, there are three major management classes which are more likely to follow present practices and techniques. See Figure 2.

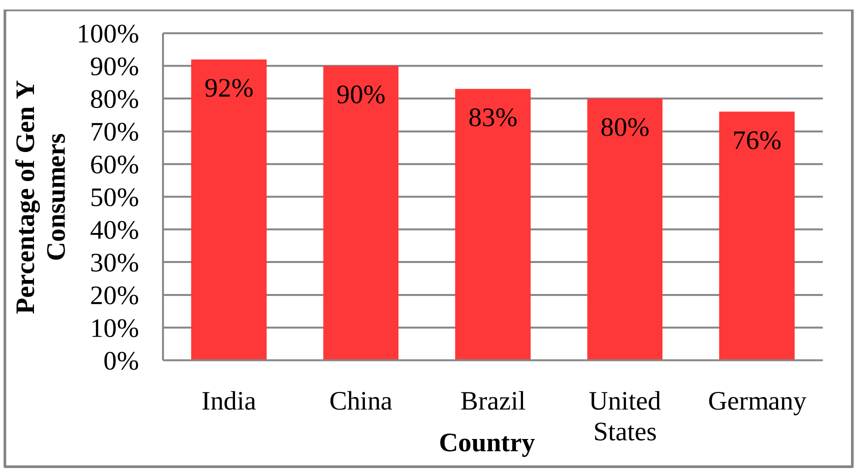

Figure 1: Generation Y Consumers Expecting to Buy a Car in the Next 5 years (in %) (Deloitte, 2014)

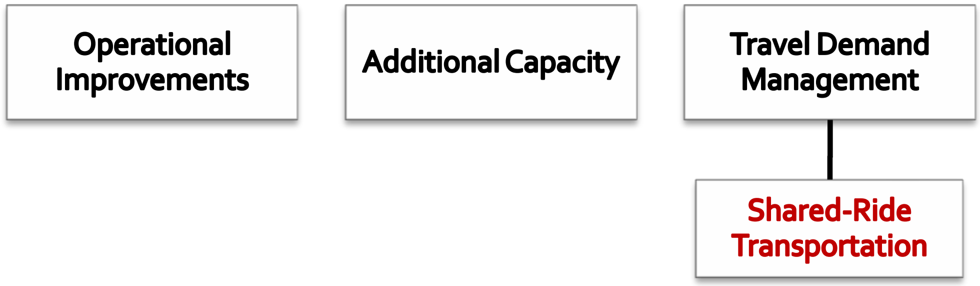

Figure 2: Major Traffic Improvement Strategies

Operational improvements incorporate the upgrade of traffic engineering practices or facilities such as introducing an improvement to the transit system operations, enhancing intersection geometrics, as well as improving traffic control methods (Urbanik, 1998). The second option is to consider building new infrastructure to serve as a capacity improvement. This may include additional lanes on freeways, building new toll roads as well as carrying out improvements to the public transport system by introducing new bus routes or adding service on existing lines.

Another important management strategy is that related to the demand side. Demand Management Strategies involve managing the demand of highway travel. For instance, it includes congestion pricing, shifting the time of travel through providing flexible work schedules, eliminating the need for travel (e.g., telecommuting), or putting more travelers into fewer cars (e.g., shared-ride transportation).

1.2. Shared-Ride Transportation

The implementation of shared-ride transportation ever since the 1940's was a viable transportation solution in many areas of the world. Shared-Ride transport modes are a form of public transportation that involves several schemes of multiple-occupant vehicles such as ridesharing, jitneys, and demand responsive transit (DRT) which in turn encompasses a range of modes such as paratransit, Dial-A-Ride Transit (DART), and shared-ride taxis.

Ridesharing refers to carpooling and vanpooling. It is a travel arrangement whereby two or more individuals share the ride in one vehicle, sharing the driving and/or the operating costs of the trip. Carpooling involves using the private cars of participants while vanpooling users use rented vans that are often provided by non-governmental organizations, employers, or government agencies and whose costs are divided among vanpoolers. Carpooling is the most common and wide-spread form of ridesharing, despite being singled out as one of the most difficult forms of mode choice to achieve (Soltys and Buliung, 2008). Conceptually, ridesharing practices aim at relieving peak-period traffic congestion through group-riding approaches, which seek to minimize the number of vehicles used for traveling, by means of high vehicle occupancy factors.

Ridesharing and other forms of shared-ride transport differ at many levels, including user schedule flexibility, vehicle ownership, payment, and operation. Jitneys are shared taxis which operate on an ad-hoc basis. They can pick-up or drop-off passengers anywhere along their route. Paratransit and Dial-A-Ride Transit are a form of public transportation which falls on the continuum between the fixed-route public transportation systems and the private car (Scalici, 1985). These systems are designed to target travelers with special needs such as the elderly and the disabled. They are flexible, cost-effective, and provide a door-to-door service. Furthermore, they are suited to low-density areas where conventional bus service is not as viable since their service is restricted to a defined zone in which they operate. Likewise, DRTs are small or medium vehicles, characterized by flexible routing and scheduling schemes, operating in shared-ride mode. Despite that DRTs, paratransit, and DART systems are often used interchangeably, paratransit and DART tend to characterize modes that serve people with special needs, whereas DRTs refer to modes which are available to the public. As for funding, shared-ride transportation may be operated by different types of organizations be it private companies for commercial reasons or the public sector (i.e., public transport companies).

Lately, there appears to be a need to adequately cater for the growing shared-ride transportation demand of users through promoting real-time services. Real-time shared-ride transportation uses GIS and global positioning system technologies on “smart phones” which are internet-enabled in order to arrange the rides in real time, only minutes before the trip takes place (Chan and Shaheen, 2012). Drivers (who may be individual travelers seeking to match a trip with another traveler, or a ride-matching software administered by the operator) post their trip as they drive, and potential riders request rides. Automatically, the software assigns travelers to vehicles and sends notifications for riders and drivers using their smart phones.

The SMART 2020 report claims that the emissions in the United States in 2020 will be reduced from 190 to 70 million metric tons of carbon dioxide if the authorities in charge plan to employ the information and communication technology solutions to optimize road transport (Global e-Sustainability Initiative, 2008). Moreover, Eggers and Jaffe (2013) claim that currently there are 13.5 million users of rideshare modes in the United States. If twice as many travelers use ridesharing, the annual driving cost savings (registration fees, fuel, maintenance, tires, insurance, and depreciation) for those commuters would increase from $55 million to $114 billion, a value equivalent to the GDP of Bangladesh. Further, the annual savings in travel time due to the reduced congestion would double to 748 million hours, a duration that is equivalent to the average lifespan of 1000 people.

1.3. What is a Shared-Ride Taxi?

As opposed to taxis that serve an individual or a group of individuals with a single origin and destination at a time, a Shared-Ride Taxi (SRT) is a public transport mode that enables two or more individuals to be served simultaneously and to share the cost of the trip, based on spatial and temporal matching. The service is often demand-responsive, door-to-door, and requires advance reservation.

As cities become more crowded and polluted, the business of taxi-sharing is catching on with services such as Uber and Lyft, founded in San Francisco in 2009 and 2012, respectively. These services let passengers link up with taxi drivers based on their origins and destinations. Today, Uber is available in over 270 cities world-wide (Auchard and Steitz, 2015). Also, in 2009, Bandwagon was launched, a service that allows riders to find one another in one of two ways: potential travelers can either search for people nearby who are going in their way in order to match up, or they can order a car and see if someone joins before the time of departure (Bandwagon, 2014).

Figure 3 which mimics the communication scheme described and illustrated by Suen et al. (1981) shows main communication links that represent the flow between the participants, the customer (traveler), the call-taker, the dispatching system, and the driver. Link 1 takes the form of the requests coming from the public requesting a trip. Links 2 and 3 represent the dispatch system's function that corresponds to the controlling of the communication between the call-taker and the taxi drivers. Often, the call-taker and the dispatch system both correspond to a wireless receiver and transmitter providing the basis for the interaction taking place in the ride-matching system. Link 4 represents a tentative link between the computer dispatching system and the traveler, representing the notifications sent to travelers after requesting a trip (Suen et al., 1981).

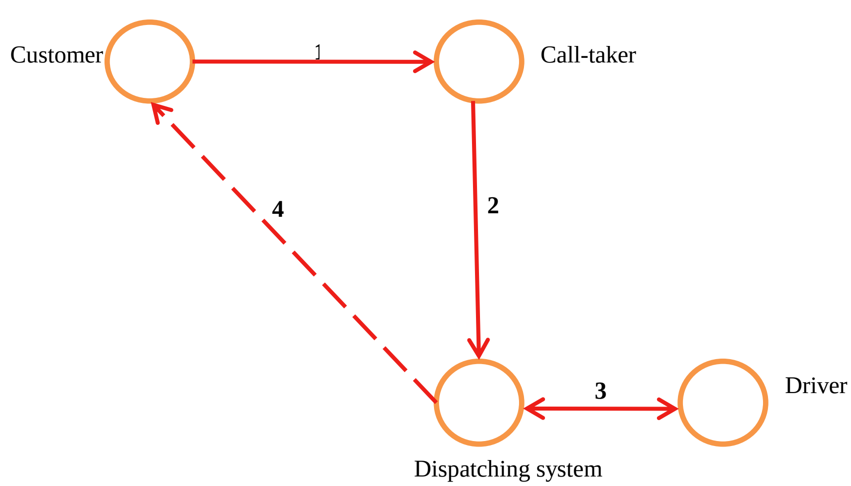

Figure 3: Shared-Ride Taxi Communication Network

Shared-Ride Taxi services unlike standard single occupancy vehicles aim to make the best use of resources. “The strongest incentive is the time incentive,” said David Mahfouda, the CEO of Bandwagon (Casey, 2014). “A lot of people think it’s an inconvenience to share rides. But there are many situations, particularly when there is more demand for taxis than supply, when it’s faster to get into a vehicle with another person.”, he said (Casey, 2014). In addition to the reduced service travel time, SRTs have many advantages on the individual level. Shared-Ride Taxi travelers experience cost savings because of the shared travel costs, as well as reduced commute stress. Also, travelers may experience travel time savings resulting from utilizing high-occupancy vehicle (HOV) lanes.

In spite of the many benefits that are perceived at the individual level, SRTs also claim advantages at the societal level and that often outweigh individual benefits. Most essentially, they reduce the number of automobiles used by travelers. As a result, reduction in energy consumption is achieved leading to fewer emissions and lower traffic congestion levels as well as reduced parking infrastructure demand.

1.5. Contributions

Travel demand modeling is a fundamental tool in the process of planning the transportation systems as it lets planners inspect the future of these systems. Besides, demand modeling allows them to undertake evaluation strategies that are the most efficient tools for solving encountered problems proactively. It is therefore a prerequisite to the provision of new transport alternatives and an essential step in the process of the systems' feasibility and design.

This thesis contributes to and develops further the literature on transportation demand modeling for shared-ride transportation. It introduces an econometric modeling framework including various Count and Discrete Choice Models such as Multinomial Logit, Poisson Regression, Negative Binomial Regression, Zero-Inflated Poisson and Negative Binomial Regressions, and Hurdle Poisson and Negative Binomial Regressions. The framework provides appropriate ways to model excess zeros in the observations. The models explore the potential of a Shared-Ride Taxi, examining the impact of different taxi attributes such as travel cost, in-vehicle time and the time window for pickup and delivery as well as travelers' attitudes toward taxi-sharing, ridesharing, and technology. The application of this modeling framework utilizes a stated preference (SP) survey which aims to reveal how the respondents' choices change with changes in the attributes of the shared-ride transportation service. Furthermore, the study contributes to the better understanding of what factors are vital for the success of alternative services or the improvement of existing ones, e.g. to what extent travelers are bothered by the low level of service and how elastic the shares of existing/new modes are to changes in the attributes of the suggested services. The thesis also offers parties in charge needed information about the system potential success or failure.

In spite of the various studies in the literature on shared-ride transportation and its demand, the research at hand is innovative because it investigates the potential ridership levels at an organization-based level, the American University of Beirut. The study contributes to the literature by evaluating the travel patterns in a university setting and presents an approach that can be followed by other universities or large institutions with similar characteristics. The outcomes are represented by the identification of the potential target groups among travelers willing to shift to the new SRT service. The methods can also be employed by other large-scale employers who might be interested in the adoption of SRTs and thereby offering their employees a stress-free commute experience. In Greater Beirut, where public transport is in deficit, this system might grant travelers a constructive way out of the stressful commute experience.

1.4. Methodology

In order to understand the market response to the provision of a Shared-Ride Taxi service in an organization-based context, this study estimates and analyzes Discrete Choice Models as well Count Models. Such demand models give insight about travelers' preferences towards the use of the new system through determining what service attributes is important to these travelers. Taxi attributes include travel cost and travel time, the time window for pick-up and delivery, vehicle size, internet and technology utilization, and individual characteristics and attitudes. Following model estimation, one model (proving the best inference and fit) is applied to predict the SRT ridership and examine how the latter varies with variation in the values of the attributes of the new taxi service.

The methodology presented in Chapter 3 is applied to the university students of American University of Beirut (AUB) in Greater Beirut Area (GBA). The effort presented in Chapter 4 is part of the Congestion Studies supported by the Neighborhood Initiative (NI) at the American University of Beirut (AUB); research previously supported by the NI includes that reported by Danaf et al. (2014) and Aoun et al. (2013). The case study targets the students of AUB and examines whether they would utilize a SRT service if it were available to them. Through an online survey, AUB students were asked questions about their current travel patterns as well as their hypothetical behavior towards the new SRT service using an SP survey consisting of 8 hypothetical choice experiments. Several models are developed including an integrated choice and latent variable model (ICLV), capturing the effect of multiple service attributes on AUB students' travel behavior (represented by the number of times they wish to utilize the SRT weekly). Analysis in Chapter 4 encompasses the following:

1. Identifying what service attributes students value most when selecting among the travel options available (including whether or not to use a SRT and how many weekly trips they decide to make),
2. Showing if there is significant heterogeneity among the users of the different modes examined. This is achieved through considering distinct data sets and estimating models accordingly, pertaining to different users (Public transport and Private cars). This makes clear whether or not students exhibit similar sensitivities to changes in the level of service of the new taxi service,
3. Defining policy scenarios that serve as a decision support tool for the provision of an SRT service in an organization-based context; this section includes defining multiple practical policy scenarios, predicting ridership (in terms of percentage of

students willing to use the SRT service 0, 1, 2, 3, 4, or 5 times per week) using each of the scenarios, and revealing what service attributes have substantial effects on the SRT ridership and what trades-offs can be made among them,

4. Establishing what Value of Time (VOT) is associated with AUB students as well as each of the different mode users considered,
5. Understanding the effect of the different service attributes along with travelers' attitudes on the SRT participation levels (i.e., through elasticity estimation).

1.6. Thesis Outline

The thesis is made up of five chapters in total and is structured as follows:

Chapter 2 presents a literature review on shared-ride transportation: how and when they emerged, the reasons of their successes and failures, a background on previous modeling frameworks implemented as well as various organization-based applications in the context of shared-ride transportation. The third chapter develops the modeling methodology and describes the data collection procedure. The modeling formulation section consists of discrete choice modeling and count data modeling methods. The last section of Chapter 3 discusses the selection criteria used to choose the best fitting model among the competing models. In Chapter 4, the methods are applied to an organization-based context, the students at American University of Beirut in Lebanon. As for the last chapter of the thesis, Chapter 5, it provides a summary of the AUB case study as well as a review of the contributions and limitations encountered throughout the process of modeling. It also provides the reader with directions for future research.

CHAPTER 2

LITERATURE REVIEW

2.1. Introduction

Shared-ride transportation services have started to become a mainstream public transport mode. DRT’s have been implemented in many areas in the world and especially in low demand areas (Enoch et al., 2006) or in relatively small areas (Takeuchi at al., 2003). In order to respond to travelers’ requests, DRTs and other shared-ride transportation services operate based on routes and timetables that may be fixed or flexible. Nowadays, such services have scheduling methods that are automated and require advance reservation. Also, there is a wide range of communication schemes including telephone and web interfaces that are provided to deploy these services.

This chapter of the thesis gives a literature review on shared-ride transportation services. The following section provides a historical overview since 1940’s until the present day. Section 2.3 discusses some successful and failing attempts of shared-ride transportation methods and pinpoints major drawbacks of the unsuccessful ones. Then, in section 2.4 the application of demand modeling in various transportation contexts including shared-ride transportation is examined. The last section of this chapter investigates several case studies in an organization-based context.

2.2. How Did Shared-Ride Transportation Services Evolve?

Shared-ride transportation services started in the early 1900's in North America as taxi-share services where trips were assigned along fixed routes (Takeuchi at al., 2013). Afterwards, according to Chan and Shaheen (2012), the following five phases took place:

2.2.1. First Phase (1942 to 1945)

As a method for conserving energy, governments during World War II encouraged ridesharing. In July 1941, a campaign was launched in the United States with a 250,000$ financial plan for advertisement; recommending lowering driving speeds, taking proper care of tires as well as sharing rides (Amey, 2010). This campaign along with other energy-conservation promotions during World War II focused on broadcasting animated posters and newspaper ads (See Figure 4). Nevertheless, there was not enough information on the success of these initiatives (Amey, 2010).

Figure 4: World War II Posters Promoting Ridesharing (Source: Ridebuzz Ridesharing and Carpooling)

2.2.2. Second Phase (1960's to 1980)

Chan and Shaheen (2012) refer to many efforts that aimed at promoting rideshares in 1960's up to 1970's and which emerged as a result of the energy crisis and the Arab oil embargo in the 1970's. Initiatives included employer-sponsored ride-matching programs, carpooling, vanpooling, HOV lanes and park-and-ride facilities.

In 1973, to minimize the energy consumption, US federal agencies became interested in investing in the employer-sponsored commuter ridesharing programs. The Federal Highway Administration conducted nationwide surveys of ridesharing programs and found a significant increase in shared-rides as well as a decrease in 23% in the vehicle-miles traveled among 197,000 employees (Chan and Shaheen, 2012). As a result of this success, 106 carpool demonstration programs were provide in 96 US metropolitan areas (Chan and Shaheen, 2012).

On the other hand, other shared-ride transportation systems emerged in North America during the 1960's and 1970's such as Shared-Ride Taxis that operated from residential areas to airports and stations and Dial-A-Ride systems which were concerned with the provision of enhanced mobility for the elderly and disabled (Takeuchi at al., 2013).

In 1970s, shared-ride transportation evolved in Japan in the form of a "Demand Bus" where routes and timetables were defined responding to each user's demand (Takeuchi at al., 2013). In Southeast Asia, paratransit services became more prominent during that era operating along fixed routes with some deviations (Takeuchi at al., 2013).

2.2.3. Third Phase (1980 to 1997)

After 1980s, the energy conservation era ceased and transportation demand management strategies began to target high congestion levels and environmental issues. Technology-based ride-matching programs began to take place, paving the way for the dynamic ridesharing systems. Nevertheless, as oil prices decreased, ridesharing lost much of its competitiveness (Chan and Shaheen, 2012). Yet, many of these systems form now the basis for many of the present shared-ride transportation systems as depicted by Chan and Shaheen (2012).

With respect to Europe, by the beginning of the 1990's, DRTs began to utilize advanced information and telecommunication technologies with flexible routes and timetables responding to users' demand. This had paved the way for the use of new transportation services with enhanced application of technologies including geographic information systems and information management systems (Benjamin et al., 1998).

2.2.4. Fourth Phase (1999 to 2004)

During this phase, services focused on fostering a bigger market especially that the ride-matching systems in the 1980's and 1990's did not provide the "critical mass" needed to achieve enough matches between users (Chan and Shaheen, 2012). The focus on mitigating traffic congestion continued in this phase. Also, online ride-matching services as well as traveler information services proliferated in North America (Chan and Shaheen, 2012) and DRTs continued to rise in Europe (Takeuchi at al., 2003).

2.2.5. Fifth Phase (2004 to present)

The last phase describes the current state of ridesharing and other shared-ride transportation methods, which are technology-enabled. As was present in the previous period, HOV lanes and park-and-ride ridesharing programs continue to operate (Chan and Shaheen, 2012). The key development occurring is the prevalent integration of the internet, mobile phones as well as social networking. Most ridesharing services employ online websites that handle communication among users and the matching interface (Chan and Shaheen, 2012). Also, social networking platforms have emerged targeting the youth; this has permitted shared-ride transportation agents to use the social network medium to match potential rides between users, effortlessly. Currently in the US, there is an annual dial-a-ride growth that exceeds 5%, a value which is anticipated to increase in the coming ten years (Markovic et al., 2013).

2.3. Successes and Failures of Shared-Ride Transportation Services

In this section, some success and failing stories in shared-ride transportation are highlighted. Success or failure of these systems is determined as per the number of rides matched and executed, that is, the number of rides completed after a match is established between the passenger and driver in the system (Siddiqi and Buliung, 2013).

Many studies in the literature brought attention to the reasons hampering the development of shared-ride transportation services. Enoch et al. (2006) show that the barriers are regulatory, fiscal, institutional and cultural, occurring at the level of the government, the operator, and the users. For example, after a two-year trial of a dial-a-bus service in Milton Keynes, UK, the system was then integrated with conventional bus services in the town due to the inflexible and the high-cost operator, the lack of political commitment, and the low fares that were not reflective of the true quality of the DRT. According to Enoch et al. (2006), DRT projects in the UK aborted due to marketing reasons; the design and the market research were based on Dutch data which were not reflective of the population in the UK. Additionally, operational problems arose due to the opposition from local taxi operators who perceived the new system as a threat to the town's existing public transport. Another problem, as depicted by Enoch et al. (2006), was the difficulty to convey the nature of the services to many users and especially the elderly and disabled. Furthermore, it was sometimes the weak coordination between operators and local authorities that impaired the development of such systems. Siddiqi and Buliung (2013) claim that there appear to be only few success stories; reasons for the systems' withdrawal include high capital and operating costs, poor service levels, technological limitations, limited ridership, and usability.

From this it can be seen that a successful shared-ride transportation system needs to cater for the market that it is intended to serve. It should have a realistic design and costing structure, which in turn ensures the cost-effectiveness of the system. Buliung et al. (2009) emphasize on the importance of a range of spatial, temporal as well as personal characteristics such as age and gender for the successful formation and use of ridesharing. They also emphasize on the significance of the technology-based arrangement which could foster the virtual forum through which the communication can happen.

2.4. Demand Modeling

2.4.1. Introduction

Demand modeling has been a key element in the assessment of transportation systems, whereby econometric models are often developed to investigate the effect of explanatory variables on the dependent variable modeled. Modeling methods and structures have been continuously improving and are being applied in many transportation applications. This section of the literature review will shed light on the different types of model structures that could be employed in the context of modeling the demand of a Shared-Ride Taxi in an organization-based context. The dependent variable represents by the number of weekly trips done by travelers, i.e. 0, 1, 2, 3, 4, or 5 trips.

2.4.2. Trip Frequency Modeling

In the literature of modeling trip frequency generation, studies utilized several models structures. Often, the traditional trip generation technique, linear regression, is used to model trip frequency (Barmby and Doornik, 1989; Kim and Susilo, 2013). Another common technique is to use Count Models such as Poisson and Negative Binomial. Barmby and Doornik (1989) conclude in their regression-based analysis that the Poisson distribution and the Negative Binomial distribution can be usefully employed in constructing a statistical model of trip frequency. In another study, Kim and Susilo (2013) aimed at finding the best model for pedestrian commuter trips in a metropolitan area. This study described the testing of linear regression as well as Poisson and Negative Binomial, and which formed the focus of their comparison. After estimating the models, results including parameter estimates and goodness of fit measures confirmed that Poisson and Negative Binomial regressions are the most appropriate for pedestrian trip generation. Furthermore, Gurmu and Trivedi (1996) developed a modeling approach for a count dataset for recreational boating trips using different count models including Zero-inflated Poisson, Hurdle Poisson, and Hurdle Negative Binomial. Analysis showed that the latter is the most satisfactory of all models considered.

2.4.3. Vehicle Ownership Modeling

In another framework, some research modeled the level of vehicle ownership (e.g., 0, 1 car, 2 cars, 2 or more cars) as a function of urban form and socioeconomic characteristics of the household. Potoglou and Susilo (2008) explained the different types of disaggregate models that are developed for that purpose. They differentiated two types of discrete choice modeling structures: the ordered and the unordered models. Ordered models include ordered Probit and ordered Logit. Multinomial Logit (MNL) and Probit models represent the unordered response models, which are based on the random utility maximization method under which the individual associates utility values to all auto ownership levels and chooses the alternative that maximizes the utility (Potoglou and Susilo, 2008). Comparing the two model structures in the context of vehicle ownership, results of three studies by Bhat and Pulugurta (1998), Potoglou and Kanaroglou (2006), and Potoglou and Susilo (2008) argued that MNL model provides a significantly improved fit over the ordered Logit model. Other model structures are also employed in this model framework such as linear regression and count models (e.g. Zhao and Kockelman, 2002; Gopisetty and Srinivasan, 2013).

2.4.4. Modeling Demand for Shared-Ride Transportation

2.4.4.1. Logit-Based Models

Logit models have been used in many transportation demand modeling contexts such as paratransit, advanced traveler information systems, mass transit technologies as well as various contexts in shared-ride transportation including carpooling and vanpooling, DRTs, paratransit, etc. A binary outcome in that context is defined as whether the respondent wants to use the targeted transport service or no. Using an SP survey, a feasibility study on DRTs in low-demand areas was undertaken by Takeuchi at al. (2003) considering a binary choice where students chose between private auto and DRT. According to this study, results of the binary logit model showed that users prefer shorter waiting time and shorter in-vehicle time. Another MNL model was developed by Benjamin et al. (1998) to forecast the ridership of paratransit services in Winston-Salem, North Carolina. The revealed preference (RP) data included information about the most recent trip made by the respondent and the SP data involved a choice scenario between the chosen mode (in the RP context) and four other alternatives including the existing bus service, a bus-route deviation service, the existing dial-a-ride service, and a dial-a-ride feeder service. A Multinomial Logit model was developed to forecast travelers' response to the introduction of the new services. The results showed that improved reservation schemes for the proposed DRT produced shifts in modal shares and that the demand for Dial-A-Ride service increases with the increase in user awareness. In fact, modeling demand using MNL is very common in the context of public transportation. Many other research studies involved Logit models in the context of mode choice including Ryley et al. (2014), Silvis and Niemeier (2009), and Buliung et al. (2009).

2.4.4.2. Count Models

On the other hand, a lot of studies employing count models appear in the literature of modeling the demand for shared-ride transportation and other transportation engineering contexts. The very popular regressions of this family are the Poisson and Negative Binomial Regressions. In a study on the impact of advanced public transportation systems on travel by Dial-A-Ride service, Ben-Akiva et al. (1996) estimated the parameters of an RP and SP Poisson Regression Model as well as an ordered Probit model. The former established a link between the count data variable (number of trips in a week) and the explanatory variables. They concluded that age, difficulties in walking, and employment status are main factors that capture users' readiness to use Dial-A-Ride service. Such framework, in fact, is common in the literature of count models in transportation applications. Benjamin and Price (2006) estimated Poisson and Negative Binomial Models that provided econometric estimates of the change in the number of trips as a function of the travel attributes influenced by the implementation of a demand-responsive mini-bus service for the elderly and disabled in North Carolina. Results indicated that trip length, number of stops, and physical comfort are important service attributes which significantly impact the number of trips chosen.

A very dominant feature of the count datasets is the presence of excess in the zero observations. In many of such cases, the data generating process leads to larger number of zero observations than would be predicted in a standard count regression. This feature may be accounted for by over-dispersion in the data resulting from unobserved heterogeneity (Gurmu and Trivedi, 1996). Studies in the context of count modeling where zero-inflation and/or over-dispersion exists are numerous (e.g. Gurmu and Trivedi, 1996; Kibria, 2006; Tait et. al, 2012; Boucher and Guillén, 2009; Hall, 2000; Lawal, 2012; Zorn, 1996). In such studies, count models like Poisson and Negative Binomial as well as those that account for the excess in the zero observation (i.e. Hurdle and Zero-Inflated models) were fitted. Then, an assessment of the performance of the models was provided and the best fitted model was concluded accordingly. Some examples of these research studies are included in the summary table below (Table 1).

2.4.4.3. Impact of Personal Attitudes on Travelers' Decisions

Other studies have examined how psychological factors influence people's decisions and choices. According to Danthurebandara et al. (2013), these factors are individual-specific and include attitudes, lifestyle, values, and perceptions. Such latent variables have been integrated with choice models in studies such as those presented by Ben-Akiva et al. (2002), Danthurebandara et al. (2013), Paulssen et al. (2013), and Temme et al. (2008). The key advantage of these models is that they assume heterogeneity among the respondents through incorporating latent variables which decrease the unexplained parts of the heterogeneity. The general framework of the model consists of two models; a choice model where the utilities of the available alternatives are specified as a function of observed variables (e.g. alternative's attributes and individual's characteristics) and latent variables, such as attitudes (Temme et al., 2008), and a latent variable model that is made up of a measurement model and a structural model. Using an integrated choice and latent variable model, Paulssen et al. (2013) argued that personal values affect people's attitudes towards different alternative attributes, which would in turn influence mode choice. In the shared-ride transportation context, some studies examined how abstract motivations such as attitudes affect traveler behavior and decisions. As cited in Chan and Shaheen (2012), a study showed that travelers see attractiveness in carpooling but refuse to carpool because it does not provide them with equivalent flexibility and convenience as a private car. Bonsall et al. (1984) revealed that there are psychological factors that influence the attractiveness of ridesharing to users e.g., desire for personal space, time, and security, and "sociability"/"having company on the journey". Also, the findings from the studies of Ben-Akiva et al. (2002) and Danthurebandara et al. (2013) have demonstrated that the implementation of the integrated choice and latent variable model framework results in an improved goodness of fit over choice models without latent variables and that these models perform very well with regard to prediction.

2.5. Organization-Based Case Studies of Shared-Ride Transportation

While there are numerous cases of ridesharing services that were open to the public use, fewer efforts focused on the provision of organization-based shared-ride transportation. A lot of research has focused on modeling the demand for carpooling services on an organizational-based level, especially for university students such as

Erdoğan et al. (2015) who studied the demand for carpooling and vanpooling in UMD University, Baltimore, Washington, D.C.

Others studied dynamic ridesharing in similar contexts without developing econometric demand models. Amey (2010) proposed a design for a technology-focused rideshare trial for faculty, staff, and students at the Massachusetts Institute of Technology, focusing on the importance of incentives and personalized marketing to overcome the "rideshare challenge". As cited by Siddiqi and Buliung (2013), examples of shared-ride transportation services in that context include: 1) Smart Traveler Program sponsored by the University of Washington and restricted to the university's faculty and students, 2) Bellevue Smart Traveler which was restricted to the employees of Bellevue working in the downtown location, and 3) Facebook-user driven goCarShare service in Edinburgh, Scotland which was open to the public but targeting Facebook users. While some of such types of services still operate until this day, many discontinued due to inefficient costing regime, poor level of service, and technological limitations. A study by Aoun et al. (2013) examined possible ways that make high-income users use high-occupancy modes rather their private cars. Google employees in San Francisco benefit from 32 free of charge Google shuttle buses (Aoun et al., 2013), which are operated by Bauer's Limousine, a private transportation company in San Francisco (Helft, 2007). As cited in Aoun et al. (2013), Google employees can track the vehicles in real-time; delays and updates are communicated to them through cell phones or emails. Another shared-ride transportation service is the focus of a study by Deakin et al. (2010). In their research, the authors assessed the potential of a dynamic ridesharing service for travel to downtown Berkeley, California, and the University of California (UC), Berkeley campus. The survey targeted UC graduate students, UC faculty and staff, and selected downtown Berkeley employers for distribution to their employees, as well as other office and retail employees in downtown Berkeley. Findings of this research revealed that users preferred to arrange their trip at least one night before, not shortly before the trip is made and that costs were a major cause for travelers being willing to consider dynamic ridesharing.

Some examples of organization-based Shared-Ride Transportation and a summary of other research studies involving demand modeling in that context are displayed in Table 1.

Table 1: A Glance at Demand Modeling in the Context of Trip Frequency, Car Ownership, and Shared-Ride Transportation

Context	Model Framework/ Structure	Author/s	Application	Dependent Variable	Model Structure	Is it an Organization -based Context?	Area
Other	Trip Frequency	Barmby & Doornik, 1989	Shopping Trips	Number of trips	P, NB	No	Sussex, UK
		Gurmu & Trivedi, 1996	Recreational Trips		ZIP, HP, HNB	No	Lake Somerville, East Texas
		Kim & Susilo, 2013	Walk trips		LR, P, NB	No	Baltimore region, USA
	Car ownership	Bhat & Pulugurta, 1998	Car ownership	Number of vehicles	Ordered Logit and MNL	No	Boston
		Zhao & Kockelman, 2002			NB	No	1995 Nationwide Personal Transportation Survey data
		Potoglou & Kanaroglou, 2008			MNL	No	Census metropolitan area of Hamilton
		Potoglou & Susilo, 2008			Ordered Probit and Logit, MNL	No	Metropolitan area of Baltimore, Netherlands, and the Osaka Metropolitan Area
		Gopisetty & Srinivasan, 2013			LR, P, Ordered Probit	No	Chennai city, India
Shared-Ride Transportation	Logit	Benjamin et al., 1998	Paratransit	Choice set: Automobile driver, automobile passenger, bus, Dial- A-Ride	MNL	No	Winston-Salem, North Carolina
		Takeuchi et al., 2003	DRTs	Choice set: Private Car and DRT	Binary Logit	No	Japan
		Silvis & Niemeier, 2009	Ridesharing	Choice set: Regular ride-sharer, Not a regular ride-sharer	Binary Logit	No	California
	Count Models	Ben-Akiva et al., 1996	Dial-A-Ride	Number of trips	P	No	Winston-Salem, North Carolina
		Benjamin & Price, 2006	Dial-A-Ride	Number of trips	P	No	Winston-Salem, North Carolina
	Organization-Based Examples	Amey, 2010	Real-time Ridesharing	(No Modeling application)		Yes	MIT, Cambridge, Massachusetts
		Deakin et al., 2010	Dynamic ridesharing	(No Modeling application)		Yes	Downtown Berkeley, and the University of California, Berkeley
		Erdoğan et al., 2015	Carpooling/ Vanpooling	A 5-point ordinal scale indicating students' interest in carpooling/vanpooling to campus as a driver or as a passenger	Ordered Probit	Yes, students are targeted.	UMD University, Baltimore–Washington, D.C

LR: Linear Regression; P: Poisson; ZIP: Zero-Inflated Poisson; HP: Hurdle Poisson; NB: Negative Binomial; HNB: Hurdle Negative Binomial; MNL: Multinomial Logit

CHAPTER 3

METHODS: MODELING THE DEMAND FOR A SHARED-RIDE TAXI

3.1. Introduction

The demand modeling procedure involved in this thesis entails the estimation of several econometric models. In order to capture travelers' readiness to use a Shared-Ride Taxi (SRT), this chapter proposes a modeling framework that involves data collection by means of a Stated Preference (SP) survey. It presents different models including count models, Multinomial Logit, and Integrated Choice and Latent Variable Models. The last section describes the model selection criteria used to arrive at the best fitting model.

Figure 5 below illustrates the different stages involved in the modeling methodology of this thesis.

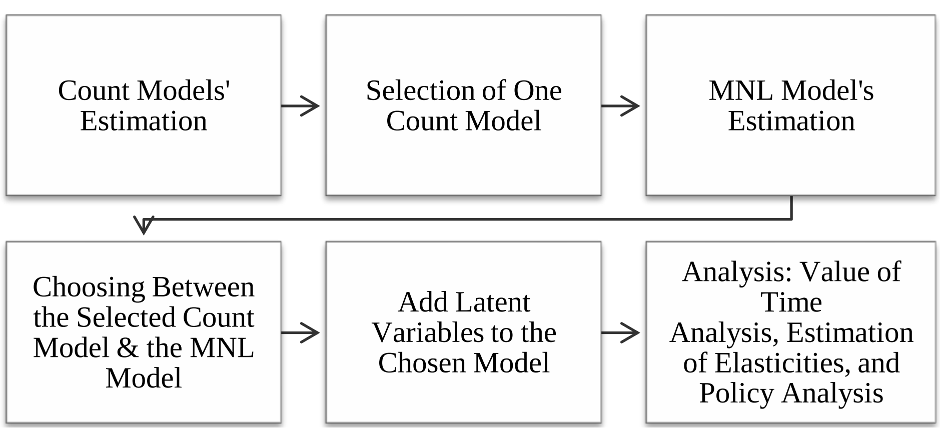

Figure 5: Chapter Map

3.1.1. Data Needs

The data required for undergoing the modeling methodology described in this chapter include the following three categories:

1. Travelers' travel and socioeconomic characteristics where respondents include current trips' mode to and from the school or workplace, travel time, travel cost, parking location and expenses. It also incorporates a question about how many times per week the respondents commute to the targeted organization. Socioeconomic characteristics include questions on gender, family size, family income, number of cars available in the family, and number of licensed drivers in the family, etc.
2. Stated Preference Data where each respondent answers a set of hypothetical choice scenarios. In each question, the respondent chooses how many times he/she will use the new Shared-Ride Taxi per week if it were implemented. The SP survey is of a fractional factorial design, where the values of different attributes related to the SRT service are varied.
3. Attitudinal Indicators of attitudes and perceptions towards transport options and technology where respondents indicate their level of agreement with a set of 7-point scale statements.

3.1.2. Modeling Framework

A Multinomial Logit model will be developed having six discrete choices being the number of weekly trips. The framework of the model incorporates a latent variable model that is based on the premise that travelers' attitudes towards ridesharing, taxi-sharing and/or technology influence their choice (See Figure 6). Additionally, several count models are going to be estimated. In both types of models, the number of trips made per week is assumed to be a function of the SRT service attributes and of the traveler's socioeconomic characteristics. Using multiple criteria, an assessment of model performance is undergone where the best model fit is deduced.

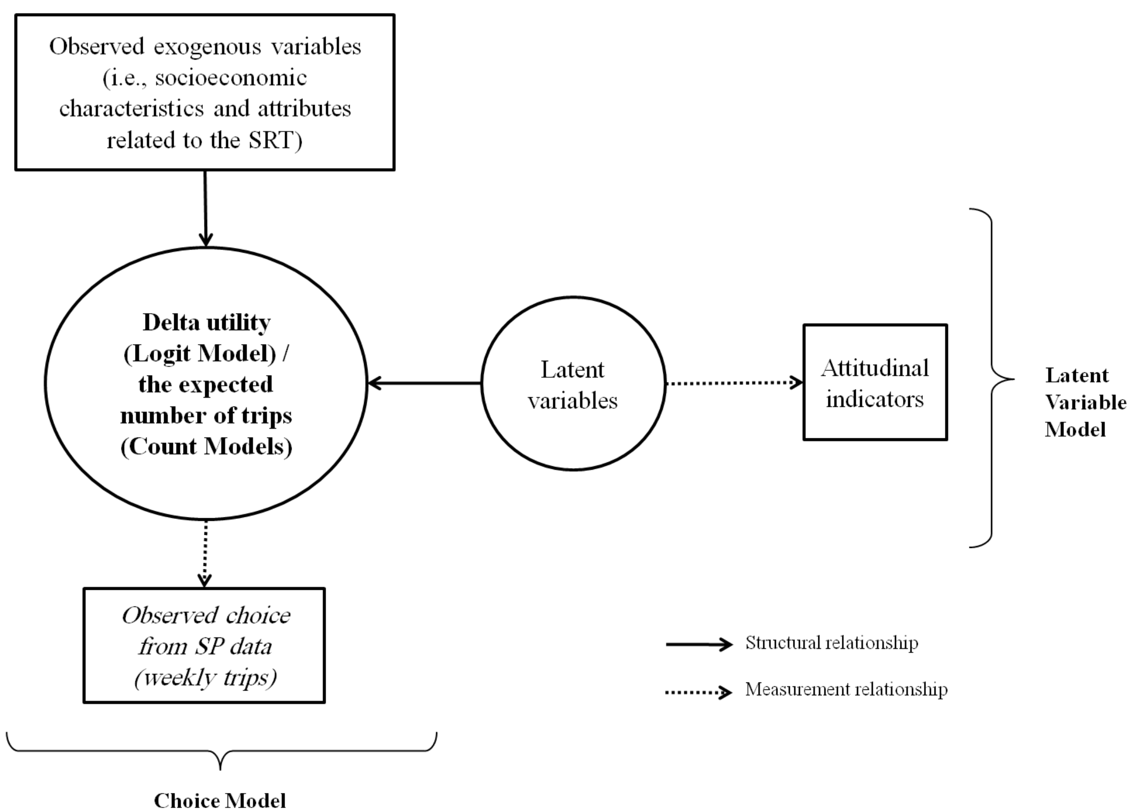

Figure 6: Framework of Modeling the Demand for an SRT Service

3.2. Count Data Models

Many outcomes in the fields on biomedical, social sciences, environmental, and transportation engineering are non-negative and discrete in nature (i.e., 0, 1, 2, …). Poisson and Negative Binomial regressions are two common regression models used in count data modeling. Examples in transportation engineering utilizing count data models include modeling recreation and walking trip frequencies (Barmby and Doornik, 1989; Ben-Akiva et al., 1996; Kim and Susilo, 2013), accident frequency (Anastasopoulos and Mannering, 2009; Kibria, 2006), and household auto ownership (Zhao and Kockelman, 2002).

Poisson regression assumes that the mean and the variance are constrained to be equal. However, this assumption is often violated and the data is considered to be over-dispersed (meaning that the variance is greater than the mean). In such cases, the Negative Binomial distribution is used as it provides an extension to the natural Poisson, allowing for over-dispersion. Sometimes, the over-dispersion in the Poisson regression is caused by the excess in the zero observations relative to what is expected from the Poisson. The increased variance (over-dispersion) caused by the latter phenomenon can be accommodated by using zero-inflated distributions such as Hurdle Poisson (HP) and Zero-Inflated Poisson (ZIP).

As cited in Tait et al. (2012), Hurdle Models were developed in 1986 by Mullahy. These models are called Hurdle Models because attaining a non-zero outcome is thought of as crossing a hurdle (Tait et al., 2012). As the latter research study also cites, Mullahy had developed a Hurdle Poisson Model along with an extension to it, Hurdle Negative Binomial. Nowadays, Hurdle models are becoming prominent as they are being widely applied to data with zero-inflation occurrences (e.g., Kibria, 2006; Tait et al., 2012).

Zorn (1996) also examined zero-inflated distributions and referred to them as a “dual regime” data generating processes which relates to two states, the first state representing the probability that the count will move from a “zero-only” state to another state where it might be a count other than zero, and the second state that represents an event-count process. Zorn (1996) discussed Hurdle models and Zero-Inflated models which were also developed by Mullahy, as cited by Tait et al. (2012). These models divide the count outcomes into two sub-populations, one including only zeros and another one including any count, including zero. Zero-Inflated Poisson and Zero-Inflated Negative Binomial (ZINB) are also very common count data models which have been previously utilized by many research studies (Zorn, 1996; Kibria, 2006; Lawal, 2012; Boucher and Guillén, 2009).

In some cases, repeated data measures for on an individual might lead to correlation in the count data, which is referred to as longitudinal data (i.e., the response is a count of a quantity that is measured in multiple occasions for the same individual). It is very important to account for the correlation; otherwise, incorrect standard errors and inaccurate inference are encountered. One approach to accommodate the correlation present in the data is to incorporate random effects as done in the two research studies by Tait et al. (2012) and Boucher and Guillén (2009). The former researchers were interested in modeling longitudinal count data where an excess in the zero observations was present. Random Effect (RE) longitudinal count models were developed such as RE Poisson, RE Negative Binomial, RE Hurdle Poisson, RE Hurdle Negative Binomial, RE Zero-Inflated Poisson, and RE Zero-Inflated Negative Binomial. With an application to insurance, the latter research study sought to explore panel count data models to determine expected number of claims per year. In their study, Boucher and Guillén (2009) argued that the panel data models allow for time dependence between observations and that the RE count models have the potential to provide a good understanding on insurance applications when companies accumulate data of clients along several years.

In this thesis, RE Poisson, RE Negative Binomial, RE ZIP, RE HP, RE ZINB, and RE Hurdle Negative Binomial (HNB) will be applied to model the number of weekly trips using a Shared-Ride Taxi in an organization-based context. The need for accounting for correlation emerges as respondents are faced with an SP survey composed of multiple hypothetical choice experiments. The outcome of the choice process is the stated number of weekly trips that will be made by the respondents.

In all regression models described in this section, let Y_{nt} represent the count for the n^{th} (n = 1, …, N) individual at the t^{th} (t = 1, …, T) choice experiment, where N is the sample size and T is the number of choice experiments presented for every individual.

3.2.1. RE Poisson Regression

The Poisson regression is used to model the relationship between the number of weekly trips made and the explanatory variables. The form of the standard Poisson regression is shown in Equation 1.

$$Pr(Y_{nt} = y|x_{nt}) = \frac{exp(-\lambda_{nt})\lambda_{nt}^{y}}{y!} for\ y \geq 0 \qquad (1)$$

where λ_{nt} is the mean of the Poisson distribution and $x_{nt} = (1, x_{nt1}, \dots, x_{ntK})$ is a vector of the covariates (explanatory variables) with size *K*. It is assumed that:

$$E(Y_{nt}) = Var(Y_{nt}) = \lambda_{nt} \qquad (2)$$

According to Booth et al. (2003), the mean λ_{nt} can be allowed to depend on explanatory variables as shown in Equation 3.

$$\log(\lambda_{nt}) = x_{nt}\beta \qquad (3)$$

where $\beta = (\beta_0, \beta_1, \dots, \beta_K)$ is a vector of regression coefficients. In order to account for the random effect present in the choices of an individual, a scalar random effect b_n is added to Equation 3. Equation 4 is used instead:

$$\log(\lambda_{nt}) = x_{nt}\beta + z_{nt} b_n \qquad (4)$$

It is assumed that b_n is independent and identically distributed random variable (*iid*) and follows a normal distribution $(0,\ z^2)$ with a mean of zero and variance of z^2. Other research studies used different popular distributions for the random effect such as gamma, or the lognormal distributions (Boucher and Guillén, 2009). As far as this thesis is concerned, the random effect should be considered as a random intercept in which z_{nt} equals one. A different context of mixed count models is employed in the literature whereby (random) parameters β are considered to vary among respondents and are considered to follow a known distribution (e.g., Anastasopoulos and Mannering, 2009).

Ben-Akiva et al. (1996) also modeled the number of weekly trips using a Poisson regression whereby they established a link between the number of trips made using a Dial-A-Ride system and characteristics of the service. Unlike the case considered in this thesis, each respondent was presented with one choice experiment; hence there was no need to account for any correlation in their data.

3.2.2. RE Negative Binomial

An alternative to the Poisson regression is the Negative Binomial regression (also referred to as Gamma-Poisson distribution) which is often used when the equi-dispersion assumption is violated. The Negative Binomial regression takes the form shown in Equation 5.

$$\Pr(Y_{nt} = y|x_{nt}) = \frac{\Gamma(y+\alpha^{-1})}{\Gamma(y+1)\Gamma(\alpha^{-1})} \times \frac{\alpha\lambda_{nt}{}^{y}}{(1+\alpha\lambda_{nt})^{y+\alpha^{-1}}} \qquad (5)$$

where λ_{nt} is the mean of the NB distribution that also takes the form shown in Equation 4, $\Gamma(.)$ is the gamma function, and α is the over-dispersion parameter. The variance of the NB distribution is expressed as:

$$Var(Y_{nt}) = \lambda_{nt} + \alpha\,\lambda_{nt}{}^{2} \qquad (6)$$

When the over-dispersion parameter α is zero, this distribution becomes equivalent to the Poisson distribution. It is this parameter that allows the Negative Binomial regression to accommodate the over-dispersion.

3.2.3. RE Zero-Inflated Models

In Zero-Inflated Models, the process generating the counts has two states: 1) a zero state from which only zeros are generated, and 2) a Poisson state (thus a ZIP regression) or a Negative Binomial state (thus a ZINB regression) from which all non-zeros and few zeros are observed. Other types of zero inflated distributions are also found in the literature such as Zero-Inflated Binomial regression which was applied by Hall (2000) on an example from Biometrics. Zero inflated distributions produce zero outcomes that can be subdivided into two subgroups: structural zeros and sampling zeros. Structural zeros are generated when the individual's only possible outcome is a zero count, while sampling zeros are an outcome of the usual Poisson or Negative Binomial process which assumes that the zero outcomes happened by chance.

The probability of a structural zero is derived from a binary distribution whereas those of the sampling zeros as well as the positive counts are governed by Poisson (or Negative Binomial). Let $Q_1(q_1)$ be a probability mass function (pmf) of a binary process (0/1) such as binary logit and $Q_2(q_2)$ be the pmf of a Poisson or a Negative Binomial distribution. The probability of observing a zero count in the ZIP or ZINB is expressed as follows:

$$P\,r(Y_{nt} = 0|x_{nt}) = Q_1(q_1 = 0) + Q_1(q_1 = 1) \times Q_2(q_2 = 0) \qquad (7)$$

The probability of obtaining a non-zero count is calculated as follows:

$$P\,r(Y_{nt} > 0|x_{nt}) = Q_1(q_1 = 1) \times Q_2(q_2 > 0) \qquad (8)$$

Suppose that the probability of observing a positive count in the first stage (where $q_1 = 1$) is denoted by φ_{nt}. The Zero-Inflated Poisson is thus described using the form below:

$$P\,r(Y_{nt} = y|x_{nt}) = \begin{cases} (1-\varphi_{nt}) + \varphi_{nt}\, exp(-\lambda_{nt}) & for\ y = 0 \\ \varphi_{nt} \exp(-\lambda_{nt})\ {\lambda_{nt}}^{y} / y! & for\ y > 0 \end{cases} \qquad (9)$$

where λ_{nt} is the mean of Poisson, $(1-\varphi_{nt})$ is the probability of the structural zeros and $\varphi_{nt}\, exp(-\lambda_{nt})$ is the probability of the sampling zeros. This leads to a total probability of the zero count greater than that observed in a standard Poisson distribution where $Pr(Y_{nt} = 0)$ is $exp(-\lambda_{nt})$.

Accordingly, the probability mass function of the Zero-Inflated Negative Binomial takes the following form shown in Equation 10.

$$P\,r(Y_{nt} = y|x_{nt}) = \begin{cases} (1-\varphi_{nt}) + \varphi_{nt}\left[1/(1+\alpha\,\lambda_{nt})^{y+1/\alpha}\right) & for\ y = 0 \\ \varphi_{nt}\dfrac{\Gamma(y+\alpha^{-1})}{\Gamma(y+1)\Gamma(\alpha^{-1})} \times \dfrac{\alpha{\lambda_{nt}}^{y}}{(1+\alpha\lambda_{nt})^{y+1/\alpha}} & for\ y > 0 \end{cases} \qquad (10)$$

where λ_{nt} is the mean of the Negative Binomial distribution, $(1-\varphi_{nt})$ is the probability of observing a structural zero and α is the over-dispersion parameter.

As is the case with standard Poisson and NB distributions, ZIP and ZINB can accommodate random effects that capture the correlation due to unobserved personal characteristics as a function of the covariates and the random intercept, Tait et al. (2012) express φ_{nt} analogously to Equation 11 shown below. Also, Equation 4 presented in section 3.2.1 applies for this model.

$$logit(\varphi_{nt}) = log(\varphi_{nt}/[1-\varphi_{nt}]) = x'_{nt}\,\gamma + z'_{nt}\,b'_{n} \qquad (11)$$

where $x'_{nt} = (1, x'_{nt1}, x'_{nt2}, \dots, x'_{ntR})$ is a row vector of the covariates in the Binary Logit model. This leads to models with covariates, in which those affecting the binary stage might not be the same as those affecting stage two (Poisson or NB).

$\beta = (\beta_0, \beta_1, \dots, \beta_K)$ and $\gamma = (\gamma_0, \gamma_1, \dots, \gamma_R)$ still represent the regressions' coefficient vectors where *K* and *R* represent the number of covariates in each of the count and Binary Logit models, respectively. z_{nt} and z'_{nt} are such that $z_{nt} = z'_{nt} = 1$; hence the vector of random effects $B_\mathrm{n} = \begin{pmatrix} b_n \\ b'_n \end{pmatrix} \sim$ MVN $(0, \Sigma)$ represents two jointly normal random intercepts, b_n and b'_n with a 2 x 2 covariance matrix Σ.

3.2.4. RE Hurdle Models

The Hurdle Models essentially consist of two stages. In the first stage, the binomial probability governs the binary choice of whether the count is zero or positive. If it is positive, the hurdle is crossed and thus in the second stage, the positive outcomes are modeled by a zero-truncated count distribution.

Suppose that stage one is governed by a binary process with pmf $S_1(s_1)$ and that stage two where Y_{nt} is strictly greater than zero, follows a zero-truncated Poisson or NB with pmf $S_2(s_2)$. Thus, the probability of observing zero and non-zero outcomes in a Hurdle Model are expressed in Equations 12 and 13, respectively.

$$Pr(Y_{nt} = 0|x_{nt}) = S_1(s_1 = 0) \qquad (12)$$

$$Pr(Y_{nt} > 0|x_{nt}) = [1 - S_1(s_1 = 0)] \times \frac{S_2(s_2 > 0)}{[1 - S_2(s_2 = 0)]} \qquad (13)$$

Suppose that $Pr(Y_{nt} > 0) = \pi_{nt}$ where π_{nt} follows a binary regression with a logistic link function. As in previous parts, assume that λ_{nt} is the mean of the second stage which is governed by a zero-truncated count model. Accordingly, the pmfs of the Hurdle Poisson and Hurdle Negative Binomial are given in Equations 14 and 15, respectively.

$$Pr(Y_{nt} = y|x_{nt}) = \begin{cases} (1-\pi_{nt}) & for\ y = 0 \\ \pi_{nt}\ \exp(-\lambda_{nt})\ \lambda_{nt}{}^{y} / \left[y! \left(1 - exp(-\lambda_{nt})\right)\right] & for\ y > 0 \end{cases} \quad (14)$$

$$Pr(Y_{nt} = y|x_{nt}) = \begin{cases} (1-\pi_{nt}) & for\ y = 0 \\ \pi_{nt} \frac{\Gamma(y+\alpha^{-1})}{\Gamma(y+1)\Gamma(\alpha^{-1})} \times \frac{\alpha\lambda_{nt}{}^{y}}{(1+\alpha\lambda_{nt})^{y+1/\alpha}} \times \frac{1}{1-(1+\alpha\lambda_{nt})^{-1/\alpha}} & for\ y > 0 \end{cases} \quad (15)$$

Hurdle models were also extended to accommodate correlation where longitudinal data is employed (e.g., Min and Agresti, 2005; Boucher and Guillén, 2009; e.g., Tait et al., 2012). This is achieved by the inclusion of random effects as done in Equations 4 and 11.

3.2.5. Maximum Likelihood Estimation for Count Models with Random Effects

In this section, the likelihood function for each of the previously discussed count models is shown. Conditional on B_n, the choices made by a respondent over multiple choice experiments are independent. Taking the product over all choice experiments, the Poisson or Negative Binomial probability of a respondent's choicesy_1 through y_T conditional on x_n and b_n is:

$$p_n(y_1, y_2, \dots, y_T|x_n, b_n; \beta) = \prod_{t=1}^{T} \Pr(Y_{nt} = y|x_{nt}, b_n) \quad (16)$$

Assuming that the random intercept is *iid* and normally distributed, the respondent's unconditional probability is thus given by integrating over the density function of $f(.)$ which denotes the pdf of the random effects:

$$P_n = \int_b p_n(y_1, y_2, \dots, y_T | x_n, b_n; \beta) \times f(b_n)\, db_n \qquad (17)$$

Taking the product over all individuals (n=1, 2, …, N) leads to the following expression for likelihood function (*Likelihood*):

$$Likelihood = \prod_{n=1}^{N} P_n \qquad (18)$$

Integrating over the random effect distribution and taking the product over all individuals, the RE Poisson and RE Negative Binomial Likelihood functions are therefore expressed as follows:

$$Likelihood_{RE\ Poisson} = \prod_{n=1}^{N} \int_{b_n} \left(\prod_{t=1}^{T} \frac{\exp(-\lambda_{nt})\lambda_{nt}^{y}}{y!} \right) \times f(b_n)\, db_n \qquad (19)$$

$$Likelihood_{RE\ NB}$$

$$= \prod_{n=1}^{N} \int_{b_n} \left(\prod_{t=1}^{T} \frac{\Gamma(y+\alpha^{-1})}{\Gamma(y+1)\Gamma(\alpha^{-1})} \times \frac{\alpha\lambda_{nt}{}^{y}}{(1+\alpha\lambda_{nt})^{y+1/\alpha}} \right) \times f(b_n)\, db_n \qquad (20)$$

Fitting zero-inflated distributions with random effects is analogous. Suppose that $I(.)$ is an indicator function such that if $Y_{nt} = 0$, $I(Y_{nt} = 0) = 1$. The general likelihood function of Zero-Inflated and Hurdle Models is expressed as:

$$Likelihood_{RE\ Hurdle\ \ or\ Zero-Inflated\ \ Models}$$

$$= \prod_{n=1}^{N} \int_{B_n} \left(\prod_{t=1}^{T} P\,r(Y_{nt} = 0|x_{nt})^{I(Y_{nt}=0)} \right.$$

$$\left. \times P\,r(Y_{nt} > 0|x_{nt}, b_n)^{1-I(Y_{nt}=0)} \right) \times f(B_n)\, dB_n \qquad (21)$$

Likelihood functions for Poisson and Negative Binomial in either of the Zero-Inflated or Hurdle Models can then be easily applied utilizing the probability mass functions shown in sections 3.2.1 and 3.2.2. As a demonstration, the probability corresponding to individual *n* and the log likelihood function of the RE Hurdle Poisson model are shown in Equation 22 and 23.

$$Log\_Likelihood_{RE\ HP} = \sum_{n=1}^{N} \log(P_{n_{RE\ HP}}) \qquad (22)$$

Where,

$$P_{n_{RE\ HP}} = \int_{B_n} \left\{ \prod_t (1-\pi_{nt})^{I(Y_{nt}=0)} \times \left\{ \pi_{nt} S_2(s_2 > 0) \times \left(1 - S_2(s_2 = 0)\right) \right\}^{1-I(Y_{nt}=0)} \right\} f(B_n)\, dB_n \qquad (23)$$

RE Hurdle Models possess a useful property over Zero-Inflated Models; parameters of hurdle models can be estimated by fitting the two components pertaining to the two stages of the model separately. As the two random effects b_n and $b_n{}'$ are assumed uncorrelated, the maximization of the log-likelihood function of model is achieved in two steps. Yau and Lee (2001) used random effects to explain the within-subject dependence in both components in the hurdle model where the random effects in both components were independent. They argue that this model has the advantage of a separate parameterization, which is simple to interpret. Such separate parameterization permits an efficient way to fit the two components separately (Yau and Lee, 2001).

In fact, maximizing the likelihood in the presence of the random effects was addressed extensively in the literature as it presents a challenge to modelers due to the complex form of the likelihood function. Hence, several authors discussed its computational algorithm in the context of mixed[1] count models. To estimate the parameters of count models with random effects, parametric and non-parametric methods

[1] A mixed model is a statistical model that integrates fixed and random effects (mixed effects).

have been used. Unlike the case with the parametric approach for Maximum Likelihood (ML) model fitting, the non-parametric approach leaves the random effects completely unspecified. Booth et al. (2003) and Min and Agresti (2005) illustrated both parametric approaches (specifying a normal distribution) and non-parametric approaches for random effects. While Booth et al. (2003) examined Negative Binomial log-linear mixed models, Min and Agresti (2005) considered both approaches in the context of zero-inflated count data.

In this thesis, the parametric approach is employed (maximum likelihood model fitting with normal random effects). Many methods have been applied in the literature to estimate generalized linear mixed models by means of this approach; methods include Gauss-Hermite quadrature, the Monte Carlo Expectation-Maximization algorithm (adopted by Hall, 2000 and Min and Agresti, 2005), Laplace Approximation, Markov Chain Monte Carlo, and PQL (Penalized quasi-likelihood). To implement RE Count Models, the glmmADMB package (generalized linear mixed model AD model builder) is utilized. The package is provided by R software (R Core Team, 2014) and developed by Fournier et al. (2012). glmmADMB handles zero-inflation as well as hurdle count models and allows for the inclusion of random effects. To estimate the models' parameters, the package considers Laplace approximation, a technique used to approximate integrals of the exponential form. This method appears to have significant advantage as a highly accurate and fast approximation to ML for such contexts (Raudenbush et al., 2000).

3.3. Multinomial Logit Model with Random Effects

The choice model that was used to analyze travelers' responses as to how many times per week they may switch to the Shared-Ride Taxi is a Multinomial Logit model. The model relates the alternatives' utilities with observed variables related to the alternatives and the decision maker. Alternatives' utilities (denoted as ΔU) are expressed as the utility of the SRT minus the utility of the traveler's current mode (CM) of commute such that:

$$\Delta U_{nit} = U_{nit\ (SRT)} - U_{n\ (CM)} \qquad (24)$$

where *n* (*n* = 1,…, *N*) represents the traveler and *t* (*t* = 1,…,*T*) represents the choice experiment. The random utility maximization model assumes that the respondent *n* is faced with a set of six alternatives *i* (*i* =0,…,5) which indicate how many times the respondent is willing to use the new Shared-Ride Taxi per week. The random utility maximization model assumes that the respondent chooses the alternative that provides the maximum utility.

As seen in Equation 25, each of the utilities of the 6 alternatives is described as a function of two main components:

$$\Delta U_{nit} = V(X_{nit}, a_n; \beta) + \epsilon_{nit} \qquad (25)$$

- *V*, the systematic component of the utility function, representing the observable attributes (the Shared-Ride Taxi service attributes or differences between its attributes and the attributes of the current mode and the socioeconomic characteristics) and the random effect representing unobserved personal characteristics related to the respondent. X_{nit} is a vector containing the values of observed variables of alternative *i* for respondent *n* and a particular choice scenario *t*. a_n represents an unobserved random component that is individual specific; this component accounts for the panel (agent) effect by fixing its value for a given individual across the choice scenarios. β is the utility coefficient vector that indicates the weight of the observed variables (β_X) on the utility.
- Disturbance $\in_{nit}$, which is the stochastic component of the utility. In fact, the Logit model assumes the unobserved components are independently and identically distributed as extreme value Type I.

The conditional probability of individual *n* choosing alternative y_i in choice scenario *t* is given in equation 26.

$$\Pr(y_i)_{nt} = \begin{cases} \dfrac{e^{\mathrm{X}_{\mathrm{nit}}\beta + \mathrm{a}_{\mathrm{n}}}}{\sum_{m=0}^{5} e^{\mathrm{X}_{\mathrm{nmt}}\beta_{\mathrm{m}}}} & for\ y_i = 0 \\ \dfrac{e^{\mathrm{X}_{\mathrm{nit}}\beta}}{\sum_{m=0}^{5} e^{\mathrm{X}_{\mathrm{nmt}}\beta_{\mathrm{m}}}} & for\ y_i = 1, 2, 3, 4, 5 \end{cases} \tag{26}$$

The random component was only included in the utility equation of the zero alternative to test its effect on the likelihood to switch or not.

The systematic utility is specified to be linear in parameters. The observed variables *X* include the service attributes that were varied in the SP survey, specifically the values of the SRT service minus the values of the attributes which the student exhibits using the current mode. Socioeconomic characteristics of the traveler are also included.

The probability of a respondent's choicesy_1 through y_T conditional on (a_n) is:

$$p_n(y_1, y_2, \ldots, y_T | X_n, a_n; \beta) = \prod_{t=1}^{T} \frac{e^{V(X_{nc_{nt}t}, a_n; \beta)}}{\sum_{m=0}^{5} e^{V(X_{nmt}, a_n; \beta)}} \qquad (27)$$

where c_{nt} is the choice made in each scenario *t*. Assuming that a_n follows a normal distribution, the unconditional joint probability of a respondent's choices can be expressed as shown in Equations 28 and 29.

$$P_n = \int_a p_n(y_1, y_2, \ldots, y_T | X_n, a_n; \beta) \times f(a_n)\, da_n \qquad (28)$$

$$P_n = \int_a \left(\prod_{t=1}^{8} \frac{e^{V(X_{nc_{nt}t}, a_n; \beta)}}{\sum_{m=0}^{5} e^{V(X_{nmt}, a_n; \beta)}} \right) \times f(a_n)\, da_n \qquad (29)$$

where $f(a_n)$ is a normal density function. Therefore, integrating over the joint distribution random effects and taking the product over all individuals lead to the following likelihood function:

$$Likelihood = \prod_{n=1}^{N} \int_{a_n} p_n(y_1, y_2, \dots, y_T | X_n, a_n; \beta) \times f(a_n)\, da_n \qquad (30)$$

$$Likelihood = \prod_{n=1}^{N} \int_{a_n} \left(\prod_{t=1}^{T} \frac{e^{V(X_{n c_{nt} t}, a_n; \beta)}}{\sum_{j=0}^{5} e^{V(X_{njt}, a_n; \beta)}} \right) \times f(a_n)\, da_n \qquad (31)$$

3.4. Integrated Choice and Latent Variable Model

The model described in this study incorporates latent factors as explanatory variables in the choice models using an integrated choice and latent variable model (ICLV). The model consists of two main components: a choice model and a latent variable model. This model structure has been developed in Ben-Akiva et al. (2002) and later extensively adopted by other studies as mentioned in section 2.4.4.3.

3.4.1. The Choice Model

In an ICLV model, the choice model is a mixed Logit model that is used to analyze travelers' responses as to how many times per week they may switch to the Shared-Ride Taxi. The model relates the alternatives' utilities with observed and latent variables related to the alternatives and the decision maker. Utilities of the alternatives are expressed as ΔU, in accordance with the utilities described in the section 3.3.

The addition in the ICLV models lies in the components of the systematic utility which is defined according to the following equation:

$$\Delta U_{nit} = V(X_{nit}, F_n; \beta) + \epsilon_{nit} \qquad (32)$$

where a_n is substituted by F_n which is a vector of latent variables for individual n. β is the utility coefficient vector that indicates the weight of the observed variables (β_X) and the weight of the attitudes (β_F) on the utility. In this case, the model accounts for the panel (agent) effect by fixing the values of the attitudes for a given individual across choice scenarios. As was the case in the RE MNL model, the latent variables were only included in the utility equation of the zero alternative to test their effect on the likelihood of using the new SRT service.

3.4.2. The Latent Variable Model

The latent variable model is a standard Structural Equation Model that is made up of two sub-models. The first one is a measurement model which is a confirmatory factor model relating the latent attitudes to their corresponding indicators (manifest variables). A linear model is specified to describe the mapping of the latent variables on the indicators (assumed to be continuous). It is expressed as follows:

$$I_n = \lambda F_n + \nu_n \qquad (33)$$

where *I* is a vector of the manifest variables, λ is a matrix of the factor loadings, and *v* is a vector of the measurement errors that are *iid* multivariate normally distributed.

As mentioned earlier, a section of the survey is designed to capture travelers' attitudes towards the internet and technology and the SRT service. Accordingly, the second sub-model is a structural model that represents the interrelationship between the endogenous latent variables and the observed explanatory variables:

$$F_n = BX_n + \omega_n \qquad (34)$$

where *B* is a matrix of the unknown regression coefficients and ω is a vector representing the random disturbances which are *iid* multivariate normal.

Assuming that $\in_{\mathrm{nit}}$ is also *iid* Extreme Value Type I, the probability of a respondent's choicesy_1 through y_T conditional on (F_n) is:

$$p_n(y_1, y_2, \dots, y_T | X_n, F_n; \beta) = \prod_{t=1}^{T} \frac{e^{V(X_{n c_{nt} t}, F_n; \beta)}}{\sum_{j=0}^{5} e^{V(X_{njt}, F_n; \beta)}} \quad (35)$$

where c_{nt} is the choice made by individual *n* in scenario *t*. The unconditional joint probability of an individual's choices and indicator values can be expressed as follows:

$$P_n = \int_F p_n(y_1, y_2, \dots, y_T | X_n, F_n; \beta) \times f(I|\ F; \lambda) \times g(F|X_n; B)\ dF \qquad (36)$$

where$f(I|X_n, F)$ and $g(F|X_n)$ are the joint probability density functions of the indicators of the latent variables and of the latent variables, respectively, and are products of normal density functions given the assumptions made earlier about the error terms in the latent variable model. Therefore, integrating over the joint distribution of the latent variables and taking the product over all individuals lead to the following likelihood function:

$$Likelihood\ = \prod_{n=1}^{N} \int_F p_n(y_1, y_2, \dots, y_T | X_n, F_n; \beta) \times f(I|F; \lambda) \times g(F|X_n; B)\ dF \qquad (37)$$

3.5. Model Selection Criteria

The selection procedure described in this section includes two steps:

Step 1:

After estimating the six count models, a number of criteria are used to select the best model fit among competing count models. A prerequisite for model selection is the assessment of the data and its needs. Accordingly, the selection will be based on variables' signs and statistical significance, ease of interpretation, the nature of the data, and the models' fit.

Poisson model assumes that the variance of the dependent variable is equal to its mean. Yet, it is not always the case. The negative binomial introduces a dispersion parameter that allows for the given occurrence and therefore accounts for the heterogeneity in the data, as suggested by Lawal (2012). The Negative Binomial regression is equivalent to the Poisson when α is zero. A larger value of α indicates that the variance is much greater than the mean, and therefore, more over-dispersion. To check if the NB regression provides a better fit than Poisson, the significance of the dispersion parameter α can be assessed by carrying out the Likelihood Ratio Test (LRT) such that:

$$H_0: \alpha = 0 \quad and \quad H_1: \alpha \neq 0$$

The LRT test statistic which is based on 1 degree of freedom equals $-2[LL(P) - LL(NB)]$ where LL is the log-likelihood.

A standard method to assess the fit of count models and to compare nested and non-nested models is to calculate the information criteria Akaike Information Criteria (*AIC*) or the Bayesian Information Criteria (*BIC*). See equations 38 and 39. In fact, equation 38 shows the formula for the corrected AIC (AIC_c) which presents an extra penalty for extra parameters.

$$AIC_c = (-2LL + 2k) + \frac{2k(k+1)}{N-k-1} \qquad (38)$$

$$BIC = -2LL + 2\log(N)\,k \qquad (39)$$

where LL is the log-likelihood, k is the number of parameters and N is the sample size. For the best fitted model, one expects a lower AIC/BIC value. The AIC criterion was derived in 1973 by Hirotugu Akaike whose measure provides a paradigm for model selection in the inference and analysis of empirical data (Burnham and Anderson, 2002). It provides an effective and objective basis for the ranking of candidate models and the selection of a best estimated model for data inference and analysis (Burnham and Anderson, 2002).

Step 2:

After selecting one count model, the chosen count model is compared to the Multinomial Logit model.

The first criterion in Step 2 is the R-squared measure. This statistic measure shows how well a model fits the data; the model with the highest R-squared indicates the best fit. While several R-squared measures are proposed and applied in the literature (Cameron and Windmeijer, 1996; Kramer, 2005), in the context given in this thesis, the McFadden's R-squared measure is used and defined as:

$$R^2 = 1 - \frac{LL_M}{LL_0} \qquad (41)$$

where LL_M represents the log likelihood of the model being estimated and LL_0 is the log-likelihood of the null model. The null log-likelihood is the log-likelihood for observing the choices given that all respondents choose at random, meaning that the six alternatives are equally likely to be chosen (Johansen, 2013). In the context of discrete choice, R^2 is not meaningful. Instead, ρ^2 is used and calculated similarly, as documented in Domencich and McFadden (1975). In this thesis, both measures are used analogously.

Another model selection criterion in the second step is the market segment prediction test which examines the goodness of fit of the model by conveying direct information about the differences between observed probabilities and models' predicted probability. The test examines the ability of the model to replicate observed shares of the alternatives of each market segment (Ben-Akiva and Lerman, 1985). Here, the terminology employed in (Ben-Akiva and Lerman, 1985) is considered. The expected number of students willing to use the SRT service *i* number of times in market segment *g* is N_{gi} such that $\sum_{all\ i} N_{gi} = N_g$. The share $\frac{N_{gi}}{N_g}$ is compared to the share predicted using the model which equals $\frac{1}{N_g}\sum_{N_g} P_n(i)$. The closer the observed to the predicted shares are, the better the model fit.

CHAPTER 4

MODELING THE DEMAND FOR A SHARED-RIDE TAXI: THE CASE OF AUB

4.1. Introduction

This chapter presents an application of the modeling approach developed in Chapter 3. It examines the provision of a Shared-Ride Taxi (SRT) for the students at the American University of Beirut, a private university in a developing country whose students mostly come from wealthy families (Danaf et al., 2014). This application is an extension to a thesis dissertation done by Hani Al-Naghi in September 2014 and which presents an evaluation framework for organization-based ridesharing with an application to the students at AUB. The university is located inside Greater Beirut Area (GBA) and at the heart of Ras Beirut, Lebanon. It extends in an area of 250,000 square meters and caters for 7920 students from Lebanon and abroad, according to 2013 university statistics. The campus overlooks the Mediterranean Sea and its upper campus is adjacent to Bliss Street. Figure 7 shows the location of AUB campus and Bliss Street with respect to the study area.

The second section of this chapter describes the current transport system in the study area. Sections 4.3 and 4.4 elaborate on the design procedure of the mobility survey employed in this application which was administered in November 2013 and the socioeconomic characteristics of the respondents, respectively. Then, in Section 4.5, the SRT demand modeling framework is illustrated in detail. The latter section integrates the methods developed in Chapter 3 and handles the gaps in the literature of demand modeling of shared-ride transportation.

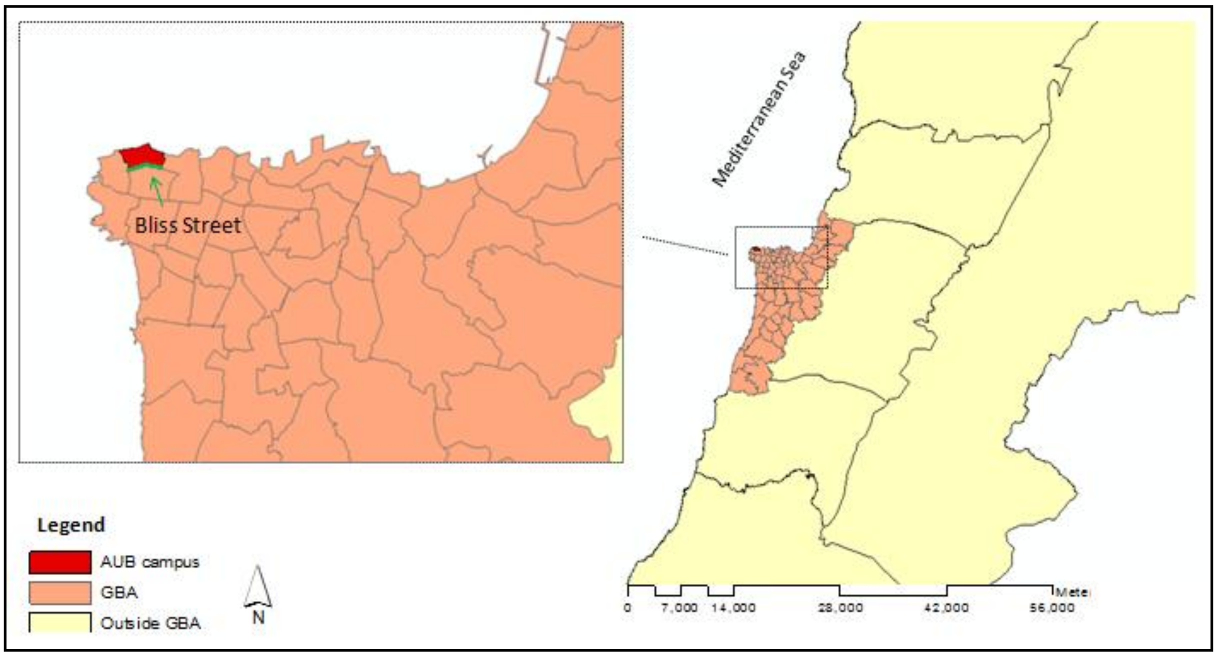

Figure 7: Greater Beirut and American University of Beirut Location Map

4.2. Existing Transport System

4.2.1. In Greater Beirut

A 200-km^2 geographical area, framing the heart of Beirut, Greater Beirut Area is the focus of this study. The population in GBA (approximately 2 million) comprises half of the Lebanese population (Aoun et al., 2013). This area embraces major service-based businesses and includes one of the well-known universities of Lebanon, the American University of Beirut. According to the Lebanese Ministry of Public Works and Transport, there are about 2.8 million daily automobile trips that took place in GBA in 2007, a value that is expected to increase to 5 million in 2015 (Aoun et al., 2013). Even though Greater Beirut is a dense area, commuters do not often rely on public transport modes or non-motorized modes (i.e., walking and biking). The limited dependency on non-motorized modes is due to the lack of the needed infrastructure such as wide sidewalks, exclusive bicycle paths, crossing facilities, etc. (Danaf et al., 2014) which is necessary for a safe and pleasing walkability and cycling in Greater Beirut Area when short-distance commuting is an option for travelers. In fact, commuters heavily rely on private cars.

Numerous attempts have been made from authorities in charge as well as non-governmental organizations to develop effective traffic control systems, reduce congestion in critical corridors, and create an efficient management system for on-street parking. Multiple studies are also on-going to promote integrated mass transit systems that will further limit congestion and improve the quality of life of the travelers inside GBA.

Public transportation is provided by many transport operators in the city. There are two main types of public transportation. First, there are the buses which run designated routes and schedules and are operated by two companies: Lebanese Commuting Company (LCC) which is a private company that runs urban services in GBA and le Office des Chemins de Feret des Transports en Commun (OCFTC) that runs the governmentally owned buses. According to the Ministry of Public Works and Transport, buses and minibuses account for 6% of the transport demand in Lebanon (Aoun et. al, 2013). They operate based on a fixed charge of 1,000 Lebanese Liras (L.L) (0.7$). Second, there are the shared taxis (jitneys) which are privately owned and whose trips cover Greater Beirut and outside Greater Beirut areas. As stated by Danaf et al. (2014), the jitneys do not operate based on fixed routes or schedules but rather on travelers' ad-hoc demand. This mode as well as that of private taxis serve 15% of the demand (Aoun et. al, 2013) and travelers benefit from the improved level of service compared to that of the buses and the relatively cheap fare of 2,000 L.L. (1.3$).

The rest of the demand, representing the majority of the trips (68%) is covered by private cars. This mode of transport is highly used due to the unreliability of public transportation which leads to an unbalanced modal split among modes in the city. Private car users benefit from the inexpensive daily parking fare which ranges between 3,000 L.L. (2$) and 5,000 L.L (4$) in GBA (Danaf et al., 2014).

4.2.2. In the AUB Neighborhood

In the neighborhood of AUB, the case of congestion exacerbates as students and employees create intense traffic, in particular in the area from Bliss Street to Hamra Street. Students continuously express their dissatisfaction with their commute be it private cars or public transportation. This dissatisfaction is caused by congestion, long travel time, and insufficient parking spaces. The area witnesses severe demand in the morning and afternoon peaks. Neighborhood parking facilities and the free parking spots experience demand that exceeds capacity up to 16% (Aoun et al., 2013). Some students use the curbside of the road to park, which is free of charge next to the lower campus exit. Others use the curbside in Bliss Street which is charged as well as several nearby parking facilities.

In the absence of public transport initiatives that address the congestion and the environmental consequences which it imports, shared-ride transportation is assessed in an attempt to identify a potential market for AUB students.

4.3. Survey Design

The data used in this study were collected in November 2013 and targeted the students of American University of Beirut. A survey was carried out by a research team in the Civil and Environmental Engineering Department at AUB as part of the Neighborhood Initiative Congestion studies. It was web-based and remained active for three weeks. Two particular aspects were evaluated: commute to AUB travel patterns and pedestrian activity in the neighborhood of AUB. While another study focused on modeling satisfaction with the walking environment, the data utilized in this research sheds light on the AUB students' commute travel patterns.

The survey collected data on the following aspects:

- *Students' Travel and Socioeconomic Characteristics:*

Travel characteristics include questions on students' mode of commute to and from AUB, door-to-door travel time, travel cost, and parking location and expenses. Socioeconomic characteristics include questions on gender, academic year, faculty, major, family size, family income, number of cars available in the family, and number of licensed drivers in the family.

- *Attitudinal Indicators:*

This section involves a subjective evaluation of attitudes towards transport options and technology. Students' answers were measured using a 7-point scale where 1 represents "strongly disagree" and 7 represents "strongly agree". The indicators capture students' attitudes towards the new Shared-Ride Taxi service, ridesharing, internet and technology, and transport modes in the context of commuting to AUB.

- *Stated Preference (SP) Data:*

Every student is provided with 8 choice scenarios/experiments. In each question, the student chooses how many times he/she will use the new SRT service per week if it were implemented. Values of the variables are presented to students in the manner shown in Table 2. The question posed was phrased as follows:

“Based on this scenario, and considering your current commuting pattern to AUB, how many days per week will you use the shared-ride taxi service?”

Table 2: Choice Scenario Example as Presented in the Survey

One-way fare	**Change in travel time**	**Maximum allowable waiting time for pick-up & early drop-off**	**Maximum number of passengers sharing a ride in a vehicle (including you)**	**Mobile application for reservation and tracking & free Wi-Fi connectivity**
2500 L.L.	10 min. more than your current travel time using your current travel mode	0 to 5 min	4 to 6 (Minivan)	Not available

The SP survey design considered three regions inside GBA in accordance with the categorization shown in Table 3 and Figure 8.

Table 3: Details on Regions A, B, and C Considered in the SP Survey Design

Region	**Name**	**Zones**	**Area**
A	Municipal Beirut	1 -24	From AUB to a radius of 5 km
B	Beirut Inner Suburbs	25- 45	5 km to 10 km away from AUB
C	Beirut Outer Suburbs	46-63	More than 10 km away from AUB

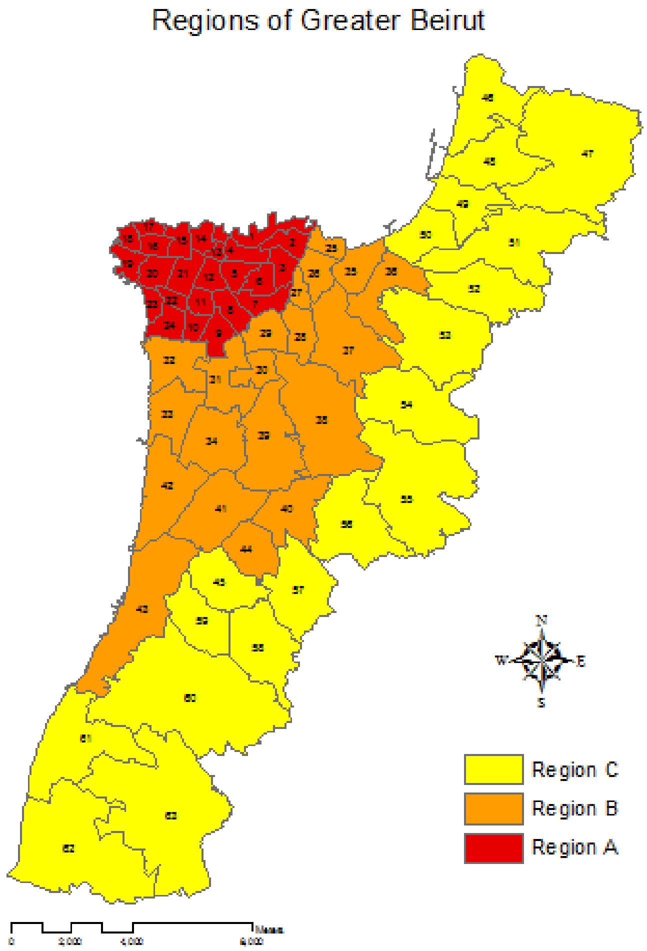

Figure 8: Geographical Map Showing Regions A, B, and C

As such, the 8 choice scenarios were not presented to all AUB students. The target population for the SP survey consisted of students who are eligible to use the Shared-Ride Taxi: students who live inside GBA and use a motorized mode (private car or public transport) to commute to AUB. Note that students who use jitney and live in Region C are excluded from the target population since these students are expected to commute using more than one jitney to complete their trip to/from AUB while the SP survey design targets only those who only use one jitney.

The 8 hypothetical scenarios were presented to students depending on their place of residence (Region A, Region B, or Region C) and their current mode of commute. These scenarios differ in the values of the following five variables:

- One-way fare of the Shared-Ride Taxi,
- Change in travel time; it is the difference between the travel time using Shared-Ride Taxi and travel time using the student's current commute mode,
- Maximum allowable waiting time for pick-up and the maximum allowable time for early drop-off; these variables represent the maximum allowable time the student may have to wait for the taxi/minibus after the assigned pick-up time and the maximum allowable difference between the actual drop-off time and the assigned drop-off time, in case of early drop-off of the student, respectively,
- Maximum number of passengers sharing a ride in a vehicle including the respondent,
- The availability of mobile application for reservation and tracking and the presence/absence of Wi-Fi connectivity in the taxi.

The variables and their levels are shown in Figure 9. Different levels of the one-way fare and change in travel time apply based on the current commute mode and the place of residence. While the combinations of levels of attributes in the SP survey are hypothetical, they represent achievable levels of service. In fact, most scenarios presented to students are practical cases. A primary consideration when designing the SP survey was to introduce reasonable attribute levels, and to include a range of levels for the fare, time, and MAWT attributes in order to further our understanding of the choice behavior and increase the explanatory power of the choice model.

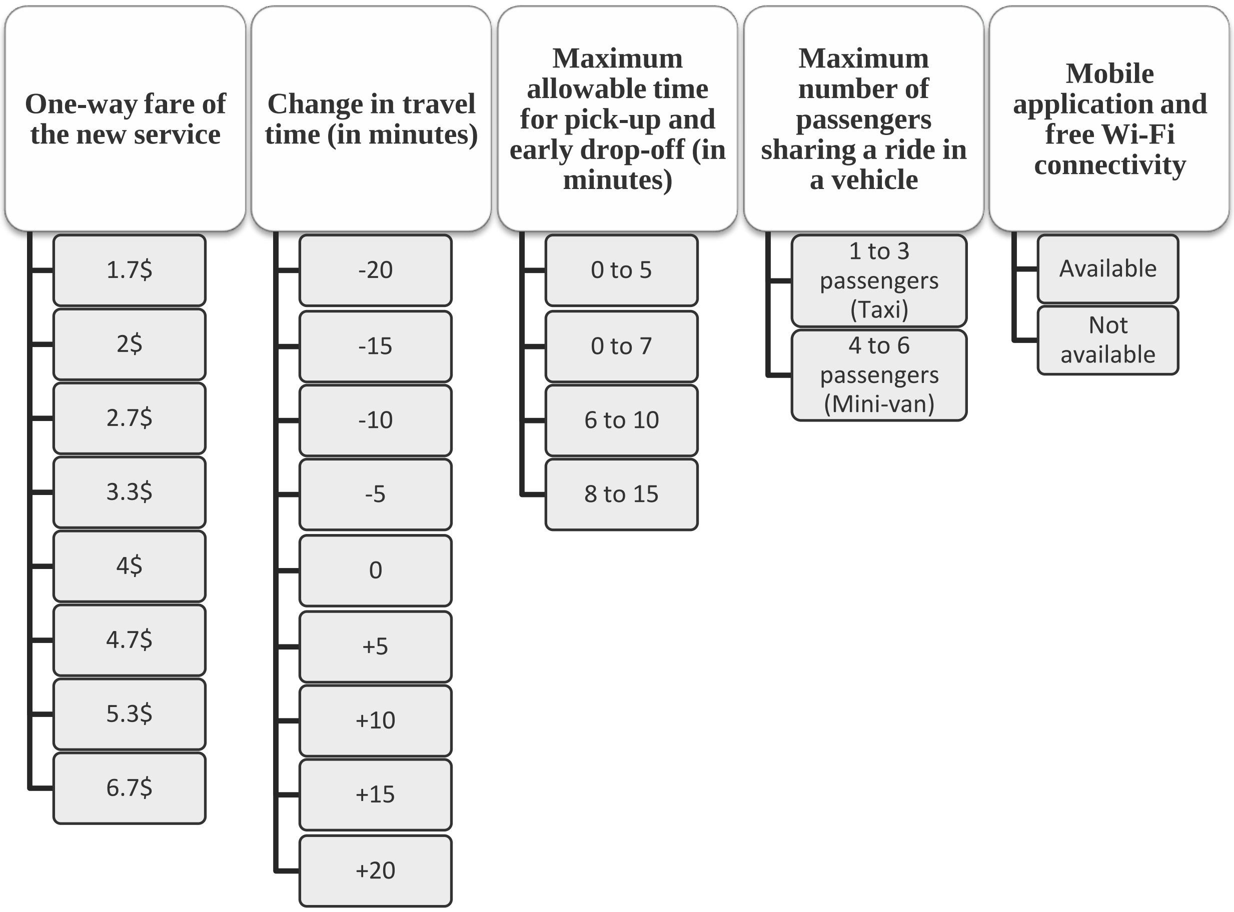

Figure 9: Service Attributes and Their Levels

4.4. Socioeconomic Characteristics and Travel Patterns

The survey response rate was 29% with a total of 2291 students participating in the survey; 1393 responses were complete and 898 were partial. The sample description below is based on the students who completed the survey.

First, it is important to mention that results show that the sample that responded to the survey is rather representative of the AUB population. While 47.7% of AUB students are males, 45.15% of the responses correspond to males in the survey. Figure 10 shows comparable percentages of respondents in each faculty to the AUB data.

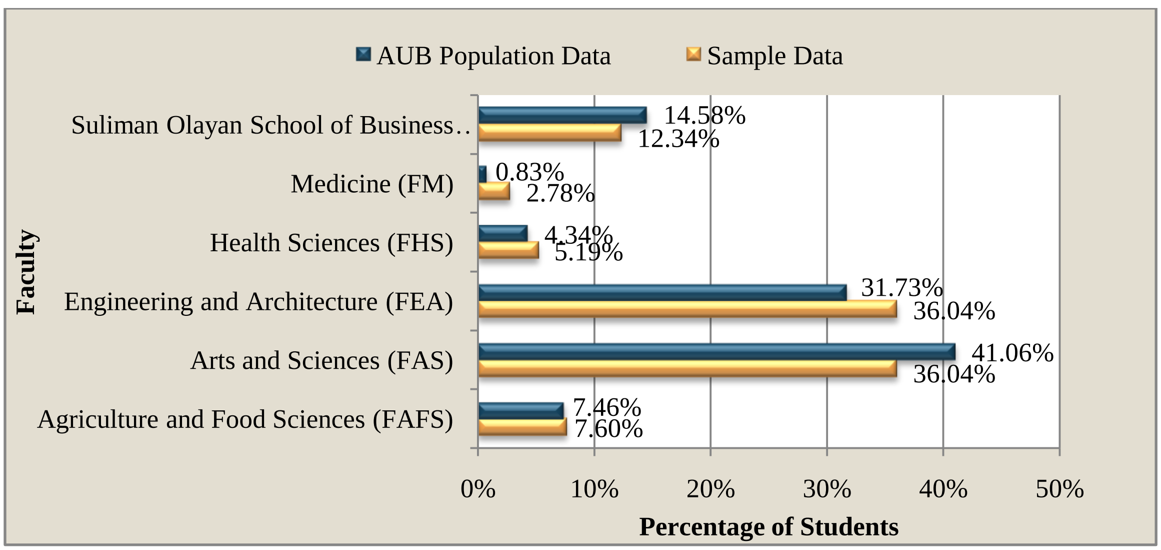

Figure 10: Percentage of Students in Each Faculty (Sample Data vs. AUB Population Data)

Second, statistics show that 1138 students (81.69%) live inside Greater Beirut Area. In Figure 11, statistics of the shares of the different modes show that 45.52%

commute via auto modes[2], 20.17% commute via non-motorized modes[3], 24.11% use public transportation[4], and the rest live on campus.

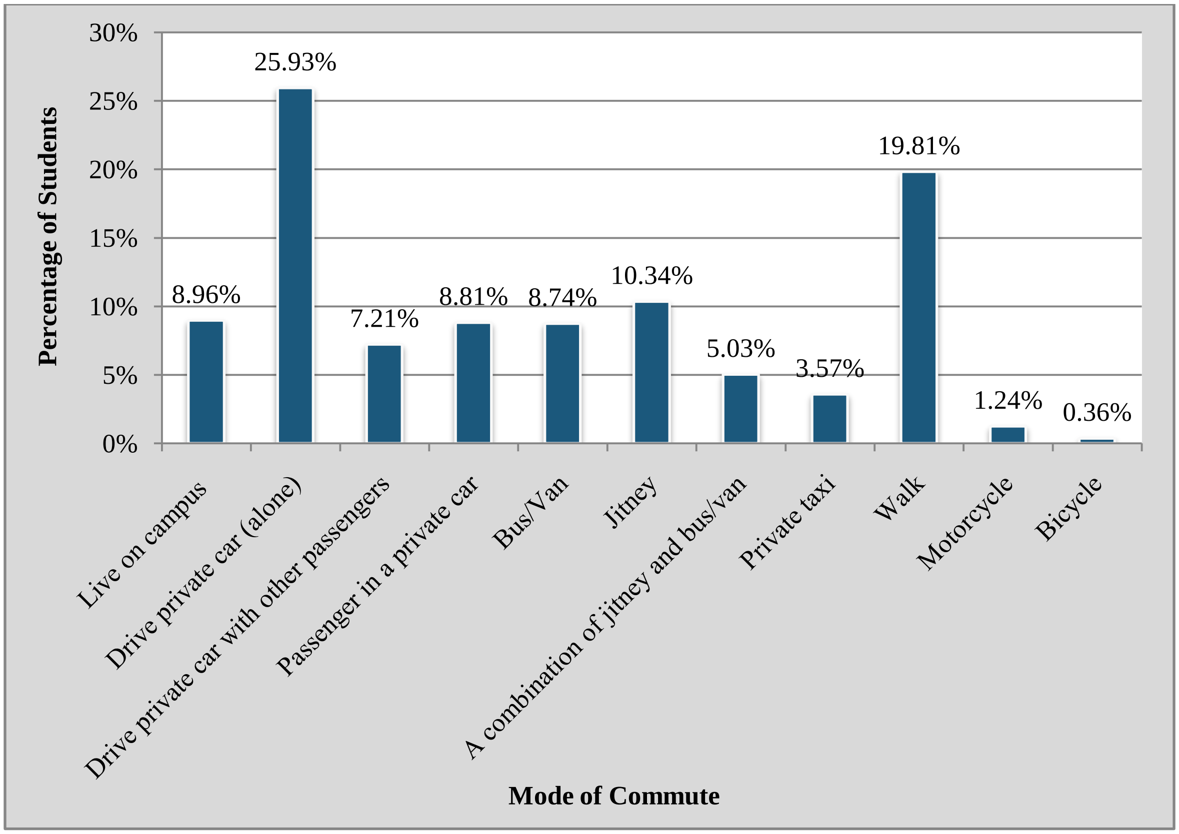

Figure 11: AUB Student Mode Split (2013)

Table 4 presents important socioeconomic characteristics and travel characteristics of the surveyed sample. It can be seen that those who commute using non-

---

[2] This category includes those who drive private care (alone), students who drive private car with other passengers in the car, those who commute as passengers in a private car, and students who commute via private taxi. Motorcycles commuters are excluded from this category.
[3] Non-Motorized modes include bicycles and walking all the way to campus.
[4] This category includes bus/van, service, and a combination of bus and service.

motorized modes of commute are the most satisfied with their commute to AUB. On a scale from 1 to 7, with 1 being very dissatisfied and 7 being very satisfied, the average satisfaction is 5.58. Looking at the other variables' values of this mode category, one notice a low travel time compared to other categories. Also, students of this category do not incur any commute cost. The statistics therefore suggest that travel time and travel cost are major indicators of students' satisfaction with the commute. On the other hand, the satisfaction of the public transport and auto commuters are 3.38 and 3.21, respectively. In a similar context, St-Louis et al. (2013) uses a large-scale travel survey to compare commuter satisfaction across different modes of transport and to examine the determinants of satisfaction across modes. The study revealed that the most satisfied commuters are those who commute by walking, while the least are metro and bus users.

The average monthly income of the families of AUB students who reported their income (55% of the sample) is 7,600,000 L.L. (5100$), with a median value of 5,000,000 L.L. (3,333$) .Table 4 also reveals that students who commute by auto modes come from wealthier families than those who commute using the public transport and non-motorized modes. On the other hand, and as expected, those who commute using auto modes have the highest average car ownership value of 2.62 vehicles whereas, on average, non-motorized modes' commuters' families own 1.20 vehicles.

The average travel time to and from AUB is 35 minutes and 41 minutes, respectively. In an attempt to understand the factors affecting mode choice, Corpuz (2007) concluded that car users are primarily concerned with speed as well as with the comfort and convenience which are associated with shorter travel time and the flexibility of the trip-making. Public transportation on the other hand, is mostly viable where parking capacity is problematic for car users, when the vehicle is not available or the mode is cheaper (Copuz, 2007).

Table 4: Travel Patterns of the Different Mode Categories

Mode of Commute	**Average Satisfaction**	**Average Travel Time (minutes)**		**Average Travel Cost (L.L.)**	**Average Car Ownership**	**Average Income (Millions L.L.)**
		To AUB	From AUB			
N/A: I live on campus	-	-	-	0	1.35	6.28
Driving private car (alone)	3.08	43.63	51.20	4,980	2.92	10.77
Driving private car with other passengers in the car	3.07	42.87	55.61	2,690	2.51	9.31
Passenger in a private car	3.65	37.96	48.75	1,160	2.14	7.68
Bus/Van	3.15	54.03	59.82	2,630	1.44	4.67
Jitney	3.58	27.81	34.45	2,890	1.64	4.07
Private taxi	3.30	31.50	45.13	14,520	1.80	7.79
Walking all the way from residence to AUB	5.56	11.00	11.56	0	1.21	3.00
Motorcycle	4.69	11.92	22.88	409	1.71	2.00
Bicycle	6.60	5.00	15.00	0	0.80	6.31

Note: The combination of bus and jitney mode was excluded due to the limited survey information pertaining to the travel time and cost.

Moreover, statistics show that most of AUB students who commute using their cars pay daily for the parking. The average monthly fee for those who pay daily and pay monthly is 113,000 L.L. (75$).

When asked about their concerns with regard to their commute to AUB, many students complained about the wasted time they spend searching for a parking spot.

Others proposed that concerned authorities should work on upgrading the current public transport system for it will be the mere solution to congestion. Notably, a lot of students showed interest and anticipation for having the new Shared-Ride Taxi put into service.

4.5. Modeling Demand for a Shared-Ride Taxi

4.5.1. Data Sets

Out of 686 observations eligible for answering the choice scenarios, a set of 508 records (sample individuals) are used in the modeling procedure after a data cleaning process was done. The process included exclusion or amending of a number of observations after the identification of inaccurate observations from the database. For instance, some of these records include unreasonable information of travel time, place of residence, and/or travel cost. Another important source of inconsistencies was the dominating scenarios in the 8 choice experiments of the SP data. A dominating scenario is defined as that which has all attributes equal those of the other scenarios with the exception of one or more that are more favorable. For example, if scenario A is dominating scenario B, then all attributes of scenario A were the same as the attributes of scenario B with the exception of the fare attribute which is higher in B than A. In that case, if the respondent chooses a higher number of trips as an answer to Scenario B than A, the whole observation pertaining to the respondent is excluded from the modeling procedures developed subsequently. As a result, 104 observations in total were excluded.

Two students' data sets were analyzed separately: students who currently use public transport modes (PT), i.e., jitney or bus (188 respondents) and students who currently commute using a private car (PC) (320 respondents). The reasons for the given division are: 1) to understand hidden dissimilarities in the behavior of the users of private cars and public transportation and 2) the limited number of records in the public transport data set when jitney and bus users are considered separately (jitney data set includes 114 records while the bus data set includes 74 records).

The mean number of weekly trips (using the SRT service) over all choice scenarios in the PT data set is 1.25 with a variance of 3.37. In the private car data set, the mean is 1.33 while the variance is 2.87. Figure 12 shows the percentage of times each of the six choices provided to students in the SP survey was selected. Fifty-two percent of the Private Car observations and 62% of the choices pertaining to the users of Public Transport were zeros. Each of alternatives 1, 2, 3, 4, and 5 was chosen 4% to 13% of the times. Another significant observation was that alternative4 had the least number of observations. This might be due to the fact it is more common for students to come one time (e.g., graduate students), two times (Tuesday- Thursday schedule), or three times weekly (Monday-Wednesday-Friday), but rarely are the schedules consisting of four days per week.

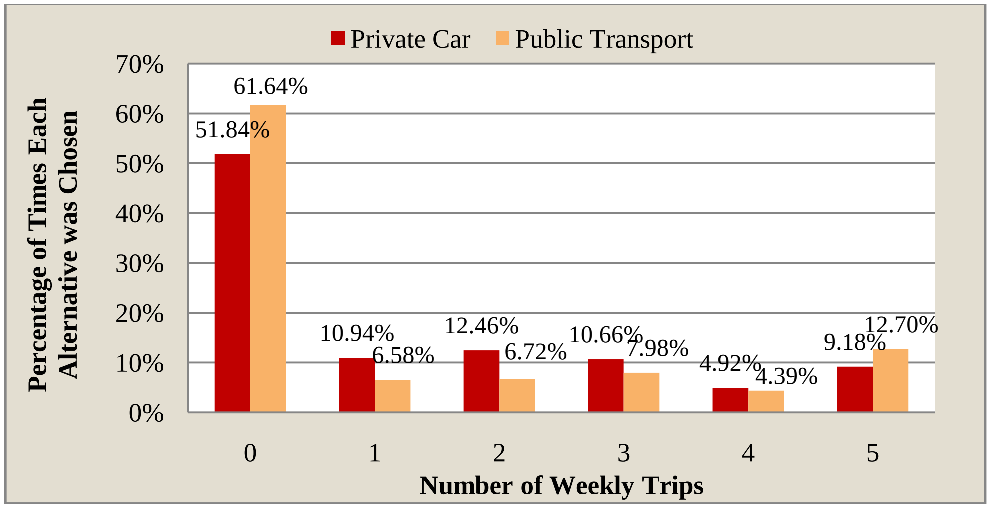

Figure 12: Choice Experiments Response Summary

4.5.2. Results

4.5.2.1. Count Models

a) Model Formulation and Selection

The first step in the modeling procedure is the estimation of the count models. Six Count Models are examined as per the methods described in Chapter 3. Then, the best model fit is chosen. In the count regression models, it is assumed that the dependent variable represents the number of weekly trips using the SRT service if it were available and that it follows a count distribution.

The six count models estimated are Poisson, Negative Binomial, Zero-Inflated Poisson, Zero-Inflated Negative Binomial, Hurdle Poisson, and Hurdle Negative Binomial. In all of the count models, the mean λ_{nt} of weekly trips by SRT (chosen by respondent *n* in choice experiment *t*) is allowed to depend on explanatory variables such that $\log(\lambda_{nt}) = x_{nt}\beta + z_{nt}b_n$. x_{nt} is a vector of the covariates, β is a vector of regression coefficients, z_{nt} is the standard deviation of b_n , and b_n is a normally distributed random variable that is included to account for the correlational effect induced due to the measured data over the 8 choice scenarios.

Five variables are included as explanatory variables in the six count models. An exception is the Zero-Inflated and Hurdle models which are capable of handling the zero observations in stand-alone distribution, assumed to be logistic. The latter distribution includes the five variables described in Table 5 as well as the socioeconomic variables- most of which proved to be insignificant.

One limitation associated with R package glmmADMB used to estimate the count models is that the ZIP and ZINB account for zero-inflation through a zero-inflation parameter that is assumed to be fixed for all observations and hence, the first stage in the model is replaced by a constant probability of zero associated with every student. Hurdle Models are expected to be favored over ZIP and ZINB due to their ability to generate a two-stage model fitting.

Table 5 describes all covariates incorporated as explanatory variables in the count models. The only socioeconomic variable that was significant is the gender dummy variable, which was added to the binary logistic model.

Table 5: Types of Explanatory Variable Employed in Models

Variable	Description
Delta Fare: One-way fare of the new service minus the one-way fare of the mode which the student currently uses (in 1,000 L.L.)	Continuous Variable
Delta Time[5]: Travel time using the SRT service minus the travel time incurred using the current mode of commute (in hours)	Continuous Variable
MAWT: Maximum allowable waiting time for pick-up and early drop-off (in hours)	Continuous Variable
Minivan Dummy: Maximum number of passengers sharing a ride in a vehicle	Dummy Variable: A value of 1 indicates a high number of passengers, and a value of 0 indicates low number of passengers
Wi-Fi Dummy: Mobile application and free Wi-Fi connectivity	Dummy Variable: A value of 1 indicates internet and Wi-Fi availability, and a value of 0 indicates the unavailability of Wi-Fi and internet in the SRT service
Male Dummy	Dummy variable: A value of 1 indicates that the respondent is a male, while a value of 0 indicates that the respondent is a female

Count data can sometimes be modeled as rates and therefore can specify an offset (or exposure) variable that indicates the maximum number of times the event could happen such that the rate is $\frac{count}{exposure}$. Upon multiplying both sides of the equation ($\frac{count}{exposure} = e^{x_{nt}\beta + z_{nt}b_n}$) by the exposure, the latter moves to the right hand side of the equation. Hence, when both sides are logged, the new variable added to the equation is thus $log(exposure)$. The coefficient of the offset variable is specified as 1. In the

[5] This value might be negative or positive. A low *Delta Time* corresponds to an improved level of service as compared to SRT with high *Delta Time*. For example, a *Delta Time* of 6 minutes is higher than that of -3 minutes and the latter is higher than a *Delta Time* of -10 minutes. Considering these 3 cases, the SRT with the best level of service (with respect to travel time) is that with -10 minutes, while the SRT with the worst level of service is that with *Delta Time* of 6 minutes, compared to the same current mode.

estimated count models, the exposure variable was the current number of trips the student makes using their current mode of commute.

The likelihood ratio test was conducted to compare the standard Poisson and Negative Binomial regression models and to infer if over-dispersion exists. The likelihood ratio test statistic values are 57.86 and -0.04 for the PT and Private Car data sets, respectively, and the critical value ${\chi^2}_{1,0.05}$ is 3.84. Therefore, for the PT data set, the null hypothesis (no over-dispersion exists) is rejected at the 5% level of significance leading to the conclusion that the Negative Binomial provides a better fit than Poisson. As for the PC data set, the null hypothesis is not rejected, meaning that there are no underlying reasons to claim that the data is over-dispersed, or that the NB regression is a better fit. This is confirmed as the log-likelihoods of the Poisson and NB are equal (-3512.9) and therefore, the latter should not contribute to any improvement to the standard Poisson. Despite the data being not over-dispersed, the high number of zero outcomes justifies the need to examine models that handle the occurrence of excessive zeros such as Zero-Inflated and Hurdle models as they are expected to provide a better fit than the standard Poisson distribution.

In the current model selection process, assessment shall focus on comparing the models' performance based on the AIC and BIC criteria. These information criteria are extensively used in the literature of count modeling for the purpose of model comparison e.g. Gurmu and Trivedi (1996), Hall (2000), Kibria (2006), and Tait et al. (2012). From

Table 6 shown below, one concludes that the Hurdle Poisson provides the best fit as compared to the rest (lowest AIC and BIC values).

Table 6: Summary of Count Models Statistics for Comparison

<table>
<tr><th colspan="2"></th><th colspan="2">Log-Likelihood</th><th colspan="2">Parameters</th><th colspan="2">AIC</th><th colspan="2">BIC</th></tr>
<tr><td colspan="2"></td><td>PC</td><td>PT</td><td>PC</td><td>PT</td><td>PC</td><td>PT</td><td>PC</td><td>PT</td></tr>
<tr><td colspan="2">Poisson</td><td>-3512.9</td><td>-1968.1</td><td colspan="2">7</td><td>7039.8</td><td>3950.3</td><td>7080.8</td><td>3987.4</td></tr>
<tr><td colspan="2">NB</td><td>-3512.9</td><td>-1939.2</td><td colspan="2">8</td><td>7042.0</td><td>3894.4</td><td>7088.7</td><td>3936.9</td></tr>
<tr><td colspan="2">ZIP</td><td>-3487.8</td><td>-</td><td colspan="2">8</td><td>6991.8</td><td>-</td><td>7038.4</td><td>-</td></tr>
<tr><td colspan="2">ZINB</td><td>-3488.7</td><td>-1884.9</td><td colspan="2">9</td><td>6995.5</td><td>3787.9</td><td>7048.0</td><td>3835.7</td></tr>
<tr><td rowspan="2">HP</td><td>BL</td><td rowspan="2">-3260.5</td><td rowspan="2">-1691.5</td><td>7</td><td>8</td><td rowspan="2">6553.0</td><td rowspan="2">3413.3</td><td rowspan="2">6625.7</td><td rowspan="2">3478.5</td></tr>
<tr><td>TP</td><td colspan="2">7</td></tr>
<tr><td rowspan="2">HNB</td><td>BL</td><td rowspan="2">-3262.0</td><td rowspan="2">-1692.6</td><td>7</td><td>8</td><td rowspan="2">6558.2</td><td rowspan="2">3417.4</td><td rowspan="2">6635.9</td><td rowspan="2">3486.9</td></tr>
<tr><td>TNB</td><td colspan="2">8</td></tr>
<tr><td colspan="10"><u>Notes</u>: - Values in bold correspond to the smallest values among competing models in a given data set.
- BL is Binary Logit, TP is truncated Poisson, and TNB is truncated Negative Binomial.
- Empty cells corresponding to the ZINB are because of the inability to reach model convergence.</td></tr>
</table>

b) Hurdle Models' Formulation

The superiority of the Hurdle Poisson Model suggests that there are two processes at work, one determining whether there are zero observations or other observations (i.e., whether or not the student is willing to consider using the SRT service), and another process (truncated Poisson) which determines how many trips will be made, given that the student accepts to use the SRT service at least once per week. Model formulation and estimation of the best fitting models (HP) will be shown below.

Equation 41 expresses the logarithm of the mean number of trips λ that is assumed to be an exponential as a linear function of the predictors.

$$log(\lambda_{nt}) = Intercept + \beta_{Delta\ Fare} \times Delta\ Fare_{nt} + \beta_{Delta\ Time} \times Delta\ Time_{nt} + \beta_{MAWT} \times MAWT_{nt} + \beta_{WIFI} \times WiFi\ Dummy_{nt} + \beta_{Minivan} \times Minivan\ Dummy_{nt} + + log(current\ trips_n) + z_{nt} b_n \quad (41)$$

Equation 42 shows the equation of the Binary Logit model used to model the zero-inflation occurrences. φ_{nt} is the probability of observing a zero count in choice scenario *t* for individual *n*.

$$logit(\varphi_{nt}) = Intercept + \beta_{Delta\ Fare} \times Delta\ Fare_{nt} + \beta_{Delta\ Time} \times Delta\ Time_{nt} + \beta_{MAWT} \times MAWT_{nt} + \beta_{WIFI} \times WiFi\ Dummy_{nt} + \beta_{Minivan} \times Minivan\ Dummy_{nt} + \beta_{Gender} \times Male\ Dummy_n + z'_{nt} b'_n \quad (42)$$

In order to account for the random effect present in the choices of a student, two scalar normally distributed random effects b_n and b'_n are added to the above equations where z_{nt} and z'_{nt} are set to 1.

c) Hurdle Models' Estimation Results

Parameter estimates, Wald standard errors (S.E.), p-values for the covariates, and the variance of the random intercept b ($\sigma_b{}^2$) in the two RE HP Models are shown in Table 7. Models were estimated using the glmmADMB package in R. By default this package in R uses the Laplace approximation for maximum likelihood estimates, which is superior to other methods used by other mixed model routines in R. Therefore, the likelihood values are used to construct the AIC and BIC criteria and to perform other statistical tests such as the likelihood ratio test.

Table 7: Estimation Results of Hurdle Poisson Models

	RE HP for Private Car Users			**RE HP for Public Transport Users**		
Binary Logit (Stage 1)						
Variable	Estimate	S.E.	Pr(>|z|)	Estimate	S.E.	Pr(>|z|)
Intercept	-3.29	0.366	<0.0001	-2.45	0.402	< 0.0001
Delta Time (Hours)	9.33	1.08	< 0.0001	6.71	1.11	< 0.0001
Delta Fare (1,000 L.L.)	0.676	0.0503	< 0.0001	1.37	0.102	< 0.0001
Minivan Dummy	0.990	0.124	< 0.0001	0.921	0.174	< 0.0001
MAWT	5.08	1.00	< 0.0001	4.25	1.61	0.0082
Wi-Fi Dummy	-0.740	0.122	< 0.0001	-0.416	0.171	0.015
Male Dummy	-			0.894	0.456	0.050
$\sigma_b{}^2$	14.6			7.78		
$S.E._{\sigma_b{}^2}$	2.08			1.42		
Observations	320			188		
LL	-1279.7			-680.615		
AIC	2573.4			1377.2		
BIC	2614.31			1412.441		

	RE HP for Private Car Users			RE HP for Public Transport Users		
Truncated Poisson (Stage 2)						
Intercept	-0.134	0.0846	< 0.0001	-0.281	0.855	0.0010
Delta Time (Hours)	-1.51	0.366	< 0.0001	-0.689	0.427	0.11
Delta Fare (1,000 L.L.)	-0.0891	0.0136	< 0.0001	-0.136	0.0298	< 0.0001
Minivan Dummy	-0.127	0.0446	0.0043	-0.120	0.0720	0.096
MAWT	-0.834	0.374	0.026	-0.491	0.0298	0.47
Wi-Fi Dummy	0.0747	0.0455	0.10	0.0505	0.0720	0.48
${\sigma_b}^2$	0.309			0.331		
$S.E._{{\sigma_b}^2}$	0.0443			0.0596		
Observations	268			150		
LL	-1980.8			-1010.9		
AIC	3975.5			2035.8		
BIC	4011.36			2066.06		

In both binary Logit models, all covariates were significant at the 95% confidence level except the gender variable, which was significant in the PT users' model only. In the truncated Poisson models, MAWT and Wi-Fi Dummy were not significant in the PT users' model. On the other hand, in the Private Cars users' model, all variables were significant, knowing that the Wi-Fi Dummy was only significant at the 90% significance. The signs of the coefficients make sense and the gender variable indicates that males are less likely to shift to the new SRT service. There is a negative sign associated with the *Delta Fare*, *Delta Time*, and MAWT for both data sets' models in stage 2, meaning that the high values associated with each of these variables lead to less weekly trips made by the students. Also, students are not in favor of a minivan as compared to the use of a small vehicle as a negative value is also associated with the Minivan Dummy coefficient. Wi-Fi Dummy coefficient is positive indicating that the Mobile application and free Wi-Fi connectivity in the Shared-Ride Taxi encourage PT and PC users to use the new taxi service. In stage 1, the Hurdle Poisson model generates the zero observations; this justifies the signs associated with the coefficients of the Binary Logit model, which appear to be opposite to those in the truncated Poisson model.

4.5.2.2. Multinomial Logit Model

Two MNL models for Private Car and Public Transport users are developed with 6 alternatives (Zero, One, Two, Three, Four, and Five weekly trips). A number of different model specifications were tested until the results were reached. The presented model specifications were concluded on the basis of parameter estimates' signs and statistical goodness-of-fit measures such as the likelihood ratio test, the robust *t*-tests, and the adjusted rho-square statistics. The models were estimated by means of numerical integration (required for maximum likelihood estimation) using the software package Python Biogeme (Bierlaire and Fetiarison, 2009).

The systematic utility equations ($\boldsymbol{V_0}, \boldsymbol{V_1}, \ldots, \boldsymbol{V_5}$) of the six alternatives for the PT data set model for individual *n* and choice scenario *t* are shown below.

$$\boldsymbol{V_{0_{nt}}} = a_n \qquad (43)$$

$$\boldsymbol{V_{1_{nt}}} = ASC_1 + \beta_{Delta\ Fare_1} \times Delta\ Fare_{nt} + \beta_{delta\ Time_1} \times Delta\ Time_{nt} + \beta_{MAWT_{1,2,3}} \times MAWT_{nt} + \beta_{WiFi_{1,2,3}} \times WiFi\ Dummy_{nt} + \beta_{Minivan\ 1} \times Minivan\ Dummy_{nt} + \beta_{Gender} \times Male\ Dummy_n \qquad (44)$$

$$\boldsymbol{V_{2_{nt}}} = ASC_2 + \beta_{Delta\ Fare_2} \times Delta\ Fare_{nt} + \beta_{delta\ Time_2} \times Delta\ Time_{nt} + \beta_{MAWT_{1,2,3}} \times MAWT_{nt} + \beta_{WiFi_{1,2,3}} \times WiFi\ Dummy_{nt} + \beta_{Minivan\ 2} \times Minivan\ Dummy_{nt} + \beta_{Gender} \times Male\ Dummy_n \qquad (45)$$

$$\boldsymbol{V_{3_{nt}}} = ASC_3 + \beta_{Delta\ Fare_3} \times Delta\ Fare_{nt} + \beta_{delta\ Time_3} \times Delta\ Time_{nt} + \beta_{MAWT_{1,2,3}} \times MAWT_{nt} + \beta_{WiFi_{1,2,3}} \times WiFi\ Dummy_{nt} + \beta_{Minivan\ 3} \times Minivan\ Dummy_{nt} + \beta_{Gender} \times Male\ Dummy_n \qquad (46)$$

$$\boldsymbol{V_{4_{nt}}} = ASC_4 + \beta_{Delta\ Fare_4} \times Delta\ Fare_{nt} + \beta_{delta\ Time_4} \times Delta\ Time_{nt} + \beta_{MAWT_{4,5}} \times MAWT_{nt} + \beta_{WiFi_{4,5}} \times WiFi\ Dummy_{nt} + \beta_{Minivan\ 4} \times Minivan\ Dummy_{nt} + \beta_{Gender} \times Male\ Dummy_n \qquad (47)$$

$$\boldsymbol{V_{5_{nt}}} = ASC_5 + \beta_{Delta\ Fare_5} \times Delta\ Fare_{nt} + \beta_{delta\ Time_5} \times Delta\ Time_{nt} + \beta_{MAWT_{4,5}} \times MAWT_{nt} + \beta_{WiFi_{4,5}} \times WiFi\ Dummy_{nt} + \beta_{Minivan\ 5} \times Minivan\ Dummy_{nt} + \beta_{Gender} \times Male\ Dummy_n \qquad (48)$$

The sole difference in the specifications between the users of PT and PC is the inclusion of the gender variable. $Male\ Dummy$ was discarded from the Private Car data set model because it proved to be insignificant. Socioeconomic characteristics such as family income, student's class/grade, number of available vehicles in the family, and the number of licensed drivers in the family were not significant when introduced to the systematic equations of the two data sets' models. Notably, when the number of trips that the student makes using his/her current mode of commute is added as an explanatory variable to the utility of both data sets, it did not prove to be significant as well. For that reason, it was excluded.

The random intercept a_n in V_0 accounts for unobserved characteristics of each student that may introduce correlation across the responses of a given student in multiple choice experiments. It is only included in the Zero alternative and it is specified to follow a multivariate normal distribution. In the MNL estimation, the standard deviationσ of a will be estimated.

The descriptions coupled with the variables shown in Table 5 in the previous section still apply. The change in travel time ($Delta\ Time$), change in travel fare ($Delta\ Fare$), and maximum allowable waiting time for late pick-up and early drop-off ($MAWT$) are continuous while all other attributes are dummy variables. Table 8 summarizes the estimation results of the MNL model.

Table 8: MNL Model Estimation Results

		Public Transport Data Set			Private Car Data Set		
Variable	**Alternative (Number of Trips)**	**Estimate**	**Robust S.E.**	**P-value**	**Estimate**	**Robust S.E.**	**P-value**
ASC (Alternative Specific Constant)	*ZERO*	0	--	--	0	--	--
	ONE	-0.151	0.486	0.76	1.33	0.418	0.00
	TWO	0.803	0.463	0.08	1.65	0.392	0.00
	THREE	1.20	0.473	0.01	1.97	0.411	0.00
	FOUR	-0.292	0.556	0.60	1.50	0.460	0.00
	FIVE	1.95	0.428	0.00	2.32	0.447	0.00
Delta Time (Hours)	*ZERO*	0	--	--	0	--	--
	ONE	-5.30	1.51	0.00	-8.05	1.37	0.00
	TWO	-4.39	1.41	0.00	-8.28	1.25	0.00
	THREE	-5.19	1.41	0.00	-10.0	1.27	0.00
	FOUR	-9.08	1.47	0.00	-10.5	1.71	0.00
	FIVE	-8.69	1.84	0.00	-11.9	1.62	0.00
Delta Fare (1,000 L.L.)	*ZERO*	0	--	--	0	--	--
	ONE	-1.05	0.147	0.00	-0.615	0.0676	0.00
	TWO	-1.35	0.148	0.00	-0.678	0.0666	0.00
	THREE	-1.48	0.153	0.00	-0.708	0.0670	0.00
	FOUR	-1.19	0.189	0.00	-0.749	0.0750	0.00
	FIVE	-1.79	0.163	0.00	-0.749	0.0799	0.00
Minivan Dummy	*ZERO*	0	--	--	0	--	--
	ONE	-0.420	0.226	0.06	-0.731	0.160	0.00
	TWO	-0.905	0.259	0.00	-0.814	0.147	0.00
	THREE	-1.05	0.212	0.00	-1.05	0.152	0.00
	FOUR	-0.841	0.206	0.00	-1.39	0.213	0.00
	FIVE	-1.3	0.292	0.00	-1.28	0.148	0.00
MAWT (Hours)	*ZERO*	0	--	--	0	--	--
	ONE						
	TWO	-3.22	1.59	0.04	-5.10	1.03	0.00
	THREE						
	FOUR	-5.95	2.11	0.00	-5.25	1.14	0.00
	FIVE						
WiFi Dummy	*ZERO*	0	--	--	0	--	--
	ONE						
	TWO	0.280	0.159	0.08	0.764	0.125	0.00
	THREE						
	FOUR	0.506	0.209	0.02	0.706	0.130	0.00
	FIVE						
Male Dummy	*ZERO*	0	--	--	(Insignifi	(Insignif	(Insign

Variable	Alternative (Number of Trips)	Public Transport Data Set			Private Car Data Set		
		Estimate	Robust S.E.	P-value	Estimate	Robust S.E.	P-value
	ONE				cant)	icant)	ificant)
	TWO		0.466				
	THREE	-0.792		0.09			
	FOUR						
	FIVE						
σ_a		2.94	0.309	0.00	4.01	0.354	0.00
Summary Statistics							
Sample size		**188**			**320**		
Final log-likelihood		**-1518.51**			**-3174.18**		
Rho squared		**0.436**			**0.308**		

Note: Empty cells corresponding to the Zero alternatives indicate that the value of the parameter has been fixed to 0, for estimation purposes.

Estimation Results of the above MNL Models give reasonable inference about the effect of the variables on the propensity of the students in using the new Shared-Ride Taxi. Negative values returned for the coefficients of time, fare, and MAWT suggest that an increase in those variables lead to a decrease in the utility of each of the alternatives. This is reasonable because values of the Zero alternative are considered as the base and are fixed to zero. As is the case with the HP models, the Male Dummy variable is significant and negative, indicating that female PT users are more interested in using the SRT service than males. Moreover, it is noticed that the values of the coefficients increase by alternative, this indicate that, in general, students who are willing to use the SRT more frequently, will be more affected by the variables.

4.5.3. Model Selection and Outlier Analysis

After presenting both MNL and Hurdle Poisson Models, this section aims at selecting one of the two models, before latent variables are added to the preferred model. To do that, R^2 statistic (or ρ^2 for logit models) and the market segment prediction tests are used as a basis for comparison.

Table 9 shows the log-likelihoods (LL) of the fitted and the null models for both data sets. The null log-likelihood is that which entails that all alternatives are equally likely to be chosen, the value of which is returned by Python Biogeme as the "initial log-likelihood" whereby initial values of the parameters as well as the standard deviation of the panel effect are set to 0.

Table 9: HP and MNL Models' Statistics Compared

<table>
<tr><th colspan="2" rowspan="2"></th><th colspan="2">Car Model</th><th colspan="2">PT Model</th></tr>
<tr><th>HP</th><th>MNL</th><th>HP</th><th>MNL</th></tr>
<tr><td rowspan="2">LL (Fitted Model)</td><td>Stage 2</td><td>-1980.77</td><td rowspan="2">-3174.18</td><td>-1010.9</td><td rowspan="2">-1518.51</td></tr>
<tr><td>Stage 1</td><td>-1279.69</td><td>-680.615</td></tr>
<tr><td colspan="2">LL (Null Model)</td><td colspan="2">-4586.68</td><td colspan="2">-2694.65</td></tr>
<tr><td rowspan="2">Parameters</td><td>Stage 2</td><td>7</td><td rowspan="2">25</td><td>7</td><td rowspan="2">26</td></tr>
<tr><td>Stage 1</td><td>7</td><td>8</td></tr>
<tr><td colspan="2">Sample size</td><td colspan="2">320</td><td colspan="2">188</td></tr>
<tr><td colspan="2">R²</td><td>0.289</td><td>0.308</td><td>0.372</td><td>0.436</td></tr>
<tr><td colspan="2">Adjusted R²</td><td>0.286</td><td>0.303</td><td>0.367</td><td>0.427</td></tr>
<tr><td colspan="6"><u>Note</u>: Stage 1 is Binary Logit whereas Stage 2 is Truncated Poisson.</td></tr>
</table>

Adjusted R^2 criterion indicates that the Multinomial Logit Model contributes to an improved model fit since its values are higher for MNL models than those returned by

HP models. Note that the adjusted R^2 adjusts for the increased number of parameters. Assuming that *k* is the number of predictors in the model, adjusted R^2 takes the form shown in Equation 49.

$$adjusted\ R^2 = 1 - \frac{LL_M - k}{LL_0} \qquad (49)$$

where LL_M is the log likelihood of the fitted model and LL_0 is that of the null model.

The second measure of fit is the one that is based on the accuracy of the model to predict the shares of the six alternatives. For the estimation sample in a Logit model, the predicted and the observed shares are equal if a full set of alternative specific constants is included in the model specification. This condition is not valid to sub sets of the sample used for estimation unless market segment specific constants are also included in the model. Consequently, market segment prediction tests are performed as another model selection criterion.

For a better model, one would expect the predicted frequencies to be close to the corresponding observed frequencies. As such, the observed shares of each alternative and the predicted shares by each of the MNL models and the Hurdle Poisson models are examined using two market segment tests. The first market segment inspected is the student's class (year of study) and the second is the number of vehicles available for the student's family. In Figures 13, 14, 15, and 16, the shares of each alternative for each market segment (New and Old Students) are illustrated. The shares' estimation (in percent) considers the eight choices pertaining to the eight different scenarios by every student. As for the other test, two market segments are also defined. Results of both tests are included in Appendix A.

New Students are freshman or first-year students, whereas Old Students are those who have been enrolled at AUB for 2 years at least.

As is evident in the four figures below, the MNL models' predictions seem to replicate the observed shares in almost every case. While Hurdle Poisson was able to handle the excessive zero outcomes, this model did not provide accurate predictive power for alternatives 1, 3, 4, and 5, as compared to the Multinomial Logit model. It, however, was able to reproduce accurate predictions for alternative 2, but still not as much as that obtainable by the MNL. The good predictive power associated with the zero outcomes and alternative 2 do not justify the superiority of this model over the MNL model.

The two comparison criteria suggest that the MNL model provides a better fit in both data sets. Henceforth, model analysis shall proceed with the specifications and estimation results of the MNL models.

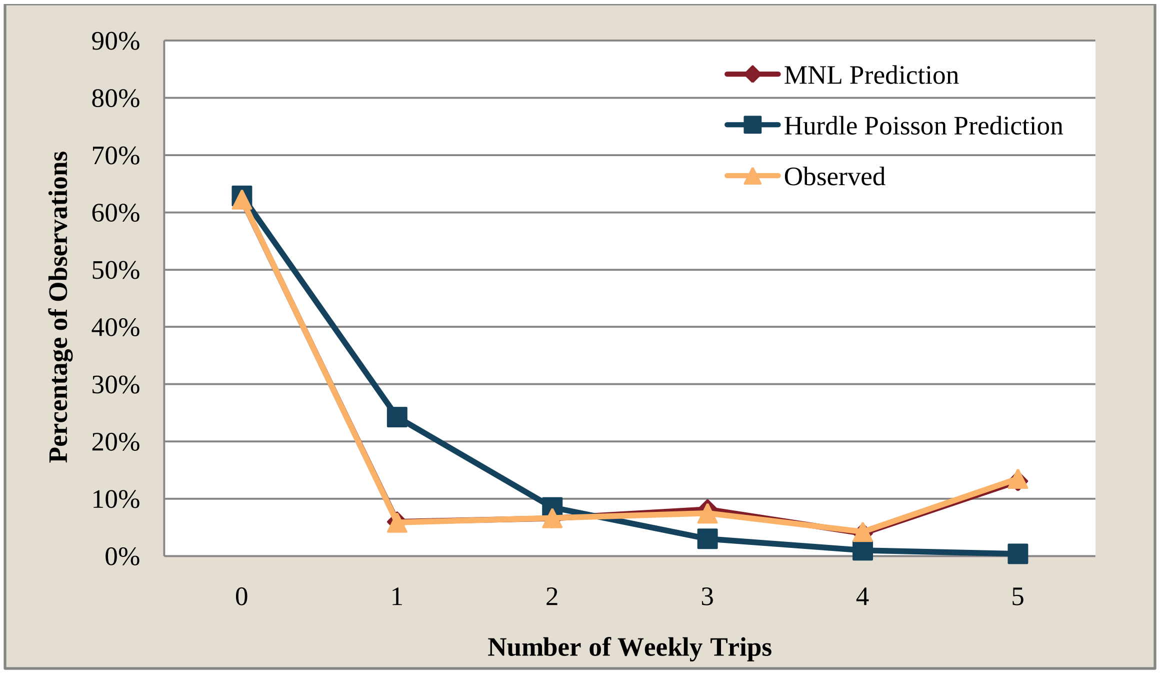

Figure 13: Percentage of Observations of New Students in the PT Data Set Willing to Use the SRT Service 0, 1, 2, 3, 4, and 5 Times Weekly

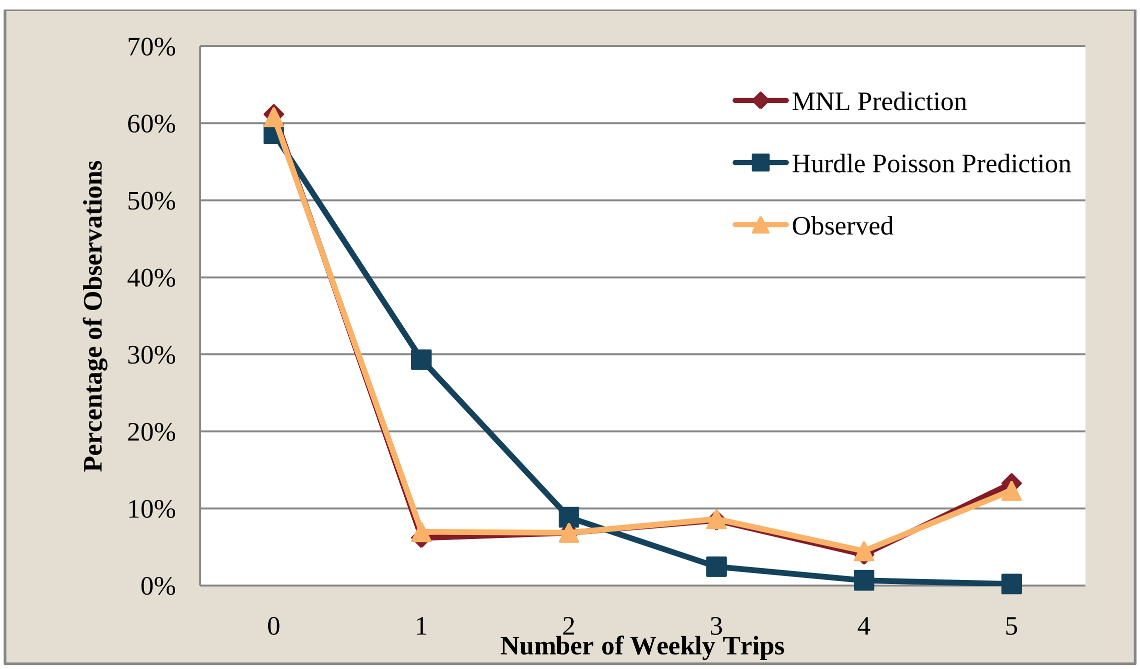

Figure 14: Percentage of Observations of Old Students in the PT Data Set Willing to Use the SRT Service 0, 1, 2, 3, 4, and 5 Times Weekly

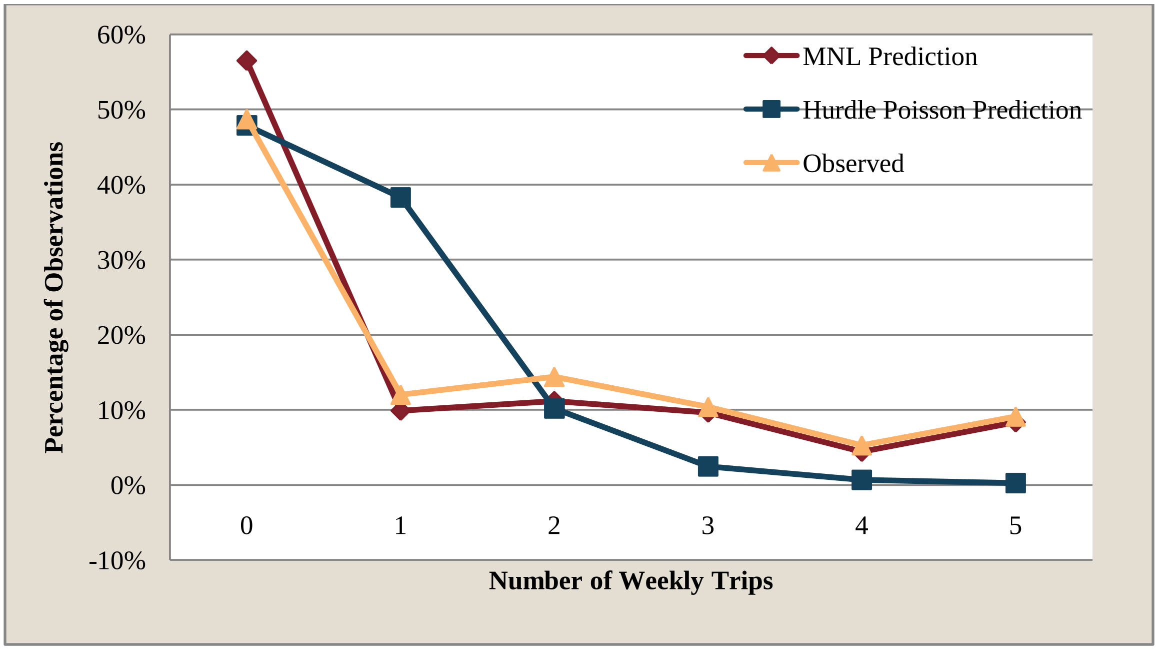

Figure 15: Percentage of Observations of New Students in the PC Data Set Willing to Use the SRT Service 0, 1, 2, 3, 4, and 5 Times Weekly

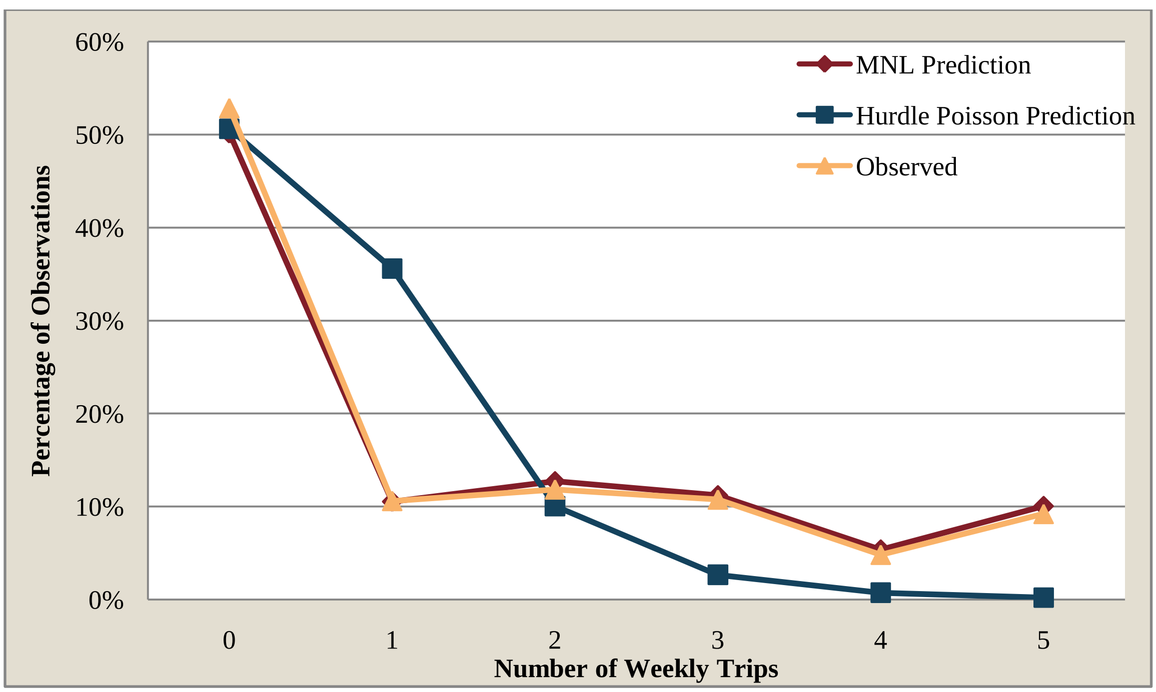

Figure 16: Percentage of Observations of Old Students in the PC Data Set Willing to Use the SRT Service 0, 1, 2, 3, 4, and 5 Times Weekly

Upon choosing to proceed with the MNL model, an important prediction test is the outlier analysis. For all observations in the estimation data sets, the predicted choice probabilities of the chosen alternative are calculated. Next, the predicted values are inspected whereby those that are extremely low are scrutinized. Since every student answered 8 choice scenarios, each of them is thus associated with 8 different choice probabilities. Analysis began with a low limit of 0.01. Students associated with at least four predicted choice probabilities less that 0.01 are selected for further inspection. No observations met this condition in the PT data set, while 2 observations in the PC data set did. The search for less serious outliers continued, where 9 PC users and 7 PT users were associated with predicted probabilities of 0.05 or less in at least 4 choice experiments. Checking for data errors in the suspicious observations, no substantive information suggested that they should be eliminated, or that edits to the models' specifications should be done. Also, regressions were performed with and without them to examine the sensitivity of the estimation results to the presence of these observations. Eliminating the 16 questionable observations from the two data sets led to only minor influence on the estimation results as the parameter estimates, variables' significance, and the values of time did not vary substantially. The PT and PC data sets remain with 188 and 320 observations, respectively.

4.5.4. Integrated Choice and Latent Variable (ICLV) Model Estimation Results

A section of the survey was designed to capture students' attitudes (as latent or unobserved variables) towards the internet and technology, the Shared-Ride Taxi service, and ridesharing. The ICLV models were estimated using the software package Python Biogeme and maximizing the likelihood function was done through numerical integration. Students indicated their level of agreement with statements related to the aforementioned attributes on a scale from 1 to 7, 1 being strongly disagree, and 7 being strongly agree. Using several of these statements, two attitudes were measured. Ridesharing (RS) attitude was measured as part of the PT Model's specifications, and Taxi-sharing (TS) attitude was included in the PC users' model. The indicators (manifest variables) and their corresponding descriptions are shown in Table 10. In the last column of the table, the latent variables which each of the indicators captures are specified. The latent variables were only included in the utility equation of the zero alternative to test their effect on the likelihood to switch or not, rather than the actual frequency of usage. The panel effect a_n that was previously integrated in the MNL models is not included in ICLV models because it is assumed that the latent variable will account for individuals' personal characteristics as their values are fixed for each individual. The systematic utility equations ($\boldsymbol{V_0}, \boldsymbol{V_1}, \dots, \boldsymbol{V_5}$) of the 6 alternatives in the ICLV model for the PT data set model are shown below. PC data set's model specifications are only different at the level of the gender dummy variable, the number of passengers in the vehicle (Minivan Dummy Variable), and the latent variable included in $\boldsymbol{V_0}$.

$$\boldsymbol{V_{0_{nt}}} = \beta_{Rides\,haring} \times F_{Rides\,haring_{\ n}} \qquad (50)$$

$$\boldsymbol{V_{1_{nt}}} = ASC_1 + \beta_{Delta\ \ Fare_1} \times Delta\ Fare_{nt} + \beta_{delta\ \ Time_1} \times Delta\ Time_{nt} + \beta_{MAWT_{1,2,3}} \times MAWT_{nt} + \beta_{WiFi_{1,2,3}} \times WiFi\ Dummy_{nt} + \beta_{Minivan\ 1,2,3} \times Minivan\ Dummy_{nt} + \beta_{Gender\ 1,2,3} \times Male\ Dummy_n \qquad (51)$$

$$\boldsymbol{V_{2_{nt}}} = ASC_2 + \beta_{Delta\ \ Fare_2} \times Delta\ Fare_{nt} + \beta_{delta\ \ Time_2} \times Delta\ Time_{nt} + \beta_{MAWT_{1,2,3}} \times MAWT_{nt} + \beta_{WiFi_{1,2,3}} \times WiFi\ Dummy_{nt} + \beta_{Minivan\ 1,2,3} \times Minivan\ Dummy_{nt} + \beta_{Gender\ 1,2,3} \times Male\ Dummy_n \qquad (52)$$

$$\boldsymbol{V_{3_{nt}}} = ASC_3 + \beta_{Delta\ \ Fare_3} \times Delta\ Fare_{nt} + \beta_{delta\ \ Time_3} \times Delta\ Time_{nt} + \beta_{MAWT_{1,2,3}} \times MAWT_{nt} + \beta_{WiFi_{1,2,3}} \times WiFi\ Dummy_{nt} + \beta_{Minivan\ 1,2,3} \times Minivan\ Dummy_{nt} + \beta_{Gender\ 1,2,3} \times Male\ Dummy_n \qquad (53)$$

$$\boldsymbol{V_{4_{nt}}} = ASC_4 + \beta_{Delta\ \ Fare_4} \times Delta\ Fare_{nt} + \beta_{delta\ \ Time_4} \times Delta\ Time_{nt} + \beta_{MAWT_{4,5}} \times MAWT_{nt} + \beta_{WiFi_{4,5}} \times WiFi\ Dummy_{nt} + \beta_{Minivan\ 4,5} \times Minivan\ Dummy_{nt} + \beta_{Gender\ 4,5} \times Male\ Dummy_{nt} \qquad (54)$$

$$\boldsymbol{V_{5_{nt}}} = ASC_5 + \beta_{Delta\ \ Fare_5} \times Delta\ Fare_{nt} + \beta_{delta\ \ Time_5} \times Delta\ Time_{nt} + \beta_{MAWT_{4,5}} \times MAWT_{nt} + \beta_{WiFi_{4,5}} \times WiFi\ Dummy_{nt} + \beta_{Minivan\ 4,5} \times Minivan\ Dummy_{nt} + \beta_{Gender\ 4,5} \times Male\ Dummy_n \qquad (55)$$

The first sub-model of the latent variable model is a confirmatory factor model relating the latent attitudes to their corresponding manifest variables. A linear model is specified to describe the mapping of the latent variables on the indicators (assumed to be continuous). It is expressed as follows:

$$I_n = \lambda F_n + \nu_n \qquad (56)$$

where *I* is a vector of the manifest variables that are individual specific, λ is a matrix of the factor loadings, and ν is a vector of the measurement errors that are *iid* multivariate normally distributed. Table 11 presents the estimation results of the ICLV choice models. Note that, for the purpose of model identification, the variances of the latent variables are set to 1. For more details on the results of the measurement equations, the reader is referred to Table 13 in Appendix B.

As is the case with the previously estimated MNL models, model specifications were established on the basis of parameter estimates' signs and statistical goodness-of-fit measures. In fact, the signs of the factor loading estimates indicate that all measured attitudes are positive (i.e. a higher value indicates more positive attitudes towards Ridesharing or Taxi-sharing).

Table 10: Latent Variables and their Indicators

Designation	**Description of Indicator (Manifest Variable)**	**Latent Variable *F***
$\boldsymbol{I_1}$	I will feel annoyed if the shared-ride taxi makes a large deviation from the direct route between my residence and AUB gate.	RS and TS
$\boldsymbol{I_2}$	I don't mind if the shared-ride taxi makes several stops to serve other students while I am on board.	TS
$\boldsymbol{I_3}$	I will use the shared-ride taxi more if I can reserve close to my time of departure and not strictly on the day before.	RS and TS
$\boldsymbol{I_4}$	I like sharing rides with others.	RS
$\boldsymbol{I_5}$	I prefer to share rides only with people of the same gender.	(Insignificant)
$\boldsymbol{I_6}$	I am willing to try ridesharing because it allows me to meet new people.	RS
$\boldsymbol{I_7}$	I will pay more to get more technologically advanced products.	(Insignificant)
$\boldsymbol{I_8}$	I often use the internet to plan my daily activities.	(Insignificant)
$\boldsymbol{I_9}$	I use the internet for chatting and entertainment on a daily basis.	(Insignificant)
$\boldsymbol{I_{10}}$	I always rely on the internet to read news and know weather conditions.	(Insignificant)

Table 11: ICLV Model Estimation Results

Variable	Alternative (Number of Trips)	Public Transport Data Set: Estimate	Public Transport Data Set: Robust S.E.	Public Transport Data Set: P-value	Private Car Data Set: Estimate	Private Car Data Set: Robust S.E.	Private Car Data Set: P-value
ASC (Alternative Specific Constant)	*ZERO*	0	--	--	0	--	
	ONE	-0.053	0.487	0.91	1.29	0.418	0.00
	TWO	0.642	0.466	0.17	1.61	0.392	0.00
	THREE	0.964	0.483	0.05	1.93	0.402	0.00
	FOUR	-0.209	0.556	0.71	1.46	0.449	0.00
	FIVE	1.85	0.449	0.00	2.28	0.443	0.00
$Delta\ Time$ (Hours)	*ZERO*	0	--	--	0	--	--
	ONE	-5.59	1.47	0.00	-8.14	1.37	0.00
	TWO	-4.94	1.42	0.00	-8.37	1.24	0.00
	THREE	-5.97	1.41	0.00	-10.1	1.26	0.00
	FOUR	-9.30	1.44	0.00	-10.6	1.70	0.00
	FIVE	-9.45	1.72	0.00	-12.0	1.60	0.00
$Delta\ Fare$ (1,000 L.L.)	*ZERO*	0	--	--	0	--	--
	ONE	-1.00	0.141	0.00	-0.612	0.0669	0.00
	TWO	-1.32	0.149	0.00	-0.675	0.0662	0.00
	THREE	-1.45	0.155	0.00	-0.705	0.0664	0.00
	FOUR	-1.13	0.186	0.00	-0.745	0.0745	0.00
	FIVE	-1.75	0.163	0.00	-0.745	0.0794	0.00
$Minivan\ Dummy$	*ZERO*	0	--	--	0	--	--
	ONE				-0.741	0.160	0.00
	TWO	-0.797	0.177	0.00	-0.823	0.146	0.00
	THREE				-1.06	0.151	0.00
	FOUR	-1.15	0.224	0.00	-1.40	0.212	0.00
	FIVE				-1.29	0.148	0.00
MAWT (Hours)	*ZERO*	0	--	--	0	--	--
	ONE						
	TWO	-3.38	1.57	0.03	-5.04	1.03	0.00
	THREE						
	FOUR	-6.17	1.99	0.00	-5.20	1.14	0.00
	FIVE						

		Public Transport Data Set			Private Car Data Set		
Variable	**Alternative (Number of Trips)**	**Estimate**	**Robust S.E.**	**P-value**	**Estimate**	**Robust S.E.**	**P-value**
WiFi Dummy	*ZERO*	0	--	--	0	--	--
	ONE	0.304	0.160	0.06	0.758	0.125	0.00
	TWO						
	THREE						
	FOUR	0.532	0.194	0.01	0.700	0.130	0.00
	FIVE						
Male Dummy	*ZERO*	0	--	--	(Insignificant)	(Insignificant)	(Insignificant)
	ONE	-0.893	0.506	0.08			
	TWO						
	THREE						
	FOUR	-1.09	0.515	0.03			
	FIVE						
Attitudes as Exogenous Variables in the Utility of the *ZERO* Alternative							
Ridesharing		-2.96	0.316	0.00	(Insignificant)		
Taxi-sharing		(Insignificant)			-4.05	0.360	0.00
Summary Statistics							
Sample size		**188**			**320**		
Final log-likelihood		**-2853.77**			**-4933.31**		
Rho squared		**0.451**			**0.394**		
Note: Empty cells corresponding to the Zero alternative indicate that the value of the parameter has been fixed to 0, for estimation purposes.							

4.5.5. Discussion

In the ICLV models of both data sets, the variables which were found significant are the change in travel time, the allowable time for late pick-up and early drop-off, the high number of passengers in the vehicle dummy (when the vehicle used is a mini-van), and the availability of internet and Wi-Fi dummy. All parameter estimates have the right signs. This suggests that the users will make fewer trips using the SRT service as the

Delta Time and the time for late pick-up and early drop-off increase and when the vehicle used is a mini-van. The change in fare variable is significant for both users meaning that they tend to be less attracted to the new taxi service under high fare.

In the PT data set model, a variable which exhibits importance is the gender. The negative values associated with the SRT alternative signify that females who currently are bus or jitney users have a higher propensity to utilize the SRT service than males. Some studies in the literature of shared-ride transportation corroborate this finding. For example, a study by Buliung et al. (2009) aims at broadening the understanding of carpool use and formation. They discuss number of different factors that were found to be associated with a successful carpooling formation. One finding was that females are 1.3 times more likely to carpool than males. Nevertheless, other studies suggest there are some factors that have potentially negative effects on the females to carpool such as household responsibility, scheduling issues, and the type of employment available to females e.g., Sermons and Koppelman (2001) and Cristaldi (2005). The results in this thesis suggest that female (who currently use public transport modes) are more likely to be in favor of the shared-ride transportation service, signifying that some of these hypotheses do not apply to the females of this particular sample, as the targeted population are students who are very unlikely to be employees or responsible for a household.

As we hypothesized in the model formulation, there are underlying latent variables that do exert important effects on travelers' choices. This is valid as two attitudes were identified that impact the choice model (Ridesharing and Taxi-sharing attitudes with P-values=0.00). However, results show that not all students have the same attitudes and not all attitudes influence students' choices in the same manner. Results imply that the attitude towards ridesharing and the attitude towards taxi-sharing influence the choices of PT and PC users, respectively. The positive values of the estimated I_2, I_3, I_4, and I_6 reveal that the measured attitude towards ridesharing is positive. Also, the signs associated with the manifest variables of the TS variables indicate a positive attitude towards taxi-sharing. The significant (and negative) coefficients of these attitudes when included as exogenous variables in the utility of the zero utility equation reveal that the more positive the value of these variables is, the more likely the student is expected to commute by SRT (or the less likely he/she is to choose the 0 alternative). Therefore, there are two attitudes that are significantly affecting the choice; students who are more accepting of the new taxi service and of sharing a ride with others are more willing to shift. Moreover, other examined manifest variables such as those related to the internet and technology were insignificant (indicators I_7 through I_{10}), revealing that students' choice behavior is independent from their attitudes toward the internet and the extent of involvement in the use of Wi-Fi and technologically advanced products.

Attitudes in the two models were purely exogenous, entering the utility equation of the Zero trip alternative, since no significant explanatory variables were found to explain these attitudes.

4.5.6. Value of Time Analysis

The value of time is an important concept in transport planning that helps in understanding the tradeoff between travel time and cost in travel demand models. This value is used to allocate a monetary value of money to the savings in the travel time when alternative transport projects are assessed (Ben-Akiva and Lerman, 1985). The standard method of calculating the VOTs is used. It involves using the tradeoff ratio obtained by the coefficients of time and cost estimated in the models. In this application, the values of time are derived as the ratio of change in travel time coefficient to the change in fare coefficient. Since the *Delta Time* and *Delta Fare* coefficients were specified as alternative specific, five different values of time (in L.L./hour) were estimated pertaining to the five alternatives 1, 2, 3, 4, and 5 weekly SRT trips. Figure 17 illustrates the values of time associated with both groups of users.

First, it is evident that the VOT of car users is significantly greater than those who currently commute by public transport modes. This observation is in conformity with the observation which indicates that car users come from wealthier families than those who commute by public transport. Wardman (2004) suggests that it is expected that the value of time varies across users of different modes, not only due to income differences, but also because of the mode in which time is spent which involves differences in the comfort and conditions of travel. Second, there isn’t an obvious distinction in values of time by alternatives. That is, even though there is a tendency for the private car users who are willing to use the SRT more frequently (i.e., alternative 5 is associated with a VOT of 16,107 L.L./hour) to have a high VOT as compared to those willing to use the SRT once or twice (13,301 L.L./hour and 12,400 L.L./hour), it is less apparent when it comes to the users of public transport.

Danaf et al. (2014) estimated the VOT of AUB students in 2010 and found it to be equal to 10,144 L.L. (6.8$/hour). Accounting for 2% inflation in 2014 (according to the Ministry of Finance in Lebanon), the more recent VOT according to Danaf et al. (2014) is 10,980 L.L./hour (7.3$/hour), which seems to be close to that obtainable using the VOT values presented in this thesis.

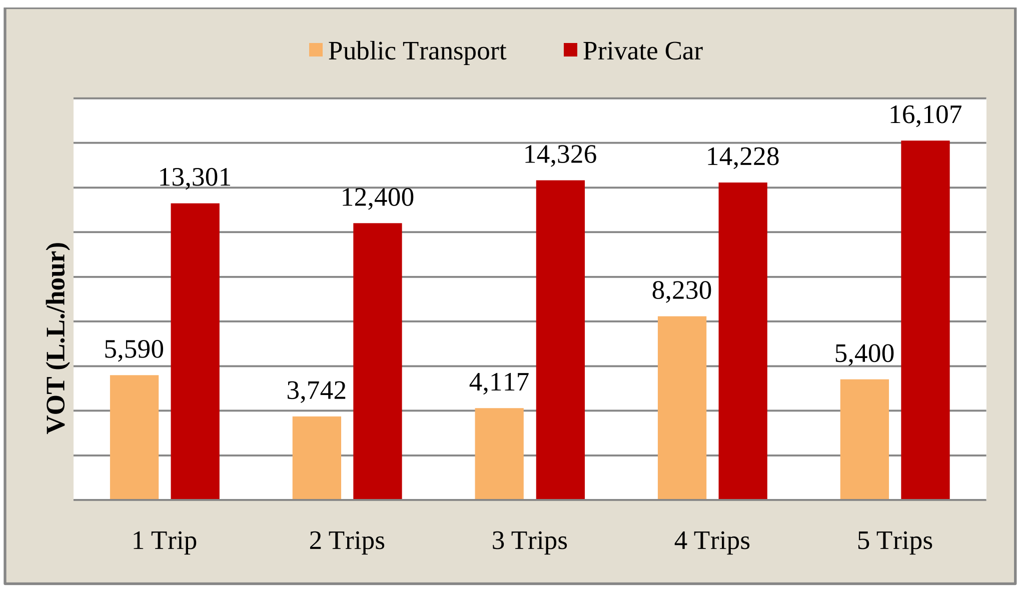

Figure 17: Values of Time (in L.L./hour) for PT and PC Users Derived Using ICLV Models (By Alternative)

4.5.7. Elasticities

Aggregate level elasticities are estimated in order to examine the effect of continuous independent variables in the ICLV Models. As reported by Bhat and Pulugurta (1998) the disaggregate level elasticity of MNL is computed as follows:

$$E_{x_k}^{P_n(i)} = x_k \left[\beta_{ki} - \sum_j P_n(i).\beta_{kj} \right] \quad (57)$$

where x_k is covariate *k*, $P_n(i)$ is the unconditional probability that individual *n* chooses alternative *i*, and β_{kj} is the coefficient of covariate *k* associated with alternative *j*. The aggregate-level elasticity is then calculated as follows:

$$E_{x_k}^{\bar{P}_n(i)} = \frac{\sum_{n=1}^{N} P_n(i).E_{x_k}^{P_n(i)}}{\sum_{n=1}^{N} P_n(i)} \quad (58)$$

Potoglou and Kanaroglou (2006) examined the influence of family structure and other socioeconomic characteristics on the number of cars owned by a household. In line with their analysis, elasticities associated with the change in SRT fare, travel time, and the maximum allowable waiting time were estimated. Since a closed-form expression of elasticities cannot be obtained for the ICLV model, elasticities can be estimated by simulation. Elasticities for the continuous exogenous variables can be viewed as the relative change in expected aggregate shares due to an increase of 1% in a given variable across all students. The elasticity value associated with SRT fare was also calculated in a similar manner, even though SRT fare, as an absolute value was not part of the models' specifications (which included delta fare instead).

As can be seen in Figure 18, elasticities of the Zero alternative are shown to capture users' likelihood to shift to the SRT service, rather than the frequency of using the SRT service.

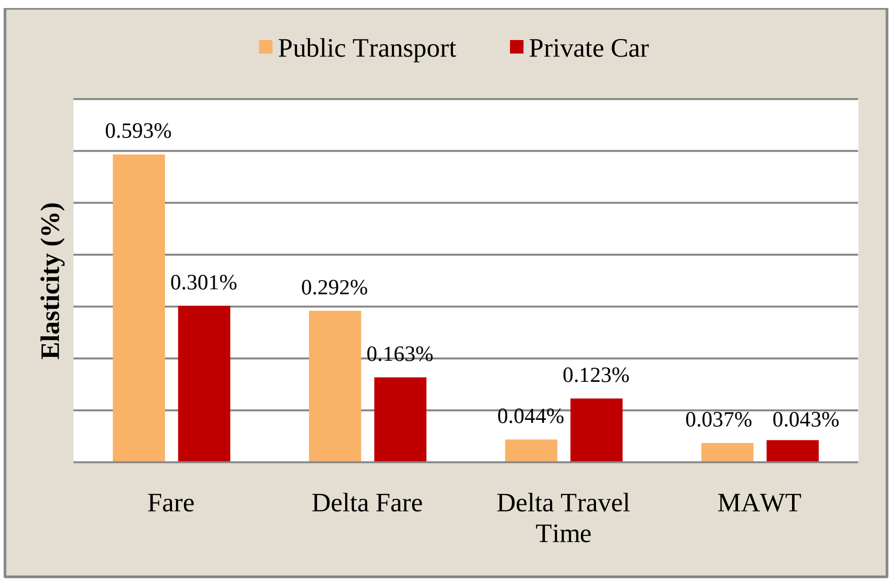

Figure 18: Aggregate-Level Elasticity Effects of Fare, Delta Fare, Delta Time, and MAWT for the Zero Weekly Trips Alternative

The following findings can be noted:

- SRT ridership in the case of PT and PC users is mostly influenced by an increase or decrease in the fare of the SRT service.
- Students are sensitive to the absolute value of the SRT fare more than the relative value of the fare with respect to that associated with their current mode.
- Students are almost not responsive to changes in the MAWT. For example, the elasticity value 0.037 for PT users means that for each 10% increase in MAWT, there is only 0.37% decrease in SRT ridership. Conversely, a 10% decrease in MAWT will increase the odds of using the SRT service by 0.37%. This result indicates that SRT ridership is relatively inelastic to changes in the maximum allowable waiting time.
- Private Car and Public Transport users are less sensitive to changes in the travel time than the fare of the SRT (and the change in the fare). Public Transport users are influenced by changes in the fare and the delta fare more than PC users, meaning that a 10% difference in the SRT fare and delta fare affect PT ridership more than that of PC users. PT users are also insensitive to travel time changes (elasticity=0.044%).

Ina research on price elasticity of vanpool by Wambalaba et al. (2004), the value of elasticity was equal to -0.61. The study concluded that PT riders for most urban areas tend to be captive riders and might not easily change modes due to changes in fares. Another study by Koffman (2007) found that a fare elasticity of demand for paratransit ridership is –0.77, meaning that a 10% change in fares corresponds to a 7.7% change in demand in the opposite direction. This study also indicated that elasticity with respect to paratransit travel time is –0.5. Such results are in conformity with the results presented in this thesis where the elasticity for trip-making is quite small with respect to time and fare, and generally, fare elasticities (and delta fare) seem to be higher than travel time (and MAWT) elasticities.

The reader is referred to Appendix C to examine the elasticity values corresponding to all alternatives.

4.5.8. Policy Analysis

This part of thesis comprises the development of robust policy scenarios that serve as a decision support tool for the provision of an SRT service for AUB students. The most popular approach is the sample enumeration which is used to compute the utilities and probabilities of all alternatives, using the estimated choice models. At this level, the choice models' estimation results are used as a basis for the forecasting. Only the five variables of interest are varied to test the effect of their change on the shares of the alternatives. Three scenarios are defined in an attempt to capture what service attributes have substantial effects on the percentage of AUB student willing to utilize SRT 0, 1, 2, 3, 4, or 5 times weekly and what trades-offs can be made among them. Figure 19 illustrates the service attribute values that are associated with each of Premium, Basic, and Economy services.

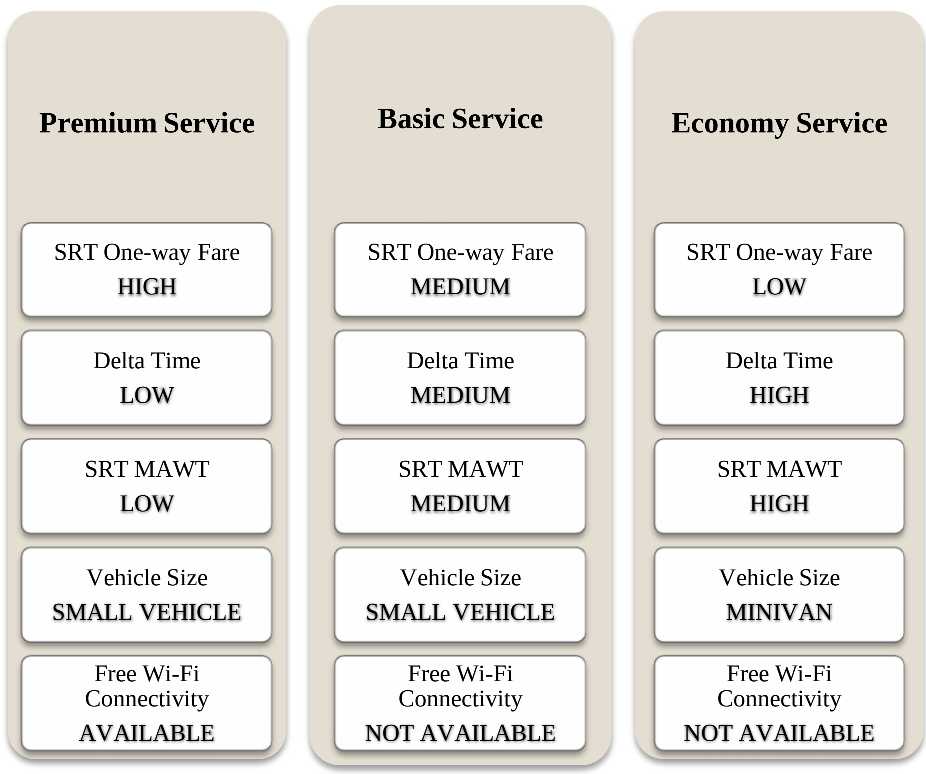

Figure 19: Premium, Basic, and Economy Services Defined

Values of the attributes for the Premium and Economy services, for every mode and every Region (A, B, or C) are included in Appendix D (Tables 15 and 16). Table 12 below shows the values associated with the attributes of Basic service.

Figures 20 and 21 show the percentage of PC and PT students willing to utilize the SRT service 0 times, 1 and 2 times, and 3 times or more (3+) weekly under the three different scenarios. Under the Premium and Basic scenarios, PC users are more willing to shift to the SRT service, unlike the PT users, who are more attracted to SRT when the Basic or Economy services are applied than it is the case with the Premium service.

Another notable observation is that the highest percentage of students willing to use SRT is associated with the PC users under the Premium scenario, whereas the lowest value is that associated with the PT users, under the same service. A Basic service representing a practical service with medium values of all attributes is the most favorable to PT users, whereas for PC users, a Premium service with very favorable but a high taxi fare is most attractive. Further, a Premium service is the least favorable for PT users while an economy service is the least favorable for car users. These findings corroborate on the heterogeneity present among users of different modes.

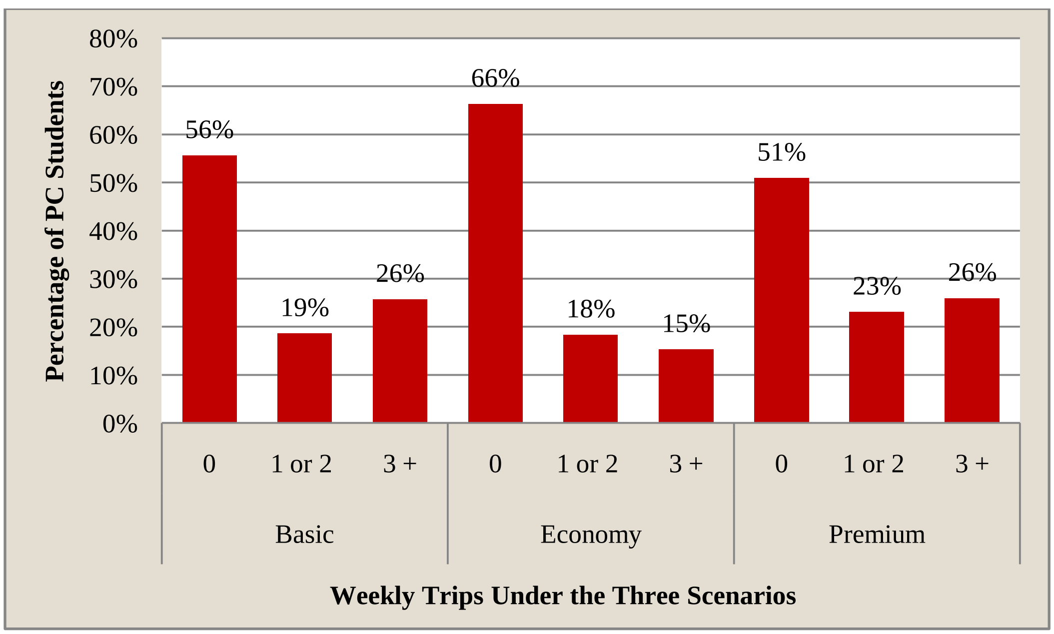

Figure 20: Percentage of PC Users Willing to Use the SRT 0, 1 or 2, and 3 Times or More Weekly as a Function of the Three Service Scenarios

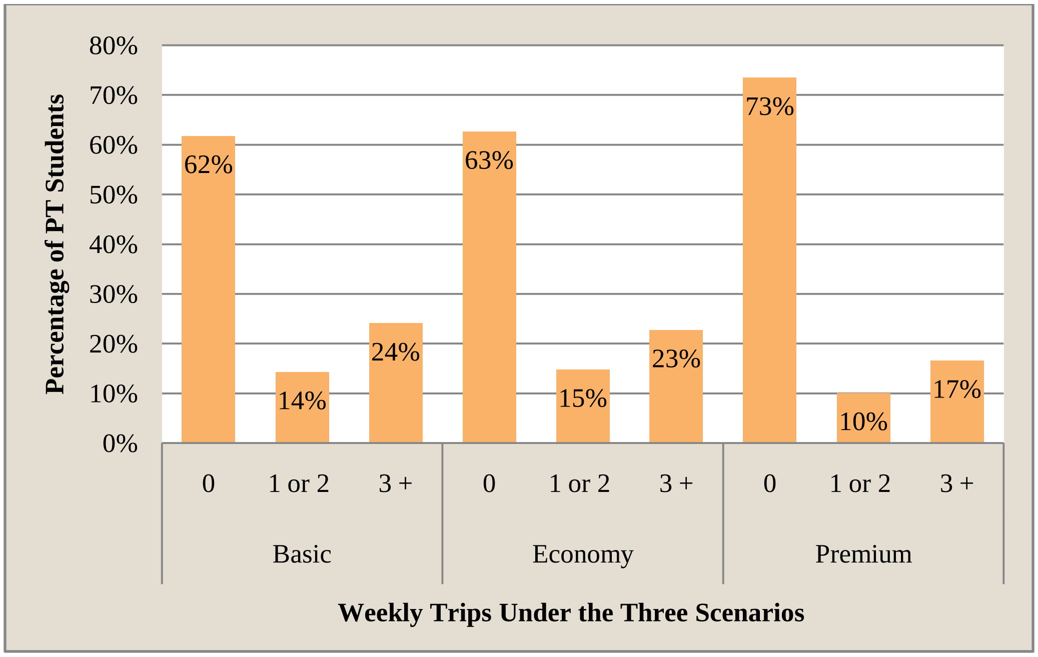

Figure 21: Percentage of PT Users Willing to Use the SRT 0, 1 or 2, and 3 Times or More Weekly as a Function of the Three Service Scenarios

The results of this model were also used to examine the effect of the variation of the level-of-service on ridership in the Basic scenario. Policy analysis of different hypothetical situations was conducted. Three levels of each attribute (low, medium, high) were tested under the Basic scenario. The values defined for the Basic scenario have medium magnitudes, as shown in Table 12. In the developed hypothetical scenarios, the aim is to evaluate how an increase and decrease in each of the three attributes (taxi fare, change in travel time, and maximum allowable waiting time for late pick-up and early drop-off) along with the use of a mini-van and internet availability impact the SRT ridership represented by the percentage of students willing to utilize the SRT service.

Low, medium, and high values were practical scenarios defined in accordance with student's current mode and his/her place of residence (See Appendix D, Tables 17 through 21). Taxi fare was increased and decreased by 1,000 L.L. (0.7$) to 2,000 L.L. (1.3$). Changes in travel time were varied by 5 minutes and allowable time for late pick-up and early drop-off was decreased to 0 minutes in the low scenario and doubled in the high scenario.

Variations are applied to each of the 5 attributes at a time where values of all other attributes were fixed to medium values as presented in the table below.

Table 12: Attributes of the Basic Service

Variable	Mode	Residence: Region A	Residence: Region B	Residence: Region C
Shared-Ride Taxi Fare	Car	4000L.L. (2.7$)	5000L.L. (3.3$)	8000L.L. (5.3$)
	Bus			
	Jitney			*(Not eligible)*
Delta Time (SRT time - current time)	Car	5	10	15
	Bus	-5	-10	-15
	Jitney	0	-5	*(Not eligible)*
Maximum Allowable Time for Pick-up and Early Drop-off	Car	5	7	7
	Bus			
	Jitney			*(Not eligible)*
Maximum Number of Passengers Sharing the Ride	Car	1 to 3 passengers		
	Bus			
	Jitney			
Internet/WiFi Availability	Car	Not Available		
	Bus			
	Jitney			

The following observations can be noted (see Figures 22 through 26):

- PT users generally exhibit the least ridership. No matter how improved the level of service is in terms of change in travel time, MAWT, and the type of vehicle, bus and jitney users are not very attracted to the new taxi service as compared to the private car users. Also, under the Basic scenario with medium values, the SRT service appeared more attractive to PC users (45% of students willing to shift to SRT) than PT users (38% of students willing to shift to SRT). Yet, when the SRT fare is decreased, public transport users tend to favor the use of the Shared-Ride Taxi over their current mode more than the PC users.

- Change in fare has the highest impact on ridership. Besides, in the case of car users, an increase in the SRT fare from 2,500 L.L. to 6,000 L.L. has little effect on ridership (19% decrease) as compared to that exhibited by PT users (40% decrease), implying that PC users are less cost sensitive. This finding is consistent with the high value of time associated with these users and the high average family income, as discussed before. A similar finding was concluded by Benjamin et al. (1998) who stated that a fare reduction would have minimal effect on automobile drivers.
- Users are more sensitive to change in travel time than a change in the time for late pick-up and early drop-off. Besides, both users are almost equally responsive to the MAWT variable.
- Although jitney and bus users are expected to be less reluctant to the use of minivans as compared to small vehicles, both groups of users are similarly sensitive to the effect of the vehicle size.
- Decreased taxi fare resulted in the highest ridership for both users (57% of PT users and 53% of PC users utilize SRT under this condition).

- Potential trade-offs between the levels of service brought by the variation in the different attributes can also be noted. In other words, results show that an improved level in one of the attributes can bring equal ridership as an improvement in another attribute. Trade-offs are not equally exhibited among the two modes. A major finding was that internet and Wi-Fi availability will induce equivalent PT and PC users ridership as providing a low maximum allowable waiting time (42% and 49%, respectively), whereas for PC users, a high *Delta Time* had similar effect as the use of minivan (~37% ridership).

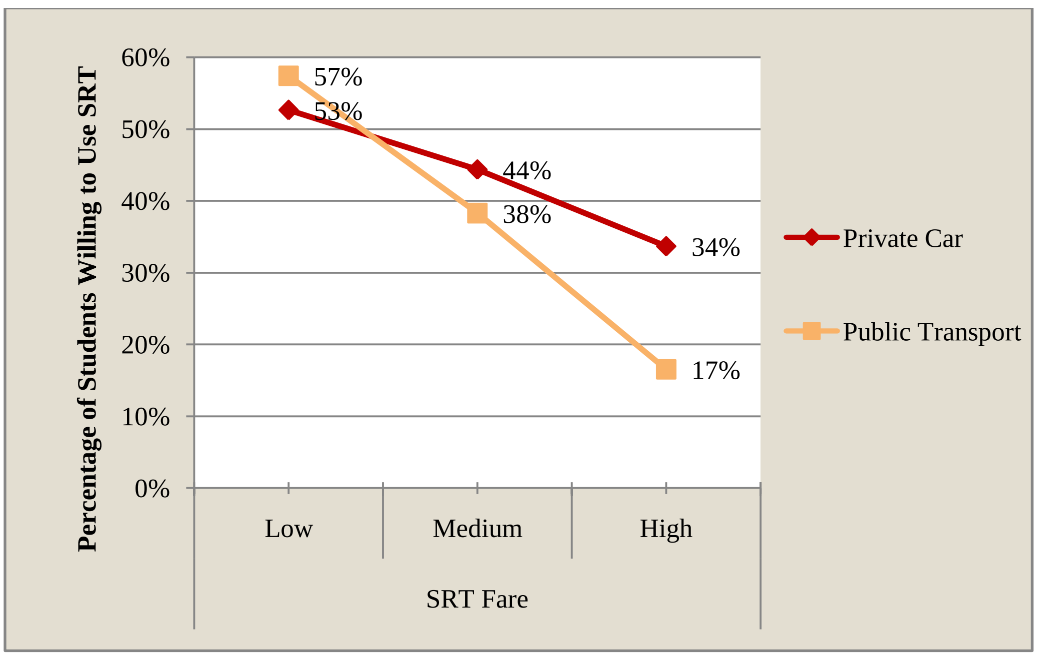

Figure 22: Percentage of Students Willing to Use SRT vs. Shared-Ride Taxi Fare

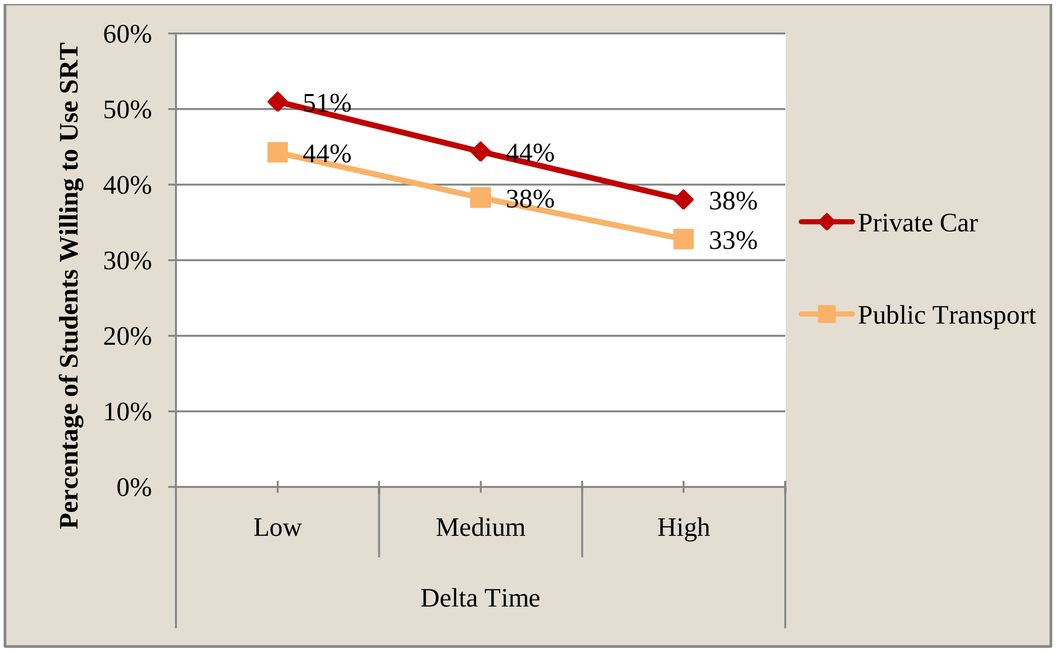

Figure 23: Percentage of Students Willing to Use SRT vs. Delta Time

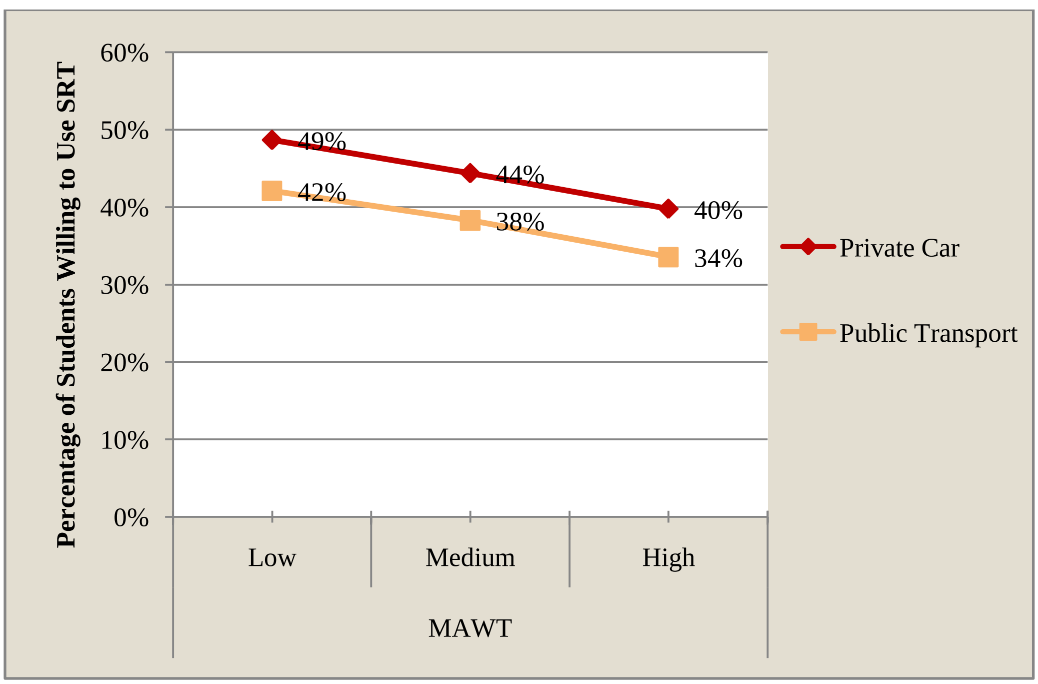

Figure 24: Percentage of Students Willing to Use SRT vs. Maximum Allowable Time for Pick-up and Early Drop-off

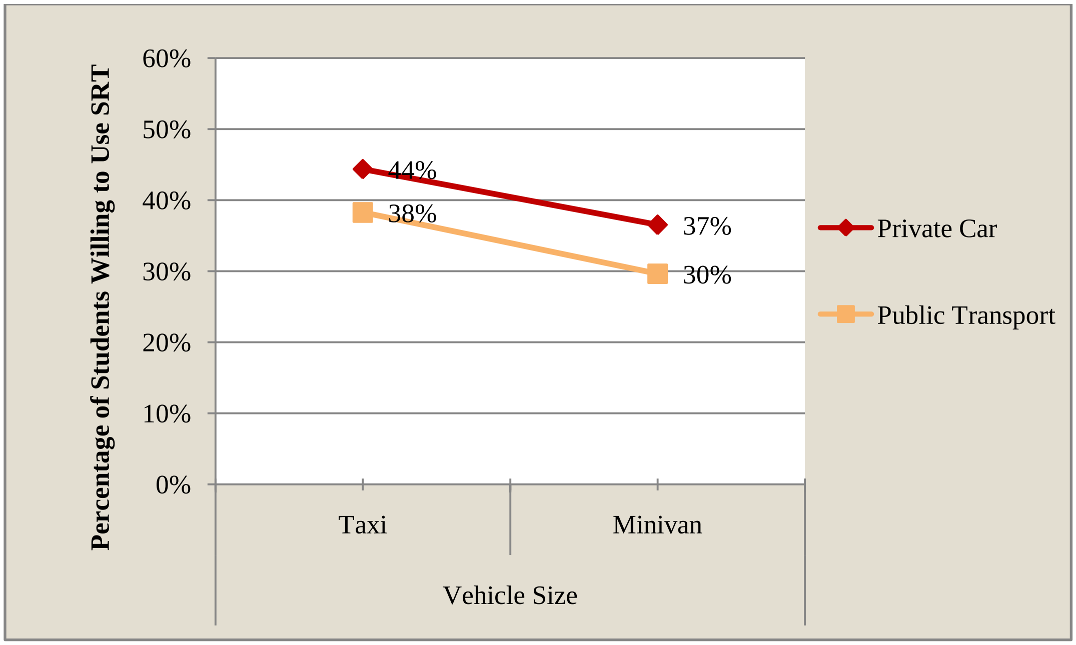

Figure 25: Percentage of Students Willing to Use SRT vs. Type of Vehicle (Number of Passengers Sharing the Ride)

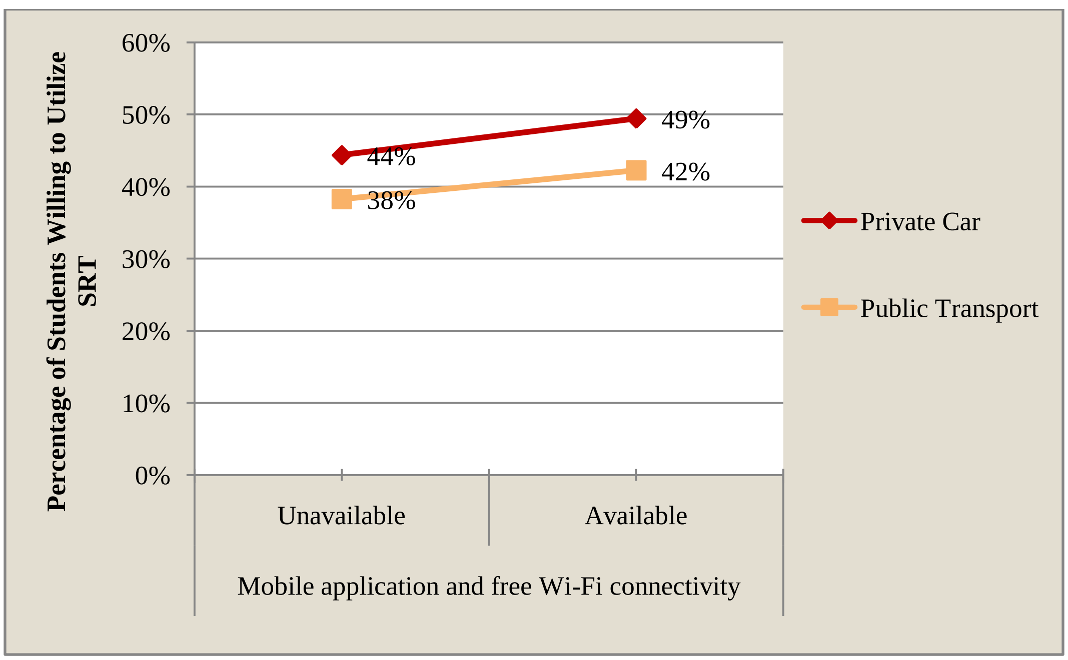

Figure 26: Percentage of Students Willing to Use SRT vs. Internet and Wi-Fi Availability

Furthermore, two levels of subsidies were tested to see their effect on the percentage of students willing to use the SRT. Two values (750 L.L. (0.50$) and 1,500 L.L. (1.00$)) were assumed to be granted to students for their one-way trip. Figures 27 and 28 below show the percentage of students expected to utilize the SRT under six scenarios including basic and economy services. For example, when a 1$ per one-way trip is offered, the total monthly subsidy granted to every student, assuming he/she goes to the university three times weekly is $1\$ \times 3\ weekly\ trips \times 2\ trips\ per\ day \times 4\ weeks\ per\ month = 24{,}000\ L.L.\ (16\$)$. The subsidy is expected to be granted by the university to encourage students to utilize the new service.

The impact of subsidies and cost incentives has been investigated in previous research studies in the context of shared-ride transportation (e.g., Concas et al. (2005) and Erdoğan et al. (2015)).

Figure 27 and 28 reveal that 750 L.L. and 1,500 L.L. given to students as a one-way trip daily will lead to a higher ridership. Also, the following can be noted:

- The effect of subsidy is greater for PT users as the difference in the percentage of students willing to utilize SRT when a subsidy of either level is given is greater for PT users than PC users. Visually, this is apparent as the slopes corresponding to the PT users are steeper.
- A subsidized basic scenario is preferred over an economy service for both users. In fact, a subsidized basic scenario increases the PT ridership by 21%.

– It is notable that unlike the unsubsidized scenarios described in previous discussions, the inclusion of a subsidy makes the SRT service more attractive to PT users, who showed a minimal propensity to shift (as compared to PC users) under unsubsidized scenarios.

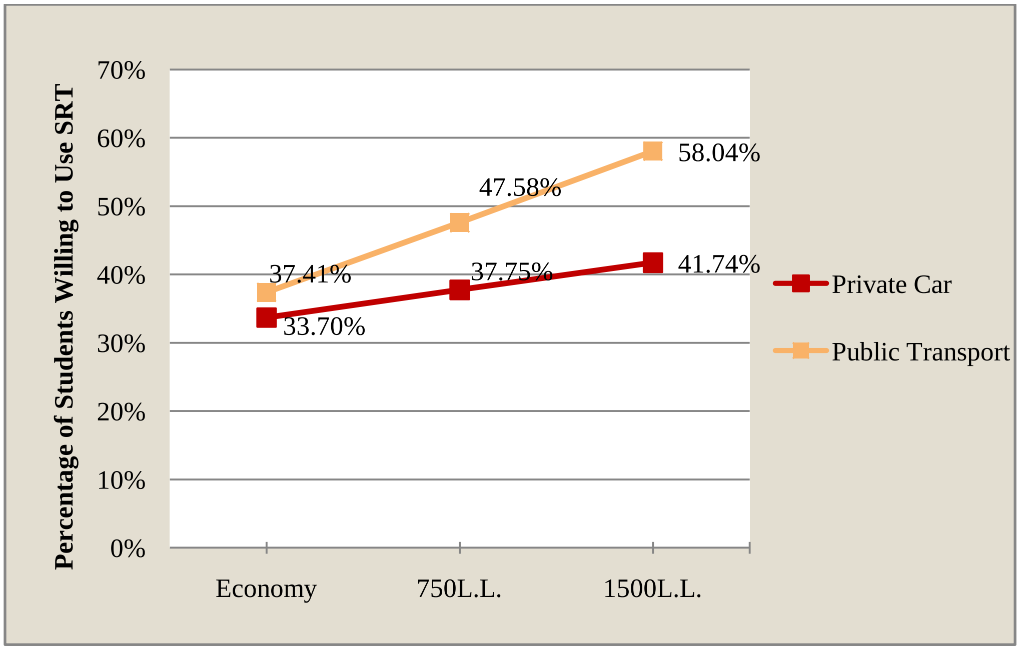

Figure 27: Percentage of Students Willing to Use SRT versus Two Levels of Subsidy on Economy Service

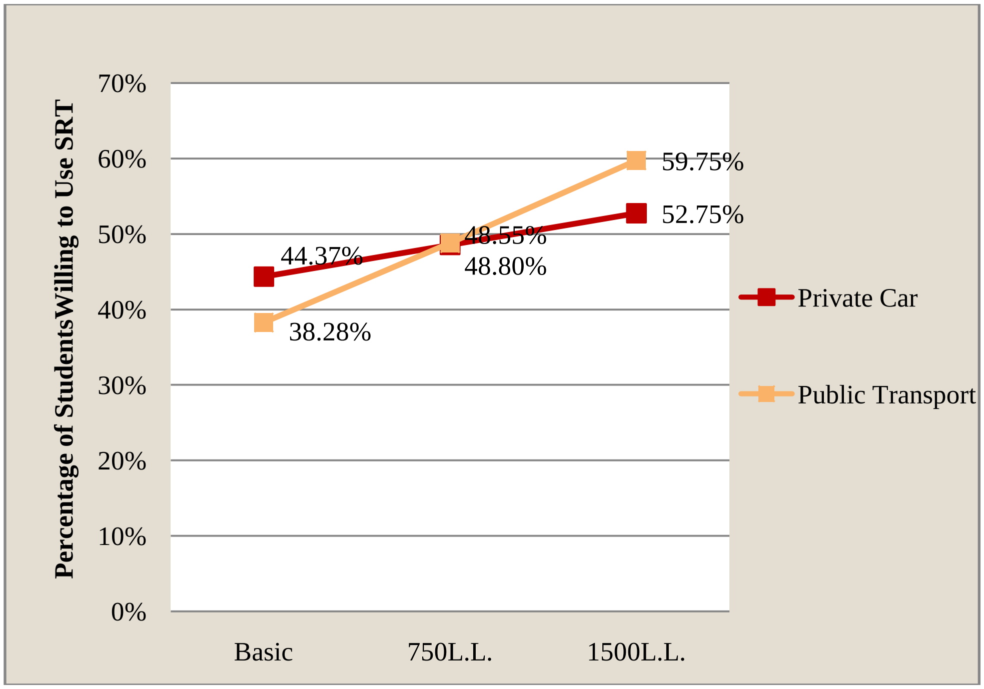

Figure 28: Percentage of Students Willing to Use SRT versus Two Levels of Subsidy on Basic Service

CHAPTER 5

CONCLUSION

This thesis presents a framework to predict the demand for a Shared-Ride Taxi (SRT) service in an organization-based context. SRTs are a form of public transportation employing multiple-occupant vehicles designed to provide flexible transport at relatively low fare rates. Travelers are often registered in a computerized scheme which employs real-time communication between travelers and the driver.

Various demand models were estimated for different data sets, corresponding to the users of different modes. The models incorporated a Random Effect (RE) that accommodates the individual-specific correlation due to the use of longitudinal data collected using an 8-question Stated Preference survey. The count models estimated include RE Poisson, RE Negative Binomial, RE Zero-Inflated Poisson, RE Zero-Inflated Negative Binomial, RE Hurdle Poisson, and RE Hurdle Negative Binomial. All models capture the effect of several SRT attributes on travelers' travel behavior-represented by the average number of trips per week using the new SRT service. The best fitting count model was compared to an RE Multinomial Logit model, which in turn, proved superior. Subsequently, an Integrated Choice and Latent Variable (ICLV) model was estimated whose structure consists of two sub-models: a choice model and a latent variable model that allow for relationships between travelers' attitudes and their choices. Model selection criteria include AIC and BIC (for count models), R^2, and market segment prediction tests.

Assuming that such service is available for travelers, the statistical models are estimated and analyzed on the students of AUB in an attempt to bring quantitative answers that would capture students' readiness to use the SRT service. Model analyses include VOT analysis, elasticity estimation, and policy analysis.

5.1. Contributions

This thesis adds to the existing literature on the feasibility of demand management strategies for congestion relief in a developing country, in general, and on modeling the demand for shared-ride transportation in an organization-based context, in particular. It provides insight into factors affecting demand for taxi-sharing for university students, which has not been done in the literature to the best of the author's knowledge. The majority of related studies focus on the service design of shared-ride transportation systems for students and employees rather than modeling its demand (e.g. Amey, 2010; Deakin et al., 2010; Erdoğan, 2015). Also, organization-based studies in the literature rarely investigated the potential of Shared-Ride Taxis; they mostly focused on the potential of carpooling and vanpooling for employees or students.

A large survey was designed with almost 2300 respondents and a response rate of 29%. This study takes as its main objective the development of a modeling approach that sheds light on the adequate ways to handle the excess zeros in the observations.

Existing studies of shared-ride transportation (utilized by the public or employed in larger scales than that of a university or workplace) incorporating SP data have focused on the impact of time window for pick-up and delivery, travel time, and price. This research is innovative as it presents a data collection procedure whose aim is to identify the impact of five different attributes on travelers' choices. In addition to travel time and cost, service attributes encompass a range of variables including internet and Wi-Fi availability, maximum allowable waiting time for late pick-up and early drop-off, and vehicle size. Further, the winning model captures two important effects: the socioeconomic characteristics of the users, as well as the latent factors that are assumed to influence the choices. Socioeconomic characteristics of travelers include the family income, cost-to-income ratio, gender, car ownership, and grade. As for the latent factors, they include students' attitudes toward the different public transport modes available, ridesharing, taxi-sharing, and internet and technology.

The modeling selection methodology for the count models was undertaken favoring the selection of a Hurdle Poison model which is capable of predicting the proportion of zeros relatively well. In the literature, Hurdle models were also preferred over other zero-inflated count models due to their simplicity and ease of interpretation (Tait et al., 2012). Further, the framework is original since it encompasses the calibration of a range of models in two families, count data models and discrete choice models. Both types are used to model the average number of trips made weekly using the SRT. Numerous ICLV models existed in the literature such as those presented by Temme et al. (2008), Ben Akiva et al. (2002), Danthurebandara et al (2013), and Paulssen et al. (2013), yet such models were mode choice models (i.e., represented by utilities of discrete choices corresponding to travel modes rather than counts). In this thesis, an ICLV model was used to represent the choice of the number of weekly trips made by SRT.

5.2. SRT Potential for AUB

The case study sheds light on the case of students at AUB. The winning model was applied to predict the SRT service participation levels. The benefits that the SRT is expected to bring to AUB students and the neighboring area in terms of lower congestion levels are considered the most important.

Key findings related to the demand of the SRT can be summarized as follows:

- Students are willing to utilize the SRT service under certain conditions. In the presence of internet and Wi-Fi and with the use of a small taxi vehicle, students are more likely to switch to the new system. Moreover, a higher travel time compared to the current travel time of the user and the high waiting time before pick-up or drop-off have negative effects on the propensity towards the use of the new taxi service. The importance of the internet variable indicates that the SRT system should employ free internet connectivity in the SRT vehicle as well as a mobile application for tracking and reservation if it is to operate competently.
- Subjective factors such as attitudes towards taxi-sharing and ridesharing impact SRT ridership. This is consistent with the results of previous research efforts

including Benjamin et al. (1998), Tsirimpa et al. (2007), Temme at al. (2008), and Paulssen at al. (2013).

- Not all students exhibit similar sensitivities to changes in the level of service of the new Shared-Ride Taxi service. These sensitivities vary depending on what mode of commute the student currently uses.
- Under three practical scenarios (Economy/Basic/Premium), results reveal that half PC users are willing to shift when the Premium scenario is applied. Only 33.7% are willing to shift under an Economy service. In the contrary, PT users are less likely to utilize the new taxi in general. Almost 38% of them are willing to use it when an Economy or Basic scenario is put into service. See Figure 29 below. Policy analysis thus reveals that students that are currently car users prefer a service with good attributes even if it is expensive. On the contrary, PT users are not willing to pay high fares represented by the premium service, as they prefer an Economy service whose attributes are rather unfavorable.

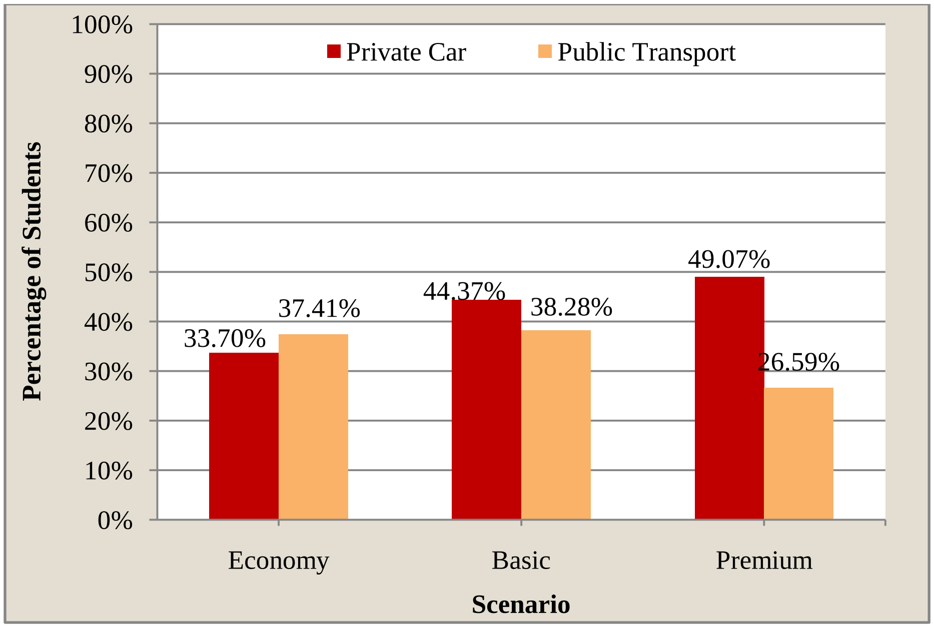

Figure 29: Percentage of Students Willing to Shift to the SRT

The findings also stress the significance of applying subsidies by the organization in which the SRT operates. It was demonstrated in this thesis that the university-based subsidies greatly influence the percentage of students willing to utilize SRT weekly. These results support the findings of Concas et al. (2005) who considered the use of discrete choice models to examine the impact of fare subsidies on the vanpool services on employees. Their study revealed that the predicted probability of choosing vanpool more than doubles when employees are offered a subsidy. Furthermore, as part of Commuter Program at MIT, multiple subsidized services are available for registered students and employees (MIT Department of Facilities, n.d.). Subsidized services include bicycles, carpools, and vanpools. MIT also promotes commuting via public transit through subsidizing 50% of the cost of a monthly pass. Accordingly, under a basic SRT service and a 1$/trip as a subsidy for every student, the annual subsidy burden on AUB would be approximately[6] 351,000$. A smaller amount (333,000$ and 296,000$) is incurred if a premium or an economy service is put into service. The implementation of a subsidized Shared-Ride Taxi service is expected to improve the impact of the university on its neighboring area by reducing auto dependency due to private car users shifting to the implemented mode of commute (results show that over 50% of them commuters are willing to utilize the SRT). Perhaps another significant improvement from which AUB can benefit is the reduced demand on parking spots and facilities as well as the decline in the quantity of emissions due to the reduction in the number of cars attracted to the area.

Universities around Lebanon and especially those present inside Greater Beirut face transportation challenges as the urban development exerts significant impact on the students' travel patterns. Results indicate that the implementation of an SRT for the students at AUB, even if unsubsidized, encourages a significant number of students (20% to 50% of motorized modes users living in GBA) to utilize the SRT service if it is appropriately priced and designed.

5.3. Limitations and Recommendations for Future Research

The methodology presented in this thesis is coupled with some challenges and limitations. A potential challenge present in the framework is the use of SP data. Even

[6] Assuming that students will utilize the SRT 3 times weekly, the total annual subsidy burden in case of a Basic scenario is $1\$ \times 3\ weekly\ trips \times 2\ trips\ per\ day \times 4\ weeks\ per\ month \times 9\ months \times 38.28\%\ of\ eligible\ PT\ users\ and\ 44.37\%\ of\ eligible\ PC\ users = 350{,}730\$$.

though SP methods are an integral part in transportation planning contexts where a new modal alternative is to be introduced to the market, the results might be associated with bias. In other words, the proportion of the new service ridership might be overestimated. However, no other data on similar types of taxi sharing is available in Lebanon. The results can therefore be interpreted as indicative of the potential switching that may occur.

Despite the goodness of fit measures associated with the winning models, refinements could still be made to the model specification so as to make it more realistic. First, the current models, in fact, does not account for the number of trips the travelers make using the current mode except for the count models which include an exposure variable. Accounting for the current number of trips in the Logit models as an explanatory variable did not prove to be significant. A possible way to approach this challenge is through examining other types of models where the response variable is modeled as a rate/proportion, e.g., beta distribution. Second, since the response variable has an upper-bound of 5, being the maximum number of weekly trips, count models can be remediated through the estimation of a right-truncated count model that sets a bound to the upper limit of the expected number of trips. Third, the model might be more realistic by employing randomly distributed coefficients of time and/or cost. Fourth, random effects in both components of the Hurdle Poisson and Hurdle Negative Binomial models are assumed uncorrelated. Another direction for future research is to allow the two random effects to be correlated and which is expected to require a more complex fitting process (Min and Agresti, 2005).

The service itself is not without some challenges. For example, challenges at the level of the SRT service might show up due to safety and privacy concerns, as the literature suggests. The organization in which an SRT operates should put an effort to mitigate possible challenges in that respect. Further, it has been shown that shared-ride transportation systems require a critical mass of travelers willing to use the SRT in order to guarantee its success. The problem is that they are only able to be set up if a large mass decide to participate within a short period of time (Ciari, 2012).

Model results suggest that an important motivation behind students' choices is their attitudes toward taxi-sharing. In fact, a taxi-sharing case study in Dublin faced a major problem which was the cultural aversion to sharing taxis causing the service to fail (Enoch et al., 2006). Casey (2014) argues that as passengers become more aware of the time and cost savings, the service will appear more lucrative for potential users. Hence, in order for the SRT system to succeed, and as far as the demand side is concerned, the service should not only be well-designed and realistically priced, but also effective marketing is necessary. In various cases, information related to the use of the system seemed too complex for the users to understand, especially the disabled and elderly (Enoch et al., 2006). The good news is that having to adequately convey the idea of the SRT with the new technological advancements that it entails might not require considerable effort if the targeted population consists of university students.

BIBLIOGRAPHY

Amey, A. M. (2010). Real-time Ridesharing: Exploring the Opportunities and Challenges of Designing a Technology-based Rideshare Trial for the MIT Community (Doctoral dissertation, Massachusetts Institute of Technology).

Anastasopoulos, C., and Mannering, L. (2009).A Note on Modeling Vehicle Accident Frequencies with Random-Parameters Count Models. Accident Analysis & Prevention, 41(1), 153-159.

Aoun, M. Abou-Zeid, I. Kaysi, C. Myntti. (2013). Reducing Parking Demand and Traffic Congestion at the American University of Beirut, Transport Policy Vol. 25, pp. 52-60.

Auchard E. and Steitz C. (2015, March, 19). German Court Bans Uber's Unlicensed Taxi Services. The Daily Star. Retrieved from http://www.dailystar.com.lb/Business/International/2015/Mar-19/291308-german-court-bans-ubers-unlicensed-taxi-services.ashx

Bandwagon (2014). How Bandwagon Works. Retrieved from: https://bandwagon.io/how-it-works

Barmby, T., and Doornik, J. (1989). Modelling Trip Frequency as a Poisson Variable. Journal of Transport Economics and Policy, 309-315.

Ben-Akiva, M., Benjamin, J., Lauprete, G. J., and Polydoropoulou, A. (1996). Impact of Advanced Public Transportation Systems on Travel by Dial-A-Ride. In Transportation Research Record: Journal of the Transportation Research Board, No. 1557, Transportation Research Board of the National Academies, Washington, D.C., pp. 72-79.

Ben-Akiva, M., and Lerman, S. (1985). Discrete Choice Analysis: Theory and Application to Travel Demand. MIT Press, Cambridge.

Ben-Akiva, M., Walker, J., Bernardino, T., Gopinath, A., Morikawa, T., and Polydoropoulou, A. Integration of Choice and Latent Variable Models (2002). In: Mahmassani, H.S. (ed.) In Perpetual Motion: Travel Behaviour Research Opportunities and Application Challenges, pp. 431–470. Elsevier, Amsterdam.

Benjamin, J., Kurauchi, S., Morikawa, T., Polydoropoulou, A., Sasaki, K., and Ben-Akiva, M. (1998). Forecasting Paratransit Ridership Using Discrete Choice Models with Explicit Consideration of Availability. In Transportation Research

Record: Journal of the Transportation Research Board, No. 1618, Transportation Research Board of the National Academies, Washington, D.C., pp. 60-65.

Benjamin, J. and Price, G. (2006). A Study of the Impact of APTS on Service Quality Perceptions of Elderly and Disabled Riders. Journal of Public Transportation, 9(1), 53.

Bhat, R., and Pulugurta, V. (1998). A Comparison of Two Alternative Behavioral Choice Mechanisms for Household Auto Ownership Decisions. Transportation Research Part B: Methodological, 32(1), 61-75.

Bierlaire, M. and Fetiarison, M. (2009). Estimation of Discrete Choice Models: Extending BIOEGEME. Proceedings of the 9th Swiss Transportation Research Conference, Ascona, Switzerland.

Bonsall, P., Spencer, A., and Tang, W. (1984). What Makes a Car-Sharer? Transportation, 12(2), 117-145.

Booth, J. G., Casella, G., Friedl, H., and Hobert, J. P. (2003). Negative Binomial Loglinear Mixed Models. Statistical Modelling, 3(3), 179-191.

Boucher, J. P., and Guillén, M. (2009). A Survey on Models for Panel Count Data with Applications to Insurance. RACSAM-Revista de la Real Academia de Ciencias Exactas, Fisicas y Naturales. Serie A. Matematicas, 103(2), 277-294.

Buliung, R. N, Soltys, K., Habel, C., and Lanyon, R. (2009). Driving Factors Behind Successful Carpool Formation and Use. Transportation Research Record: Journal of the Transportation Research Board, 2118(1), 31-38.

Burnham, K., and Anderson, D. (2002). Model Selection and Multimodal Inference: A Practical Information-Theoretic Approach. 2nd ed Springer-Verlag. New York, NY.

Cameron, C., and Windmeijer, F. A. (1996). R-Squared Measures for Count Data Regression Models with Applications to Health-Care Utilization. Journal of Business & Economic Statistics, 14(2), 209-220.

Casey, M. (2014, September 1). Share a Taxi with a Stranger? Idea is Catching in Big Cities. Fortune. Retrieved from http://fortune.com/2014/09/01/share-a-taxi-with-a-stranger-idea-is-catching-in-big-cities/

Chan, N., and Shaheen, S. (2012). Ridesharing in North America: Past, Present, and Future. Transport Reviews, 32(1), 93-112.

Ciari, F., (2012, May). Why Do People Carpool: Results From a Swiss Survey. In 12th Swiss Transport Research Conference, Ascona.

Concas, S., Winters, P., and Wambalaba, F. (2005). Fare Pricing Elasticity, Subsidies, and Demand for Vanpool Services. Transportation Research Record: Journal of the Transportation Research Board, 1924(1), 215-223.

Corpuz, G. (2007). Public Transport or Private Vehicle: Factors that Impact on Mode Choice. In 30th Australasian Transport Research Forum (p. 11).

Cristaldi, F. (2005). Commuting and Gender in Italy: a Methodological Issue. The Professional Geographer, 57(2), 268-284.

Danaf, M., Abou-Zeid, M., and Kaysi, I. (2014). Modeling Travel Choices of Students at a Private, Urban University: Insights and Policy Implications, Case Studies on Transport Policy, Vol. 2, No. 3, pp. 142-152.

Danthurebandara, V., Vandebroek, M., and Yu, J. (2013). Integrated Mixed Logit and Latent Variable Models. Marketing Letters, 24(3), 245-259.

Dargay, J., Gately, D., and Sommer, M. (2007). Vehicle Ownership and Income Growth, Worldwide: 1960-2030. The Energy Journal, 143-170.

Deakin, E., Frick, K. T., and Shively, K. (2010). Markets for Dynamic Ridesharing? Transportation Research Record: Journal of the Transportation Research Board, 2187(1), 131-137.

Deloitte (2014). Global Automotive Consumer Study: The Changing Nature of Mobility. Retrieved from http://www2.deloitte.com/content/dam/Deloitte/global/Documents/Manufacturing/gx-mfg-geny-automotive-consumer.pdf

Domencich, T. A., and McFadden, D. (1975). Statistical Estimation of Choice Probability Functions.Urban Travel Demand-A Behavioral Analysis.(pp.101-1025). Oxford England: North-Holland Publishing Company Limited

Eggers, W. and Jaffe, J. (2013). Gov on the Go. Boosting Public Sector Productivity by Going Mobile. Retrieved from http://dupress.com/articles/gov-on-the-go/

Enoch, M., Potter, S.Parkhurst, G., Smith, M. (2006). Why Do Demand Responsive Transport Systems Fail? Presented at 85th Annual Meeting of the Transportation Research Board, Washington DC.

Erdoğan, S., Cirillo, C., and Tremblay, M. (2015). Ridesharing as a Green Commute Alternative: A Campus Case Study. International Journal of Sustainable Transportation, 9(5), 377-388.

Fournier, D., .Skaug, H., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M., Nielsen, A., and Sibert, J. (2012). AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. Optimization Methods and Software. Vol. 27, Issue 2.

Global e-Sustainability Initiative (2008) SMART 2020: Enabling the Low Carbon Economy in the Information Age. United States Report Addendum. Retrieved from http://www.smart2020.org/_assets/files/Smart2020UnitedStatesReportAddendum.pdf

Gopisetty, V., andSrinivasan, K. (2013). Joint Models for Analysis of Household Trip Frequency and Vehicle Ownership in Chennai City. International Journal of Advances in Engineering Sciences and Applied Mathematics, 5(2-3), 129-144.

Gurmu, S., andTrivedi, P. (1996). Excess Zeros in Count Models for Recreational Trips. Journal of Business & Economic Statistics, 14(4), 469-477.

Hall, D. (2000). Zero-Inflated Poisson and Binomial Regression with Random Effects: A Case Study. Biometrics, 56(4), 1030-1039.

Helft, M. (2007, March 10). Google's Buses Help its Workers Beat the Rush. The New York Times. Retrieved from: http://www.nytimes.com/2007/03/10/technology/10google.html?pagewanted=print&_r=0

Johansen, G. (2013). Modelling the Effects of Personality Traits on Ridership: The Case of High Speed Rail in Norway.

Kibria, B. (2006). Applications of Some Discrete Regression Models for Count Data. Pakistan Journal of Statistics and Operation Research, 2(1), 1-16

Kim, N., and Susilo, Y. (2013). Comparison of Pedestrian Trip Generation Models. Journal of Advanced Transportation, 47(4), 399-412.

Koffman, D. (2007). Improving ADA Complementary Paratransit Demand Estimation (Vol. 119). Transportation Research Board.

Kramer, M. (2005). R^2 Statistics for Mixed Models. In Proceedings of the Conference on Applied Statistics in Agriculture (Vol. 17, pp. 148-160).

Lawal, B. (2012). Zero-Inflated Count Regression Models with Applications to Some Examples. Quality & Quantity, 46(1), 19-38.

Markovic, N., Milinkovic, S., Schonfeld, P., and Drobnjak, Z. (2013). Planning Dial-A-Ride Services: Statistical and Metamodeling Approach. In Transportation Research Record: Journal of the Transportation Research Board, No. 4205, Transportation Research Board of the National Academies, Washington, D.C.,. 72-79.

Massachusetts Institute of Technology (MIT) Department of Facilities (n.d.). Providing Links to Convenient Options. Retrieved from: http://web.mit.edu/facilities/transportation/index.html

Min, Y., and Agresti, A. (2005). Random Effect Models for Repeated Measures of Zero-Inflated Count Data. Statistical Modelling, 5(1), 1-19.

Paulssen, M., Temme, D., Vij, A., and Walker, J. (2013). Values, Attitudes and Travel Behavior: AHierarchical Latent Variable Mixed Logit Model of Travel Mode Choice. Transportation, 1-16.

Potoglou, D., and Kanaroglou, P. (2008). Modelling Car Ownership in Urban Areas: A Case Study of Hamilton, Canada. Journal of Transport Geography,16(1), 42-54.

Potoglou, D., and Susilo, Y. (2008). Comparison of Vehicle-Ownership Models.Transportation Research Record: Journal of the Transportation Research Board, 2076(1), 97-105.

R Core Team (2014). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL: http://www.R-project.org/.

Raudenbush, S. W., Yang, M. L., and Yosef, M. (2000). Maximum Likelihood for Generalized Linear Models with Nested Random Effects Via High-Order, Multivariate Laplace Approximation. Journal of Computational and Graphical Statistics, 9(1), 141-157.

Ryley, J., Stanley, A., Enoch, P., Zanni, M., and Quddus, A. (2014). Investigating the Contribution of Demand Responsive Transport to a sustainable Local Public Transport System. Research in Transportation Economics, 48, 364-372.

Scalici, A. (1985), Paratransit: An Instructional Module, Morgantown, WV: West Virginia University, Technology Education Program.

Sermons, M., and Koppelman, F. (2001). Representing the Differences Between Female and Male Commute Behavior in Residential Location Choice Models. Journal of Transport Geography, 9(2), 101-110.

Siddiqi, Z., and Buliung, R. (2013). Dynamic Ridesharing and Information and Communications Technology: Past, Present and Future Prospects. Transportation Planning and Technology, 36(6), 479-498.

Silvis, J., and Niemeier, D. (2009). Social Network and Dwelling Characteristics that Influence Ridesharing Behavior of Seniors. Transportation Research Record: Journal of the Transportation Research Board, 2118(1), 47-54.

Soltys, K., and Buliung, R. (2008). Commuting Connections: Gender, Carpooling and Cyberspace. Presented at the Association for Commuter Transportation TDM Summit, Halifax

St-Louis E., Manaugh K., Lierop D., El-Geneidy A. (2013). The Happy Commuter: A Comparison of Commuter Satisfaction Across Modes. Transportation Research Board 93rd Annual Meeting. Washington DC, United States.

Suen, L., Ebrahim, A., and Oksenhendler, M. (1981). Computerised Dispatching for Shared-Ride Taxi Operations in Canada. Transportation Planning and Technology, 7(1), 33-48.

Tait, M., Horrocks, J., and Ginis, K. (2012). Modelling Longitudinal Count Data with Zero-Inflation: An Application to Physical Activity. Journal of Statistics and Applications, 7(1), 55-79.

Takeuchi, R., Okura, I., Nakamura, F., and Hiraishi, H. (2003). Feasibility Study on Demand Responsive Transport systems (DRTS).Proceedings of the 5th Eastern Asia Society for Transportation Studies Conference, Vol. 5, pp. 388-397, Tokyo, Japan.

Temme, D., Paulssen, M., and Dannewald, T. (2008). Incorporating Latent Variables Into Discrete Choice Models—a Simultaneous Estimation Approach Using SEM Software. BuR-Business Research, 1(2), 220-237.

Tsirimpa, A., Polydoropoulou, A., and Antoniou, C. (2007). Development of a Mixed Multi-Nomiallogit Model to Capture the Impact of Information Systems on Travelers' Switching Behavior. Journal of Intelligent Transportation Systems,11(2), 79-89.

Urbanik, T. (1998). Management of Surface Transportation Systems (No.1479). Transportation Research Board.

Wambalaba, F., Concas, S., and Chavarria, M., (2004). Price Elasticity of Rideshare: Commuter Fringe Benefits forVanpools. National Center for Transit Research, Center for Urban Transportation Research, University of South Florida.

Wardman, M. (2004). Public Transport Values of Time. Transport policy, 11(4), 363-377.

Yau, K., and Lee, A. (2001). Zero-Inflated Poisson Regression with Random Effects to Evaluate an Occupational Injury Prevention Programme. Statistics in Medicine, 20(19), 2907-2920.

Zhao, Y., and Kockelman, K. (2002). Household Vehicle Ownership by Vehicle Type: Application of a Multivariate Negative Binomial Model. In 81st Annual Meeting of the Transportation Research Board. Washington, DC.

Zorn, C. (1996). Evaluating Zero-Inflated and Hurdle Poisson Specifications. Midwest Political Science Association, 18(20), 1-16.

APPENDIX A: Market Segment Prediction Test Results

Table 13: Market Segment Prediction Test for the PT Data Set (by Student's Class)

Alternative	Market Segment:	New Students		Old Students	
		Count	Percentage	Count	Percentage
Zero	Observed	309.00	62.30%	613.00	60.81%
	MNL	308.72	62.24%	616.65	61.18%
	Hurdle Poisson	311.84	62.87%	590.59	58.63%
One	Observed	29.00	5.85%	70.00	6.94%
	MNL	29.52	5.95%	62.42	6.19%
	Hurdle Poisson	120.37	24.27%	295.11	29.30%
Two	Observed	33.00	6.65%	69.00	6.85%
	MNL	32.75	6.60%	69.28	6.87%
	Hurdle Poisson	42.06	8.48%	88.96	8.83%
Three	Observed	37.00	7.46%	87.00	8.63%
	MNL	40.49	8.16%	85.33	8.47%
	Hurdle Poisson	14.97	3.02%	24.33	2.42%
Four	Observed	21.00	4.23%	45.00	4.46%
	MNL	19.61	3.95%	40.61	4.03%
	Hurdle Poisson	4.97	1.00%	6.60	0.65%
Five	Observed	67.00	13.51%	124.00	12.30%
	MNL	64.91	13.09%	133.71	13.26%

Table 14: Market Segment Prediction Test for PC Data Set (by Student's Class)

Alternative	**Market Segment:**	**New Students**		**Old Students**	
		Count	Percentage	Count	Percentage
Zero	Observed	304.00	48.72%	1023	52.84%
	MNL	352.46	56.48%	970.83	50.15%
	Hurdle Poisson	298.85	47.89%	980.18	50.63%
One	Observed	75.00	12.02%	205.00	10.59%
	MNL	61.81	9.91%	203.82	10.53%
	Hurdle Poisson	238.95	38.29%	689.21	35.60%
Two	Observed	90.00	14.42%	229.00	11.83%
	MNL	69.82	11.19%	246.07	12.71%
	Hurdle Poisson	63.55	10.18%	194.62	10.05%
Three	Observed	65.00	10.42%	208.00	10.74%
	MNL	60.24	9.65%	216.87	11.20%
	Hurdle Poisson	15.43	2.47%	51.65	2.67%
Four	Observed	33.00	5.29%	93.00	4.80%
	MNL	27.59	4.42%	103.87	5.37%
	Hurdle Poisson	4.29	0.69%	14.35	0.74%
Five	Observed	57.00	9.13%	178.00	9.19%
	MNL	52.07	8.35%	194.54	10.05%
	Hurdle Poisson	1.55	0.25%	4.21	0.22%

Table 15: Market Segment Prediction Test for PT Data Set (by Number of Vehicles Available)

Alternative	**Market Segment:**	**0 and 1 Vehicles**		**2 or more Vehicles**	
		Count	**Percentage**	**Count**	**Percentage**
Zero	Observed	455.00	61.82%	472	61.46%
	MNL Prediction	450.83	61.25%	474.55	61.79%
	Hurdle Poisson Prediction	449.45	61.12%	452.98	59.06%
One	Observed	57.00	7.74%	42	5.47%
	MNL Prediction	45.62768	6.20%	46.31	6.03%
	Hurdle Poisson Prediction	199.17	27.08%	216.18	28.18%
Two	Observed	47.00	6.39%	54	7.03%
	MNL Prediction	49.12	6.67%	52.92	6.89%
	Hurdle Poisson Prediction	62.07	8.44%	68.95	8.99%
Three	Observed	52.00	7.07%	68	8.85%
	MNL Prediction	60.86	8.27%	64.97	8.46%
	Hurdle Poisson Prediction	17.93	2.44%	20.56	2.68%
Four	Observed	39.00	5.30%	27	3.52%
	MNL Prediction	30.69	4.17%	29.53	3.84%
	Hurdle Poisson Prediction	5.24	0.71%	6.33	0.83%
Five	Observed	86.00	11.68%	105	13.67%
	MNL Prediction	98.88	13.43%	99.73	12.99%
	Hurdle Poisson Prediction	1.54	0.21%	2.036403	0.27%

Table 16: Market Segment Prediction Test for PC Data Set (by Number of Vehicles Available)

Alternative	Market Segment:	0, 1, and 2 Vehicles		3, 4, and 5 Vehicles	
		Count	Percentage	Count	Percentage
Zero	Observed	630	52.15%	697	51.55%
	MNL Prediction	660.25	54.66%	663.05	49.04%
	Hurdle Poisson Prediction	642.14	53.16%	636.89	47.11%
One	Observed	142.00	11.75%	138	10.21%
	MNL Prediction	123.11	10.19%	142.51	10.54%
	Hurdle Poisson Prediction	415.45	34.39%	512.71	37.92%
Two	Observed	164.00	13.58%	155	11.46%
	MNL Prediction	140.48	11.63%	175.40	12.97%
	Hurdle Poisson Prediction	113.12	9.36%	145.05	10.73%
Three	Observed	135.00	11.18%	138	10.21%
	MNL Prediction	122.03	10.10%	155.08	11.47%
	Hurdle Poisson Prediction	27.17	2.25%	39.91	2.95%
Four	Observed	48.00	3.97%	78	5.77%
	MNL Prediction	56.01	4.64%	75.45	5.58%
	Hurdle Poisson Prediction	6.75	0.56%	11.88	0.88%
Five	Observed	89.00	7.37%	146	10.80%
	MNL Prediction	106.11	8.78%	140.50	10.39%
	Hurdle Poisson Prediction	1.97	0.16%	3.79	0.28%

APPENDIX B: ICLV Measurement Model

Table 17: ICLV Estimation Results – Measurement Model

		Factor Loading			**Standard Deviation**		
Latent Variable	**Indicator**	Estimate	Robust S.E.	P-value	Estimate	Robust S.E	P-value
RS (PT Model)	**I2**	0.177	0.165	0.28	1.51	0.0733	0.00
	I3	0.229	0.138	0.10	1.66	0.0798	0.00
	I4	0.321	0.129	0.01	1.34	0.0715	0.00
	I6	0.318	0.131	0.01	1.45	0.0732	0.00
TS (Car Model)	**I1**	-0.33	0.111	0.00	1.47	0.0761	0.00
	I2	0.694	0.110	0.00	1.53	0.0602	0.00
	I3	0.438	0.118	0.00	1.63	0.0522	0.00
(Insignificant)	**I5**	(Insignificant)					
	I7						
	I8						
	I9						
	I10						

APPENDIX C: Elasticity Estimation

Table 18: Elasticity Values of Four Different Variables for Both Data Sets

	Alternative	**Zero**	**1 Trip**	**2 Trips**	**3 Trips**	**4 Trips**	**5 Trips**
Public Transport	Fare	0.593%	-0.019%	-0.091%	-0.142%	-0.035%	-0.306%
	Delta Fare	0.292%	-0.021%	-0.052%	-0.072%	-0.026%	-0.120%
	Delta Time	0.044%	-0.005%	0.000%	-0.004%	-0.016%	-0.019%
	MAWT	0.037%	-0.002%	0.000%	0.000%	-0.011%	-0.024%
Private Car	Fare	0.301%	-0.031%	-0.063%	-0.073%	-0.048%	-0.087%
	Delta Fare	0.163%	-0.020%	-0.035%	-0.039%	-0.025%	-0.045%
	Delta Time	0.123%	-0.007%	-0.009%	-0.032%	-0.020%	-0.055%
	MAWT	0.043%	-0.009%	-0.009%	-0.009%	-0.005%	-0.010%

APPENDIX D: Values of Attributes Varied in Policy Analysis

Table 19: Attributes of the Premium Service

Variable	User	Residence		
		Region 1	**Region 2**	**Region 3**
Shared-Ride Taxi Fare	Car	6,000 L.L. (4$)	7,000 L.L. (4.7$)	10,000 L.L. (6.7$)
	Bus			
	Jitney			(Excluded)
Delta Time (SRT time - current time)	Car	0	5	10
	Bus	-10	-15	-20
	Jitney	-5	-10	(Excluded)
Maximum Allowable Time for Pick-up and Early Drop-off	Car	0	0	0
	Bus			
	Jitney			(Excluded)
Maximum Number of Passengers Sharing the Ride	Car	1 to 3 passengers		
	Bus			
	Jitney			
Internet/WiFi Availability	Car	Available		
	Bus			
	Jitney			

Table 20: Attributes of the Economy Service

Variable	User	Residence		
		Region 1	**Region 2**	**Region 3**
Shared-Ride Taxi Fare	Car	2,500L.L. (1.7$)	4,000L.L. (2.7$)	6,000L.L. (4$)
	Bus			
	Jitney			(Excluded)
Delta Time (SRT time - current time)	Car	10	15	20
	Bus	0	-5	-10
	Jitney	5	0	(Excluded)
Maximum Allowable Time for Pick-up and Early Drop-off	Car	10	15	15
	Bus			
	Jitney			(Excluded)
Maximum Number of Passengers Sharing the Ride	Car	4 to 6 passengers		
	Bus			
	Jitney			
Internet/WiFi Availability	Car	Not Available		
	Bus			
	Jitney			

Table 21: SRT Fare Variations Done on the Basic Service

	Private Car/Bus/Jitney		
Region	A	B	C
Low Taxi Fare	2,500 L.L.	4,000 L.L.	6,000 L.L.
Medium Taxi Fare	4,000 L.L.	5,000 L.L.	8,000 L.L.
High Taxi Fare	6,000 L.L.	7,000 L.L.	10,000 L.L.

Table 22: Delta Time (in minutes) Variations Done on Basic Scenario

Data Set	**Jitney**			**Bus**			**Private Car**		
Region	A	B	C	A	B	C	A	B	C
Low SRT Time	-5	-10		-10	-15	-20	0	+5	+10
Medium SRT Time	0	-5		-5	-10	-15	+5	+10	+15
High SRT Time	+5	0		0	-5	-10	+10	+15	+20

Table 23: MAWT (in minutes) Variations Done on Basic Scenario

Data Set	**PT/ PC**		
Region	A	B	C
Low MAWT	0	0	0
Medium MAWT	5	7	7
High MAWT	10	15	15

Table 24: Vehicle Size Variations Done on Basic Service

Data Set	**PT/ PC**		
Region	A	B	C
Taxi	1 to 3 passengers		
Minivan	4 to 6 passengers		

Table 25: Wi-Fi Availability Variations Done on Basic Scenario

Data Set	**PT/ PC**		
Region	A	B	C
Available	Available		
Not Available	Not Available		