Technical University of Munich.
Department of Civil, Geo and Environmental Engineering
Chair of Transportation Systems Engineering

Master's Thesis

Why Do We Share Our Rides?

Investigating Factors Influencing the Adoption, Use Frequency, and Characteristics of Organized Pooled Rides. The Case Study of Jetty in Mexico

Author:
Mohamed Abouelela

Master's thesis submitted under the supervision of
Univ-Prof. Dr. Alejandro Tirachini, Universidad de Chile

the co-supervision of
Univ-Prof. Dr. Constantinos Antoniou, TUM

the co-supervision of
Univ-Dr. Manos Chaniotakis, UCL

in order to be awarded the Degree of
Master's of Science in Transportation Systems

May 25, 2020

Declaration Concerning the Master's Thesis

I hereby declare that all the work concluded in this research is my work, and all the material and resources used were referenced in the corresponding locations. This thesis was never submitted anywhere for assessment.

Munich, 25^{th} May, 2020

Mohamed Abouelela

Acknowledgements

I would like to express my sincere gratitude for my Supervisors Prof. Dr. Alejandro Tirachini, Prof. Dr. Constantinos Antoniou, and Dr. Manos Chaniotakis for their guidance, support, and patience during my trip to complete this work.

Secondly, I would like to praise my family, and my friends who supported me along the journey of grad school.

Finally, I would like to thank all the Jetty team, especially the founder, Mr. Onésimo Flores, for allowing and supporting us to survey their customers, which was the core of this thesis.

Abstract

Shared mobility services have recently disturbed many aspects of the urban transportation landscape, encouraged by services' popularity, ease of use, and advancement in information and communication technologies (ICT). However, the full understanding of the different shared services and their synergies with other components of urban environments is not fully understood. One of the least studied services is pooled services. We have limited knowledge about what modes are being replaced by pooled services, the interaction between pooled services and other modes of transportation, and their general demand characteristics.

This thesis tries to bridge the current knowledge gap by investigating the choice of shifting to pooled services, the choice between different pooled service setups, and the use frequency of pooled services through developing three hybrid choice models and one binary logit model. The estimation was performed using data collected via an online survey conducted in Mexico City, Mexico (CDMX) for users of a commercially organized pooled service, Jetty. Pooled service adoption process and use were modeled as a function of the users' sociodemographic, latent travel attitude, public transportation (PUT) accessibility, trip-related characteristics, reasons to use the service, and users' activities during the trips.

The estimated models show that users' sociodemographic and travel attitudes are the main factors impacting the shift from different modes to Jetty. Service-related characteristics such as multi-tasking, trip fare, and avoiding parking problems are also impacting the shift decision. Sociodemographic factors drive the choice between the different service categories in addition to the access and egress distances to the service pick-up and drop-off locations and modes used for the same tasks. On the other hand, the frequency of service use is mainly impacted by trip characteristics such as total trip distance, access, and egress distance, and the waiting time at the user's nearest Metro stations. Income and employment were the only two sociodemographic factors impacting the use frequency.

Table of Contents

Table of Contents IV

List of Figures VI

List of Tables VIII

List of Abbreviations X

1 Introduction 1

1.1 Introduction . . . 1

1.2 Motivation . . . 2

1.3 Research Objectives and Questions . . . 3

1.4 Expected Contributions . . . 4

1.5 Research Framework . . . 5

2 Literature Review 6

2.1 Sharing Economy . . . 6

2.2 Shared Mobility . . . 7

2.2.1 Taxonomy and Definitions . . . 8

2.2.2 Travel Demand . . . 12

2.2.3 Shared Mobility User's Profile . . . 13

2.2.4 Shared Mobility Trip Purpose . . . 14

2.2.5 Exogenous Factors Impacting Shared Mobility Use . . . 15

2.2.6 Reasons to Use Shared Mobility . . . 17

2.3 Shared Mobility Study Methodology Framework . . . 17

2.3.1 Data Collection . . . 17

2.3.2 Data Preparation . . . 18

2.3.3 Modeling Techniques . . . 20

3 Methodology 31

3.1 Data Collection . . . 32

3.1.1 Jetty Users' Survey . . . 32

3.1.2 Jetty Use Database . . . 35

3.1.3 GTFS files . . . 35

3.2 Data Analysis and Processing . . . 36

3.2.1 Explanatory Data Analysis . . . 36

3.2.2 Data Preparation . . . 37

3.3 Modeling Framework . . . 38

3.4 Tested Hypotheses . . . 39

4 Case Study 41

4.1 Transportation in Mexico City . . . 41

4.2 Jetty . . . 48

4.3 Data Analysis . . . 50

4.3.1 Survey Data . . . 50

4.3.2 Jetty Database . . . 73

4.3.3 GTFS Files 79
4.4 Data Preparation 80
4.4.1 Zero and Near-Zero Variance 80
4.4.2 Collinearity Detection 80
4.4.3 Numerical Variables Scaling 82
4.4.4 One-Hot Encoding 83
4.4.5 Exploratory Factor Analysis 83

5 Modeling Process **84**
5.1 Shift to Jetty 84
5.2 Services Choice 91
5.3 Use Frequency 96
5.3.1 Ordinal Logit Hybrid Choice Model 96
5.3.2 Multinomial Logit Hybrid Choice Model 101

6 Discussions **105**
6.1 Findings 105
6.1.1 Data Analysis Findings 105
6.1.2 EFA Findings 107
6.1.3 Models Findings 107
6.2 Tested Hypotheses 108
6.3 Policy implication 110
6.4 Lessons Learned from the Survey 113

7 Conclusion **114**
7.1 Conclusion 114
7.2 Study Limitation 115
7.3 Recommendations and Future work 116

Bibliography **117**

Appendices **137**

A Additional Analysis **137**

B Additional Models **144**

List of Figures

1.1 Thesis Framework (*own illustration*) 5

2.1 Ridesharing categories, source [53] 10
2.2 Ordered Probability Model, Source [269]. 24
2.3 HCM model, source [28, 135] 25

3.1 Methodology Framework (*own illustration*) 31

4.1 CDMX Federal District Boundary, and Metropolitan Area Boundary, Source: [120] 41
4.2 RTP Network, Source: [166] 43
4.3 CDMX Public Transportation Map, Source: [166] 44
4.4 Comparison between Jetty and Put Trip, Source: *Jetty*[7] 48
4.5 Jetty Routes and Stations, Source: [188] 49
4.6 Survey Responses Time Distribution 50
4.7 Percentage of Last Trip OD Occurrence to the Total Jetty Use ODs 53
4.8 Last Jetty Trip Purpose 54
4.9 Access and Egress Disaggregated Modes 55
4.10 Access and Egress Top Ten Combined Modes 55
4.11 Access and Egress Top Ten Combined Modes by Peak Hour 56
4.12 Access and Egress Time Distribution, log scale 57
4.13 Access and Egress Distances Distribution 57
4.14 Disaggregated Modes to Replace Last Jetty Trip 59
4.15 Top Ten Combined Modes to Replace Last Jetty Trip 60
4.16 Jetty Trip and Replaced Trip Time Distributions 61
4.17 Willingness to Walk Time Distribution 63
4.18 Modes Use Frequency 64
4.19 Distribution of Disaggregated Users' Activities During Jetty Trips per Gender 65
4.20 Distribution of Disaggregated Reasons to Use Jetty per Gender 66
4.21 Home and Work Locations, Background: Google Maps, (*own illustration*) . 67
4.22 A)Home and B)Work Concentration Locations, Background: Google Maps, (*own illustration*) 68
4.23 Survey Comments Word Cloud (*own illustration*) 69
4.24 Associated Words (*own illustration*) 70
4.25 Sentiment Analysis Using Finn Årup Nielsen Word List, (*own illustration*) . 72
4.26 Sentiment Analysis Using EmoLex, (*own illustration*) 72
4.27 Number of Jetty Trips per User Distribution 73
4.28 Relative Jetty Use Frequency Distribution per Gender 74
4.29 Jetty Trips Departure Time 75
4.30 Ratio of Jetty Trip Distance to Jetty Total Route Distance 77

4.31 Percentage of Total Jetty Trips per Vehicle Type 77
4.32 Percentage of User's Trips per Vehicle Type 78
4.33 Jetty Vehicles Occupancy Rate, May 2019 78
4.34 Sociodemographic Variables Correlation Coefficients 81
4.35 Jetty Use Related Variables Correlation Coefficients 81
4.36 GTFS Files Data Correlation Coefficients 82

5.1 Shift to Jetty HCM Full Path Diagram (*own illustration*) 85
5.2 Jetty Use Frequency Ordered-HCM Full Path Diagram (*own illustration*) . 97
5.3 Jetty Use Frequency MNL-HCM Full Path Diagram (*own illustration*) . . . 102

A.1 Percentage of Jetty Trips per Workday 140
A.2 Jetty Route Distance and Trip Distance Distribution 140
A.3 Ratio of Jetty Trip Distance to Route Distance, CDF 141
A.4 Jetty and Replaced Trip Travel Time Distribution by Peak Hour 141
A.5 Shift to Jetty Users' Sentiment Score 143
A.6 Jetty Service Choice Users' Sentiment Score 143
A.7 Jetty Use Frequency Users' Sentiment Score 143

List of Tables

2.1 Shared Mobility Users' Profile, Selected Studies 14
2.2 Shard Mobility Trip Purpose Summary, Selected Studies 15
2.3 Exogenous Factors Affecting the Use of Shared Mobility, Selected Studies . 16
2.4 Reasons to Use Shared Mobility, Selected Studies 17
2.5 HCM Measures of Fit, Selected Studies, Source: [173] 29

4.1 Travelling Population by Mode in Working Day, CDMX Source: [190] . . . 46
4.2 Survey Response Time Summary 50
4.3 Survey Sample Sociodemographic Summary Statistics 52
4.4 Access and Egress Time Summary Statistics 56
4.5 Access and Egress Distance Summary Statistics 58
4.6 Number of Modes to Replace Jetty Summary Statistic 58
4.7 Disaggregated Modes Replacing Last Jetty Trip 59
4.8 Jetty Trip Time Summary Statistics 60
4.9 Users' Anticipated Car Parking Cost Summary Statistics 61
4.10 Willingness to Walk Time Summary Statistics 62
4.11 Willingness to Walk Ordered Logistic Regression 63
4.12 Survey Comments Sentiment Score Summary Statistics 71
4.13 Number of Jetty Trips per User Summary Statistics 73
4.14 Percent of Morning Trips Summary Statistics 75
4.15 Trip Characteristic Summary Statistics 76
4.16 GTFS Files Data Summary Statistics 79
4.17 EFA Results 83

5.1 Combined Modes Replacing Last Jetty Trip Summary Statistics 84
5.2 Trips in Three Different Modes Summary 85
5.3 Shift to Jetty from Car-Based Modes HCM and Binary Logit Model Results 89
5.4 Shift to Jetty from Car-Based Modes Latent Variable Model 89
5.5 Shift to Jetty from Car-Based Modes Significant variables Summary 90
5.6 Most Frequently Used vehicle Type per User Summary Statistics 91
5.7 Service Choice Binary Logit Model Results 94
5.8 Service Choice Model Significant Variables Summary 95
5.9 Jetty Use Frequency Summary Statistics 96
5.10 Jetty Use Frequency Model-1 Results 100
5.11 Jetty Use Frequency Model-2 Results 103
5.12 Jetty Use Frequency Model Significant Variables Summary Statistics 104

6.1 Tested Hypotheses Summary 110

A.1 Combined Modes to Replace Last Jetty Trip 137

A.2 Access and Egress Time by Mode, Top Five Modes Summary Statistics . . 138
A.3 Disaggregated Activities During Jetty Trip Summary Statistics 138
A.4 Disaggregated Reasons to Use Jetty Summary Statistics 138
A.5 Modes Use Frequency Summary Statistics 139
A.6 Shift to Jetty Users' Sentiment Score Summary Statistics 142
A.7 Jetty Service Choice Users' Sentiment Score Summary Statistics 142
A.8 Jetty Use Frequency Users' Sentiment Score Summary Statistics 142

B.1 Shift to Jetty HCM, Model-2 144
B.2 Shift to Jetty HCM, Model-3 145
B.3 Service Choice Binary Probit Model Results 146
B.4 Jetty Use Frequency Ordered Probit Model 147

List of Abbreviations

AIC	Akaike Information Criterion
ATS	Alternative Transit Service
B2B	Business to Business
B2C	Business to Consumer
BIC	Bayesian Information Criteria
C2C	Consumer to Consumer
CDMX	Ciudad de México
CNS	Courier Network Services
Cov	Covariance
DRT	Demand Responsive Transport
EFA	Confirmatory Factor Analysis
EFA	Explanatory Factor Analysis
GTFS	General Transit Feed Specification
HCM	Hybrid Choice Model
HOV	High Occupancy Vehicle Lane
ICLV	Integrated Choice and Latent Variable Models
ICT	Information and Communication Technology
LRT	Likelihood Ratio Test
LV	Latent Variable
ML	Maximum Likelihood
MOD	Mobility on Demand
MXN	Mexican Peso
OD	Origin Destination
OLS	Ordinary Least Squares
P2P	Peer to Peer
Pax	Passenger
PC	Personal Computer
PCA	Principle Component Analysis
PUT	Public Transportation
PVS	Personal Vehicle Sharing
S.E.	Standard Error
SD	Standard Deviation
TCRP	Transit Cooperative Research Program
TNC	Transport Network Company
VIF	Variance Inflation Rate
VKT	Vehicle Kilometers of Travel
ZMVM	Zona Metropolitanadel del Valle de México

Chapter 1

Introduction

This chapter gives a comprehensive introduction to the thesis, explains the motivation behind the research, defines the research questions and the expected contributions, and explains the structure of this thesis.

1.1. Introduction

The last decade has witnessed rapid and dynamic developments of new technologies and business models, resulting in breakthroughs encouraged and supported by the advancement of information and communication technologies (ICT). These technologies helped and facilitated new economic systems and models such as e-commerce and sharing economies. Consumers can buy or share products using their PC or smartphone from virtual online shops, companies, or individual providers.

The sharing economy is one of the most influential economic trends that emerged in the last decade. Supported by goals of sustainability, social, economic, and environmental benefits, the growth of the sharing economy in its five key sectors (accommodation, passenger transport, household services, professional and technical service, and collaborative finance) is noteworthy. An example of the substantial economic gains resulting from the sharing economy can be found in the European Union (EU). In 2015, the EU revenues from sharing economies were 28 Billion EUR, with experts estimating a potential of extra (160-572) billion EUR to be added to the EU economy by the sharing economy [81].

One of the five main pillars of sharing economy investment is the transportation sector, which is also a fundamental corner of urban life. The new sharing economy transportation services, referred to as shared mobility, have introduced new options to the traditional urban transport ecosystem previously consisting of only private and public transportation.

Shared mobility can be defined[1] as the concurrent or the sequential use of the different modes of transport including, but not limited to, passenger-vehicles, bicycles, scooters, or any other travel modes. It gives the users temporary access to use the different travel modes based on a pay-per-use basis [229, 225]. Shared mobility services range from micro shared services, such as scooter sharing and bikesharing, to the replacement or the complement/integration of public transportation in the case of alternative transit service, and micro-transit. Furthermore, the use of shared mobility services has expeditiously grown on a global scale as a part of the growth of sharing economy. For example, the number of on-demand ride services has grown from one million trips in 2013 to 25 billion trips worldwide in 2016 [221]. Another example of the growth of shared mobility is Uber, the well-known on-demand ride service company. Uber started its operation in one city, San Francisco, California, USA in 2012, and now it operates in 63 countries, 700 cities, and services 91 million monthly active users worldwide [63].

The well-advertised sustainability benefits of the sharing economy, and subsequently shared mobility, present new opportunities to the urban transportation sector, as they can help reducing traffic externalities such as air and noise pollution and traffic congestion [209]. Transportation sector depends strongly on fossil fuel consumption; therefore,

[1] A detailed definition of shared mobility is discussed in Chapter-2

it is one of the top four global sources of GHG emissions [35, 74]. Transportation's share of global oil demand has risen from 33% in 1971 to 47% in 2002, and it is expected to be 54% by 2030 if no intervention takes place. Besides, transportation produces 13% of Greenhouse Gas (GHG) globally, and if the current travel trends hold, without the extensive adoption of cleaner sources of energy and higher utilization, the transportation industry's production of GHG is expected to rise to 40% by 2050 [50]. In the EU, 23.2% of GHG are produced by transportation, and 72% of transportation GHG emissions are generated from road transportation [82].

The global phenomenon of air pollution is more evident in developing countries, as they mainly depend on fossil fuel as a prime energy source, and limited resources curb their expenditure, and investments to improve air quality [171, 104, 6]. Developing countries produce a large share of GHG, and subsequently, they have a considerable impact on global climate change. Also, developing countries' residents are expected to be ten times more likely to be affected by climate change and associated disasters than wealthier countries [6]. At the same time, eight of the world's top ten most populated megacities are in developing countries. Additionally, 27 of the world's total of 33 megacities are located in developing countries [258]. These countries will experience rapid population growth accounting for 90% of the world's total population growth by 2050, placing tremendous stress on infrastructure and essential urban services [49].

One of the over-stressed sectors of urban life is the transport sector, where dwellers of developing countries mainly depend on public transport and paratransit for their daily commute [50]. Public transportation supply in developing countries is not always sufficient to meet the demand. The service quality is generally poor, and in some cases, dangerous, which pushes people to use private motorized modes [119, 197, 50]. The increase in motorization rates in developing countries will result in more pollution and energy consumption. For example, the side effects of air pollution in developing countries has increased the mortality rates when compared to countries with healthier air [236, 154].

An immediate expansion of infrastructure and changes in land-use policies are required to cope with the current rapid increase in travel demand in developing countries. Infrastructure and land-use projects require substantial investments, considering the limited resources in developing countries; such plans are hard to materialize soon [80, 33].

Shared mobility could be a quick promising solution for transit system problems in developing countries. The introduction of shared mobility to the developing countries' markets could increase transportation systems' efficiency, reduce energy consumption and related emissions, mitigate traffic congestion, and reduce the demand on the infrastructure, in addition to the economic benefits [25, 224].

1.2. Motivation

The understanding of shared mobility and its interactions with different urban environment components, such as users, transport supply, land use, accessibility, and travel demand, is not heavily studied, especially for pooled services. Moreover, the impact of new shared mobility services on long and short term travel decisions are not fully comprehended. Also, the integration of the shared mobility services in travel demand models is not regularly implemented [142, 228]. Previous needs to the research of shared mobility are due to the novelty of evolving services that attract more users and increase the provider's profitability. These factors broadly impact the population use and understanding of available or new shared services [59, 142].

The scarcity of user-level data presents another limitation to shared mobility research. Current studies largely depend on aggregated data sources due to the ethical and legal issues mobility companies face surrounding public sharing of consumer data [149].

Existing research tries to address some of the previous limitations of shared mobility studies concentrated in North America and Europe. A limited number of published papers discuss shared mobility in China and India. This shortage in research demonstrates how research in the developing countries is limited in general, not only in the case of transportation-related studies [40].

Previous limitations within existing literature motivated the development of the following research and the unique setup for the proposed case study of a commercial, third-party, organized pooled shared mobility service in Mexico. The service in question, Jetty, operates in Mexico City, Mexico, the world's fifth big city in 2018 [258]. The city has one of the highest motorization rates of developing countries. For each newborn baby, two cars are added to the city roads [123]. Moreover, according to INRIX 2019 Global Traffic Scorecard[2], Mexico City is the most crowded city in Mexico, the second most crowded city in the northern hemisphere, and the world's third most crowded city in 2018. Also, it was possible to collect the service users' travel data and detailed interviews for the same users.

1.3. Research Objectives and Questions

The main objectives of this research are to investigate characteristics of shared mobility users in developing countries, explore the interaction between the shared mobility and public transit and traffic conditions, investigate motivations pushing people from different modes to shared mobility services, the impact of people travel attitudes on their choice and use of shared mobility services, and exogenous factors impacting a user's choice for different shared services.

Exploring the various research objectives will provide insights to planners, operators, and authorities to quantify shared mobility demand, plan new services, and address potential downsides of public transport networks that reduce ridership. Moreover, this research will help identify different factors that help integrate new shared mobility services within existing public transport systems. Achieving the research objectives will be done by answering the main research questions:

1) ***Are*** shared mobility users' profiles in Mexico city, an example of a city in a developing country, the same as in Europe and in North America?

2) ***What*** are the synergies between public and private transportation and shared mobility services?

3) ***What*** factors affect a users' decision to adopt shared mobility services?

4) ***What*** factors affect a users' decision to choose between different shared mobility services?

5) ***What*** are the factors that affect the frequency of use of shared mobility?

[2] inrix.com/scorecard-city/?city=Mexico%20City&index=3

6)***How*** does the users' daily travel behaviour impact their frequency of using shared mobility?

1.4. Expected Contributions

Answering the previous research questions throughout this research methodology is expected to have the following contributions:

Theoretical Contribution

- Provide theoretical insights on the use of shared mobility and especially pooled services in developing countries.
- Exploration and definition of shared mobility users' profiles in developing countries and compare them with other users in other developed countries.
- Identify the factors affecting the acceptance of pooled services.
- Identify the impact of user travel patterns, attitude, on shared mobility use.

Methodological Contribution

- Developing a framework to study shared mobility in the context of the Global South
- Modeling the synergy between land use, shared mobility, accessibility to public transport, and users' sociodemographics
- Integrating open-source data with a survey questionnaire to identify factors affecting different modes use on the individual level.
- Modeling the factors affecting adoption of and frequency of shared services use using choice models and hybrid choice models, HCM.

Practical Contribution

- Providing practical insights about shared mobility demand characteristics.
- Identification of the factors that push users away from public transportation and the factors that help integrate new shared mobility services within public transport networks.

1.5. Research Framework

This thesis consists of seven chapters. The first chapter includes the introduction, where the thesis problem and research questions are defined. The second chapter is the literature review, which has two prime objectives; the first is to explain the concepts of the sharing economy and shared mobility. The second objective is to explore the standard methodologies to study shared mobility. The third chapter, the methodology, explains the methods used in this thesis in detail. The fourth chapter is the case study, where a case study setup, implementation, and the corresponding data analysis is performed. The fifth chapter is the modeling chapter, in which the models developed to answer the rest of the research questions are estimated. The last two chapters are the discussion and conclusion, where the general thesis conclusions and policy recommendations are developed and presented. Figure-1.1 shows the thesis framework.

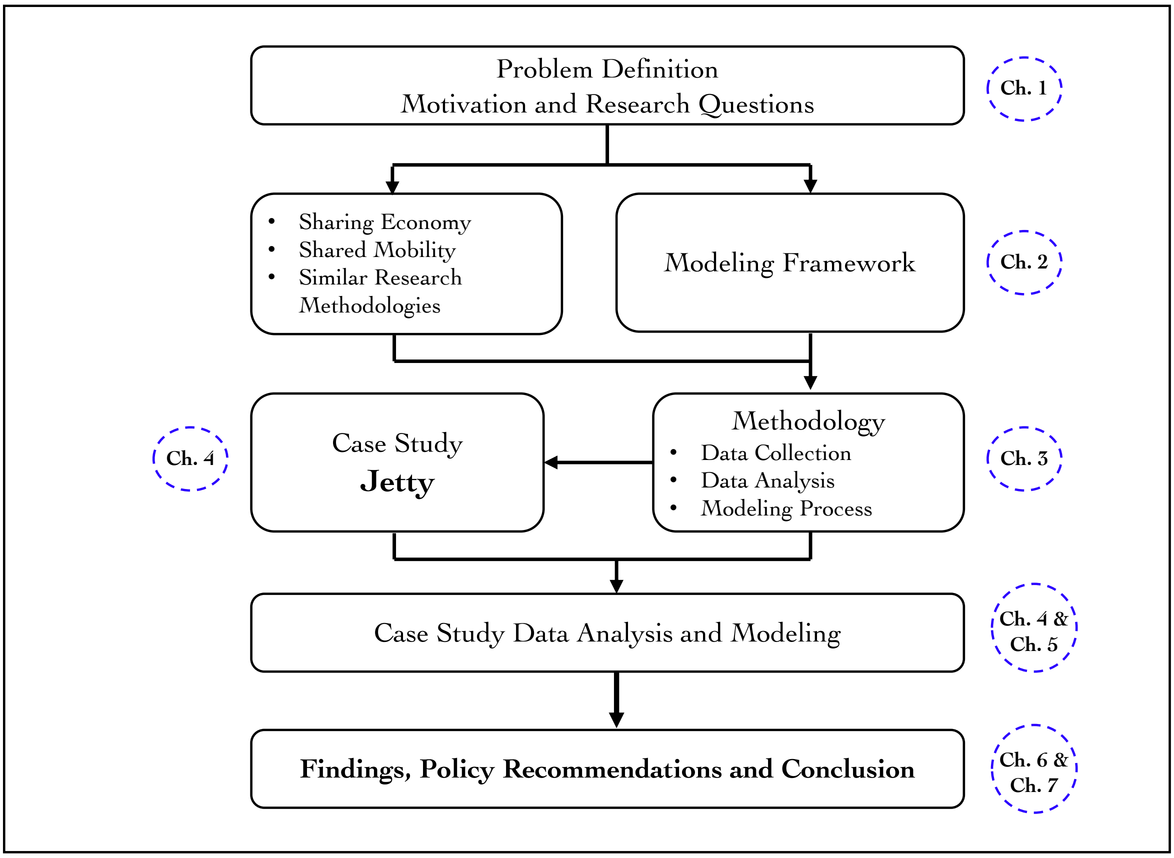

Figure 1.1 – Thesis Framework (*own illustration*)

Chapter 2

Literature Review

This chapter's objective is to provide the scientific background in terms of definitions, calculations, and methodologies that are used in this thesis. The chapter consists of two sections. The first section discusses the concept of shared mobility, its demand characterization, and the different services users' profile. The fundamental dimensions of shared mobility to establish the basis for the survey questions used in the next chapters are discussed. The second section of this chapter reviews the methodologies used in studying similar shared mobility services, focusing on data collection and data processing techniques. The data processing part thoroughly discusses the statistical modeling process used to answer the main research questions and the relevant literature.

2.1. Sharing Economy

Definition

The recent advances in information and communication technologies (ICT) has encouraged the exchange of goods, skills, space, and services through digital platforms, websites, and smartphone applications. This exchange materializes directly between the providers and consumers without the intervention of a third party in what is commonly called the sharing economy or collaborative consumption. The digital platforms do not hold physical goods or provide services. Instead, they connect or match people, owners, and providers with consumers. These steps curtail the cost of coordination processes and sometimes evade costly governmental regulations associated with obtaining operation licenses [10, 110, 132].

Business Models

The sharing economy utilizes different business models; notwithstanding, the are two primary models: Consumer-to-Consumer (C2C) and Business-to-Consumer (B2C). C2C is sometimes referred to as the Peer-to-Peer model, where the individual providers grant access to other people to use their underutilized services, goods, vehicles, or space. Examples of this model are Blablacar for ride-sharing, Peerby for tool sharing, and Airbnb for spacing sharing [196]. In the B2C model, a company owns the assets and grants access to users based on membership fees plus use-based fees. Examples of this model are Lime scooter, oBike, Zipcar, Car2go, and DriveNow [61].

Motivations

The incentives to participate in sharing economies can be encapsulated into three main groups: I) Fiscal benefits that come from the cost-cutting of operation and coordination; II) Social benefits of bringing people together to achieve social cohesion; III) Environmental benefits of raising the efficiency of the use of different goods and reduce energy consumption. Indeed, the motivations to participate in the sharing economy will differ by different sectors, sociodemographic groups, and users or providers [38, 39, 4].

Sharing Economy Sustainability Appraisal

The main challenge of the sharing economy is to achieve the motivations anticipated by users. As advertised, the sharing economy has positive impacts on the economy; however, these impacts are more complex to calculate, evaluate, and verify. For example, the hotel industry in Texas has experienced significant reductions in its earnings resulting from Airbnb's growth in Texas [278]. Additionally, wealthy providers are making the majority of profits in the sharing economy due to owning the majority of physical properties. Their domination of the sharing market results in unequal distribution within the sharing economy [89].

Environmental impacts of the sharing economy are not always as positive or eco-friendly as advertised by the platforms. Empirical evidence does not currently exist to illustrate the positive effects within all the cases of sharing economy. Adverse side effects happen, but they are categorically ignored [263, 89]. For example, bikesharing systems are advertised to reduce congestion, emissions, and fuel use [217]. However, the imperial evidence in London showed that the introduction of bikesharing had increased the total kilometer traveled (VKT) by motorized traffic, which was an unexpected adverse impact. This increase in VKT is due to two main factors. The first was that most of the replaced trips were public transportation trips. The second factor was the extra truck mileage added due to the bike redistribution process [85]. Reducing the services' cost and the associated financial savings can also increase usage rates. The low price of new motorized-ride-services encourages people to travel more frequently, thus increasing the VKT and the associated volume of carbon emissions harmful to the environment [273].

Social benefits are also doubted to last. More people are participating in sharing platforms to gain money instead of the platform's advertised targets of creating social relations. Participants have a sense of disenchantment related to the relationships created through various platforms and do not expect them to last [89]. The overall long-term-impacts of the sharing economy are inconclusive and hard to anticipate; however, even though the short-term impacts are more direct and obvious, their overall distribution can be skewed to a specific population [212, 153].

2.2. Shared Mobility

Shared mobility is a form of the sharing economy providing travelers with short-term-access usage of different travel modes on a need to travel basis. The term shared mobility is an umbrella for a wide range of services that can be, but not limited to, scooter sharing, bike sharing, carsharing, and ridesharing in different-sized vehicles (carpooling, vanpooling), and demand-responsive services also known as mobility-on-demand services (MOD). Shared mobility can provide alternatives for public transportation with different varieties of services such as paratransit, shuttle, and microtransit. This group of services can either replace and complement fixed-route public transport road-based services and public transport rail-based services. Shared mobility services also extend to the urban freight transport sector, where the delivery of parcels could be coupled with people's transportation, such as courier network services [223, 60, 241, 174]

Shared mobility can be categorized into main seven categories: [223, 61, 241]

1. Scooter Sharing
2. Bikesharing
3. Carsharing
4. Ridesharing
5. On-Demand Ride Services
6. Alternative Transit Services
7. Courier Network Services

A unified definition is vital for the different shared mobility services for both public and private sectors to eliminate confusion and discrepancies. Standard terms allow the clarification of the various policies for each service. The shared mobility policies could include but are not limited to, parking use regulation, operation license fees, smog checks, taxation, insurance, right of way, and operation zoning related to each type of these services. Furthermore, the private sector needs to clearly define individual mobility services to plan for different aspects of each service and their operations [229, 225].

2.2.1. Taxonomy and Definitions

The main groups of shared mobility are defined as:

Scooter Sharing

Scooter sharing is the newest introduction to the shared mobility family. Shared scooter systems include two types of scooters: standing electronic-scooters (kick-scooter) and Moped-style scooters. Kick-scooters are a shared dock-less (free-floating) electronic powered scooter that can be picked up and dropped off anywhere. Moped-style scooters are the most recent addition to scooter sharing. Users can sit to ride and operate the vehicle. Also, Moped-style scooters can have gasoline or electric propulsion systems, 97% of moped-style scooters are electric, and operating on a dock-less system [223, 1, 222, 60, 229, 116].

Bikesharing

Bikesharing systems can be categorized into three main system types. The first type is the free-floating systems where users pick up and drop-off the bicycle anywhere in the operating zone. The second system is the station-based. Users pick up and drop off bikes at dedicated and branded stations. The third system is the hybrid system, which is a station-based system, but in some areas of the city users are allowed to drop the bike anywhere, not within a station [223, 1, 222, 60, 229].

Carsharing

Carsharing is defined as a fleet of small size road vehicles owned by an operator and can be accessed by users of the platform. Carsharing gives users the same level of freedom that comes with using a private car without the duties of owning a car.

Carsharing systems operate in three different operational models. The first model is the round-trip carsharing system. Users acquire the vehicle from a fixed station and return the car to the same station. The second model is the one-way system, where the user picks up the car from one station and returns it to any branded location within the company. The third and most recent model is the free-floating model. Vehicles are scattered around the city, and the user can start his journey by using the car closest to his location. The transaction is finalized when the vehicle is returned within a geographically confined area.

Carsharing business models follow the general B2C and P2P models of sharing economy. P2P or personal-vehicle-sharing (PVS), where the owner of the vehicle rent his car for a short term to another user directly with, or without third-party intervention has one main advantage, which is the ability to extend the carsharing service outside of the urban dense population areas where regular carsharing services usually operate [223, 60, 241, 125, 67, 107, 229].

Ridesharing

Ridesharing is a mode of transportation where users with matching or partially matching itineraries and schedules share a vehicle and the emanated cost of the trip [91, 247, 237, 53, 214, 241, 229].

Ridesharing includes carpooling and vanpooling and their subcategories categorized by the size of the vehicle. According to the Federal Highway Administration, vanpooling is described as a group of seven to fifteen passengers traveling together; if the number of passengers is less than seven, it is considered carpooling [97]. Moreover, ridesharing can be categorized based on the relation between the users of the pooling services into three main groups: I) acquaintance based service where users know each other, such as coworkers or family members (Fampool). II) Organization-based services where users join the service through digital platforms. III) Ad-hoc ridesharing (slugging), where drivers pick up share riders to gain access to the highway's High Occupancy Vehicle lanes (HOV) [241, 233]. Figure-2.1 shows the different categories of ridesharing.

On Demand Ride Services

On Demand Ride Services, Transportation Network Companies (TNCs), ridesourcing, ride-hailing, e-ridehailing, ride-booking are the different commonly used names for the same transportation mode. The basic model of these services is that passengers use a mobile app to connect to the nearest available driver that uses his private vehicle to transport the passenger from his current point to another point [241, 229]. A well-known example of these services providers is Uber.

Another form of on-demand services is ride-splitting. Users with matching or partially matching routes share the ride to reduce the travel cost [241].

It is worth mentioning that the introduction of TNCs to the transportation market has forced a reimagining of traditional taxi services. E-hailing now enables users to hail a taxi through a mobile application. The main differences between traditional taxi services and TNCs are that professional drivers use a standard taxi auto and do not apply surging price schemes. E-hailing is not limited to conventional passenger vehicles, but it is extended to other vehicles such as Pedicab[1] or rickshaw[2] [241].

[1] stpetepedicab.com, last accessed on 13/05/20
[2] globallyspotted.com/auto-n-cab-smartest-auto-hailing-app/, last accessed on 11/12/19

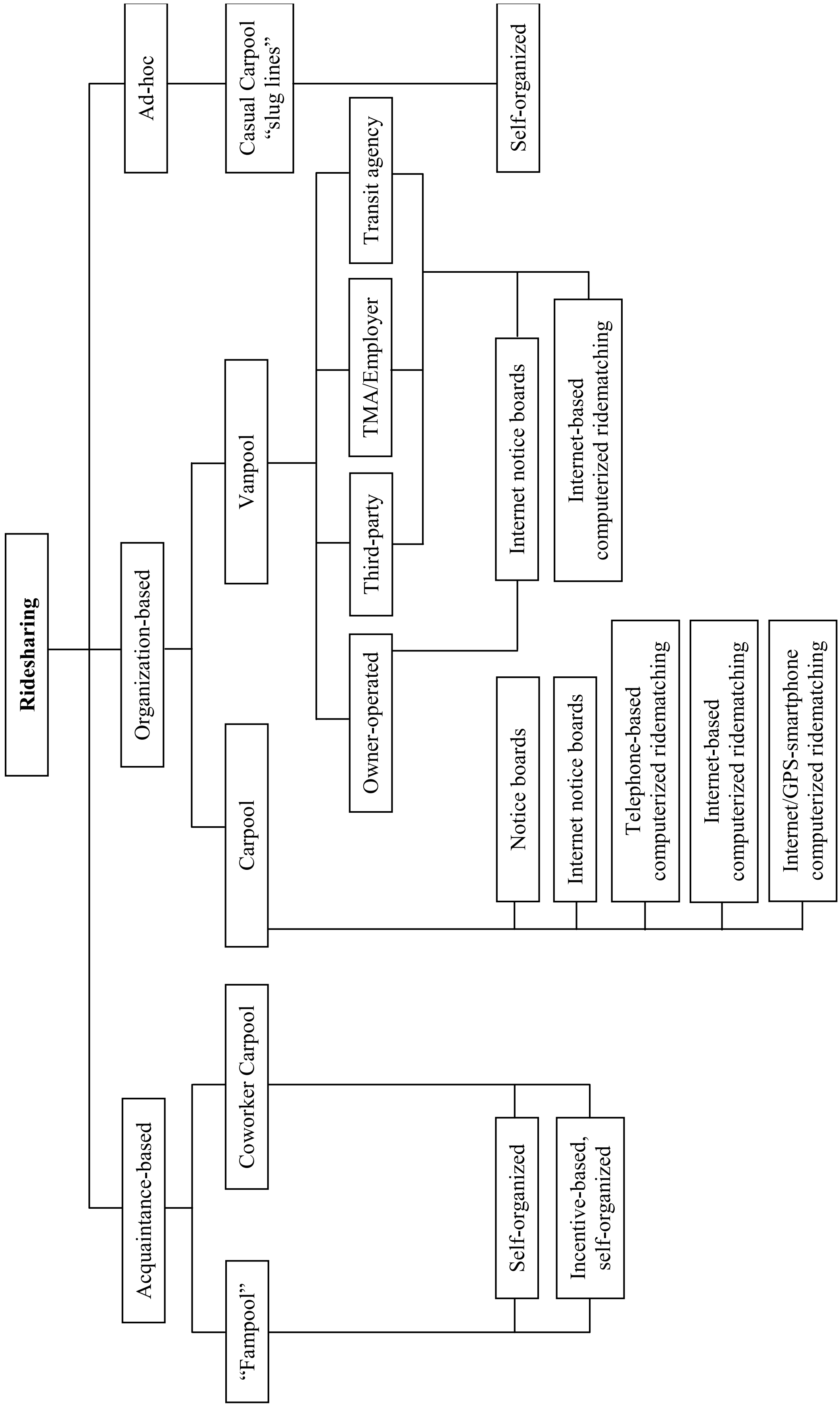

Figure 2.1 – Ridesharing categories, source [53]

Alternative Transit Services

Alternative transit services (ATS) and Demand Responsive Transport (DRT) are services that run lateral to public transit. These two formats encircle extensive types of modes such as paratransit, shuttles, microtransit, vans, dollar vans, jitneys, and small buses. Alternative transit services differ from public transportation as they are costly to the provider as it caters to a low volume of traveler per trip. ATS generally try to address the spatial or temporal gaps in public transportation coverage [229, 60, 241].

Shuttles are widely used form of ATS. They are generally vans or small size buses operated by professional drivers. Shuttles mainly connect passengers hot spots areas to public transportation stations providing first and last mile services, or they provide direct connections to job centers and hospitality centers, and in some cases, they replace public transportation lines during maintenance [229, 60, 241].

Microtransit is another form of ATS that has recently advanced due to advances in routing and tracking technologies and the growth of accessibility in information and communication technologies (ICT). Microtransit is a publicly or privately operated service that uses buses, vans or shuttles, and could be operated on fixed, flexible routes, schedules, or on-demand operations [51, 229, 241, 60]. According to the Transit Cooperative Research Program (TCRP) in North America [253], microtransit operation schemes can be categorized in six main categories:

Route Deviation: Vehicles operate on scheduled, determined routes, but can deviate from the course to respond to demand within a zone around the route. The zone could be predefined or flexible.

Point Deviation: Vehicles provide a demand responsive service within a zone without fixed routes.

Demand-Responsive Connections: Vehicles provide a demand responsive service within a delineated area, with one or more connections to a fixed-route-network.

Request Stops: Vehicles operate along defined routes and schedules, and a finite number of stations near to the route based on users' demand requests.

Flexible-Route Segment: Vehicles operate along defined routes and schedules; however, the route could be flexible in response to the demand for limited parts.

Zone Route: Vehicles operate along a corridor; the users define its alignment with a scheduled arrival and departure times to one or more endpoints.

Courier Network Services

Courier Network Services (CNS) or flexible goods delivery is a service enabled through a digital platform, mobile application, or a website, where the freight providers (food or goods) are connected to individual couriers using private vehicles, scooters, or bikes to perform the delivery process. CNS is still evolving in different shapes and business models. The two primary business models are: I) Peer-to-Peer services that allow any participant to perform a delivery using his private transport mode. II) Paired On-Demand Passenger Ride and Courier Services allow TNCs to provide freight delivery in separate trips or within the same trips with passengers [216, 241, 60, 229].

2.2.2. Travel Demand

The main research questions, section-1.3, are mainly concerned with shared mobility demand characterizations. Therefore, understanding the travel demand and the factors influencing it is fundamental in constructing a comprehensive overview of travel demand.

Definition, and Influencing Factors

The efficiency of transportation infrastructure has played a significant role in a society's economic success beginning in the early days of civilization [32]. Travel demand and the factors influencing it are essential for determining the efficiency of the current transportation infrastructure and the need for future investments. Moreover, the accurate prediction of travel demand supports long term planning and investment decisions [256].

Travel demand can be defined as the volume and type of travel that people perform under certain conditions influencing their travel choices. Travel choices could be long term choices like home and work locations and car ownership, or short term decisions like travel frequency, trip time, and destination [44]. Travel demand can be categorized based on a trip's purpose as home-based trips or leisure trips, based on the used mode like a private car or public transportation, or based on a trip's distance, such as long-distance and short-distance trips.

On the other hand, the factors affecting the travel demand can be categorized under, but not limited to, five main groups:

1. Sociodemographic [150, 64, 179, 57]
    - Gender
    - Age
    - Education level
    - Ethnicity
    - Birth rate
    - Household size, and number of children in household
    - Car ownership rates
    - Personal income
2. Travel Cost [95, 22, 268, 11]
    - Gas prices and tax rates
    - Private vehicle cost and registration fees, vehicle tax, and insurance cost
    - Road tolls and congestion charges
    - Parking cost
    - Public transportation fares and governmental subsidies.
3. Available Travel Options [19, 48, 280, 148]
    - Walking
    - Cycling
    - Public transportation
    - Different ride sharing services
    - Car sharing services
    - Ridehailing, and e-ridehailing
    - Private vehicles

4. Travel Options Characteristics [7, 24, 139, 147, 128, 91]
   - Relative speed and delay
   - Reliability of the service
   - Convenience and comfort
   - Safety and security
   - Waiting times and waiting conditions
   - Parking availability
5. Land Use, Built Environment [37, 281, 51, 156]
   - Density
   - Diversity
   - Design of pedestrian and bicycle friendly streets

The factors influencing travel demand of shared mobility modes is explored in the next subsections.

2.2.3. Shared Mobility User's Profile

The user's profile or sociodemographic of shared mobility users impacts the different services demand following the general notion of travel demand. The first step to analyze the user's profile is to recognize the various user groups of different services. The first observation from multiple studies is that some users' demographic characteristics are notably consistent across the different services and even across different countries. Shared mobility users are more likely to be highly educated, full-time-employed, young with high-income level compared to the average population, as revealed by some of the selected studies conducted in the United States, Great Britain, Germany, Canada, and Australia [226, 85, 177, 200, 186, 41, 71, 116, 133, 218]. Impacts of other demographic factors were observed in some instances. For example, males are frequent users in the cases of scooter sharing, bikesharing, carsharing, and ride-sourcing services compared to female users [71, 116, 226, 85, 177, 98, 186, 200, 133, 228]. In London, the majority of the ATS users are females [266]. Also, the role of ethnicity was evident in cases of bikesharing, TNC, and ATS, where people from white ethnicity were the most frequent users of the mentioned services compared to people from other ethnicity [226, 93, 227].

Exceptions of the previous user's profile have also been noticed. In China, 53% of Didi's users, a ride-hailing service, did not have a high education compared to the average population [244], Another case where the service user's profile is not matching with the general pattern of shared mobility users is the case of carpooling services. Immigrants living in an immigrant-centric neighborhood in Southern California, USA, tend to use more carpooling trips than the average population [36]. Shaheen, stocker, and Mundler found that in France, the majority of BlaBla car, an online carpooling service, passengers are low-income students, while the drivers are more likely to belong to higher income groups [228].

Table-2.1 Summarizes the characteristic of the users of the different shared mobility services across selected studies. The attributes in Table-2.1 are compared to the average population; for example, young means the users tend to be younger than the average population age.

Table 2.1 – Shared Mobility Users' Profile, Selected Studies

Characteristics	SS	BS	CS	RS	TNCs	ATS
High Education		[226, 85, 177, 200]	[175, 72]	[228, 72]	[101, 277, 248]	[227]
High Income level		[226, 85, 177, 186, 200]	[175, 72]	[118, 72]	[101, 248]	[227]
Males	[71, 116]	[226, 85, 177, 98, 186, 200]	[133]	[228]		
Young	[71, 116]	[226, 85, 177, 200, 218]	[133, 175, 72]	[228]	[101, 277, 202, 248, 93, 72]	[227, 266]
White Ethnicity		[226, 41, 218]			[93]	[227]
Employed	[116]	[226]	[133, 72]	[118, 72]	[93]	
High Car Ownership		[84, 219]	[133]			[227]
Low Car Ownership					[101, 202, 93]	

SS= Scooter Sharing, BS = Bikesharing, CS= Carsharing, RS= Ridesharing

2.2.4. Shared Mobility Trip Purpose

An essential dimension of travel demand estimation is the demand for different trip purposes. Shared mobility customers use the available services to fulfill their daily travel needs. The review of some selected studies for the various services shows that, in general, different shared mobility services are used for all the different trip purposes.

The most common use for shared scooters is to perform work and school trips in Austin, Texas, Portland, Oregon, and San Francisco, California [54]. According to Noland, scooters are primarily used to perform recreational activities in Louisville, Kentucky [185]. Ricci claims that work-related trips are the most frequent trips for bikesharing systems. This use pattern is evident in four cities in North America (Washington DC, Minneapolis-Saint Paul, Montreal, Toronto), and in London, Uk [205]. In New York, USA, using bikesharing for going to eating, shopping, and commuting to transit stations were found to be most performed trip types [23].

According to Mueller et al., in Germany, carsharing is used to commute to home and for leisure activities [175]. Daejin, Ko, and Yujin conducted a survey to study carsharing in Seoul, South Korea, revealing the service is used mainly for leisure, followed by business and personal activity [133].

Carpooling users in France use the service for different purposes based on their income, where low-income users use the service for work and school trips more than the other income groups; however, the higher income groups use the service for leisure trips [228]. In Houston, Texas, casual carpooling, or slugging, users mainly use the service for commute trips [43].

Tirachini and Gomez-Lobo studied ride-hailing in Chile, and they surveyed the users' trip purposes. The investigation showed that the top three trip purposes were leisure, visiting someone, and work or commute trips [249]. In Brazil, the prime three trip purposes for ride-sourcing were social and going out trips, work and commuting, and shopping trips [234].

Forty-four percent of RideKC users, the ATS service in Kansas City, MO, US, use the service for commute trips, and one-third of the users use the service for work-related trips [227]. In the greater Manchester, UK, area, the analysis of ATS trip purposes showed that the top three trip purposes are leisure, work, and shopping, where leisure and shopping trips represented one-half of the trips. Commutes to work represented 28% of the total trips [266].

Table 2.2 – Shard Mobility Trip Purpose Summary, Selected Studies

Trip Purpose	SS	BS	CS	RS	TNCs	ATS
Work/commuting Business and related		[205]	[133, 175]**	[228]	[249, 234]✓	[227, 266] •
School	[54]			[228]		
Recreational	[185]	[23]*				
Leisure			[175, 133]**	[228]	[249, 234]✓	[266]•
Shopping		[23]*			[234]	[266]•
Commute to Transit Stations		[23]*				

SS= Scooter Sharing, BS = Bikesharing, CS= Carsharing, RS= Ridesharing
* in [23] the order of purposes is going for eating, shopping and commute to transit stations
** in [133] the order of purposes is leisure, business, and personal activities
✓in [249] the order of purposes is (restaurant, bars, parties), to visit someone, work/commuting
✓in [234] the order of purposes is social and going out trips, work and commuting, and shopping trips
•in [266] the order of purposes is leisure, work, and shopping trips

2.2.5. Exogenous Factors Impacting Shared Mobility Use

Different external factors affect travel decisions, such as weather, service accessibility, modes availability, and land use. In the same direction, estimating the impact of the various exogenous factors on the use of shared mobility is essential for demand prediction and estimation. External factors affecting mobility use can be grouped into five main categories across the different services;

I) service-related factors such as presence and stations' location, and the available number of bikes in the station in case of bikesharing services [20, 158, 200], or the available number of vehicles and their conditions in the case of carsharing systems [69].

II) Infrastructure also affects the use of shared mobility. In the case of bikesharing, the presence of bike lanes is strongly affecting the use of the service [16, 146, 240, 45]. Also, The high intersection density is affecting the use of carsharing and bikesharing [16, 56]. Intuitively the availability of parking spaces and a higher road density influence the use of carsharing and ride-sourcing services [176, 117, 56, 99]. Moreover, the availability of high occupancy vehicle lanes (HOV) promotes the use of ridesharing services [42, 96].

III) Land use and built environment are significant factors impacting the trip generation process. In the case of shared mobility, the impact of the different land use is not present for the different services. Mixed land use is found to impact carsharing and TNC services [117, 9], and commercial land use impacts the use of bikesharing carsharing, and TNC [134, 146, 129, 117, 99]. The case of single land use associated with a specific service is also observed, parks only found to increase the use of bikesharing [134, 146, 240], POI are impacting the use of carsharing [272, 56], also, residential land use is only associated with carsharing use [134, 146]. Educational related land use, universities and schools, is the most observed land use in the reviewed studies that have impact on bikesharing and carsharing use [16, 158, 240, 134, 240, 117, 56, 238]. Intuitively, the increase in population density increases the use of all services [16, 240, 45, 21, 117, 239, 99].

IV) Shared mobility services are new travel options in the urban environment. The synergy between shared mobility and other available travel options impacts the use of shared mobility. For example, the reduced accessibility to PUT increases the use of bikesharing, carsharing, ridesharing, and TNCs [16, 146, 45, 231, 21, 117, 99, 17]. Also, the availability of bikesharing stations and carsharing stations around metro stations increases the use of the services [134, 129, 56].

V) Weather plays a significant role in influencing the use of shared mobility, especially bikesharing services [16, 134, 146, 158, 231]. Moreover, adverse weather conditions increase the use of carsharing, ridesharing, and TNC services [99, 94, 243].

Table-2.3 summarize the exogenous factors affecting the use of shared mobility services from selected studies.

Table 2.3 – Exogenous Factors Affecting the Use of Shared Mobility, Selected Studies

	BS	CS	RS	TNCs
Service Related Factors				
Presence of Docking Station at Origin and Destination	[20, 158, 200]			
Distance from Home to Downtown	[20]			
Membership of the Service	[20]			
Number of Vehicles Parked		[69]		
Vehicle Age		[69]		
Infrastructure Related Factors				
Bike Infrastructure	[16, 146, 240, 45]			
Intersection Density	[16]	[56]		[99]
Parking Availability		[176, 117, 56, 275]	[259, 2, 239]	
Availability of HOV Lanes			[42, 96]	
High Road Density		[117]		
Land Use				
Mixed Land use		[117]		[9]
Land use, Commercial or Business	[134, 146]	[129, 117]		[99]
Land use, Parks	[134, 146, 240]			
Land use, Residential		[134, 146]		
Land use, Schools/Universities	[16, 158, 240, 134, 240]	[117, 56, 238]		
POI		[272, 56]		
population Density	[16, 240, 45]	[21, 117]	[239]	[99]
PUT Related Factors				
Put Accessibility	[16, 146, 45, 231]	[21, 117, 131]	[259]	[99, 17]
Presence of Put station, subway	[134]	[129, 56]		
Presence of Put station, Bus		[56]		
Weather				
Adverse Weather Conditions		[275]	[243]	[94]
Weather, Temperature	[16, 158, 231]			
Weather, snow	[16, 146, 158]			
Weather, Precipitation	[16, 134, 146, 158, 231]			
Weather, Wind Speed	[146, 158]			

SS= Scooter Sharing, BS = Bikesharing, CS= Carsharing, RS= Ridesharing

2.2.6. Reasons to Use Shared Mobility

The next important aspect of shared mobility demand characterization is identifying the reasons why users prefer each of the different services. Users' motivations to use shared mobility are in line with the general incentives to engage in sharing economy. These motivations can be categorized into the same main categories: economic, social, and environmental incentives. Comparing user motivations to use shared mobility reveals that across the different services, economic incentive is the most common factor [86, 151, 265, 15, 215, 13, 201, 58, 112, 250] followed by the time saving [124, 210, 143, 215, 201, 112, 58, 250]. Also, convenience in terms of safety and comfort is evident in multiple studies for multiple services [84, 124, 210, 143, 215, 201, 58]. Health benefits is only evident in the case of bike-sharing [47, 184, 183]. Table-2.4 shows the summary of the reasons to use shared mobility for the different services from several selected studies.

Table 2.4 – Reasons to Use Shared Mobility, Selected Studies

Reasons	BS	CS	RS	TNCs
Ease of Use	[86, 152]			[201, 58, 265, 14]
Ease of Payment				[201, 58, 250]
Economic Incentive	[86, 151, 265, 184, 183]	[114, 271]	[15, 215, 13, 207]	[201, 58, 112, 250]
Time saving	[184]	[124, 210, 143]	[215]	[201, 112, 58, 250]
Convenience	[84]	[124, 210, 143, 114]	[215, 267]	[201, 58, 14]
Environmental Awareness	[47, 183]	[210, 271]	[15, 207, 267]	
Lifestyle	[47]	[143, 114, 271]		
Health Benefits	[47, 184, 183]			

SS= Scooter Sharing, BS = Bikesharing, CS= Carsharing, RS= Ridesharing

2.3. Shared Mobility Study Methodology Framework

Different methods are applied to study shared mobility services demand, the services adaptation process, and the factors impacting this process. These methods rely heavily on data collection to be used in the different analysis, processing, and modeling stages. Firstly in this section, the different data collection methods are discussed; then, processing techniques that were implemented to prepare the data for the modeling process were explained. Finally, the details of the modeling process are described.

2.3.1. Data Collection

The data collection process comes after defining the study objectives and research questions and formulating the study hypotheses stags. In the context of shared mobility studies, data collection methods depend mainly on the study's objective and not directly related to the service itself. Five primary sources of data are commonly used in shared mobility studies: surveys, open-source data, mobile phone, GPS, and the combination of any of the previous resource [55].

When specific individual-level information is in question, such as user's demographic, travel habits, and motivation to use different services, surveys, both online or face to face interviews, and travel diaries are used. For example, online surveys were deployed to study bikesharing user's demographics [226, 85, 177, 200], and also carsharing programs member's demographics [175]. Motivation to use the different services were also identified using online surveys in case of ridesharing and carsharing [133, 212, 15]. Face to face interviews were also used to study shared mobility, where Tirachini and del Río used street interviews to investigate ride-hailing users' travel behavior [250], and Shaheen et al.

investigated the motivations to use casual carpooling in the San Francisco Bay Area [215].

Although, surveys are useful tools to investigate user-level information, especially this information are not always publicly available, due to the rise in data privacy concern. Surveys have disadvantages that affect their use. Personal interviews are costly to perform, and their responses are not always easy to valid [108, 18]. Online surveys do not grant the representation of the general population resulting in non-coverage bias, where marginalized groups such as households with no internet access and elderly are not accessible with such surveys. Also, some users avoid the use of online surveys as the fear that their private data is leaked [103, 8].

The advances in information and communication technologies have impacted the data and data collection process in positive ways, such as adding new sources of information, social media, that was not available to use in the classical transportation studies. Also, the increased availability of GPS units, and mobile-phones, increased the volume of the generated data, creating what is commonly named as Big Data [55].

Big data are increasingly being used in shared mobility studies, such as investigating scooter sharing use patterns using trip-level data provided by the operators in the USA [185, 162]. Also, open-source data use is evident in the study of built environment factors affecting the use of bikesharing systems using built environment locations and POI data from Google Maps[3], and OpenStreetMap[4] [75, 16]. Big data use also raises concerns regarding the data quality and the data completeness, which are not always granted nor easy to verify [55].

2.3.2. Data Preparation

The next step after the data collection process is data preparation for the modeling process. Three main pre-modeling techniques were applied to the collected data: collinearity detection, zero and near-zero variance variable detection, and explanatory factor analysis (EFA)[5].

Zero and Near-zero Variance Variables

The problem of zero variance or near-zero variance emanates when there is no difference within a variable distribution, and the variance of the variable is zero or almost zero. These variables cause problems in the models' estimation process and, in some cases, make the models unidentifiable [140, 111]. There are some criteria applied to identify these variables. Max defined the criteria to define near-zero variance as I) The ratio of the unique values within the variable is less than 20%, and II) The ratio between the top two most frequent values is greater than 20 [140]. One solution to the previous issue is to remove the problematic variables [111].

Collinearity

The statistical phenomenon of collinearity arises when there is a high correlation between at least two of the used predictor variables in regression models. Collinearity is mostly expected when a high number of covariates are used in the model, thus preventing the estimation of a stable model, leading to unstable and biased standard errors, affecting

[3] Googlemaps.com

[4] planet.openstreetmap.org

[5] Under the scope of this thesis the EFA results were used as indicators for the latent variables models; therefore, EFA was considered as apart from the data preparation process.

the estimation of an accurate variance, and resulting in a wrong interpretation for the estimated model coefficients [260, 169, 165].

There are several statistical methods used to diagnose collinearity. The most common methods are the pairwise correlation coefficients between the different variables and the variance inflation factor (VIF); however, there are no cutoffs values for both criteria to determine collinearity [180, 260, 169, 165, 30, 26].

Under the scope of this thesis, the pairwise correlation was calculated using Pearson's correlation coefficient for continuous numerical variables [136]. For ordinal Likert scale data [145], a Polychoric correlation coefficient was used to calculate similarities between the different ordinal variables, and in explanatory factor analysis as explained in the next subsections.

Polychoric correlation is more suitable for ordinal variables than Pearson's correlation, as the latter assumes the interval measure nature of the data, which is not the case for Likert data [115].

After diagnosing collinearity, several techniques are applied, such as using principal component analysis (PCA), combining correlated factors into one factor if the combination of factors makes sense. The other frequently used method is removing one of the correlated variables from the model [260, 169, 140]. The latter approach is the one that will be used in the scope of this thesis modeling process.

Factor Analysis

Factor analysis was initially used at the beginning of the twentieth century by Spearman to measure the underlying relation between the psychical tendencies, and psychical activities [235]. Factor analysis is commonly used in transport-related studies to reduce the number of variables by estimating a lesser number of common factors that capture most of the covariance in the original data and to capture the underlying latent construction between the different variables [269]. There are two main types of factor analysis: explanatory factor analysis (EFA) and confirmatory factor analysis (CFA). CFA tests the hypothesis regarding the latent construct of the data. Before implementing CFA, a hypothesized relationship between the factors and the different variables should be established. The subsequent explanation concentrates only on EFA as it is used in the modeling process under the scope of this thesis. EFA tries to capture common factors affecting the variables and the influence of the variables on each factor [70, 269]. In other words, EFA estimates the factors discern the latent construction underlying within the original data [70].

Karlaftis, Matthew, and McCarthy used factor analysis to investigate the public transportation performance in Indiana, USA [130]. Syed, Sharfuddin, and Khan used factor analysis to determine the factors affecting public transportation ridership in Canada [242]. Le-Klähn, Hall, and Gerike used factor analysis to explore visitors satisfaction with public transportation in Munich, Germany [137].

EFA depends mainly on the correlation between the different variables; therefore, it is suitable for ordinal and ratio data. Factor analysis can be obtained from solving the set with the linear equations for each variable (x), as shown in the below equation system equation-2.1, were (F_i) are the factors, (ℓ_{im}) is the factor loadings for factor (m) and variable (i), (μ_i) is the population mean and (ϵ_n) is the associated random error. In general, (m) number of factors are $\lll$ original number of variables. Factor loading near to one indicates that the (X_i) is highly influenced by (F_j). On the contrary, if the factor loading is near to zero, this indicates the (X_i) is less influenced by (F_j)

$$\begin{aligned}
X_1 &= \mu_1 + \ell_{11}F_1 + \ell_{12}F_2 + \dots\dots + \ell_{1m}F_m + \epsilon_1 \\
X_2 &= \mu_2 + \ell_{21}F_1 + \ell_{22}F_2 + \dots\dots + \ell_{2m}F_m + \epsilon_2 \\
&\vdots \\
X_n &= \mu_n + \ell_{n1}F_1 + \ell_{n2}F_2 + \dots\dots + \ell_{nm}F_m + \epsilon_n
\end{aligned} \tag{2.1}$$

EFA Model Assumptions

- All the random error terms have a mean value of zero; E(ϵ_i) = 0 for i =1,2,...,n
- The mean of the factors is zero E(f_i) = 0 for i =1,2,...,m
- The variance of all the common factors is one; $\sigma^2(f_i)$ = 1; i =1,2,...,m
- specific variance is the variance of the error term; $\sigma^2(\epsilon_i) = \Psi$, where Ψ is a diagonal matrix
- There is no correlation between any of the factors; $Cov(f_i, f_j) = 0 \quad \forall \quad i \neq j$

Four main methods estimate factor scores. I) Ordinary least square, II) Weighted least square, III) Regression method and IV) Maximum likelihood, which assumes that the data are almost normally distributed [68].

With (n) equations and (m) unknowns, additional information or restrictions are needed to solve the previous set of equations-2.1. The type of restriction imposed defines the type of factor analysis model. There are two main types of factor models: orthogonal and oblique models. Orthogonal factors satisfy the model assumptions mentioned before; however, the oblique models relax the assumptions where the obtained factor loadings can be correlated. The common orthogonal rotation methods are Varimax, Quartimax, and Equamax. The common oblique rotation methods are direct Oblimin, Quartimin, and Promax [68]. The factor rotation process's main purpose is to improve the interpretability and the fit of the estimated model. It is to be noted that factor rotation does not change the amount of the explained variability of the models; however, it improves the model fit.

The next step in EFA is to decide on the number of factors to retain. One of the first used methods is to maintain the factors that have an eigenvalue of more than one. This method resulted in over retention of factors and inaccurate results [261]. According to Costello, and Osborne, the scree test is the best practice to decide on the number of extracted factors, where the eigenvalues are printed against the number of factors and the natural bend in the data should decide on an approximate number of variables [68].

2.3.3. Modeling Techniques

Summary of Selected Studies

The three main research questions, section-1.3, investigate the factors affecting the use of shared mobility, the shift from different models, and the choice between the various shared mobility services. The answers to these questions have a universal discrete nature, that is modeled using econometric tools such as discrete outcome models (choice models). The type of the used models is mainly decided by the investigated factor, or commonly named dependent variable.

A wide range of shared mobility studies investigate the factors impacting the adoption of shared mobility, or the factors that lead to shifting from the different modes use to shared mobility use. In this case, the modeling process is estimating the factor impacting the choice between two or more options that are mutually exclusive[6]. In such cases, binary and multinomial probit and logit models are used. For example, the adaptation of Uber and Lyft in California was investigated using a binary logit model [8], and a multinomial logit model was used to explore the factors impacted the shift to ride-hailing from the different modes in Boston, USA [94].

In other studies, the explored factors, dependent variables, have an ordered nature such as ordered scale responses or ordered frequency of use, which entitle the use of models that account for the ordered nature of the investigated factors. In this situation, ordered logit and probit models (OLM) are widely used. Some examples of the ordered model applications are: I) the investigation of the factors causing differences in trip duration between ride-hailing trips and public transportation trips, where the dependent variable, time difference, was a three-level ordered categorical variable represent the time difference between the two modes [276]. II) Generalized ordinal logit models were used to check the factors impacting the ride-hailing frequency of use. The dependent variable was the ordered levels of use frequency [250]. III) The factors impacting the attitudes of electrical-carsharing program members were investigated using ordered probit models; the dependent variables were an ordered five-point scale representing the different attitudes [133].

Other modeling techniques such as generalized additive mixed models, multiple regression, structural equation and partial least squares structural equation models (PLS-SEM) were used to investigate carsharing and ridesharing use and motivation to use [117, 124, 143, 15, 14].

A modern technique that was induced and supported by the proliferation of social media use is sentiment analysis or opinion mining. Sentiment analysis is the investigation of the attitude, opinion expressed by a body of text; these attitudes could be categorized as negative or positive attitudes. Recently, the usage of sentiment analysis in transport-related studies has increased, especially by using data extracted from social media [178]. For example, Twitter feeds were used to investigate transit riders' satisfaction, to evaluate real-time transit network performance, and to assess transit performance [62, 149, 105]. Also, survey responses were used to investigate the behavior of the commuters who combine cycling and public transportation [163].

In the next subsections, the mathematical formulation and estimation details of the discrete outcome models are explained

Discrete Outcome models

Definition

The discrete outcome, choice, models are the group of models that can be used to model the situation where an agent (individual, company, or any decision-makers) needs to choose between a set of alternatives, products or a series of options over the time, or an ordered response on a scale. These choices have a general discrete nature, meaning

[6] Next section includes a detailed discussion for discrete choice models

the choice of one option is limited to this option because of its discrete countable nature [251, 27].

Choice Set Properties

The discrete nature of the alternatives, choices demonstrates three main properties:

- The options are mutually exclusive. Choosing one option means no other options are chosen
- Options must be exhaustive, meaning that the choice set should include all the possible options
- The number of options must be limited and countable

The first two attributes are not confining as the experiment design process can grant that all the included alternatives are mutually exclusive and exhaustive. The third attribute, the finiteness of the choice set, is a restrictive condition. The dependant variable needs to be discrete and not a continuous variable, which is the case in regression models [251, 27].

Models Formulation

The underlying assumption of choice models is that the agent, or decision-maker, is a natural optimizer trying to choose the option that provides him with the maximum utility [251]. Following Marschak's notion that the respondent differentiates between the different levels of the utility, discrete choice models can be driven as a derivative of utility maximization process, usually called random utility models (RUMs) [157]. RUMs can be described as a decision-maker (n) chooses between a set of options (J = *1.. j*). There is a level of gain, or utility, associated with each of the alternatives in (J). Based on these assumptions, the choice model for choosing an alternative (i) can be formulated as

$$U_{ni} > U_{nj}\ ; \forall j \neq i \tag{2.2}$$

The choice utility (U) consists of two parts; the first part is the observed part and can be related to both the choice alternative's attributes (x_{nj}), which could change between the different alternatives, and the decision maker's attributes (s_n), which are fixed for all the attributes. This part of the utility is generally called the "representative utility" [251], and it is formulated as:

$$V_{ni} = V(x_{nj}, s_n);\ \forall j \tag{2.3}$$

The second part of the utility is the unobserved part, and the total utility function consisting from both parts can be formulated as:

$$U_{ni} = V_{ni} + \epsilon_{nj} \tag{2.4}$$

In equation-2.4, (ϵ_{nj}) is the term represents all the unobserved utilities not captured by V_{ni}, in addition to the error. The assumption of (ϵ_{nj}) defines the model type as explained in the subsequent sections. Based on the previous utility formulation equation-2.4, the probability of choosing alternative (i) from a choice set contains (*i, j*) could be expressed as:

$$\begin{aligned} P_{ni} &= Prob(U_{ni} > U_{nj}; \quad \forall i \neq j) \\ &= Prob(V_{ni} + \epsilon_{ni} > V_{nj} + \epsilon_{nj}; \quad \forall i \neq j) \\ &= Prob(\epsilon_{nj} - \epsilon_{ni} < V_{ni} - V_{nj}; \quad \forall i \neq j) \end{aligned} \tag{2.5}$$

According to Train [251], the probability that the difference between the unobserved terms $(\epsilon_{nj} - \epsilon_{ni})$ is less than the observed terms $(V_{ni} - V_{nj})$ is a cumulative distribution and it can be formulated as:

$$\begin{aligned} P_{ni} &= \int_{\epsilon} I(\epsilon_{nj} - \epsilon_{ni} < V_{ni} - V_{nj}; \quad \forall i \neq j) f(\epsilon_n) d\epsilon_n; \\ \text{where } I &= \begin{cases} 1 & \text{if (expression is correct) is true} \\ 0 & Otherwise \end{cases} \end{aligned} \tag{2.6}$$

Multinomial Models

Logit Models

As discussed in the previous sections, the assumption of the distribution of the error term (ϵ_{ni}) defines the choice model type. The multinomial logit model is the most common, and the most popular type of the RUMs family. It is built around two main assumptions. The first assumption is that the error term ϵ_{ni} is independent and identically distributed (IID) with Gumbel distribution and zero expectation value. The second assumption is the independence of the irrelevant alternative, or the choice between any alternatives is independent of the properties of other alternatives not present in the choice set [251, 27]. The density distribution of the error term can be formulated as:

$$f(\epsilon_{nj}) = f(\epsilon_{nj}) = e^{-\epsilon_{nj} - e^{-\epsilon_{nj}}} \tag{2.7}$$

and the cumulative distribution is

$$F(\epsilon_{nj}) = f(\epsilon_{nj}) = e^{-e^{-\epsilon_{nj}}} \tag{2.8}$$

considering the probability distribution of the error term, the probability to choose an options from a choice set (J) as a closed form expression is formulated as:

$$P_{ni} = \frac{e^{V_{ni}}}{\sum_j e^{V_{nj}}} \tag{2.9}$$

Probit Models

Probit models overcome the main limitations of the logit model, which are the inability to produce random taste change, IIA assumptions, and it cannot be used with panel data with temporally correlated unobserved factors. Probit models assume that the unobserved utilities are normally distributed, which could be a limitation is specific cases [251]. Equation-2.10 shows the formulation of a binary probit model, where (Φ) is the standard normal cumulative distribution function [29].

$$P_{ni} = \Phi\left(V_i - V_j\right) \tag{2.10}$$

It is to be noted that if the choice options are more than two, the model is named multinomial logit, or multinomial probit model, and if there are only two options in the choice set, the model is named a binary logit, or binary probit model.

Ordered Responses Models

Ordered response models are extensions for the multinomial probit and logit models. The outcome of these extended models is ordered, which is not accounted for initially in the case of multinomial logit and probit models [269]. Moreover, the assumptions of independent errors, as in the case of the logit model, do not stand in the case of the ordered outcome. In ordered models, one alternative is closely similar to the values near to it and irrelevant to the alternative far from it [251].
Figure-2.2 shows the distribution of the utility (U) around ($\grave{\beta}x$) following the distribution of the error term (ϵ). The probability of choosing the first level of the outcome ($y = 1$) is the probability of ($-\beta X$) when ($\mu_0 = 0$), (μ) is referred to as the thresholds, or the values differentiate between the different choice levels. The probability of each of the choice levels can be formulated as:

$$\begin{aligned} y &= 1 \qquad if \quad U \leq \mu_0 \\ y &= 2 \qquad if \quad \mu_0 < U \leq \mu_1 \\ &\vdots \\ y &= I \qquad if \quad U \geq \mu_{I-1} \end{aligned} \tag{2.11}$$

Based on the assumptions of the error term (ϵ), the exact probability of each choice level is formulated as

$$P(y = I) = \frac{e^{I-\grave{\beta}x}}{1 + e^{I-\grave{\beta}x}} - \frac{e^{(I-1)-\grave{\beta}x}}{1 + e^{(I-1)-\grave{\beta}x}} \tag{2.12}$$

It is to be noted that if the responses are nominally ordered, it should be transformed to numerical values in the same ordinal order.

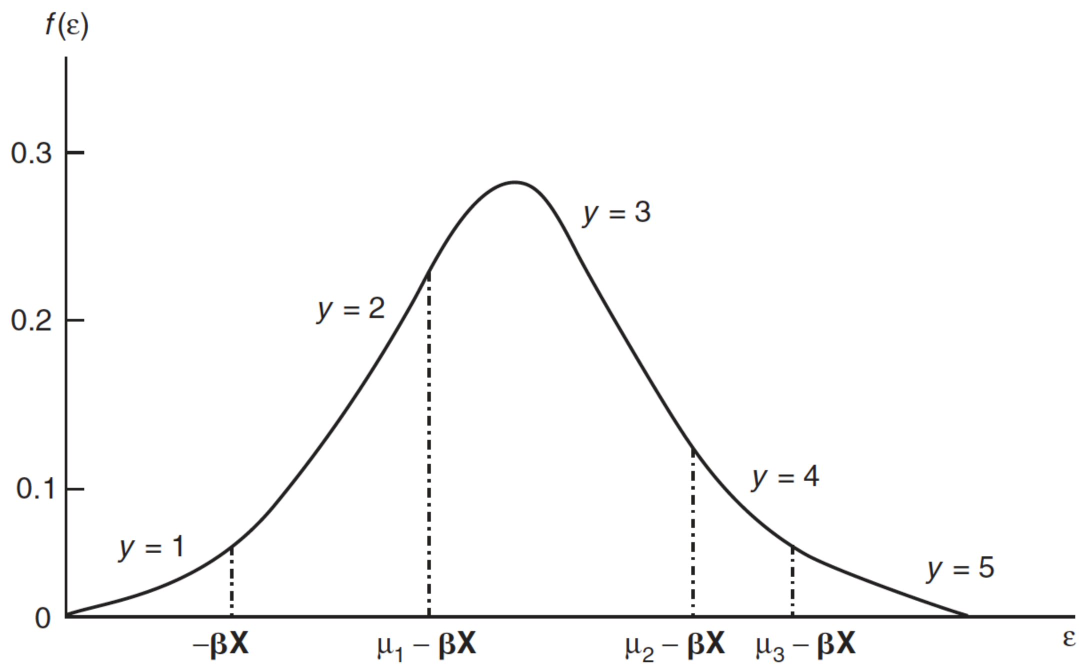

Figure 2.2 – Ordered Probability Model, Source [269].

Integrated Choice and Latent Variables Models

Integrated Choice and Latent Variables Model (ICLV) or Hybrid Choice Model (HCM) are different names for the recent extension to the rational discrete choice models. HCM was initially proposed by MacFadden in 1986 and by Train et al. in 1987 [252, 160]. ICLV models integrate the latent variable model to the choice model. The primary purpose of this integration is to improve the model capabilities of interpreting the choice process by integrating the user's cognitive behavior, attitude, and psychological factors into the choice model, and improving the model goodness of fit when applicable [245, 28, 135]. Figure-2.3 shows the overall modeling framework for the HCM. The right part represents the latent variable part, which cannot be directly observed. However, it can be determined through attitudinal indicators measuring the personal attitudes for the different users [28, 135].

Personal attitudinal indicators are typically measured through survey responses that answer questions to express or evaluate different behavioral approaches such as perception, comfort, environmental concerns, motivation, and convenience on a specified scale [28, 135]. Table-2.5 shows examples of the latent variable used in some selected studies.

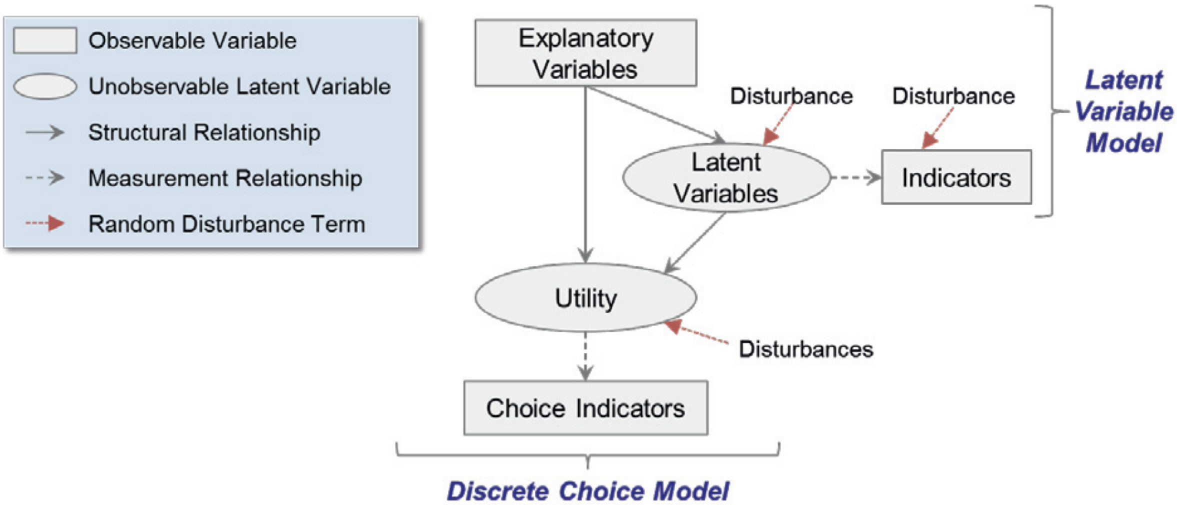

Figure 2.3 – HCM model, source [28, 135]

HCM Formulation

As stated earlier, the main two components of HCM is the latent variable model and the discrete choice model. The latent variable model consists of two parts; the Structural equation part, and the measurement equations part. The structural equation part is the part that represents the relation between the explanatory variable and the latent variable. Following the notion of Ben-Akiva et al. [28], the structure equation part can be formulated as:

$$X^* = h(X;\gamma) + \eta \qquad and \qquad \eta \sim D(0, \textstyle\sum_\eta) \tag{2.13}$$

(X^*) is the vector of the latent unobserved variable that includes (S^*) the cognitive attributes of the decision-maker, and (Z_i^*) the latent attributes of the option (i) as recognized by the agent, decision maker. (X) is the vector of observed variables, that contains (S) observed attributes of the decision-maker, and (Z_i) the observed attributes of alternative(i). (η) is a random disturbance parameter, and its covariance is $(\sum_\eta)$. (γ) is an unknown term.

The second part of the latent variable model is the measurement equations part. This part represents the relation between the indicators and the latent variables, and it can be formulated for each of the indicators as:

$$I = g(X, X^*; \alpha) + \upsilon \qquad and \qquad \upsilon \sim D(0, \textstyle\sum_\upsilon) \tag{2.14}$$

(υ) is a random disturbance parameter where its covariance is ($\sum_\upsilon$), and (α) is an unknown term. Indicators are not exclusively measured from attitudinal survey questions, but they could be measured from latent attributes of the decision-maker, latent sociodemographic measures, or latent attributes of the choice options [28].

Under the scope of this research, the indicators of the measurement model are estimated from 5-levels Likert scale questions, with (I_r) refers to the attitudinal question number (r), and (α_l) is the associated latent attitude. An ordered logit model was used to represents the likelihood of the observed indicator $I_{n,s}$ for respondent n, and it is formulated as:

$$LI_{n,r} = \sum_{p=1}^{5} x_{In,r,p} \left(\frac{e^{\tau_{I_{r,p}} - \zeta_{l,s}\alpha_{n,l}}}{1+e^{\tau_{I_{r,p}} - \zeta_{l,r}\alpha_{n,l}}} - \frac{e^{\tau_{I_{r,p-1}} - \zeta_{l,r}\alpha_{n,l}}}{1+e^{\tau_{I_{r,p-1}} - \zeta_{l,r}\alpha_{n,l}}} \right) \tag{2.15}$$

Where ($x_{In,r,p} = 1$) if person (n) choose level (p) for question number ($r = 1, \ldots, r$), and ($\tau_{I_{r,p}}$) are thresholds between the different Likert scale levels of the ordered logit model normalized as $\tau_{I_{r,0}} = -\infty$ and $\tau_{I_{r,5}} = +\infty$, and the estimated ($\zeta_{l,r}$) measures the impact of latent attitude (α_l) on question (I_r).

From the previous formulated equations, the final integrated model will consist of equation-2.4 representing the choice model and equations (2.13) to (2.15) represent latent variable model, with choice probability of $P(y|X, X^*; \beta, \sum_\epsilon)$ depending on both observed and latent variables [29, 28].

Subsequently, the likelihood function for the HCM joint choice probability, assuming the independence of the error terms (α, β, γ) is formulated as:

$$\begin{aligned} &L(y, I|X; \alpha, \beta, \gamma, \textstyle\sum_\epsilon, \sum_\upsilon, \sum_\eta) = \\ &\qquad \int_{X^*} P(y|X, X^*, \beta, \textstyle\sum_\epsilon) f_1(I|X, X^*; \alpha, \sum_\upsilon) f_2(X^*|X; \gamma, \sum_\eta) dX^* \end{aligned} \tag{2.16}$$

The first part of equation-2.16 resembles the choice model, the second part resembles the measurement equation from the latent model, and the last part resembles the structural equation from the latent model [28, 29].

Models Estimation

Maximum Likelihood

The objective of fitting regression models is to find the unknown parameters to minimize the sum of squared errors for the fitted values [109]. Maximum likelihood (ML) is a standard method to estimate the model parameters for regression models. Assuming a sample is randomly drawn, the probability of a decision-maker (n) to choose an alternative (i) from a sample can be formulated as:

$$P(i) = \prod_i (P_{ni})^{y_{ni}} \tag{2.17}$$

In equation-2.17 if decision maker (n) choose alternative (i) $y_{ni} = 1$ and $y_{ni} = 0$ otherwise. For the total sample (N), and based on the assumption that all the choices are independent from each other, the probability of choosing alternative is formulated as:

$$L(\beta) = \prod_{n=1}^{N} \prod_{i} (P_{ni})^{y_{ni}} \tag{2.18}$$

In equation-2.18 (β) is the vector of the model's parameters that is needed to be estimated. The log likelihood for equation-2.18 is formulated as:

$$LL(\beta) = \sum_{n=1}^{N} \sum_{i} y_{ni} \; ln \; P_{ni} \tag{2.19}$$

The estimators for equation-2.19 are the values of (β) that maximize equation-2.20.

$$(\beta) = argmax_{\beta} \; \sum_{n=1}^{N} \sum_{i} y_{ni} \; ln \; P_{ni} \tag{2.20}$$

According to McFadden [159], equation-2.20 is a global concave function for linear parameters utility. The ML estimation can be formulated as shown in equation-2.21, with respect to each of the parameters as zero:

$$\frac{dLL(\beta)}{d\beta} = 0 \tag{2.21}$$

Reference to equation-2.21, the ML estimator is the value of (β) that satisfies the first derivative condition. By considering the linear parameters utility in equation-2.4 , and combining equation-2.19 and the logit probability formula, the first order condition can be formulated as in equation-2.22, and subsequently the ML can be estimated.

$$\sum_{n} \sum_{i} (y_{ni} - P_{ni}) x_{ni} = 0 \tag{2.22}$$

HCM Estimation

Similar to previous equations set for ML estimation for the multinomial models, and based on likelihood equation-2.16 of HCM, the ML function for the HCM is formulated as:

$$argmax_{\alpha,\beta,\gamma,\sum} \sum_{n=1}^{N} lnf(y_n, I_n | X_n; \alpha, \beta, \gamma, \textstyle\sum) \tag{2.23}$$

Equation-2.23 is a complex equation with multi-dimension integrals. These dimensions correspond to the integral of the choice model plus the integral of the latent variable. Three main methods are used to estimate such complex integral; sequential estimation, simultaneous numerical estimation, and simulation estimation [29, 28, 251].

Models Statistical Evaluation

Different statistical tests are used to determine how good the model fits the data or to evaluate the significance of the estimated model's parameters. In the subsections, the tests that are used in the modeling section are demonstrated.

Hypothesis Testing

Following the same procedures for OLS regression model, a standard one-tailed t-statistics is used to evaluate if the estimated parameter (β) differ from zero or no. In equation-2.24 the S.E. is the standard error for the estimated parameter [269, 251].

$$t_{statistic} = \frac{\beta - 0}{S.E.(\beta)} \tag{2.24}$$

It is to be noted that the use of t-statistic is an approximate use, as t-statistic assumes that the estimated parameters are normally distributed; however, the MNL assumes that the estimated parameters are derived from extreme value distribution [269].

Likelihood Ratio Index

An overall measure for the estimated model goodness of fit is the likelihood ratio index; refer to equation-2.25. $LL(0)$ is the log-likelihood for the empty model, setting all the parameters to zero, and the $LL(\widehat{\beta})$ is the log-likelihood for the estimated model. The value of (ρ^2) is between zero and one, where zero is the lowest value that indicates the model does not fit the data at all, and one indicates that the model is correctly predicting the choices [251].

$$\rho^2 = 1 - \frac{LL(\widehat{\beta})}{LL(0)} \tag{2.25}$$

One of the problems of using (ρ) as an indication for the model's fitting is that the value of (ρ) increases with the increase of the number of parameters estimated for the model. To deal with this situation, another version of (ρ) named as ($corrected \rho$), is calculated accounting for the number of estimated parameters (K) [269].

$$corrected\ \rho^2 = 1 - \frac{LL(\widehat{\beta}) - K}{LL(0)} \tag{2.26}$$

Likelihood Ratio Test

Another more general form of hypothesis testing is the likelihood ratio test (LRT). LRT is a test used to determine if adding a specific parameter to a model improves the model's overall fit. In other words, the null hypothesis (H_0) of the LRT is that the restricted model is the improved model. The log-likelihood of the model with the extra parameter ($LL(\beta_U)$) unrestricted model and the log-likelihood of the restricted model ($LL(\beta_R)$) are calculated. Subsequently, the (χ^2) statistics distribution with a degree of freedom equal to the difference in the number of parameters between the two models is calculated, as shown in the following equation-2.27. If the value of (χ^2) exceeds the critical value (α^2) with a certain degree of freedom, the null hypothesis is rejected. It is to be noted that to calculate the LRT, and the restricted model should be nested in the unrestricted model [251, 269].

$$H_0 :\ -2[LL(\beta_R) - LL(\beta_U)] \sim \chi^2_{df,\alpha^2} \tag{2.27}$$

AIC and BIC

Akaike information criterion (AIC) [5] and Bayesian information criteria on (BIC) [213] are two methods to evaluate the model's selection process [251]. In AIC (LL) is the value of the log-likelihood and (K) is the number of parameters. In BIC, (N) is the sample size. The lower the values of AIC, and BIC the better the model fit, both of AIC and BIC measures the trade-off between complexity and goodness of fit, noting that BIC penalizes the models heavier than AIC by $\log(N)$.

$$AIC = -2LL + 2K \tag{2.28}$$

$$BIC = -2LL + 2log(N)K \tag{2.29}$$

HCM Goodness of Fit

The model likelihood is mostly the base for calculating the different models' evaluation techniques. According to Vij, Akshay, and Walker [264], there are two methods to test the goodness of fit for HC models. These methods depend mainly on the study objective and the model's primary functions. The first method is used when the model's purpose is prediction, and if the indicators are not available, the model likelihood should only consider the choice model part. The second method is suitable when HCM is used to gain insight from the latent variable part, and if the indicators are expected to be available, the likelihood of the HCM should consider the latent model part.

No specific method is observed across some selected studies to evaluate the fit of HC models, and even some studies did not consider evaluating the model fit at all. However, the most common method across selected studies is using (ρ^2) for the choice model part to evaluate the HCM fit. Table-2.5 summarizes the used measures in some selected studies.

According to Yutaka and Ricardo [173], the question of the best estimator for the goodness of fit of an HCM is not fully answered and needs further studies.

Table 2.5 – HCM Measures of Fit, Selected Studies, Source: [173]

Author	Study Context	Used Latent Variable	Used Model Fit Assessment
Morikawa et al. [172]	Mode choice	Comfort and convenience	ρ^2 for choice model only
Fukuda and Moriuchi [90]	Parking choice	Risk attitude and public morality	χ^2, RMSEA, GFI
Karmargianni et al. [127]	Mode choice	Propensity for active transport and parental willingness to escort their children	ρ^2 for choice model only
Tsirimpa et al.[257]	Travel behavior	Risk attitude	ρ^2 for choice model only
Yafriez et al. [274]	Mode choice	Reliability, comfort-safety, and accessibility	ρ^2 for choice model only
Fleischer et al. [87]	Flight choice	Fear of flying	None
Prato et al. [195]	Mode choice	Comfort and convenience	ρ^2 for choice model only
Olaru et al. [187]	Residential location choice	Surroundings, social contracts, transport access, facilities within cycling distance, and facilities within walking distance	χ^2 , RMSEA
Paulssen et al. [193]	Mode choice	Power, hedonism, security, comfort and convenience, ownership, and flexibility	None
Galdames et al. [92]	Mode choice	Attitude	None
Bhat et al. [31]	Route choice	Cyclist attitude and safety conscious personality	Likelihood ratio tests for the integrated HCM model (versus alternative specifications)

RMSEA is root mean square error of approximation; GFI is goodness-of-fit index.

Resampling Methods

Resampling techniques are tools that facilitate measuring the accuracy of the estimated model by repeatedly drawing subsamples of the original data set and refitting a model on each of these subsamples. These techniques expand the amount of information gained from the model compared to the one-time classical estimation process by providing an unbiased estimate for the model performance. Two of the most common resampling techniques are the bootstrap technique and the cross-validation technique [122, 109].

Cross-validation

Cross-validation is dependant on splitting a data set into two parts. The first is the training part, which is used to fit the model. The second is the validating, hold-out-set, or testing part, in which we use the estimated model to examine the model fit [122]. There are three main methods for cross-validation;

The Validation Set Approach: The data set is split into two subsamples, one sample for model training purposes and one for model estimation. Model accuracy is calculated from validating the model using the testing subsample.

Leave-out-one Cross-validation, (LOOCV): One observation is omitted from the data set and the model is fitted on the remaining (*n-1*) observations. The process is repeated (n) times, and the average accuracy is calculated. The high number of the fitting repetitions, increasing with sample size, increases the computational efforts for this technique substantially.

k-Fold Cross-validation: The data set is split into k equal or almost equal k folds, where (*k-1*) folds are used for modeling, and the remaining fold is used for validation purpose. The process is repeated at least (k) times and the model's average accuracy is calculated [122, 109].

Ten fold cross-validation is proven to outperform the computationally expensive LOOCV, even if it is possible to increase the number of folds more than ten [138]. Moreover, according to Gareth et al., k-fold cross-validation is more accurate than LOOCV [122].

Bootstrapping

Efron [76] introduced bootstrap as a new method to enhance the models' uncertainty estimation measures. Bootstrap is used to assess the uncertainty related to the estimation of a model's parameters, such as the estimation of standard error, confidence intervals, or probability distribution for the regression model coefficients [3].

To produce a bootstrap data set, the data set of size (n) is sampled several times with replacement, and each time the new data set size is (n). The probability of any observation not to be selected after (n) samples is $(1 - \frac{1}{n})^n$ [138]. The model's coefficient is estimated for each of the new bootstrap data sets, and measures of accuracy such as standard error and confidence interval can be calculated accordingly [122, 109].

Chapter 3

Methodology

This chapter demonstrates the methodology used to answer the research questions, consisting of four main parts. The first part explores the data collection process, including the survey design and the data collection stage. The second part of the chapter demonstrates the data analysis and processing stage. The third section explains the modeling framework used in the next chapters, and the last part shows the hypotheses tested by the different analysis methods. Figure-3.1 shows the methodology framework used for this thesis.

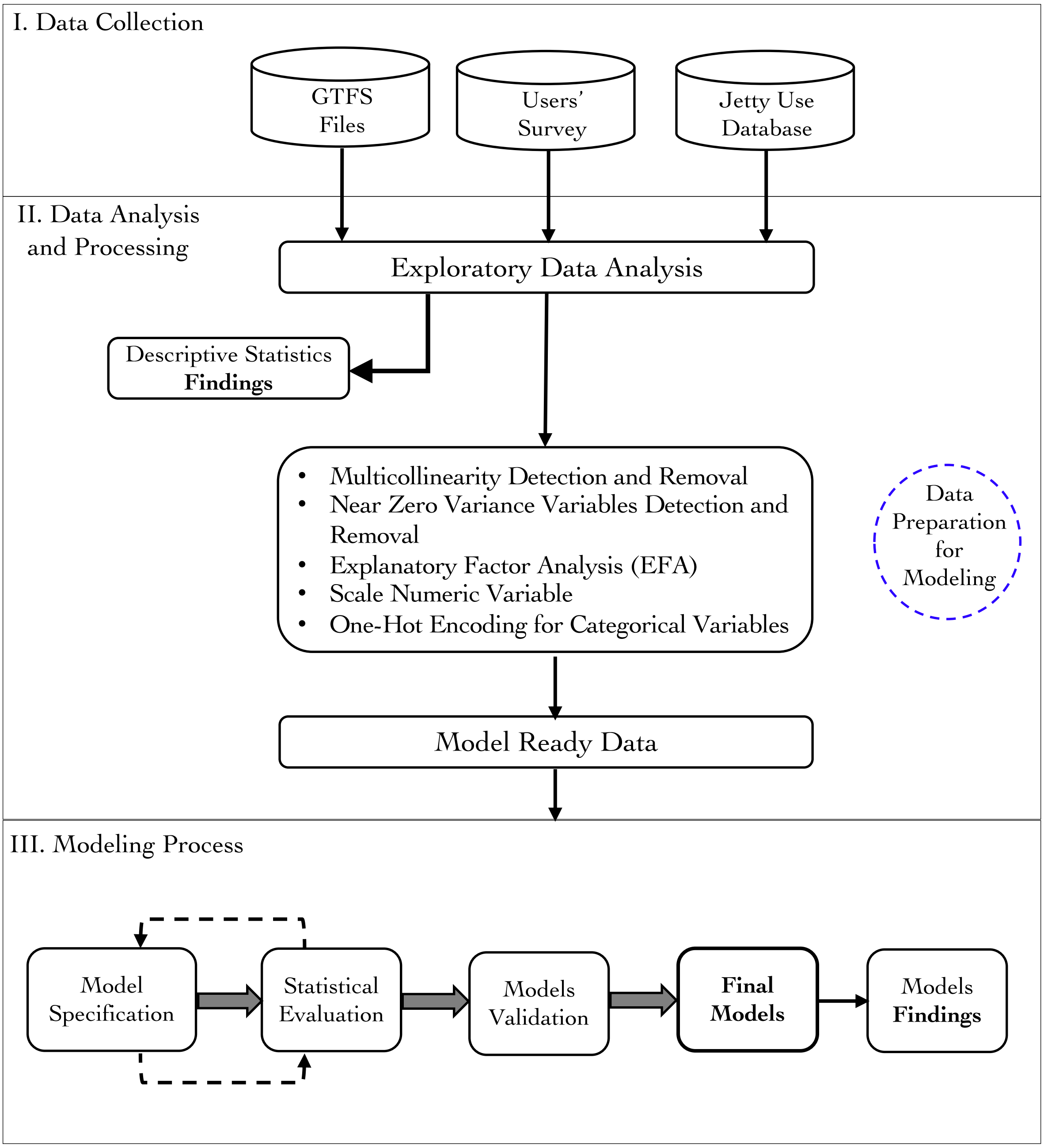

Figure 3.1 – Methodology Framework (*own illustration*)

3.1. Data Collection

3.1.1. Jetty Users' Survey

Initial Agreement

The chair of Transport System Engineering[1] (TSE) in the Technical University of Munich (TUM) contacted Jetty[2], a commercial pooled service platform, to arrange a collaboration between the department and Jetty to study the service properties, its impact on congestion, the service attraction factors, the sociodemographic characteristics of the users, users' general travel behavior, the factor driving the demand of Jetty (frequency of use), and the factors affecting the choice of the different Jetty services types.

The two entities concluded an agreement based on the condition that Jetty will not interfere in the research process in any manner that would affect the research results, and conclusion. Jetty's role was limited in distributing the survey questionnaire and providing any extra information needed by the research team.

Based on the literature review for factors affecting the adaptation and use of shared mobility services discussed in detail in section-2.2, and the objective and motivation of this thesis, a survey was designed to address the research objectives, and it consisted of three main parts:

Survey Structure

Survey Part One

The first part of the survey investigated the details of the users' last Jetty trip and how they would have replaced it if Jetty did not exist. This part consisted of seven questions. The next subsection shows part one questions and the proposed answers for each question.

Q1.1) Trip Purpose: This question asked the users to specify their last Jetty trip purpose; the user had to choose one of the six options provided:

1. Work
2. Study
3. Leisure
4. Formalities
5. Shopping
6. Other

Q1.2) Access Mode: This question asked the users to specify the modes they used to access Jetty from their trip origin point. Users were able to specify up to two modes that were used in combination to access Jetty. The options to choose from them were:

1. Walking
2. Microbus
3. Bus, (Combi)
4. Metrobus
5. Metro
6. Car
7. Bicycle
8. Taxi (without App)
9. E-hailing (Uber, Cabify, Taxify)
10. Others

Q1.3) Access Time: This question asked the users to specify the time in minutes they needed to access Jetty from the origin of their last trip.

[1] tse.bgu.tum.de
[2] jetty.mx

Q1.4) Egress Modes: This question asked the users to specify the modes they used to egress from Jetty to their final destination. Users were able to specify up to two modes that were used in combination to egress from Jetty. The modes to choose from them were the same modes offered in the used modes for access question (Q1.2).

Q1.5) Egress Time: This question asked the users to specify the time; in minutes, they needed to egress from Jetty to their final destination at their last Jetty trip.

Q1.6) Jetty's Replacement Modes: This question asked the users to specify the modes they would have used to replace their last Jetty trip if Jetty did not exist. Users were able to specify up to three modes combined to answer this question from the following options:

1. Car as a Driver
2. Car as a Passenger
3. Taxi
4. E-hailing
5. Shared Taxi
6. Metro
7. Suburban Train
8. Metrobus
9. Ecobus
10. Camion, Bus
11. Microbus
12. Combi
13. Motorcycle
14. No Trip
15. Others

Q1.7) Replacement Trip Duration: This question asked the users to specify the total trip time in minutes they would have needed to replace their last Jetty trip using the alternative modes specified in question (Q1.6).

Q1.8) Replacement Trip Parking Cost: This question asked the users to specify the parking cost they would have bore if they used private cars to replace their last Jetty trip in MXN[3]. This question choice set consisted of six price ranges:

1. Zero
2. Between 10 to 20 MXN
3. Between 21 to 40 MXN
4. Between 41 to 60 MXN
5. More than 60 MXN
6. Others

Survey Part Two

The second part of the survey investigated the users' general travel behavior, the activity users perform during Jetty trips, the reasons to use Jetty, and the willingness to walk to the Jetty access point. This part consisted of four questions.

Q2.1) Travel Modes Use Frequency: This question asked the users to specify the frequency of using different transport modes, available in Mexico City, on a Likert [145] scale. The scale consisted of five points describing the user frequency of use for the different services:

1. Never
2. Less than once a month
3. 1 - 3 times per month
4. 1 - 3 times a week
5. 4 or More times a week

The modes that were under examination in this question were:

1. Car, as a Driver
2. Car, as a Passenger
3. Taxi
4. E-hailing
5. Motorcycle
6. Metro
7. Metrobus
8. Suburban Train
9. Light Rail

[3] One US Dollar = 19 Mexican pesos (MXN) in July 2019, source: xe.com, accessed Feb 1^{st}

10. Trolleybus
11. RTP
12. Bus
13. Microbus
14. Bus, (Combi)
15. Jetty
16. Shared-App-Vehicle
17. Bicycle
18. Sharedbicycle
19. Shared Scooter
20. Walking

Q2.2) Activity During Trips: This question asked the users to specify the activities they perform while using Jetty. Users had the option to specify up to three combined activities from the offered activity list:

1. Socialize with other passengers
2. Work
3. Study
4. Use of smartphone
5. Talk on phone
6. Reading for pleasure
7. Sleeping
8. Look out of the window
9. Other

Q2.3) Reasons to Use Jetty: This question asked the users to specify which reasons attract them to use Jetty. Users could specify up to six combined reasons from the offered list:

1. Travel time (trip duration)
2. Access and Egress time
3. Travel time reliability
4. Booking of seat
5. Quality of vehicle
6. Fare
7. Traffic safety
8. Security against theft
9. Security against harassment
10. Use of time while travelling
11. Socializing with other passengers
12. Avoid parking problems
13. Ease of payment
14. Fare transparency

Q2.4) Willing to Walk: This question asked the users to specify how long they are willing to walk to the Jetty access point from the origins of their trips. The answers were a six interval time scale:

1. Less than 4 minutes
2. Between 4 and 6 minutes
3. Between 7 and 10 minutes
4. Between 11 and 15 minutes
5. Between 16 and 20 minutes
6. More than 20 minutes

Survey Part Three

The last part of the survey was designed to explore the sociodemographic characteristics of Jetty users. This part consisted of eleven questions that asked about the users' age, gender, employment status, education level, number of cars in the household, availability of a driving license, household size, personal and household income, and home and work zip code.

At the end of the survey, an optional question was provided to the users to add any comments regarding Jetty use or Jetty service in general.

Survey Administration and Sampling

The survey was implemented using the open-source survey software *Limesurvey*[4], and it was implemented in Spanish on a website administrated by the chair of transport system engineering.

Jetty distributed the survey to 50 users as a pilot stage to test the survey set up. Afterward, the final survey was deployed to Jetty users who used it at least once in the six months before May 2019. The final survey was collected in the period between the 30th of May to the 11th of June 2019, and around 3000 responses were received.

Note: It is worth noting that Jetty did not have any access to the collected data

3.1.2. Jetty Use Database

Jetty provided the use data for the survey participants for the last seven months before the survey launching date in order to assess users' use frequency, users' trip characteristics, and users' services preference. The received data covered the period from November 2018 to June 2019.

The main goal of requesting the user database was to get more insights from the service use behavior and to include them in the models to measure the trip characteristic's impact on the different research questions. The received database included the following trips' details:

1. Trip ID
2. Route ID
3. Pick-up and Drop-off Coordinates
4. Trip Distance
5. Number of booked Tickets
6. Fare
7. Type of used Vehicles
8. Vehicle Capacity
9. Departure and Arrival times
10. Trip total route length

A separate database for Jetty demand for May 2019 was provided to assess the occupancy of the different vehicles. The database contained the information related to the demand per each vehicle type, the number of commuters in each vehicle type, and the various vehicles' capacities.

3.1.3. GTFS files

General transit feeding specification (GTFS) files are open-source data files developed by Google[5] to incorporate transit-related information to the Google Map platform. The files contain scheduled information for public transportation and geographic information on the routes and station locations of the different transport modes. GTFS data have been published in Mexico City since June 2013 [79].

GTFS files for all the public transportation services were downloaded from the open mobility data platform[6] for Mexico City. The different services' official operators provide the data. The reason for including GTFS data in the analysis and the modeling process was to examine the synergy between Jetty use and the accessibility to the other public transport services and how they, in combination with users' home and work locations, impact them Jetty use.

[4] limesurvey.org
[5] developers.google.com/transit/gtfs
[6] transitfeeds.com

3.2. Data Analysis and Processing

The second stage in this research methodology is the data analysis and processing stage. After collecting the previous three sources of information, the next step was to explore the collected data and prepare it for the final stage of modeling. The details of the subsequent stages are explored in the below subsections.

3.2.1. Explanatory Data Analysis

Data Cleaning

The first step before starting the analytical process was to check the quality of the received data. This process was applied for the three sources of information mentioned earlier.

The survey data was checked, and uncompleted answers and duplicated entries were removed. The check of the duplicated answers was done by checking the users' ID. When two answers were utterly identical with the same user ID, the two answers were considered duplicated, and only one of them was kept in the data set.

The Jetty use database and GTFS files were checked, and no problems were noticed except that some variables have no variably, zero variance, as explained in the next subsections.

Data Analysis

Three primary methods were used for data analysis. The first method is the descriptive statistics that included calculating the minimum, maximum, mean, median, and standard deviation for the numeric variables. For categorical and ordinal variables, counting was applied, and market segmentation for the different counts was implemented when needed. The previous calculations were done using the core functions in the statistical software R [199].

The second method used for data analysis was data visualization techniques. Histograms, box plots, bar charts, and line charts were used to visualize and explore the different variables by plotting them on parallel coordinates or plotting the different variables under investigation against each other. The visualization task and data preparation for the plotting task were done using the *tidyverse* package [270] in the statistical software R.

Text Analysis

The third part of the analytical techniques was the text analysis for the comment question at the end of the survey. First, responses were cleaned from unnecessary words, also referred to as stop words such as (or, and, for). Afterward, two techniques for text analysis were performed. The first technique was the word counts and the associated word count. All responses were combined, and the individual words were counted to show the most frequent words. The associated word count was done to investigate which words are associated and if this association gives insight into the service properties. This task was performed using *tidyverse* package [270], and *tidytext* package [232] in R [199]. The second technique is the sentimental analysis for each users' comments. The cleaned comments were associated with attitudes using attitude dictionaries. They used dictionaries have a score, or attitude, associated with each word. Attitudes were aggregated per response, and the overall attitudes for all responses were evaluated.

3.2.2. Data Preparation

After cleaning and exploring the available databases, the next step was to prepare the data for the modeling process. This step's primary purpose was to remove the variables that could cause a problem in estimating the models by introducing instability or false information in the estimation process.

Multicollinearity

The detection of multicollinearity was done by using Pearson's [235] correlation coefficient for the numeric data using the core function (*cor*) in R [199]. The Polychoric correlation coefficient was calculated for the categorical and ordinal data using *polycor* package [88] in R [199]. Removal of some of the correlated variables were used to solve the correlation problem when detected.

Near Zero and Zero Variance Detection

The detection of near-zero and zero variance variables was done using the (*nearZeroVar*) function in *caret* package [141] in R [199]. Variables with less than 5% variability were removed. It is worth mentioning that this method was not applied to ordinal and binary data; as such, data is expected to have less variability; removing near-zero variance categorical-variables would result in losing important information.

Explanatory Factor Analysis

EFA was applied to the second part of the survey concerning the travel modes use frequency. The use levels ordered in ascending order, and the polychoric correlation matrix was calculated using *polycor* package [88], in R [199]. After calculating the correlation matrix, the EFA was calculated using *psych* package [203] in R [199], applying varimax rotation technique. The scree test was used to define the number of factors. EFA calculation was done in an iterative technique, where variables with factor absolute loading value less than (0.4) were removed until the EFA estimation was stable [106].

Numerical variable Scaling

All numerical variables were scaled to have a mean value of zero. Scaling was done in two steps. The first step was subtracting the mean of each column from all the corresponding observations. The next step was dividing the variables by their standard deviation. The core function (*Scale*) in R [199] was used for the scaling [34].

One Hot Encoding

The categorical variables were one-hot encoded. In this procedure, the categorical variable is converted to (**n**) number of variables equivalent to its original number of levels. For each of the new variables; if the original observation belongs to this category, it is coded as one and zero otherwise. Some of the software packages require this procedure to process the categorical data involved in the modeling process.

3.3. Modeling Framework

The final part of the thesis methodology is the models' estimation process. One of the main objectives of the thesis is to investigate the impact of the user's general travel pattern, or users' travel attitude, on the choice, and frequency of using shared mobility services. Attitudes are cognitive characteristics of the user that are gained through a long time, and it is reflected, under the scope of this research, on the people's travel choices [29]. The integration of people's attitudes on the different choice options was done by integrating the latent variable model into the different choice models using HCM. The answers for question (Q2.1), section-3.1.1, were used as the indicators for the measurement equation of the latent variable.

The additional knowledge and improvement in the models gained from using HCM are not always questioned. In some cases, the reduced choice model performs better than the HCM [264, 8]. Therefore, the previous point was addressed during the estimation process, as explained in the next procedure. The following procedures were applied for all the estimated models in the following sequence.

The dependent variable for each model is identified according to the research question it answers.

Afterward, the purposeful selection technique was implemented using the Zhang notion [279]. The selection was made by regressing each of the covariates individually against the dependent variable. Next, the covariates are ordered according to their significance in predicting the dependent variables. A cut-off value of 70% significance (p-value ≤ 0.3) was considered, and all the variables within this significance threshold were prioritized in the model building process. The models were estimated in iterative processing, starting from an empty model and adding variable by variable until a stable model was concluded. Coefficients with significant levels equal to or higher than 95% were kept. However, coefficients with significant levels less than 95% but higher than 90% were retained when they provided intuitive interpretation of the models. Also, the lower level of significance could be because of the limited sample size [142].

After estimating the choice model, the HCM was estimated guided by the choice model results, and the EFA[7] results. Several hypothesized specifications for the structure equation part were tested for each model, and the specifications that demonstrated the best results for the model in terms of coefficient interpretation and the overall model fit were kept. Subsequently, the overall fit of the HCM and the reduced choice model is compared to decide on which model fits the data better.

In some cases, two slightly different models were estimated. In these cases, statistical evaluation techniques like likelihood ratio test, AIC, and BIC were used to evaluate which model to be kept.

Note: It is worth noting that although a purposeful selection technique was used, the insignificant variables were tested. They were proven to be insignificant when added to the model, in most of the cases. This step was done to verify the purposeful selection technique.

[7]Two latent variables were suggested by the EFA results, and the details are discussed in the next chapters

After the model estimation process, model validation using the bootstrap technique was implemented for choice models only when they prove to fit better than the HCM. The re-sampling process used 1000 samples, and 95% confidence intervals (CI) for the estimated coefficients (β) are calculated using the first order normal approximation CI formulated as:

$$CI_{UL,LL} = \bar{\theta} \pm z_{\alpha/2} \cdot \hat{\sigma} \tag{3.1}$$

In equation-3.1, ($\bar{\theta}$) is the mean of the bootstrap estimated coefficient, ($z_{\alpha/2}$) is the Z-score corresponding to (α) level of significance, and $\hat{\sigma}$ is the margin of error [198].

All the choice models were estimated twice, one time using *Apollo* [113] package in R [199], and the *PandasBiogeme* package [34]. The main intent of using two different estimation packages was to make sure that the estimation optimization algorithm did not stop at a local optimum point, and the estimated models are stable regardless of the used software. HCM were only estimated using *Apollo* package.

Note: All observations with missing attributes were excluded from the modeling process. The final number of observations used to build each model is indicated in the models summary tables.

3.4. Tested Hypotheses

The following hypotheses were considered before the data analysis and the model estimation process based on the literature review, the evaluation of the current conditions of the public and private transport in CDMX, and this research objective. The verification and rejection of each of the following hypothesis is discussed in details in chapter-6.

Frequent User's Profile Hypotheses

Hypothesis 1: Gender-based violence impact: Females users are more likely to use Jetty more than males.

Hypothesis 2: Age impact: Young age group users are more likely to use Jetty compared to older groups.

Hypothesis 3: Income: High-income group are more likely to use Jetty compared to lower-income groups.

Hypothesis 4: Car ownership: Users with high cars ownership rates are more likely to use Jetty.

Hypothesis 5: Household size: Small-sized households are more likely to use Jetty compared to larger households.

Hypothesis 6: Employment: Full-time employed users are more likely Use Jetty compared to other users groups.

Hypothesis 7: Education: People with high education level are more likely to use Jetty compared with people with low education level.

Service and Trip Related Hypotheses

Hypothesis 8: Access and egress: Shorter access and egress distances will increase Jetty use.

Hypothesis 9: Trip distance: Longer trip distances will increase Jetty use rate.

Hypothesis 10: Fare: Jetty trip cost will reduce Jetty use.

Hypothesis 11: Multi-tasking and convenience: Multi-tasking and convenience provided by Jetty use will increase use rate.

Hypothesis 12: Safety: Safety and security offered by Jetty will increase use rate.

Transportation and Traffic Related Hypotheses

Hypothesis 13: PUT accessibility: People with poor accessibility to PUT are more likely to use Jetty .

Hypothesis 14: People are more likely to use Jetty in the mornings .

Modeling Hypothesis

Hypothesis 15: Including latent variable to the models will increase the model fit.

Chapter 4

Case Study

This chapter consists of two parts. The first part explains the situation of public transportation services, and traffic in Mexico City and a brief description of Jetty service. The second part of this chapter is the second part of this research methodology, where all the collected data are analyzed and explored. Finally, the collected data is prepared for the modeling process.

4.1. Transportation in Mexico City

Introduction

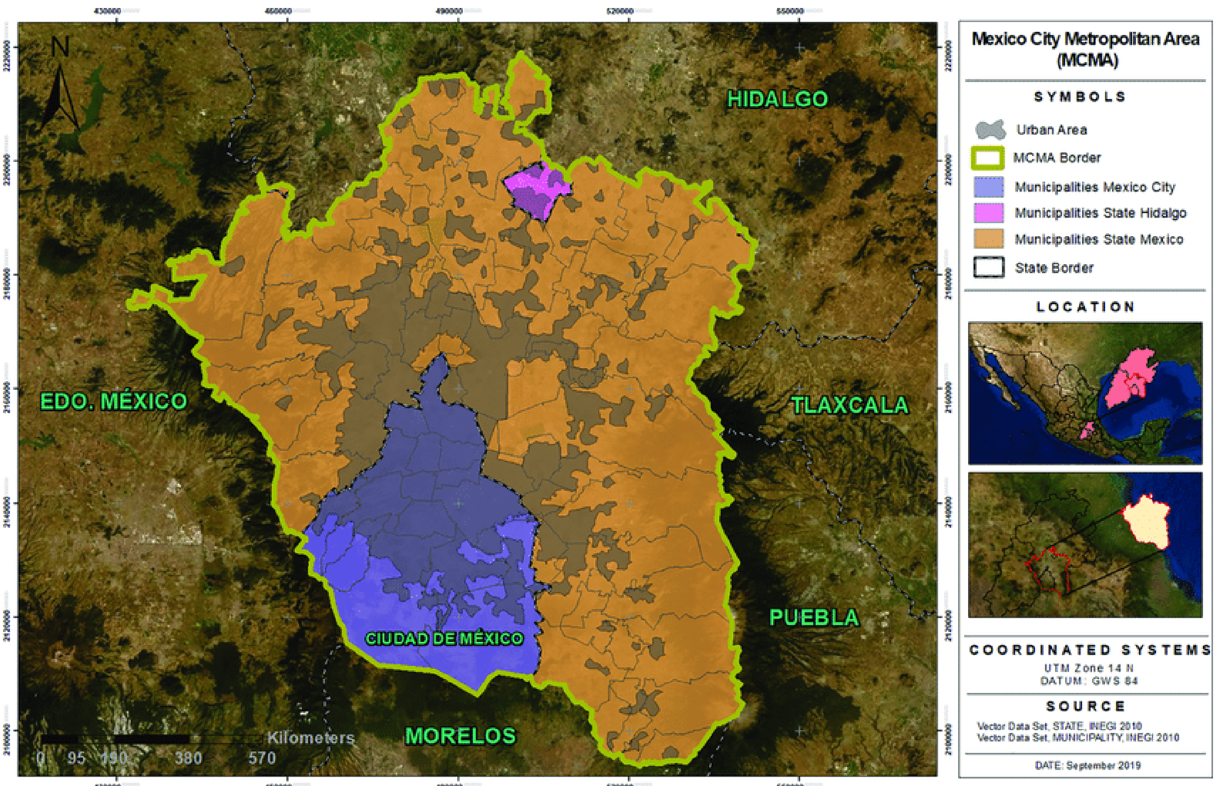

Figure 4.1 – CDMX Federal District Boundary, and Metropolitan Area Boundary, Source: [120]

Mexico City, Ciudad de México (CDMX), is the capital city of Mexico, and it is a part of the Valley of Mexico Metropolitan Area, Zona Metropolitana del Valle de México (ZMVM). ZMVM is the most populated area in Central and North America, and the world's fifth-largest metropolitan area with a total population of 21.58 million inhabitants [258]. The Population of CDMX is around 9 million inhabitants. ZMVM consists of CDMX, which has 16 counties, and 57 municipalities (17 counties) in the adjacent states of Mexico and Hidalgo [164, 120].

Public Transportation in CDMX

Public transportation plays a significant role in the life of the people of CDMX. Around half of the daily trips are performed by public transportation [190]. The city's public transportation network is well-connected in the central part of the city, with a lesser extent to the suburban areas, especially in the north of the city [188]. The main component of the public transportation system are:

Metro: The Metro or the subway network consists of 12 lines with a total length of network in service equal to 226.488 km, and 195 stations. The metro network is connected to other transportation modes such as Metrobús, Mexibús, and the light rail system. In 2018, the metro transported 1.647 billion users. The metro service during the working days is from 5:00 am to 12:00 am with a ticket price of one trip equal to five MXN [166, 189].

BRT, (Metrobús): BRT is a bus rapid transit system (BRT), consists of seven lines with a total length of 140 km, and 217 stations spaced on average of 645.2 meters. The average daily demand at the busiest corridor, line one, was 480,000 passengers per day in 2016, and the average daily demand for the least busy corridor, line four, was 65,000 passengers per day in 2016. The Metrobús network is connected to the other modes such as Metro, Suburban Train, and Ecobici. The Metrobús service during the working days is from 4:30 am to 12:00 am with a ticket price of one trip equal to six MXN [168, 167].

Light Rail, (Tren Ligero): Tren Ligero, or the electric light rail, is a 13.04-kilometer single line with 16 stations and two terminals serving the south part of CDMX. The light rail is connected to the metro network. The service runs from 5:00 am to 12:00 am during the working days with the single trip cost of three MXN [254].

Trolleybus, (Trolebús): The trolleybus network consists of eight lines with a total operating length of 203.64 kilometers. The trolleybuses' fleet consists of 290 electric vehicles that run with a four-minute average headway, and the network operating hours start from 5:00 am to 1:00 am during the working days. The trip cost is four MXN per trip. The bus network is connected through street transfer stations to the metro, trolleybus, and the Metrobús network [255].

RTP Bus: RTP, or the bus services network is the network with the widest geographical coverage in CDMX. The bus network consists of six main services [155]:

- *Ecobus Service*: the service runs on two routes and the trip costs five MXN.
- *Express Service*: the service runs on 22 routes and the trip costs four MXN.
- *Athena Service*: the service runs on 53 routes and the trip costs two MXN.
- *Ordinary Service*: the service runs on 91 routes and the trip costs two MXN.
- *Night Bus Service*: the service runs on 7 routes and the trip costs seven MXN.
- *School Bus Service*: this service runs based on the different schools schedules

Cable Car, (Cablebús): Cablebús, or the cable car, is the latest member of the public transportation network in CDMX. The cable car runs in the north of the city between the station of Indios Verdes and Cuautepec [46].

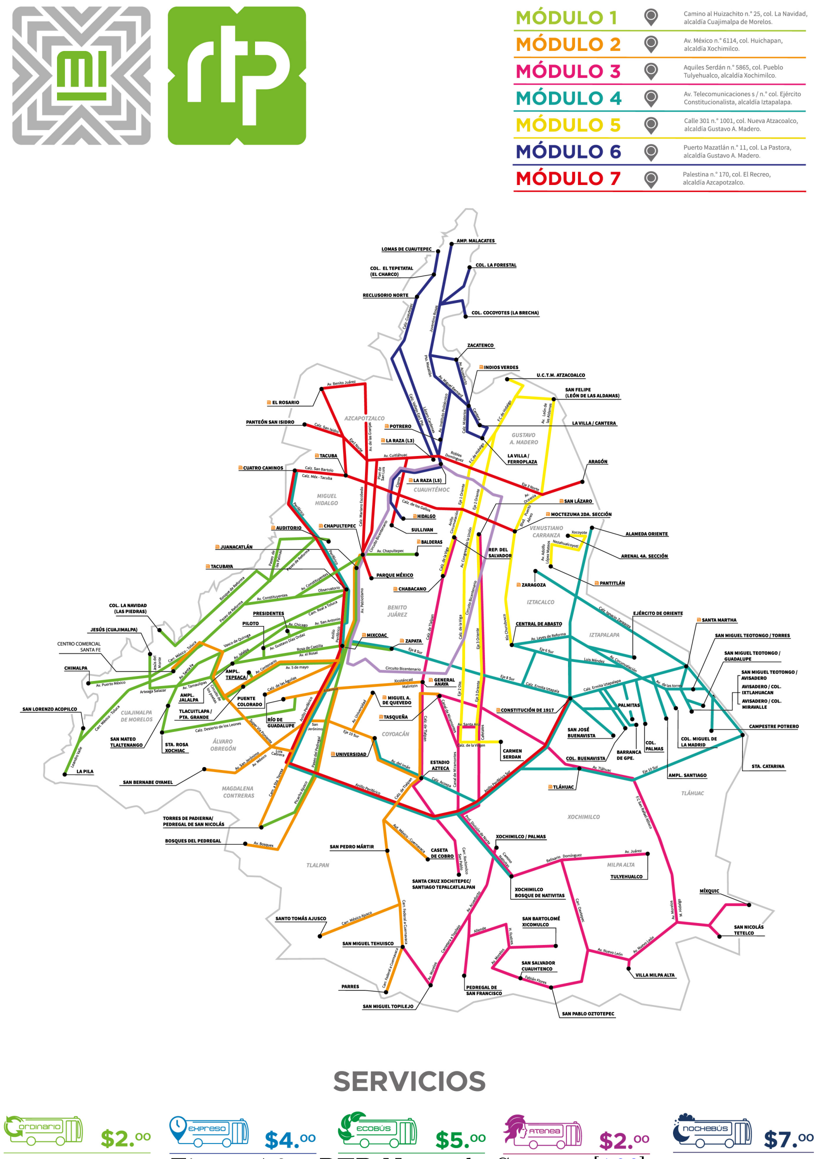

Figure 4.2 – RTP Network, Source: [166]

Suburban Train, (Tren Suburbano): The suburban train is the first mass rail system to connect CDMX to ZMVM, serving the three municipalities of Tlalnepantla, Tultitlán, and Cuautitlán. The suburban train runs from the north of the city (Cuautitlán station) to the Buenavista station within the city's boundaries, where there is a possibility to transfer to the Metro and the Metrobus. The total line length is 27 kilometers, and it has five stations between the origin and destination terminus. The train operating hours start at 5:00 am and end at 0:30 am the next day, with a ticket cost of 6.5 MXN. The suburban train ticket is exclusive for the use of the train and cannot be used to access the other public transportation systems. The suburban train transports 100 million passengers annually. 4.8 million person profits from the system 1.8 million of them are from the CDMX side [83].

Mexibus, (Mexibús): Mexibús is a BRT system that consists of three corridors with a total length of 52 kilometers and 94 stations, and it is connected to the metro network in different locations. Mexibús connects CDMX with the municipalities of Ecatepec, Tecámac, Nezahualcóyotl, Chimalhuacán, Coacalco de Berriozábal, Tultitlán, and Cuautitlán Izcalli. The system runs with ten minutes average headway, and its operating hours on working days start at 4:00 am to 1:00 am in the next morning, and one trip ticket costs seven MXN [121].

Collective Services, (Servicios de Transporte Colectivo): Collective services are forms of paratransit that are loosely regulated services with no governmental subsidies. The vehicles are owned and operated by individuals who obtain a governmental license for operation. Collective services can be hailed from almost anywhere in CDMX, and they do not have fixed routes or stops. The used vehicles are generally in unfavorable conditions, the drivers are mostly aggressive towards the passengers, and their driving style results in frequent crashes involving pedestrians and onboard passengers. The trip cost in these collective services varies from five to ten MXN depending on the trip distance [188].

Figure 4.3 – CDMX Public Transportation Map, Source: [166]

Shared Mobility in Mexico City

Being the second-largest nation in Latin America, Mexico attracted several shared mobility providers. This attraction is evident in the case of TNC companies where Cabify started its operation in 2012, and Uber started its operation in 2013 in the city as one of the earliest operation sites in Latin America [77, 188, 63]. The current shared mobility landscape in CDMX can be briefed as follows:

Scooter Sharing: Multiple international and national companies operate different networks of shared scooters in the city. For example, the Mexican company Grin[1], established in 2018, operates a fleet of e-scooters in CDMX with an average cost of 2 MXN a minute after the initial subscription cost. Also, in 2018, the American micro-mobility Lime-scooter[2] extended its service to Mexico. E-scooter operators face different challenges in CDMX that vary from the absence of a formal operational framework to the robbery of the scooters[3].

Bike Sharing: One of the oldest sharing mobility schemes in CDMX started in February 2010 by the local government launching ECOBICI. The funding of the ECOBICI bikesharing system was handled by a private operator[4]. ECOBICI is a station based bikesharing system with regulated working hours from 5:00 am to 00:30 am the next day. In 2020 there are around 324 thousand users registered in ECOBICI[4]. Other bikesharing companies like Mobike[5] operates in the city.

TNC: Several TNC companies operate in the city such as Uber, Cabify, Easy Taxi, and Yaxi [77]. The city authorities are facilitating the operating TNC companies in order to provide an alternative option for the public transportation services in exchange for 1.5% of the companies' gross revenues directed to a mobility fund. The success of the TNC in CDMX could be due to the poor quality, unreliability, and the hazards encountered when using taxi services [188]. Although TNC is considered a successful service in CDMX, its high price compared to other services limits its use. According to the 2017 OD survey, the percentage of average daily trips performed in TNC compared to other paratransit services is 1.4% [190, 188].

Car Sharing: Carrot carsharing provider is considered the first and the only service provider in the country. It started its operation in 2012[6].

[1]https://ongrin.com/, accessed Feb 1^{st}, 2020

[2]https://www.li.me, accessed Feb 1^{st}, 2020

[3]https://www.bloomberg.com/news/articles/2019-07-30/mexico-city-s-e-scooter-push-runs-into-thefts-tough-regulation, accessed Feb 1^{st}, 2020

[4]https://www.ecobici.cdmx.gob.mx/en/service-information/what%20is%20ecobici, accessed Feb 1^{st}, 2020

[5]https://mobike.com/global/blog/post/mobike-for-mexico, accessed Feb 1^{st}, 2020

[6]carrotcargo.mx

Insights from the 2017 Household Origin-Destination Survey

Household travel surveys are performed every ten years in CDMX; the following observations were summarized from the latest survey conducted in 2017 to evaluate the general travel pattern in CDMX.

Forty-four percent (44%) of the city's households own at least one car, 4.4% own a motorcycle, and 14% have a bicycle. Eighty-four percent (84%) of CDMX residents perform at least one trip per week during working days. The average number of trips per household in CDMX during the weekdays is 4.2 trips per day, and if walking trips are excluded, the average plunges to 2.1 trips per day [190].

Table-4.1 shows that public transportation is the primary mode of transportation in CDMX, moving around 53% of the total passengers. Two-thirds of the public transportation trips are made by collective services (combi, and microbus) in CDMX, and collective services make three out of four public transportation trips in ZMVM [190].

The morning peak hour is from 7:00 to 8:00 am, with more than 4 million trips on an average weekday. The evening peak hour is from 18:00 to 19:00. The morning peak hour is the highest of the two peaks [190].

Table 4.1 – Travelling Population by Mode in Working Day, CDMX Source: [190]

Mode	**Million Pax***	**Pct. %**
Public Transport	**3.71**	**53.4**
Collective (Microbus or Combi)	2.55	68.7
Ride Hailing and E-Hailing	0.59	16
Metro	1.33	36
Metrobus or Mexibus	0.35	9.4
Other modes	0.58	15.6
Private Transport	**1.75**	**25.2**
Car	1.62	92.7
Motorcycle	0.07	3.8
School Bus	0.07	3.9
Walking	**4.68**	**67.4**
Bicycle	**0.11**	**1.6**
Other	**0.01**	**0.2**
Total	**6.93**	

The summation of of Modes percentage is greater than 100 percent as the same person could use more than one mode in the same trip.

Pax* = Passenger

Public Transportation Situation in CDMX

Public transportation users in CDMX face various challenges in their daily commute. The most critical problems are personal safety and gender-based violence. For example, 90% of public transportation users in CDMX feel unsafe while using transit services. Sexual harassment and petty crimes are frequent on overcrowded buses. Gender-based violence is a growing problem for both users and operators, where 23% of women reported avoiding public transportation [188]. Rivadeneyra et al. conducted a survey to investigate the sexual harassment incidents in public transportation against female users. The study disclosed that collective services (Combi and Microbus) are the most unsafe modes with 44% of the incidents occurred on board, followed by metro 26%, BRT 14%, and trolleys 6% [206]. The metro system is considered the least safe in central and south America, and the second least safe in the world's largest fifteen cities [164].

Moreover, the fleet of public transportation is outdated, with several problems affecting the travel time and causing persistent delays. For example:

- In 2018, 101 of the 384 metro trains were out of service due to their bad conditions
- In 2017, there were 22,195 system failures in the metro
- Three hundred trolleybuses are over 20 years old
- In 2017, only 63% of the trolleybus fleet was in operation
- One third of the light rail trains are out of service for different reasons
- Twenty seven percent (27%) of the RTP fleet is out of service.
- Metrobus is always overcrowded, reducing the service quality
- Increases in travel time in public transportation compared to the private car quantified as (39% in the metro, 54% in collective service, 33%in autobus, and 22% in RTP) [230, 188]

4.2. Jetty

Jetty[7] is an application-based service that runs within the vicinity of CDMX. Jetty allows users to book a seat, share a ride, and the cost in different-sized vehicles, varying from three-seat taxis to 45 seat buses. The application matches the user's origin-destination (OD) with the nearest available routes and vehicles, providing several travel options for the users to choose from. Once the user chooses travel options, the ticket is issued and sent, followed by live-updates of Jetty's time and location. The beta version of the application was launched in July 2017 [188].

Jetty does not have a specific definition according to the shared mobility taxonomy. If Jetty service is classified based on the used vehicle size, Jetty could come under the category of carpooling and vanpooling. If Jetty is classified based on the operation scheme, Jetty is entitled to be ATS as it operates on a fixed geographic and temporal schedule with a finite number of stations near to the route, similar to the request stops ATS service category, discussed in detail in section-2.2.1. The received Jetty trip database, analyzed in detail in sec-4.3.2, shows that 97% of the trips are performed in vans or buses with vehicular capacity ranges from 13 to 45 seats.

Jetty does not own the vehicles but cooperates with collective-services (Jitney) to provide vehicles. Jetty makes sure the used vehicles are in good quality, new natural gas-powered vehicles are used, and insured. The drivers of the vehicles are insured, and they have the proper work contracts with the necessary background check-ups performed [188].

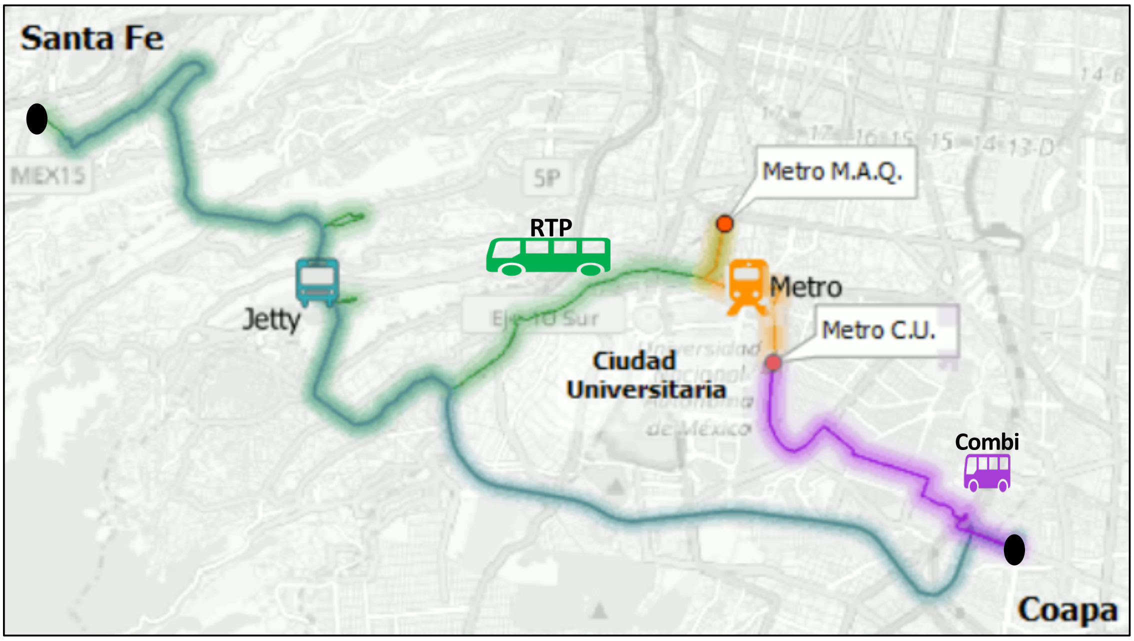

Figure 4.4 – Comparison between Jetty and Put Trip, Source: *Jetty*[7]

The company business model can be described as users pay the service fees to Jetty, Jetty keeps marginal revenue and pays the rest to the vehicle owner who pays the driver's salaries. In Jetty's case, this is a fixed salary and not a commission from the revenues as the case of regular collective services [188].

[7] www.jetty.mx

Jetty routes are concentrated in locations with sparse public transportation services in the north-west of the city, providing a direct connection to the job centers and shopping center locations in Santa Fe business area and Polanco districts. Figure-4.5 shows Jetty routes and pick-up and drop-off locations. Jetty direct connections save travel time, with a less number of transfers compared to using public transport and paratransit [188]. Regarding the travel cost and travel time, Jetty is more expensive than paratransit and PUT, but cheaper than e-hailing, a 23 kilometer trip from Mundo E, the north of the city, to Santa Fe costs in average 24 MXN in paratransit, 215 MXN in e-hailing, and 69 MXN in Jetty, the corresponding travel time for the same trip is on average 138 minutes in paratransit, 50 minutes in e-hailing, and 65 minutes in Jetty [188]. Figure-4.4 shows a comparison between a Jetty trip and a trip done on Public transport between Coapa, and Santa Fe. This trip reveals the potential in time and number of transfers saved when replacing public transportation with Jetty.

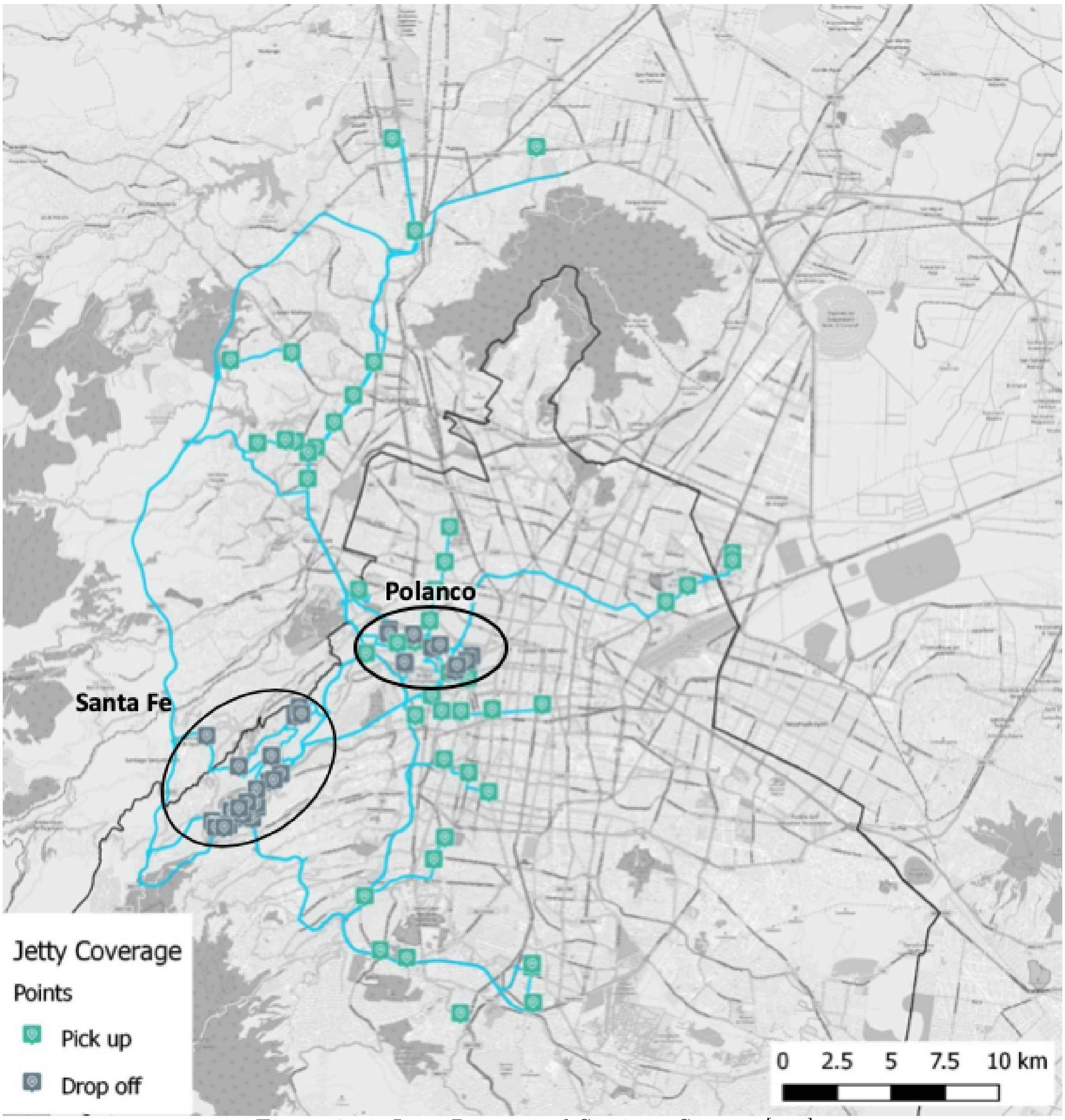

Figure 4.5 – Jetty Routes and Stations, Source: [188]

4.3. Data Analysis

4.3.1. Survey Data

Data Cleaning

The step before data processing is to assure the data quality and to clean and remove faulty or duplicated observations. Initially, 3000 responses were collected. From them, 2513 answers were completed; however, 29 of the completed surveys were identical, considering that they came from the same user ID. The duplicated observations were removed adjusting the total sample size to (2484) responses.

Afterward, the survey completion time was checked. This step is performed to detect any irrationality in the interview times, considering that users could have saved the survey and answered it on multiple sessions. Table-4.2 and Figure-4.2 show the distribution of the response times for the different users. No response was removed based on the survey completion time.

Table 4.2 – Survey Response Time Summary

Statistics	Mean	SD	Min	1^{st} Q.	Median	3^{rd} Q.	Max
Minutes	13.54	25.27	2.59	6.69	8.89	12.38	463.16

N = 2484

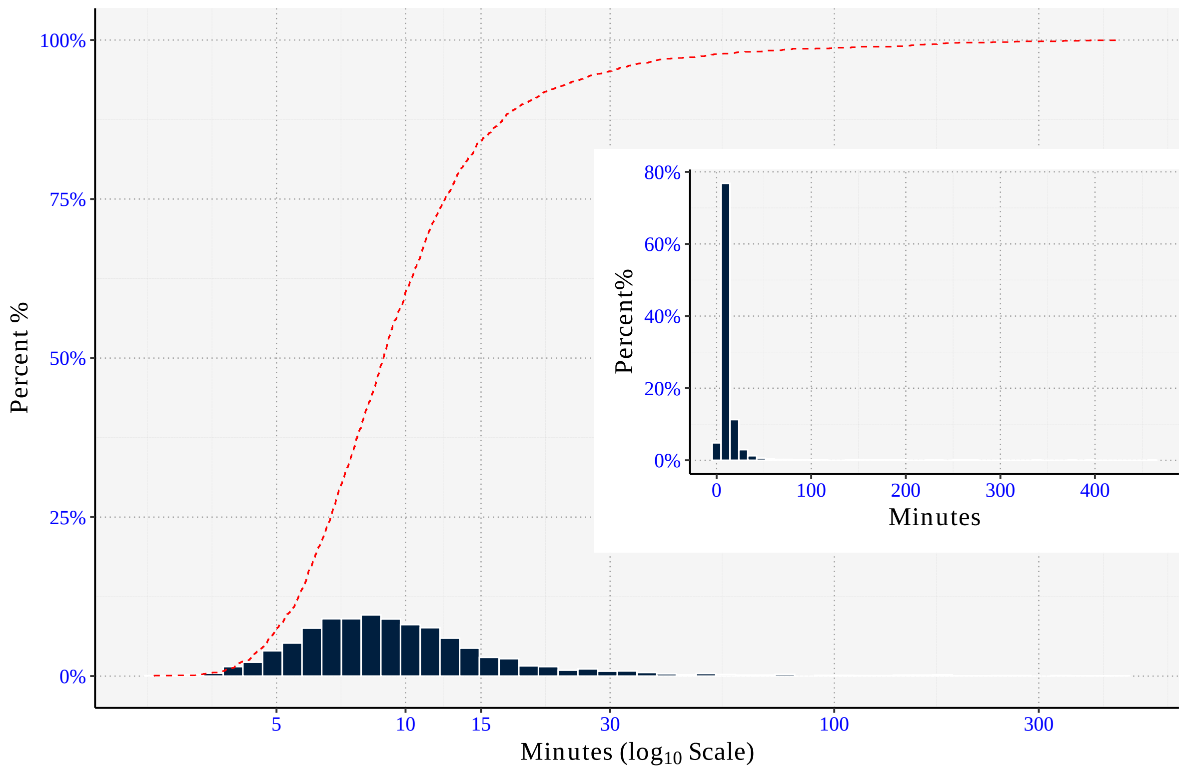

Figure 4.6 – Survey Responses Time Distribution

Users' Profile

After the data cleaning process, the first part analyzed was the sociodemographic questions, the third section in the survey. The understanding of the users' demographics is essential to conclude the Jetty's user characteristics and to compare these characteristics with the city's general population and with other users in different cities in other countries.

Table-4.3 shows the survey sample sociodemographic statistics summary and their comparison with the CDMX population. The sample is a strong representation of the average population of CDMX in terms of age, gender ratio, and household size distribution. The survey sample does not follow CDMX population distribution of average income, car ownership rate, or education level. The sample is skewed towards higher income groups, higher education levels, and higher car ownership rates. Only 32.1% of the population of CDMX has a higher education compared to the 88% of the sample, with at least a bachelor's degree. Moreover, the average income per person in CDMX is 10,000 MXN[8], while in the survey sample, the average income is around 20,000 MXN. Also, 80% of the survey sample reported at least having one car per household with an average of 1.25 cars per household. This is 2.5 times the city average of 0.53 cars per household [120, 192].

The differences between the survey sample, Jetty users, and the city average demographics rates match with the finding of the literature regarding the shared mobility users Characteristics, shown in Table-2.1. Although Mexico is a developing country, interestingly, the characteristics of Jetty users closely resembles the characteristics of other users in developed, industrial countries such as the United States, Great Britain, Germany, Canada, and Australia. The only significant difference between Jetty users and the other shared mobility users is that there is a strong presence of women among users. This presence could be because of increased violence against women in the different transportation modes in CDMX [206, 164].

[8] One US Dollar = 19 Mexican pesos (MXN) in July 2019, source: xe.com, accessed Feb 1^{st}

Table 4.3 – Survey Sample Sociodemographic Summary Statistics

Variable		Survey Sample	CDMX Population
	Levels	No (Pct%)	Level
Age			
	Between 18 and 25	376 (15.14%)	28,2% Between 6 and 24 year
	Between 26 and 35	1143 (46.01%)	54.3% Between 25 and 59 years
	Between 36 and 45	635 (25.56%)	
	Between 46 and 55	239 (9.62%)	
	More than 56	64 (2.58%)	17.5% 60 years and more
	Missing	27 (1.09%)	
Gender			
	Female	1212 (48.79%)	Ratio of Male:Female 1:1.11
	Male	1262 (50.81%)	
	Other	10 (0.4%)	
	Missing	00 (0%)	
Household Size			
	1	122 (4.91%)	Average Household Size, 3.2 unit
	2	556 (22.38%)	
	3	582 (23.43%)	
	4	630 (25.36%)	
	5 and more	495 (19.93%)	
	Missing	99 (3.99%)	
Personal Income, Pesos			
	Less than 10,000	320 (12.88%)	Average Monthly Income 10,000MXN
	10,000 - 20,000	878 (35.35%)	
	20,000 - 30,000	484 (19.48%)	
	30,000 - 40,000	216 (8.7%)	
	40,000 - 50,000	90 (3.62%)	
	More than 50,000	94 (3.78%)	
	Missing	402 (16.18%)	
Household Income, Pesos			
	Less than 10,000	150 (6.04%)	
	10,000 - 20,000	410 (16.51%)	
	20,000 - 30,000	417 (16.79%)	
	30,000 - 40,000	323 (13%)	
	40,000 - 50,000	224 (9.02%)	
	50,000 - 60,000	179 (7.21%)	
	60,000 - 70,000	107 (4.31%)	
	More than 70,000	149 (6%)	
	Missing	525 (21.14%)	
Driving License			
	Yes	1901 (76.53%)	
	No	583 (23.47%)	
	Missing	00 (0%)	
Cars in Household			
	0	520 (20.93%)	44% Household with at least one car
	1	1130 (45.49%)	
	2	628 (25.28%)	
	3 or More	206 (8.29%)	
	Missing	0.00	
Education level			
	Masters or Doctorate	378 (15.22%)	High Education 32.1%
	Bachelor or professional degree	1813 (72.99%)	Upper Secondary 26.6%
	Technical career	134 (5.39%)	Basic Schooling 38.9%
	High School or Baccalaureate	126 (5.07%)	No specific degree 0.3%
	Other	13 (0.52%)	Illiterate 1.5%
	Missing	20 (0.81%)	
Employment Status			
	Full time job	2127 (85.6%)	95.5% Economically Active
	Part time job	99 (4.0%)	
	Full time study	32 (1.3%)	
	Other	226 (9.1%)	
	N = 2484 (100%)		CDMX Population = 8,811,266 (2017) [120]

Description of Latest Jetty Trip

The next part of the survey explores the attributes and details of the latest Jetty trips.

Frequency of Last Trip

The first item to identify was how much does the last Jetty trip represent from the total Jetty use, or in other words, what is the share of the last Jetty trip recurrence from the overall Jetty use for each user. Trips for each user, retrieved from the Jetty's use database, were ordered based on their timing and date. The most recent trip before the survey completion time, extracted from the user's survey, was considered the last trip for each user. The geographic coordinates of the OD of the last trip were compared with the other trips ODs' coordinates for each user to calculate the percentage of the last trip OD recurrence to all the trips performed by each user.

Figure-4.7 shows that for 50% of the users the last trip OD represents 50% or more of their Jetty trips. Also, for 22% of the users the last trip OD is the only OD trip they perform using Jetty.

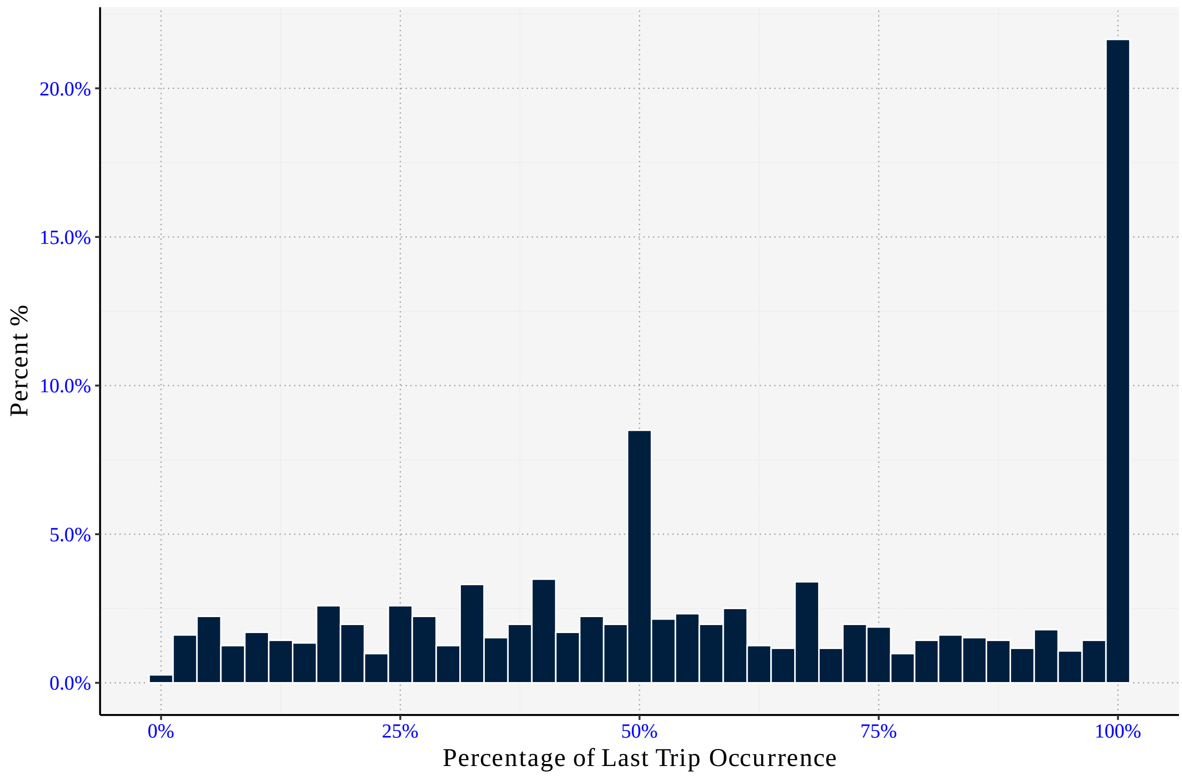

Figure 4.7 – Percentage of Last Trip OD Occurrence to the Total Jetty Use ODs

Last Trip Purpose

The users were asked to specify the purpose of their last Jetty trip. Some users reported going home under the other category answer, so a new trip purpose, going home, was created for those users. Also, as a part of the data cleaning process, the reported other category was assigned to the main five groups when possible. The primary purpose of using Jetty was to work for almost 95% of the users, followed by going to study and going home with a marginal percentage distributed between the other choices. Figure-4.8 shows that there was an almost equal distribution of trips' purposes per gender.

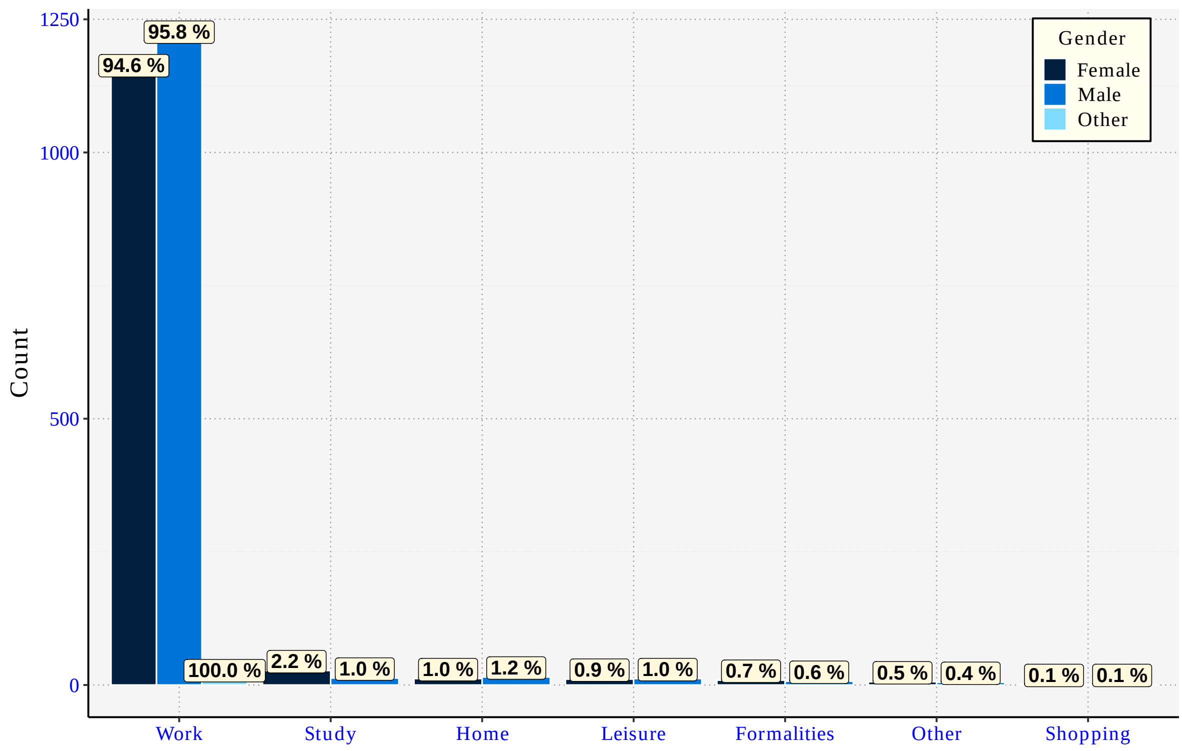

Figure 4.8 – Last Jetty Trip Purpose

Last Trip Access, and Egress Analysis

Access and Egress Modes

The next question in the survey asked the users to indicated the modes used to access and egress Jetty on their last trip. Users could choose up to two of the nine available options, or specify other unmentioned option.

First, the modes specified under other categories were allocated to the rest of the modes, when possible. The access and egress modes' analysis shows that most of the users (88%, and 95%) accessed or egressed the service using only one mode of transport, respectively. The rest of the users (12%, and 5%) combined two modes to access or egress Jetty. The desegregated access and egress modes analysis shows that more than 50% of users egress from Jetty by walking, and 34% access Jetty by walking. The most desegregated used mode to access and egress Jetty is walking, followed by microbus, car, and e-hailing.

It is worth mentioning that when the walk was specified in addition to another mode as the used modes to access or egress Jetty, the second mode was only assigned to the user, as walking is needed access any of the other modes.
The top combined modes ranking does not differ from the desegregated modes noting that taxi and e-hailing, and ride-hailing were combined under one category, taxi. The access and egress modes' analysis indicated that, overall, the pick-up and drop-off stations are well suited within the vicinity of the OD for most of the users facilitating access and egressing the service by one mode, mostly walking.

Figures 4.9 to 4.11 show the distribution of the desegregated, combined, and combined peak time access and egress modes. The mode distribution between the two peak hours shows an increase in egress by walking in the morning peak compared to the evening peak. This difference could be due to morning time constraints that do not exist during the afternoon peak, or since the morning peak is more severe than the evening peak. The rest of the modes are following the overall distribution trend of mode use.

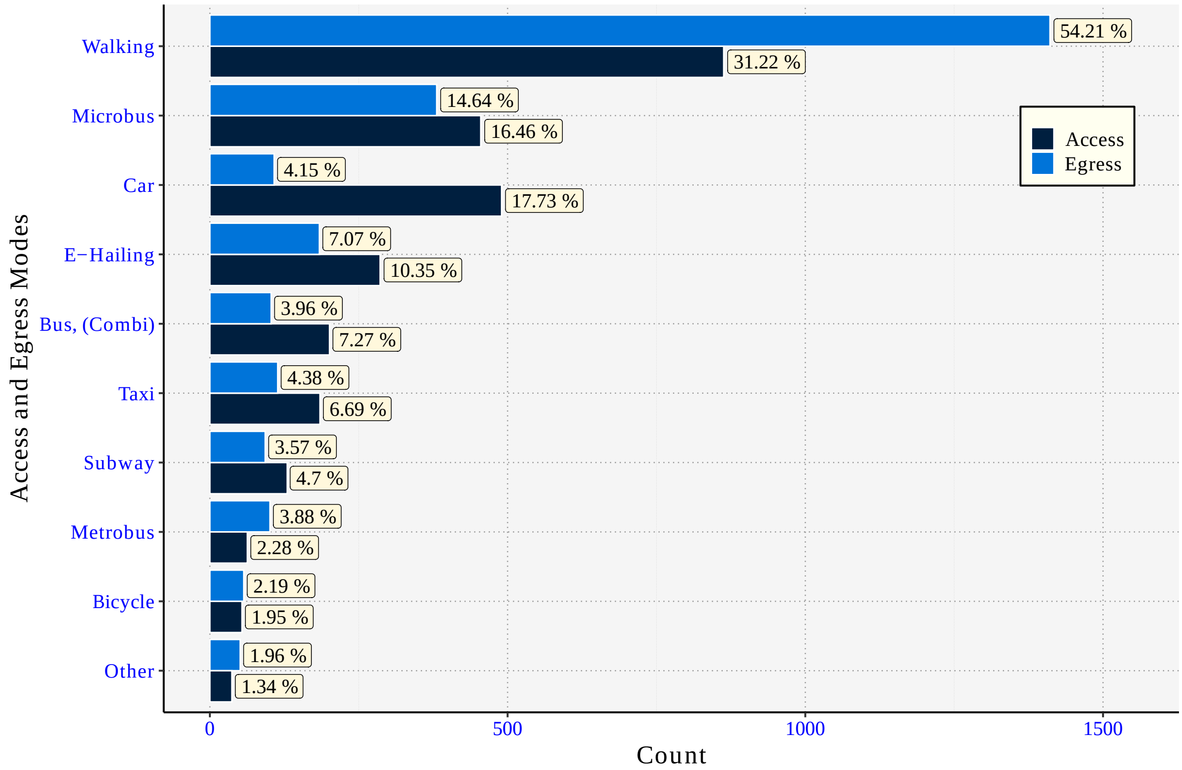

Figure 4.9 – Access and Egress Disaggregated Modes

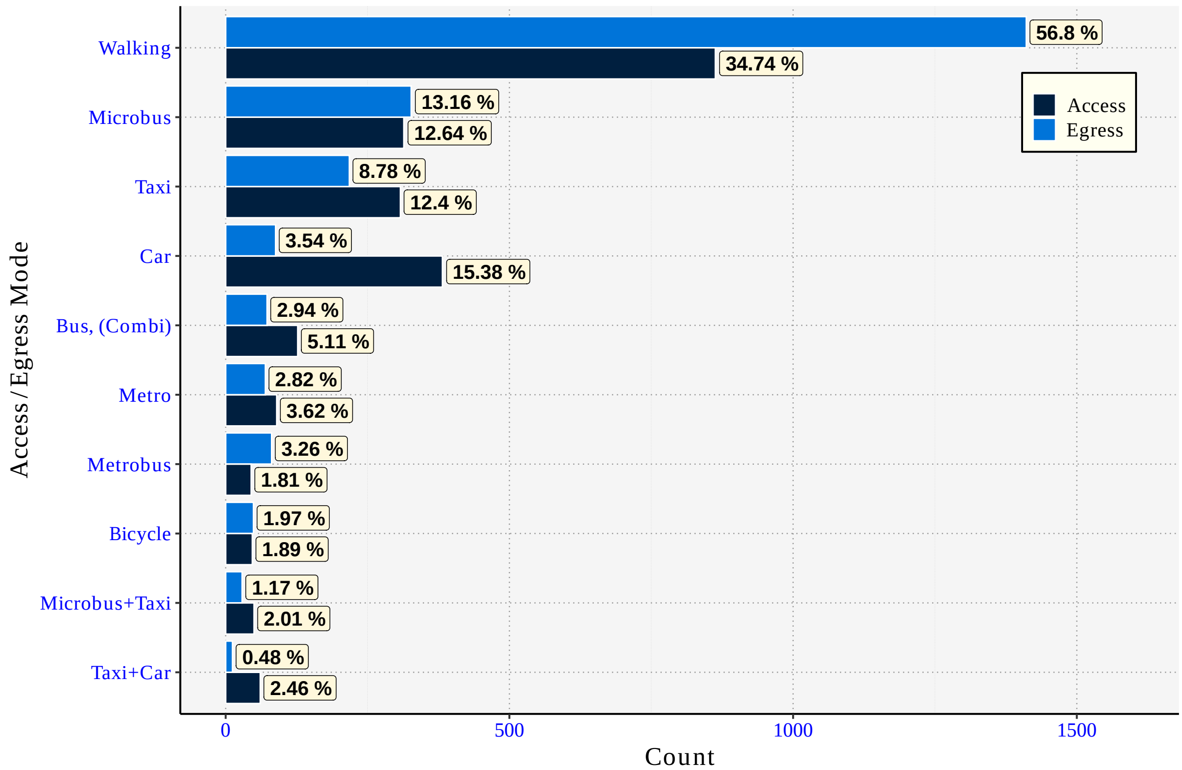

Figure 4.10 – Access and Egress Top Ten Combined Modes

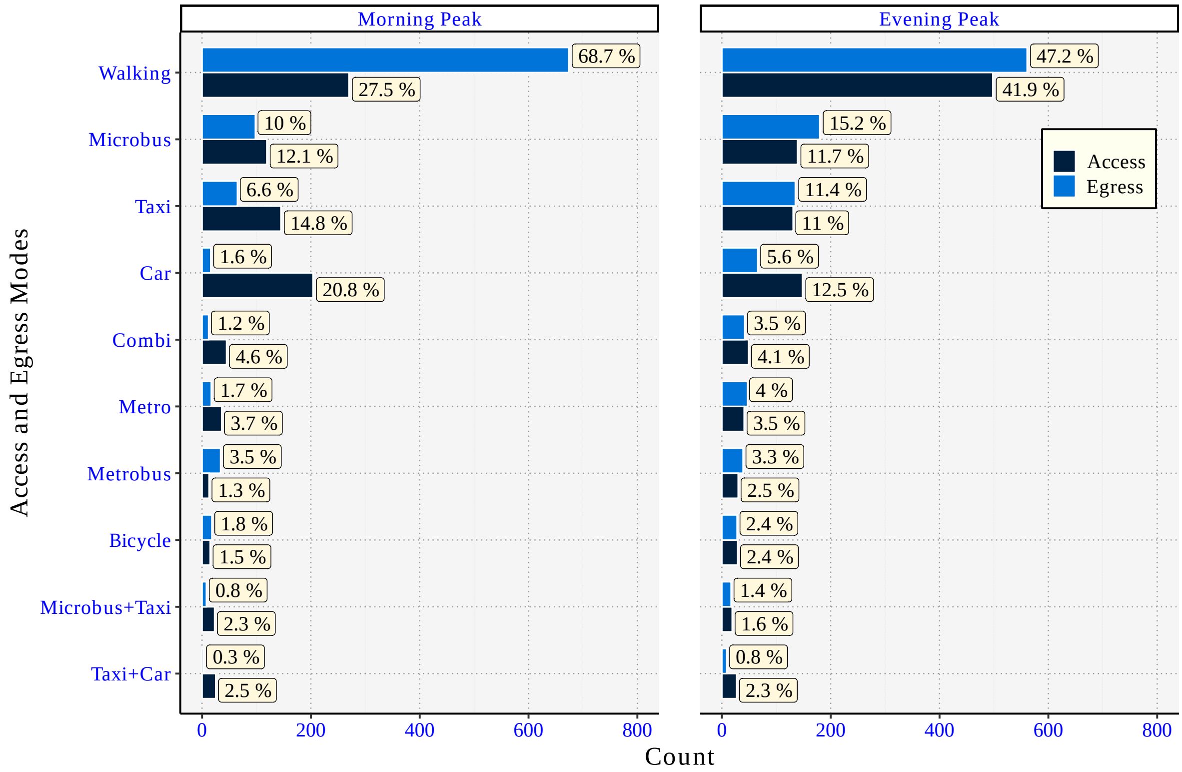

Figure 4.11 – Access and Egress Top Ten Combined Modes by Peak Hour

Access and Egress Time

Users specified the time they use to access and egress to and from Jetty. The analysis of these times shows that the average and median egress times are less than the access time with around three and a half to five minutes. This observation confirms that more users egress from the service by walking. The standard deviation for both times is almost equal.

Figure-4.12 shows four peaks for the access and egress time around 5, 11, 17, and 33 minutes with egress time being less skewed to the left. It was observed that some users reported extended access and egress times. For example, 17.8% and 12.0% of the reported access and egress times are over 30 minutes, and 3.6% and 2.3% are over 60 minutes, respectively. Table-A.2 shows the access and egress times by modes summary for the top five used modes.

Table 4.4 – Access and Egress Time Summary Statistics

Times (Minutes)	Mean	SD	Min	1^{st}Q.	Median	3^{rd}Q.	Max
Access Time	18.03	14.67	0	10	15	20	120
Egress Time	14.69	14.26	0	5	10	20	202

N = 2484

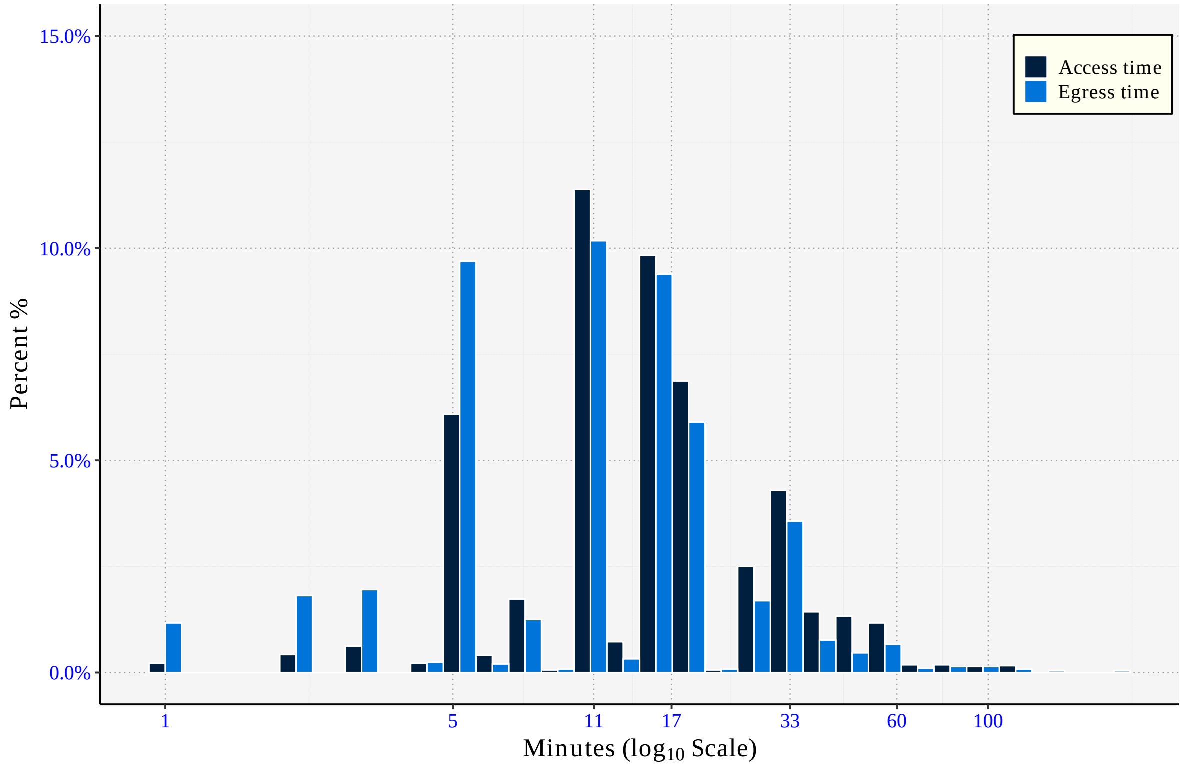

Figure 4.12 – Access and Egress Time Distribution, log scale

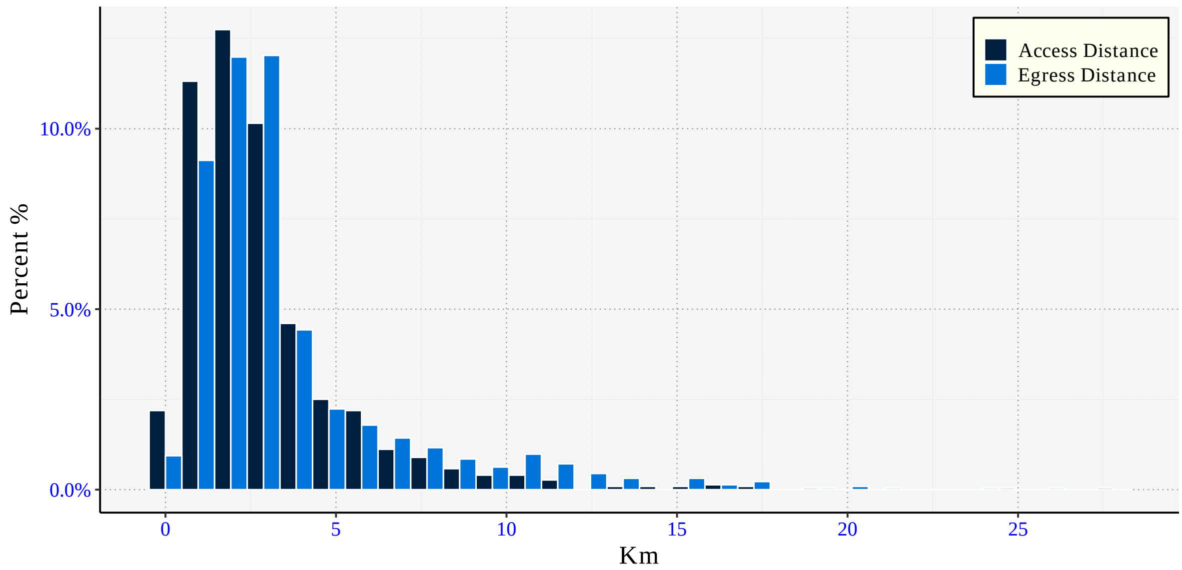

Figure 4.13 – Access and Egress Distances Distribution

Access and Egress Distance

Users provided their home and work locations zip codes, which were geocoded using the *ggmap* package [126] in R [199], and the corresponding coordinates were retrieved respectively. Jetty database included the pick-up and drop-off coordinates for each trip.

The distance between the home and pick-up location and work and pick-up location was calculated, and the shortest distance to the pick-up point was considered the access distance, and the other location distance to the drop-off point was calculated and assigned as the egress distance. Figure-4.13 shows the access and egress distances distribution and, Table-4.5 shows the summary of the access and egress distances.

Table 4.5 – Access and Egress Distance Summary Statistics

Distance (Km)	Mean	SD	Min	1^{st}Q.	Median	3^{rd}Q.	Max
Access Distance	2.98	2.68	0.10	1.37	2.29	3.53	27.90
Egress Distance	3.64	3.34	0.10	1.70	2.61	4.14	26.10

N = 1118

Modes to Replace The Last Jetty Trip

In this question, the users were asked to specify up to three modes that they would have used to replace the latest Jetty trip. There were 14 options, including no trip option. If none of the options suited the users, they had to specify that option under the other category.

Forty-two users indicated that they would not have performed the last trip if Jetty was not available. Three users did not specify any options. From the people who would not perform the last trip, thirty-six (36) users (80%) have at least one car in their household. This finding indicates the convenience of using Jetty and that Jetty induces extra travel demand.
Seventy-four percent (74%) of the users would use two or more modes to replace their last Jetty trip; this percentage goes along the fact that Jetty provided a direct connection between the different users' OD pairs. Table-4.6 shows the frequency and proportion of the number of modes that would have been used to replace the last Jetty trip.

Table 4.6 – Number of Modes to Replace Jetty Summary Statistic

No of Modes	0	1	2	3	Total
Count	45	601	800	1038	2484
Pct.%	1.80%	24.20%	32.20%	41.80%	100%

The disaggregated modes replacing the last Jetty trip shows that more than 50% of the trips have chosen the Metro to replace a link of the last trip. Also, the high number of cars per household is reflected in the users' choices as 34% of the users mark that they would use a car (as a passenger or as a driver) to perform one leg of the latest trip. Additionally, the average number of modes per trip = 2.14, which is a strong indication for the reason to use Jetty as it generally replaces multi-link trips, reference to Table-4.7. Figure-4.14 shows the analysis of disaggregated modes to replace the last Jetty trip per gender. The ratio of the use of the disaggregated modes per gender is almost balanced between males and females, except for the car as a driver, where there is a significant increase of males users compared to female users. Moreover, female users are slightly using more cars as passengers, and taking more rides using the Metrobus.

The 12 main travel options are aggregated to six main groups to reduce the complexity of the replaced trip analysis. Motorcycles and shared taxis were kept as separate groups due to their unique nature. The aggregated groups are:

1. Car
   - Car as a Driver
   - Car as a Passenger
2. Taxi
   - Taxi
   - E-hailing, and Ride-hailing
3. Microbus
   - Bus
   - Microbus
   - Combi
4. Metro
   - Suburban Train
   - Metro
5. Bus
   - Metrobus
   - Ecobus
6. Shared Taxi

Table 4.7 – Disaggregated Modes Replacing Last Jetty Trip

Rank	Mode	Count (Pct.%)	Rank	Mode	Count(Pct.%)
1	Metro	1273 (51.25%)	8	Ecobus	262 (10.55%)
2	Camion, (Bus)	800 (32.21%)	9	Metrobus	259 (10.43%)
3	Car, as a Driver	608 (24.48%)	10	Combi	215 (8.66%)
4	E-hailing and Ride-hailing	549 (22.1%)	11	Taxi	206 (8.29%)
5	Microbus	463 (18.64%)	12	Suburban Train	112 (4.51%)
6	Shared Taxi	275 (11.07%)	13	No Trip	45 (1.81%)
7	Car, as a Passenger	266 (10.71%)	14	Motorcycle	27 (1.09%)

Total Modes per trip = 5315/2484 trips = 2.14 mode/trip

The distribution of the combined modes per gender is almost balanced, with no significant difference. Figure-4.15, and Table-A.1 show the top ten combined modes to replace the last Jetty trip per gender.

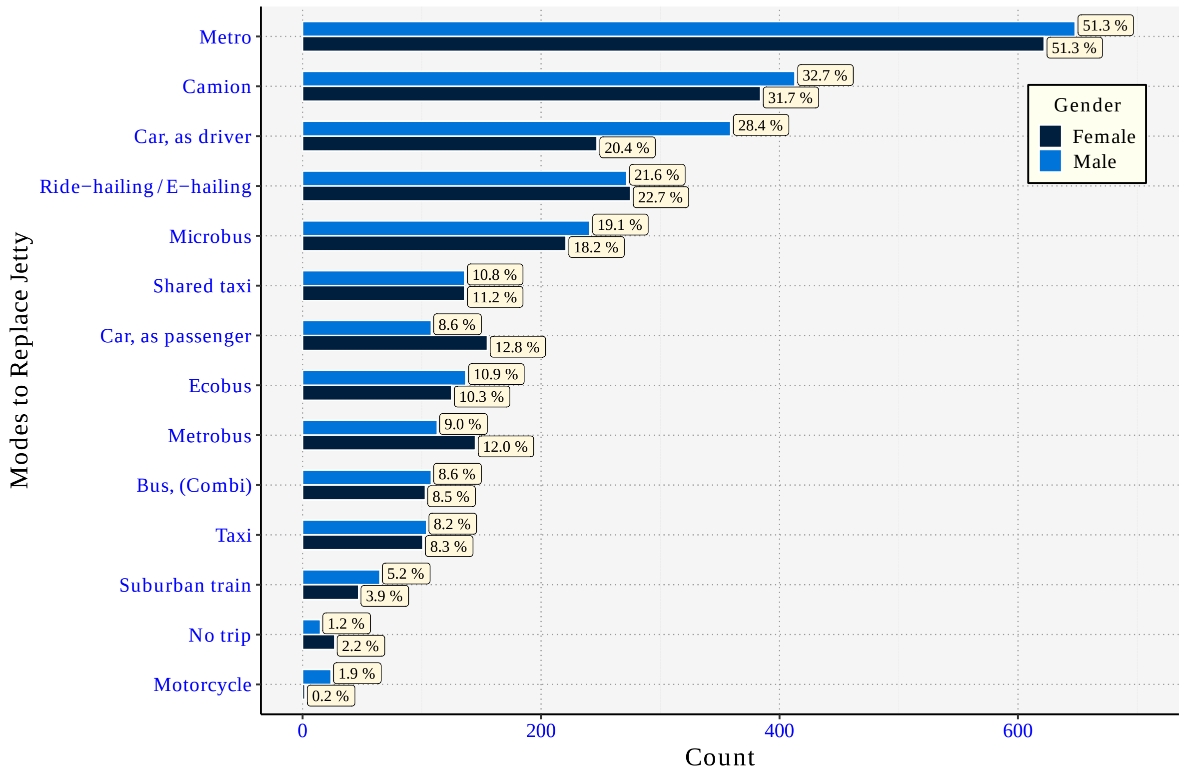

Figure 4.14 – Disaggregated Modes to Replace Last Jetty Trip

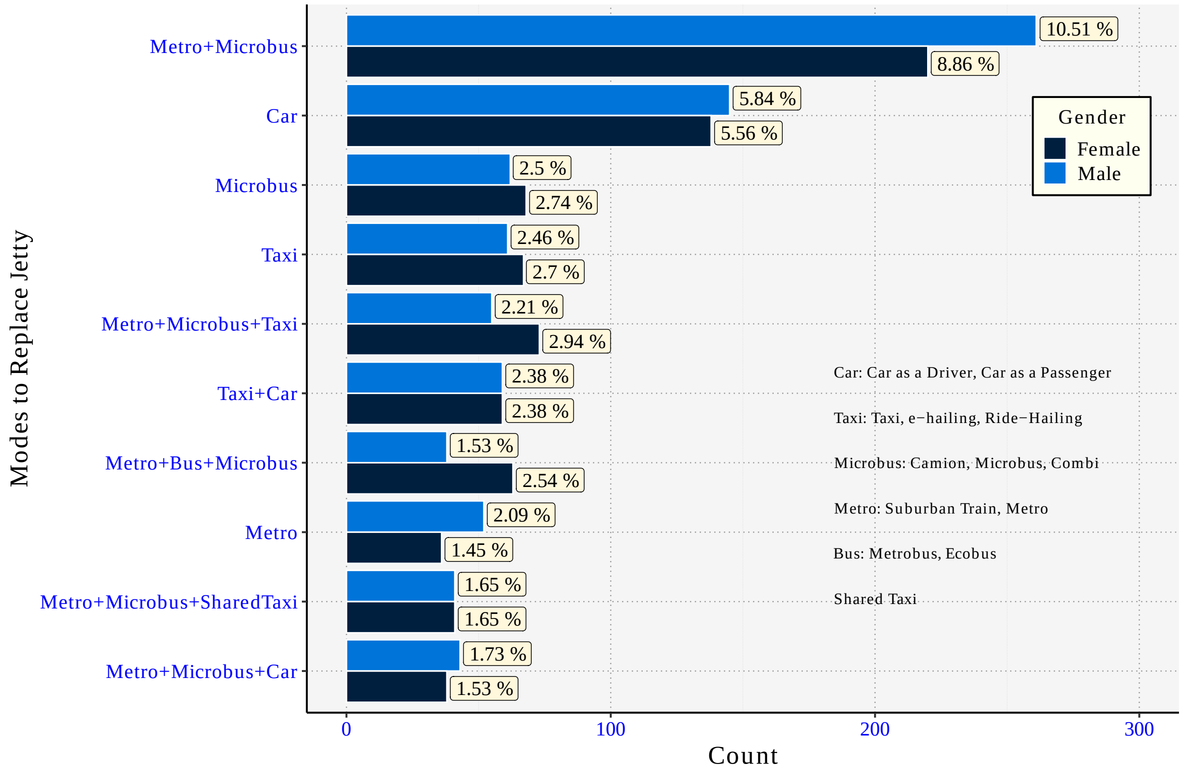

Figure 4.15 – Top Ten Combined Modes to Replace Last Jetty Trip

Time to Replace Last Jetty Trip

The users were asked to determine the time they would need to replace the last Jetty trip using the different modes specified in the previous question. The user's subjective estimated time was compared to the actual Jetty trip duration. Jetty trip travel time consists of the access and egress time to and from Jetty and the in-vehicle travel time retrieved from Jetty use database. The average last trip replace-time reported by the users was 96 minutes, compared to the average 78 minutes needed to perform the Jetty trip. The saving of travel time reflects the complexity of the trips Jetty replaced in terms of length and number of transfers.

Table-4.8 shows the summary statistics of the Jetty in-vehicle time and time to replace the last trip. Figure-4.16 show the comparison between the in Jetty trip travel time and the replaced trip travel time.

It is to be noted that users reported the last trip access, egress, and replaced trip travel time as subjective travel time, and not actual measured time, while in Jetty travel time is measured by the operator, and only provide for 2169 users only. The subjective travel time is not always accurate, and it might be overestimated or underestimated depending on different factors, such as route and destination familiarity, trip direction. Also, sociodemographic characteristics of the users impact their subjective time estimation [246].

Table 4.8 – Jetty Trip Time Summary Statistics

Variable (Minutes)	Mean	SD	Min	1^{st}Q.	Median	3^{rd}Q.	Max	N
Time to replace last trip	96.90	42.2	0	65	90	120	330	2484
Jetty In vehicle time	45.34	12.97	9	40	45	48	97	2169

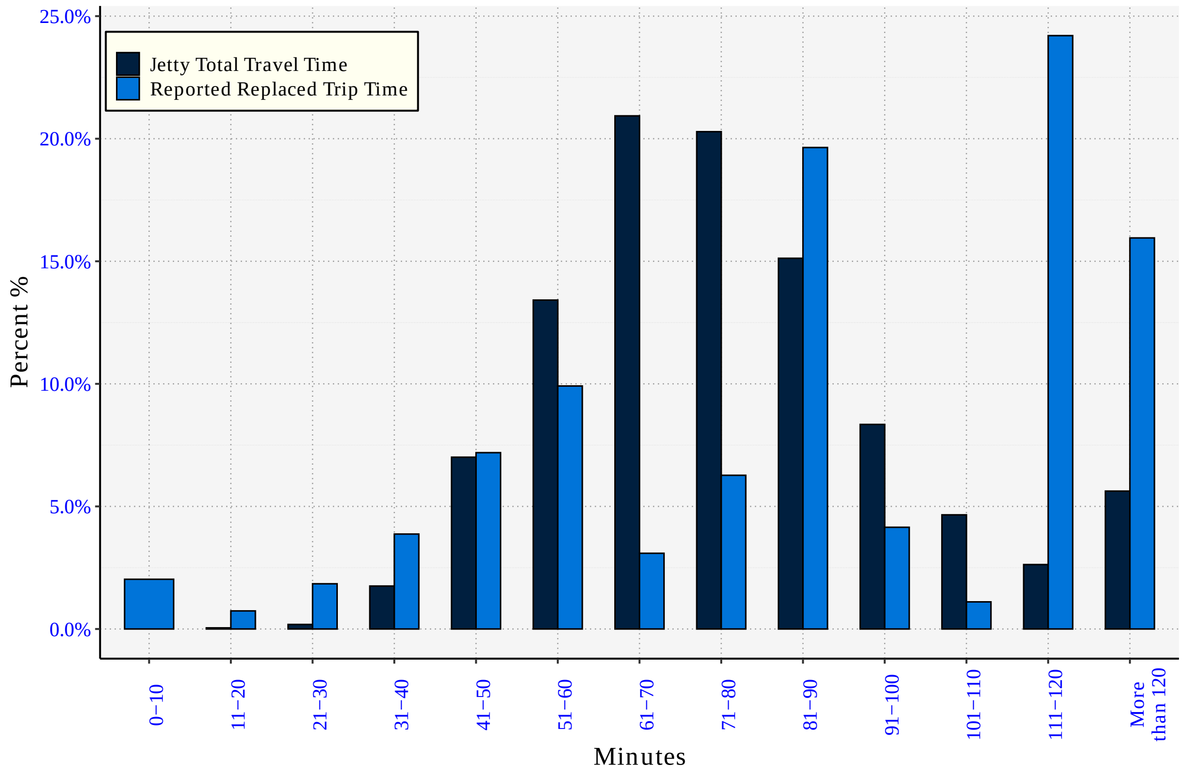

Figure 4.16 – Jetty Trip and Replaced Trip Time Distributions

Parking Cost

Users were asked to specify the expected parking cost if they used a private car to replace the last Jetty trip. Not all users in the sample have a private car or a driving license. Therefore, only people with a driving license, and with at least one car in the household, are identified (1653 users). Their estimation of the parking cost is compared to the overall sample estimation.

Table-4.9 shows the statistics for the reported cost by the total sample and the filtered users' group. There is no significant difference between the sub-sample and the entire sample. Almost 50% of the users reported that they expected zero parking costs. This factor could explain the high percentage of users that reported to use a car in at least one leg to replace the last Jetty trip (34%) compared to the city average car trips percentage 25%, refer to Table-4.7, and Table-4.1.

Table 4.9 – Users' Anticipated Car Parking Cost Summary Statistics

Parking Cost (MXN)	Count (Pct.%)	Count (Pct.%)
Zero	1149 (46.26%)	735 (44.46%)
Between 10 - 20	50 (2.01%)	32 (1.94%)
Between 21 - 40	103 (4.15%)	74 (4.48%)
Between 41 - 60	205 (8.25%)	162 (9.8%)
More than 60	763 (30.72%)	545 (32.97%)
Other	214 (8.62%)	105 (6.35%)
N	2484 (100%)	1653 (66.54%)

Willingness to Walk to Access Jetty

The users were asked to specify the time they are willing to walk to the Jetty access point. Almost 60% of the users expressed their willingness to walk up to 10 minutes to access Jetty. Only four percent (4%) reported that they would walk more than 20 minutes to access Jetty. Table-4.10 shows the summary of the stated times.

Table 4.10 – Willingness to Walk Time Summary Statistics

Willingness to Walk	Count (Pct.%)
Less than 4 minutes	111 (4.47%)
Between 4 and 6 Minutes	408 (16.43%)
Between 7 and 10 Minutes	900 (36.23%)
Between 11 and 15 Minutes	633 (25.48%)
Between 16 and 20 Minutes	328 (13.20%)
More than 20 Minutes	104 (4.19%)
Total	2484 (100.00%)

To further investigate the relationship between the willingness to walk, and the reported access and egress walking time, users were split into four main groups. The first two groups consisted of the users who access or egress the service by walking, and the other two groups consisted of the users who access and egress the service by other modes. Eight-hundred sixty-six users (866) users access Jetty by walk, and 1418 users egress the service by walk. The respondent's willingness to walk time was plotted against the reported access and egress time. Figure-4.17 shows a proportional relation between the reported access and egress time and the willingness to walk to the Jetty access point; however, the relation is not the same across all the four groups. Four ordered logistic regression models were performed using *MASS* [262] package in R [199], to confirm the previous observation. The access and egress times were used as regressors in four different models, and the reported willingness to walk time was the dependent variable.

Table-4.11 shows the regression results, which indicated the positive impact of the actual access and egress walking time, or people who reported more extended access and egress time are more likely to walk a longer distance to the access point. The positive relationship between the access and egress duration and the willingness to walk in the case of the users who access or egress by walk is more influential than the case, where other modes are used to access or egress Jetty (the value of the estimated coefficient for the first two groups are higher than the second two groups respectively). Also, the egress time coefficient for the people who egress by walk has more impact and higher coefficient value compared to the access time coefficient for the people who access by walk in the willingness to walk time. The interceptors between the different levels of the dependent variable for the different models are significant, which indicates that people are sensitive to the difference in time between the different levels.

Table 4.11 – Willingness to Walk Ordered Logistic Regression

No	Description	Access Time by Mode		Egress Time by Mode	
		Walking	Other	Walking	Other
		β(t-stat)	β(t-stat)	β(t-stat)	β(t-stat)
1	Access Duration	0.02 (4.0)	0.01 (3.6)	—	—
2	Egress Duration	—	—	0.06 (8.44)	0.003 (1.13)
Threshold 1	Less than 4 Min.|4-6 Min.	-3.28 (-15.6)	-2.66 (-21.1)	-2.52 (-18.2)	-3.05 (-18.4)
Threshold 2	4-6 Min.|7-10 Min.	-1.17 (-11.3)	-1.07 (12.3)	-0.76 (-8.6)	-1.32 (13.0)
Threshold 3	7-10 Min.|11-15 Min.	0.45 (4.8)	0.57 (7.0)	0.91 (10.4)	0.31 (3.4)
Threshold 4	11-15 Min.|16-20 Min.	1.76 (15.5)	1.84 (19.6)	2.24 (21.0)	1.60 (15.3)
Threshold 5	16-20 Min.|More than 20 Min.	3.33 (18.4)	3.43 (23.7)	3.75 (24.1)	3.32 (18.9)
Sample Size		866	1618	1418	1066
AIC		2660.0	5034.1	4347.6	3300.2

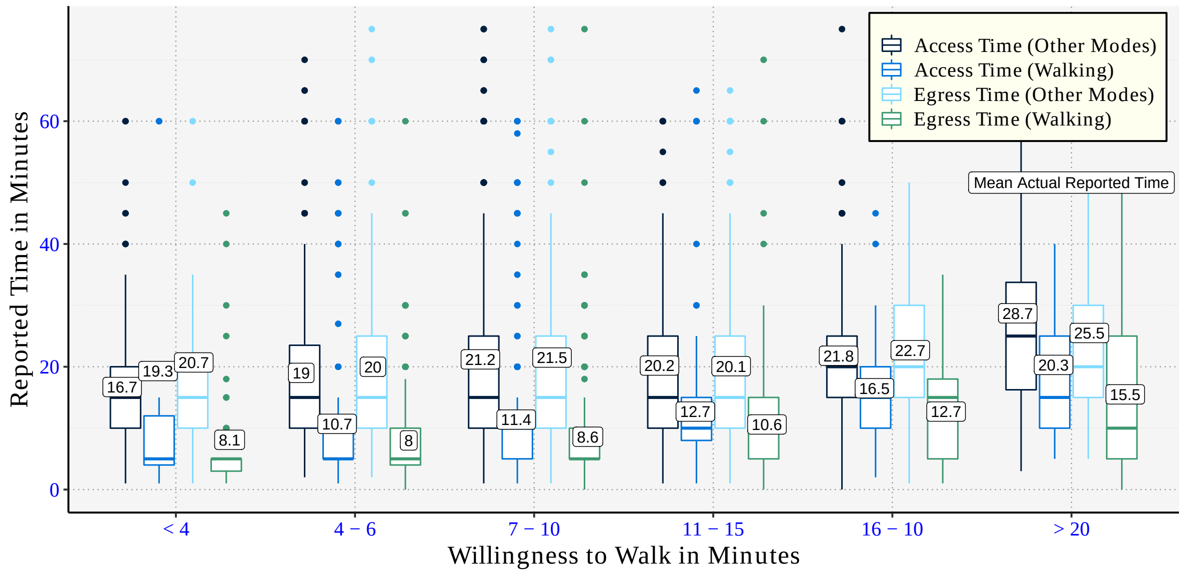

Figure 4.17 – Willingness to Walk Time Distribution

Different Modes Use Frequency

The survey asked the users to specify their use frequency for the different modes to assess their travel patterns. The users ranked their use frequency on a five-point-ordered scale that ranges from never using this mode to using this mode more than four times per week. Table-A.5, and Figure-4.18 show the reported frequencies per each of the 20 modes.

The least used modes are shared Scooter, bikesharing, and suburban train. This limitation could be because shared scooters are recently introduced to the city and the limited geographical cover for the suburban train as it runs only in the north of CDMX. The most used modes are e-hailing, metro, car as a passenger, and walking. These modes are in parallel with the highest used modes to replace the last Jetty trip. This finding is logical as most of the users (95%) reported that their last trip purpose was work, which is a frequent daily trip. People's travel behavior is habitual in general [211, 12]. Jetty users are regular e-hailing users. Thirty-five percent (35%) of Jetty users use e-hailing at least once a week compared to only 16% using a taxi at the same rate, which indicates that Jetty users are frequent e-hailing users and in general shared mobility users prefer the use of shared services. Also, this use pattern might be due to the advantages e-hailing provides for the users compared to taxis in terms of ease of payment and fare transparency [188].

The use per gender is balanced, except for two modes, where males are more frequent drivers than females, and females are more frequent car as passengers than males.

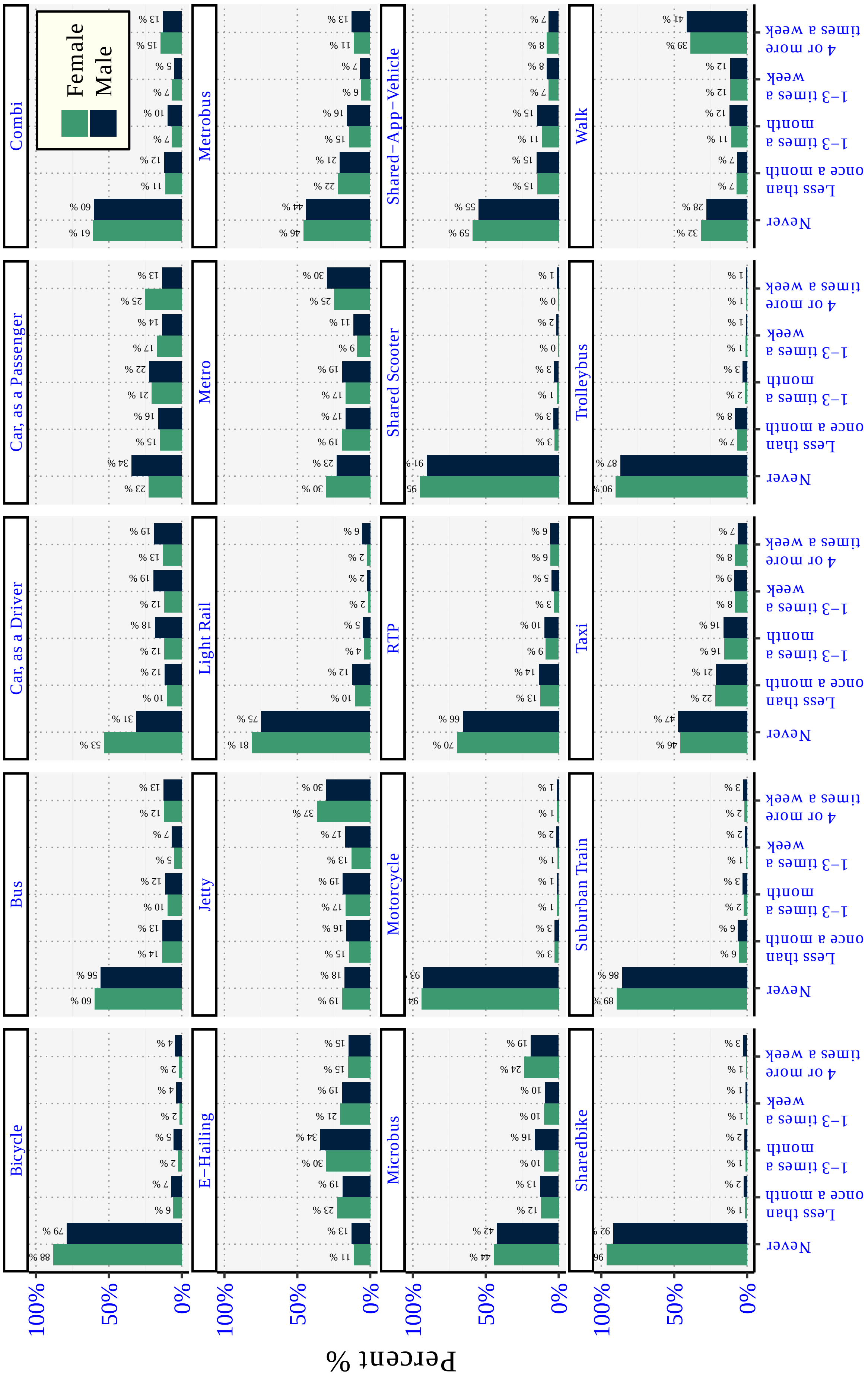

Figure 4.18 – Modes Use Frequency

Users' Activities During Jetty Trip

The survey asked the users to specify up to three activities they do during their travel in Jetty. Table-A.3, and Figure-4.19 show the summary for the disaggregated users' activities. The most specified activities are sleeping, using of the smartphone, and looking out of the window.

There is no significant difference between the genders for the different activities except for I) sleeping, where women outnumbered men by 9%, which reflects the sense of security women experience, while using Jetty, II)looking out of the window, where female outnumbered male by 4% III) working, and reading for pleasure, where males outnumber females by around 4%.

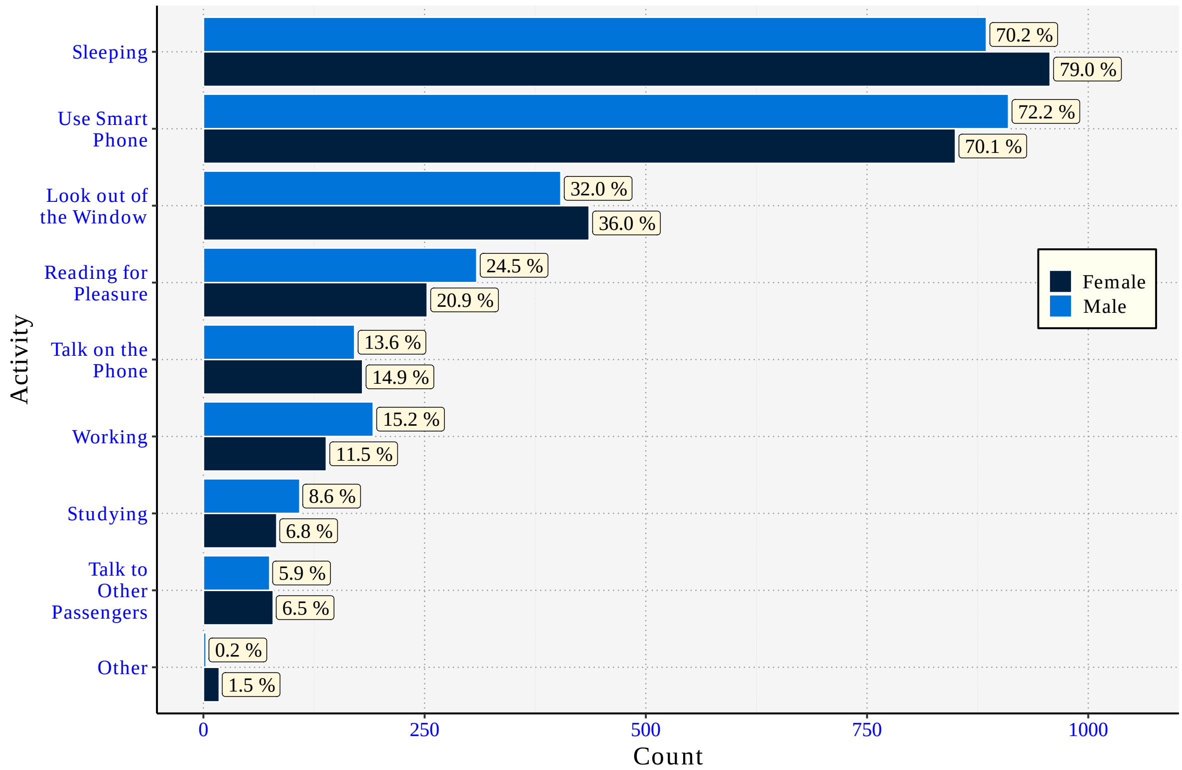

Figure 4.19 – Distribution of Disaggregated Users' Activities During Jetty Trips per Gender

Reasons to Use Jetty

Users are asked to specify up to six reasons from a choice set that consisted of fourteen options for why they use Jetty. Table-A.4, and Figure-4.20 show the summary of the disaggregated reasons to use Jetty.

The top three reasons to use Jetty are booking of the seat, security against theft, and saving in travel time. These reasons reflect the problems of public transit in a crowded city like CDMX. The gender distribution for the different reasons is almost balanced for all reasons except for two: I) security against harassment. Females reported this reason six times more than males, which reflects the increasing gender-based violence problem in public transportation in CDMX [206, 164]. II) The second difference is in avoiding parking problems; males were twice as likely as females to report this reason to use Jetty. However, this could be because males use cars as drivers more than females, as shown in Figure-4.18.

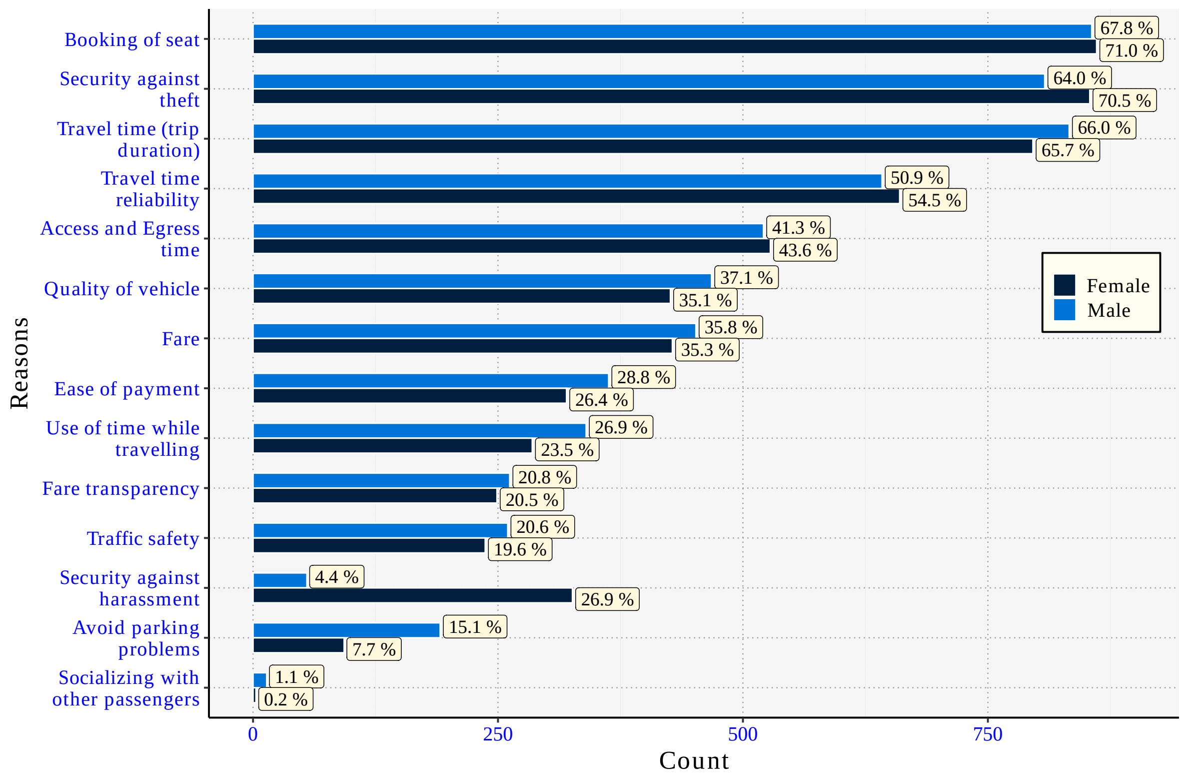

Figure 4.20 – Distribution of Disaggregated Reasons to Use Jetty per Gender

Note: It is worth mentioning that phi coefficient of correlation for binary variables [78] was calculated for the variables of reasons to use Jetty and activities during Jetty trip, 23 variables, to investigate if there is any correlation between any pairs of the different variables, and the estimated phi coefficient was less than 0.1 between all variables, which indicates that there is no association between any of the variables.

Home and Work Locations

The last mandatory question in the survey was to specify the home and work locations' zip codes. Only **1118** users specified a correct combination of home-work locations. In the next sections, these users are referred to as the partial sample. The data of the 1118 users will be used in the modeling process as their full information are available.

Home and work zip codes were geocoded using the *ggmap* package [126] in R [199], and the corresponding coordinates were retrieved respectively.*ggmap* package uses Google maps[9] as a reference for the geocoding process. The location of homes and work were plotted against the city map. Results show that home locations were scattered around the city and extended till the borders of ZMVM, refer to Figure-4.21. Work locations were concentrated in the areas around Santa Fe and Polanco, which is the main drop off areas for Jetty users. A point density analysis was done for both home and work locations, using *MASS* [262] package, to confirm the previous observations. The density of home locations shows a relatively equal distribution around the city with an evident concentration in the south. Work locations are mainly concentrated in two areas, Santa Fe and Polanco, which are the main drop-off points of Jetty. The points density analysis conforms with Jetty pick-up and drop-off locations. Figure-4.21 shows the users' home and work locations, and Figure-4.22 A) shows the density for the home locations, and B) shows the density of work locations.

[9] https://www.google.com/maps

Eight hundred ninety-three (893) users representing (79.9%) of the sample reside inside the limits of CDMX, and 225 users representing (20.1%) of the sample live outside the boundaries of CDMX, but inside the borders of ZMVM. A binary factor indicating if a Jetty user resides inside the limits of CDMX was created to be included in the different models to evaluate the various impacts of the user's home location on adopting Jetty as a travel mode.

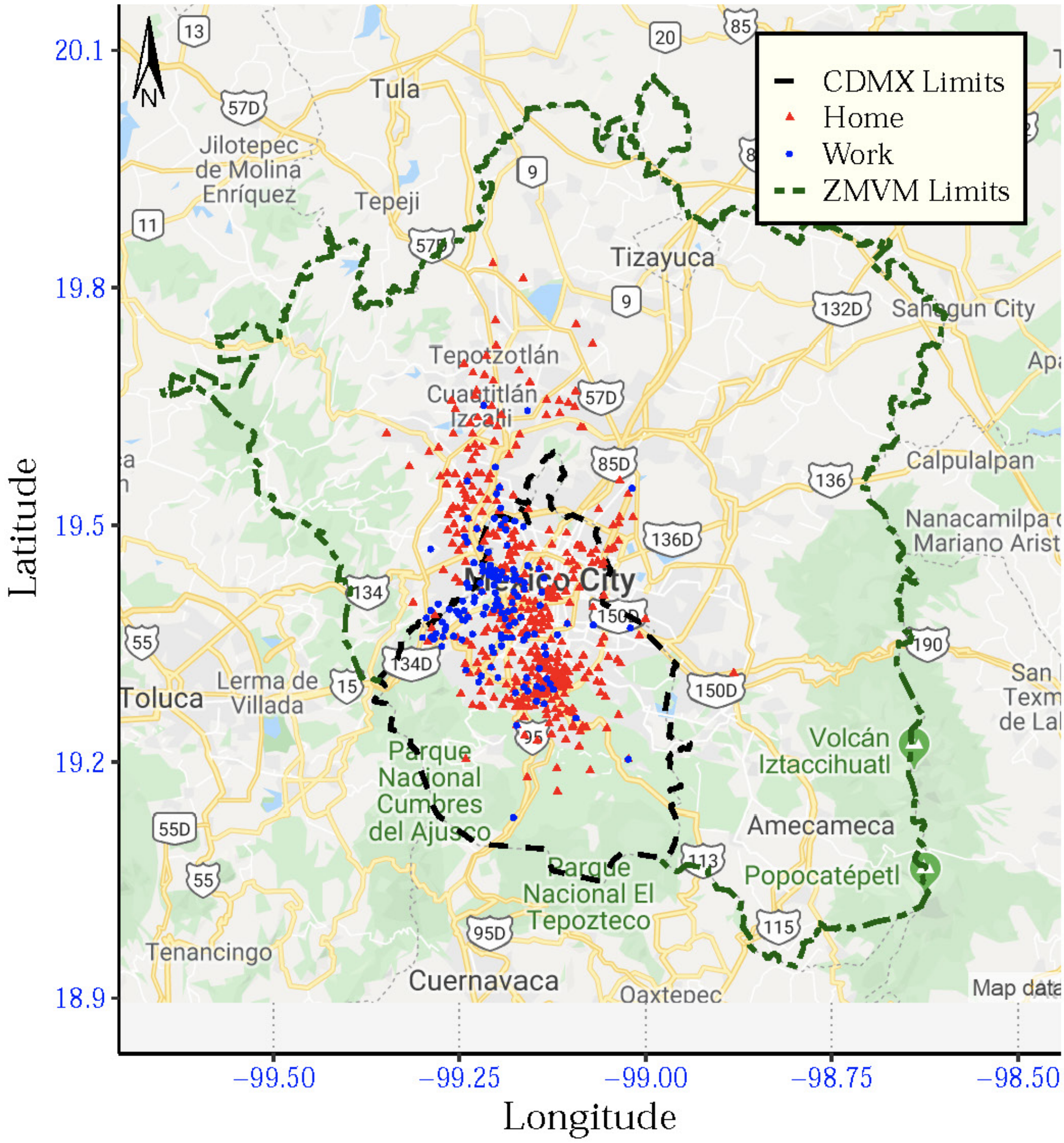

Figure 4.21 – Home and Work Locations, Background: Google Maps, (*own illustration*)

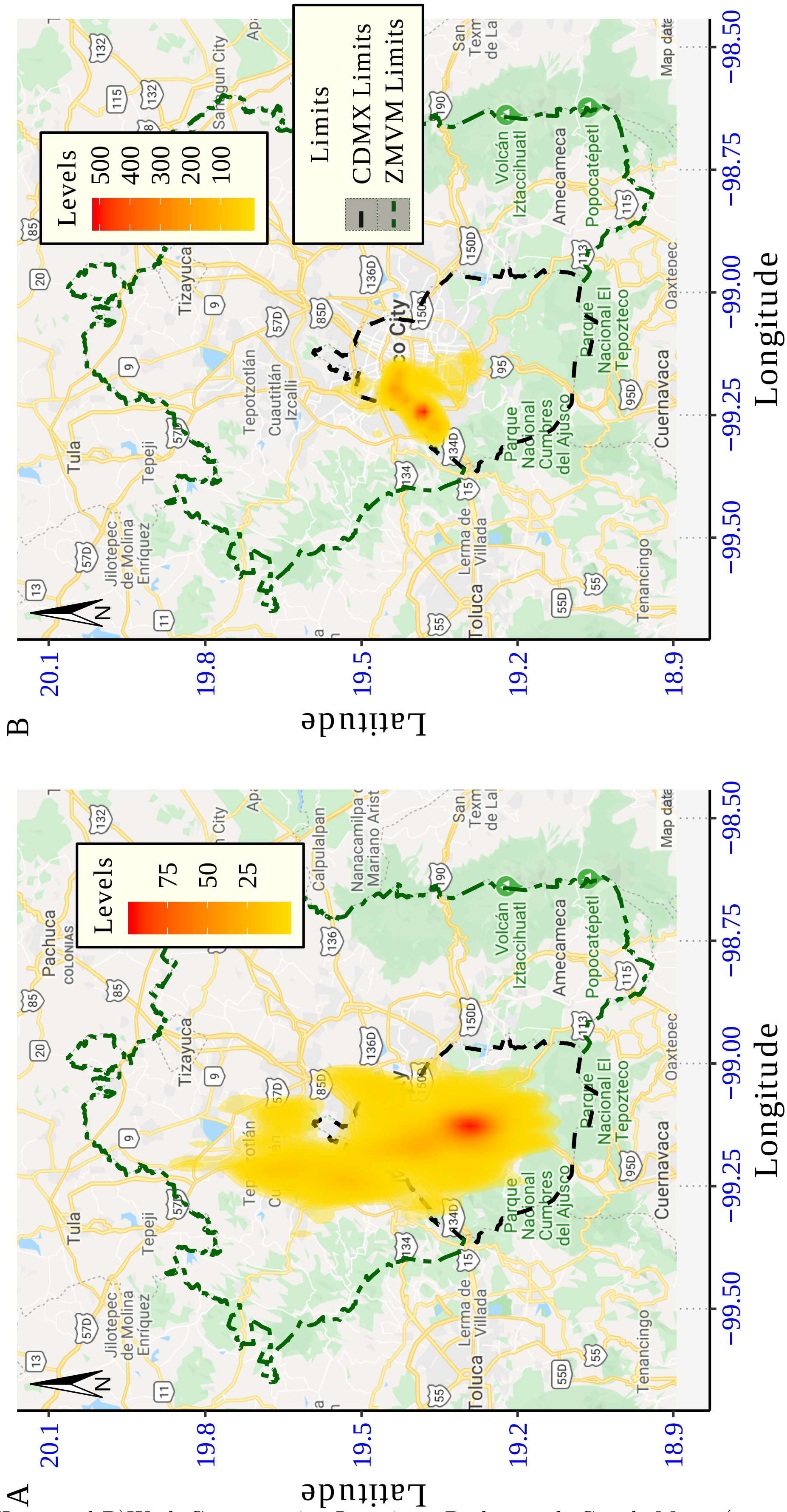

Figure 4.22 – A)Home and B)Work Concentration Locations, Background: Google Maps, (*own illustration*)

Text Analysis

The last question in the survey was an optional question, where the interviewee could write a comment on Jetty service. Only 864 users provided comments. Text and sentimental analysis were performed on the comments to investigate users' impressions about the service.

Word Cloud

The first technique used to evaluate the responses was to create a word cloud or to count the most frequently mentioned words in the different users' responses. The responses were translated from Spanish to English using Google[10] translate and the translated answers were reviewed and manually corrected when needed. Stop words were removed from the translated responses. Stop words, such as articles and prepositions; removal is a standard norm to perform during the text mining process to improve the outcome of the analysis [208, 181]. After completing the text cleaning process, the top most frequent words are plotted, and the frequency of response weighs each word size in the text. The most repeated words are Service, Routes, Time, Jetty, and Santa Fe. Figure-4.23 Shows the word cloud for the top hundred repeated words, where the word excellent is from the top five mentioned words.

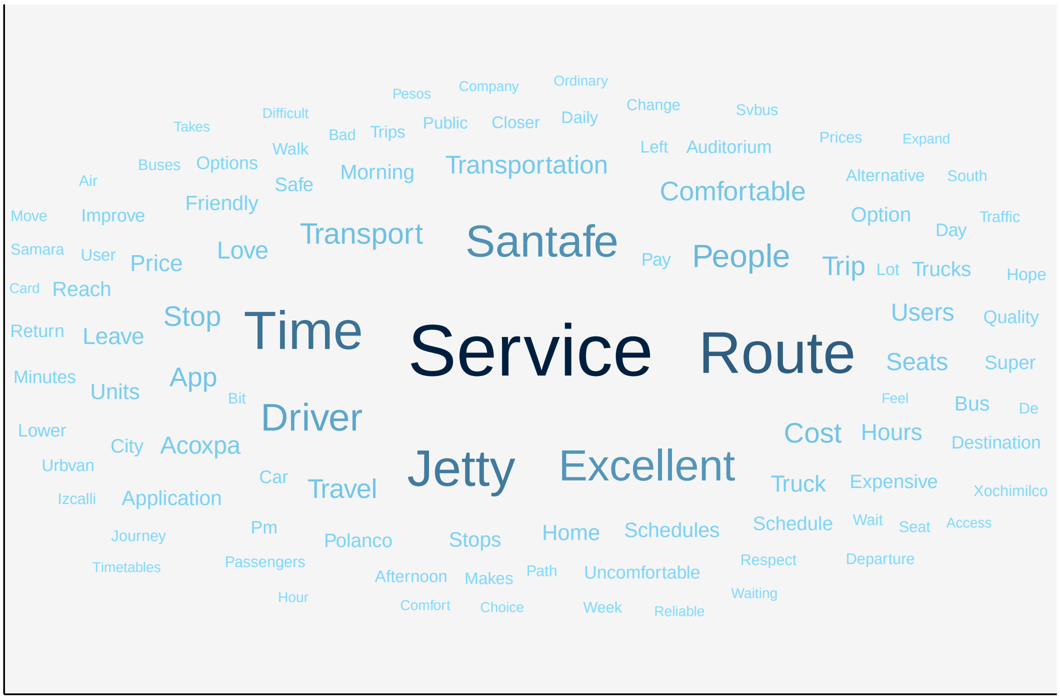

Figure 4.23 – Survey Comments Word Cloud (*own illustration*)

Words Relations

The next investigated item was the relationship between the words of the responses and how they are tied together, or in other words, what words are generally mentioned together or in consecutive order. A nigram of two words, commonly named bigram, is created using the (*tidytext*) package [232] in R [199], and their reparation is counted. Only words repeated more than two times are kept. A network graph for the correlated words was created using (*ggraph*) package [194]; the edges-links color was weighted by the

[10] https://translate.google.com/

frequency of the word correlation counts. The darker the edge, the larger the number of times the correlated words were mentioned.

The word relations revealed some interesting insights about the users' impression on Jetty service. The most related word pair is excellent service, and the word service is connected to convenience. Also, there is a request to expand the routes, where the two words expand, and routes are associated. Complaints were also noticed where the words (uncomfortable seats, and bad experience) are associated, refer to Figure-4.24.

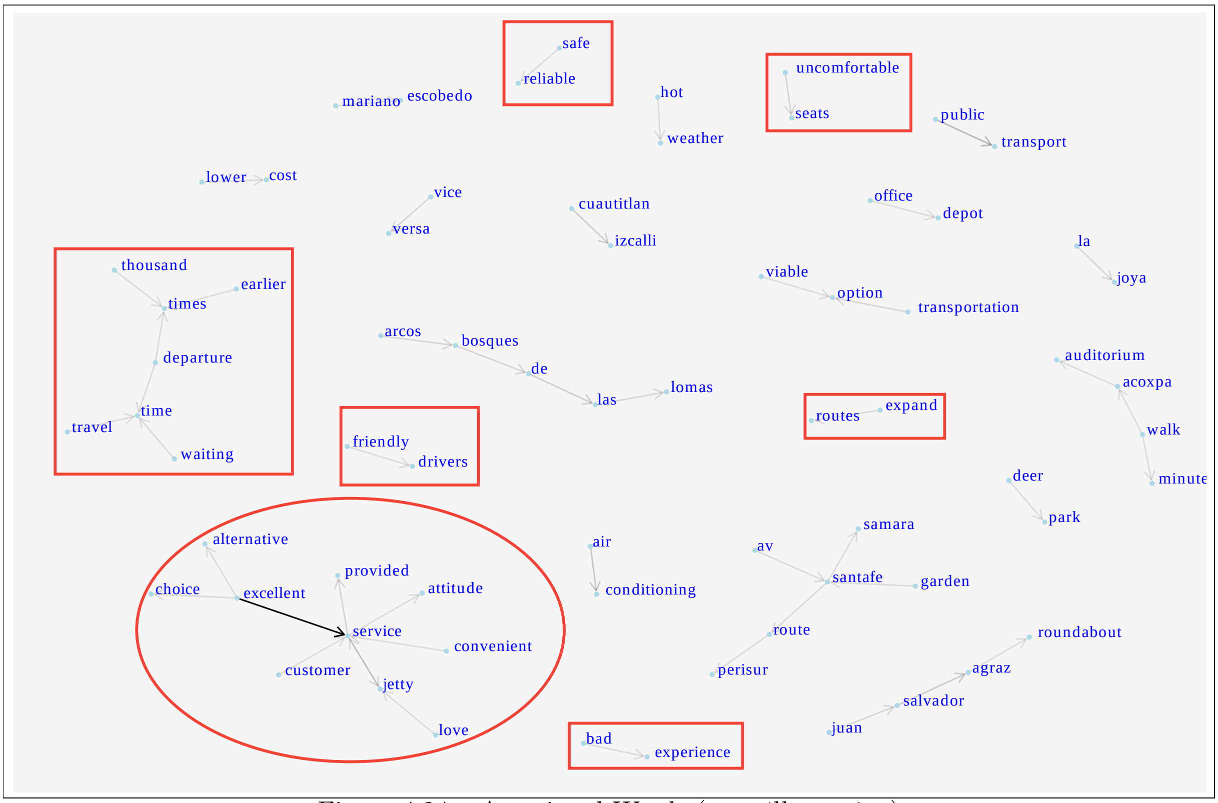

Figure 4.24 – Associated Words (*own illustration*)

Sentiment Analysis

The sentiment analysis was the last technique used to evaluate the user's opinion about Jetty using the survey comments. Sentiment analysis is widely adapted to investigate people's opinions expressed in text data. Several techniques are used to perform sentiment analysis. Two of the most commonly used techniques are machine learning algorithms, and lexicon-based analysis (word lists analysis) [144].

Under the scope of this thesis, the lexicon-based method was used to perform the sentiment analysis, and two word-lists were used. The main reasons to use word lists are I) they are simple to use; II) their predicted results are proven to be accurate when implemented on several data sets [204]. Two word-lists were used to verify and confirm the results obtained from the analysis, as the two lists use two different techniques to calculate the sentiment score.

The first used word list was the Finn Årup Nielsen Word List (AFINN) [182]. This word list assigns a score for each word that ranges between -5 and 5. Each response was disassembled to separate words, and each word was assigned a score based on the list. Some words, such as names and articles, do not have scores according to the list. The words with no-score were removed, and the sum of the score per each user was calculated.

Figure-4.25 shows the score distribution per gender. Seventy percent (70%) of the scores are positive number, with average female response score (+2.04) higher than the average male response scores (1.53%).

The second word-list was the NRC Word-Emotion Association Lexicon (EmoLex). This dictionary assigns Plutchik's eight basic emotions (anger, anticipation, disgust, fear, joy, sadness, surprise, and trust) in addition to the polar classification of positive and negative to each word [170]. Only positive and negative word-classification was considered, and it was assigned to the dismantled responses words, respectively. Words with no score were removed. Next, the positive and negative words were counted per each user, and the difference between the positive word count and the negative word count was assigned as the score per user.

Figure-4.26 shows the Emolex list score distribution per gender. Similar to the AFINN list, 70% of the Emolex scores are positive scores with female users having higher average score (+1.19) compared to the male users(+1.02).

It is to be noted that both lists' results have the same overall trend. However, in terms of absolute values, there are some differences. AFINN scores are higher than Emolex scores, but mainly because of the different calculation methods, and the number of words in each list. Similar differences were noticed, when both lists were used to perform sentiment analysis on the same data set [232, 204].

Table 4.12 – Survey Comments Sentiment Score Summary Statistics

User Group	Mean	Std.Dev	Min	$1^{st}Q.$	Median	$3^{rd}Q.$	Max	Word List
All Users	3.09	3.85	-8	1	3	5	39	AFINN
Female	3.33	3.95	-6	1	3	5	22	AFINN
Male	2.81	3.71	-8	1	3	4	39	AFINN
All Users	1.28	1.82	-4	0	1	2	10	Emolex
Female	1.37	1.85	-4	1	1	2	9	Emolex
Male	1.16	1.79	-3	0	1	2	10	Emolex

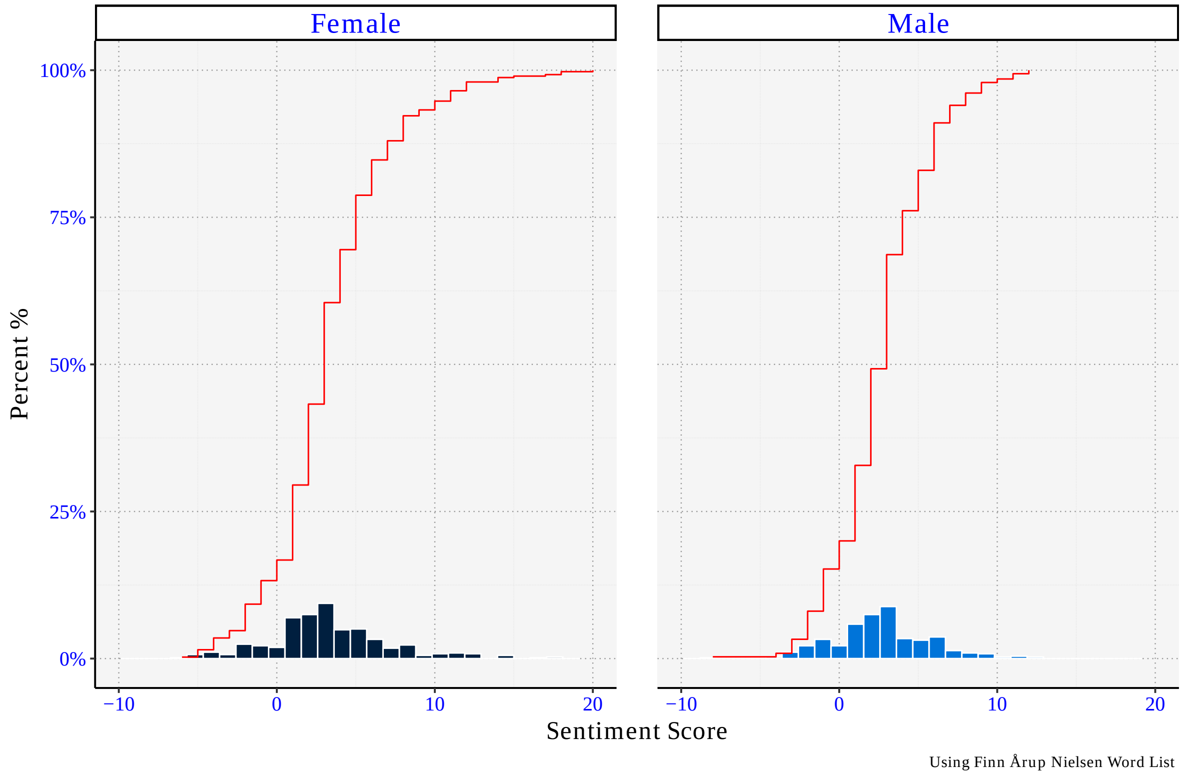

Figure 4.25 – Sentiment Analysis Using Finn Årup Nielsen Word List, (*own illustration*)

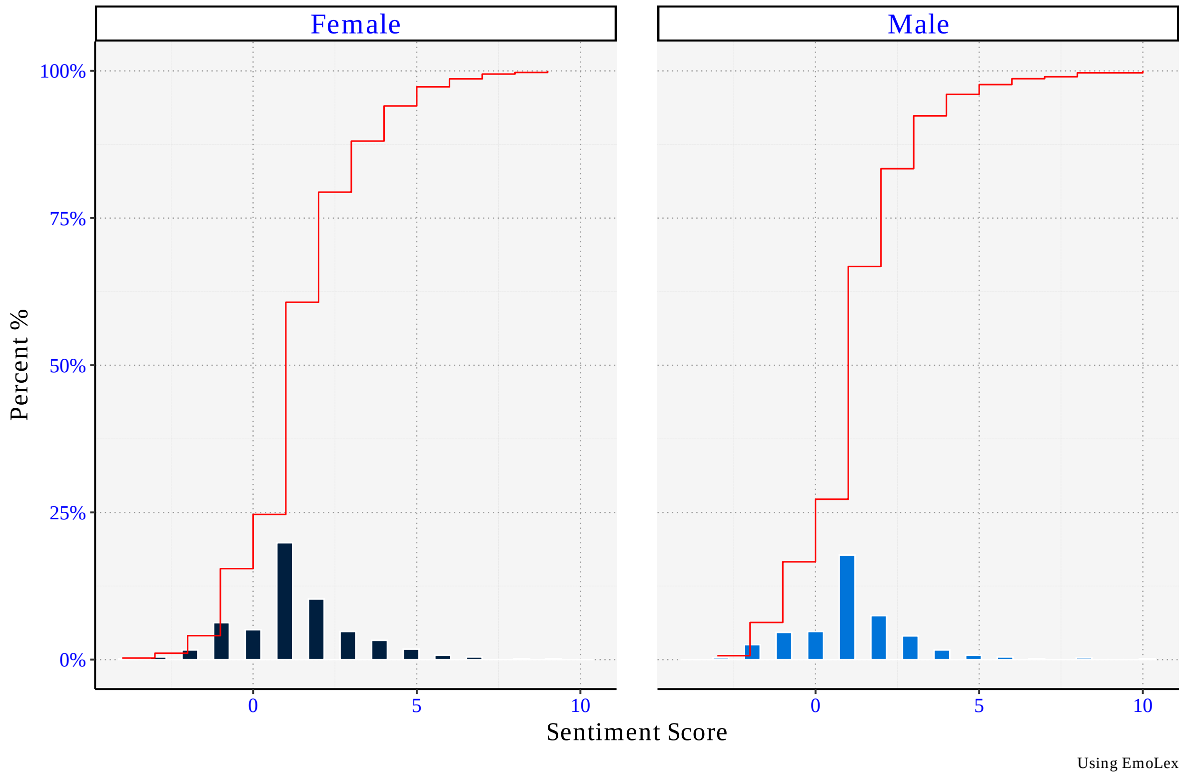

Figure 4.26 – Sentiment Analysis Using EmoLex, (*own illustration*)

4.3.2. Jetty Database

For the second source of information, the Jetty database, Jetty operators provided the use date and trip details of the survey participants for the last seven months before the survey launching date. The received user data included trip details for 96,317 trips performed by 2196 users from the total survey respondents. Also, the demand and occupancy data for the month of may were provided by Jetty for the different vehicle categories.

Use Frequency

The first item to analyze was the Jetty use frequency in the last seven months per user. A comparison between the total sample of (2169) and the partial sample containing (1118) users was made to check the adequacy of the partial sample distribution, as the latter one will be used for the modeling process, as stated earlier. There was no significant difference between the total sample and the partial sample in terms of distribution. Figure-4.27 shows the frequency of use histogram. Sub-Figures A and B show the total sample histogram and sub-Figures C and D show the partial sample histogram. Table-4.13 shows the summary of Jetty use.

On average, users use Jetty less than two times a week, with the partial sample having a slightly higher average.

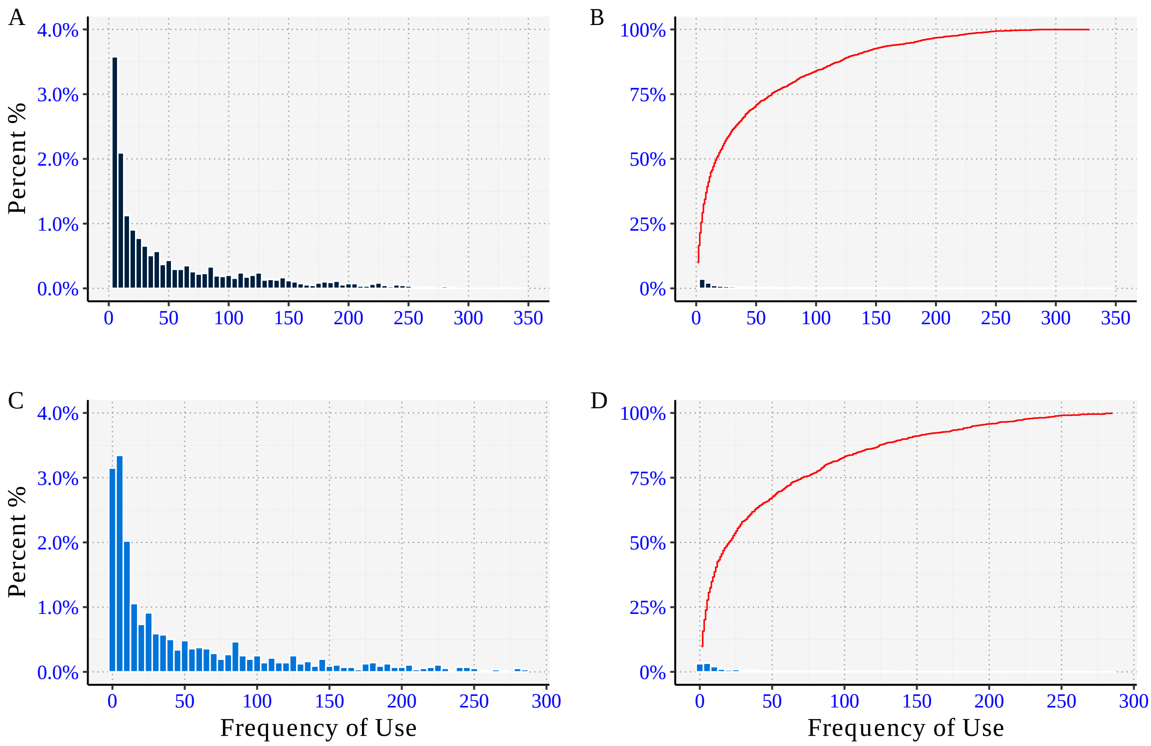

Figure 4.27 – Number of Jetty Trips per User Distribution

Table 4.13 – Number of Jetty Trips per User Summary Statistics

Frequency of Use	Mean	Std.Dev	Min	$1^{st}Q.$	Median	$3^{rd}Q.$	Max	N
Total Sample	44.4	58.4	1	4	17	63	328	2169
Partial Sample	48.5	61.8	1	5	20	71	285	1118

In order to compare the Jetty use rate with other modes reported by the interviewee in the survey, Figure-4.18, the frequency of the Jetty use rate was grouped to four categories similar to the levels used in the survey. However, relative use frequency was calculated to compensate for the different users' starting Jetty use date. For each user, the relative use frequency was calculated as the number of Jetty trips in the period between the first and last rides in the received database. The use rate categories are:

- Less Than Once a Month
- 1 - 3 Times per Month
- 1 - 3 Times per Week
- More than 3 Times per Week

Figure-4.28 shows the distribution of Jetty use frequency for the partial sample per gender. The use frequency is balanced between the gender, with females slightly using the service more frequently than men on the high use rates.

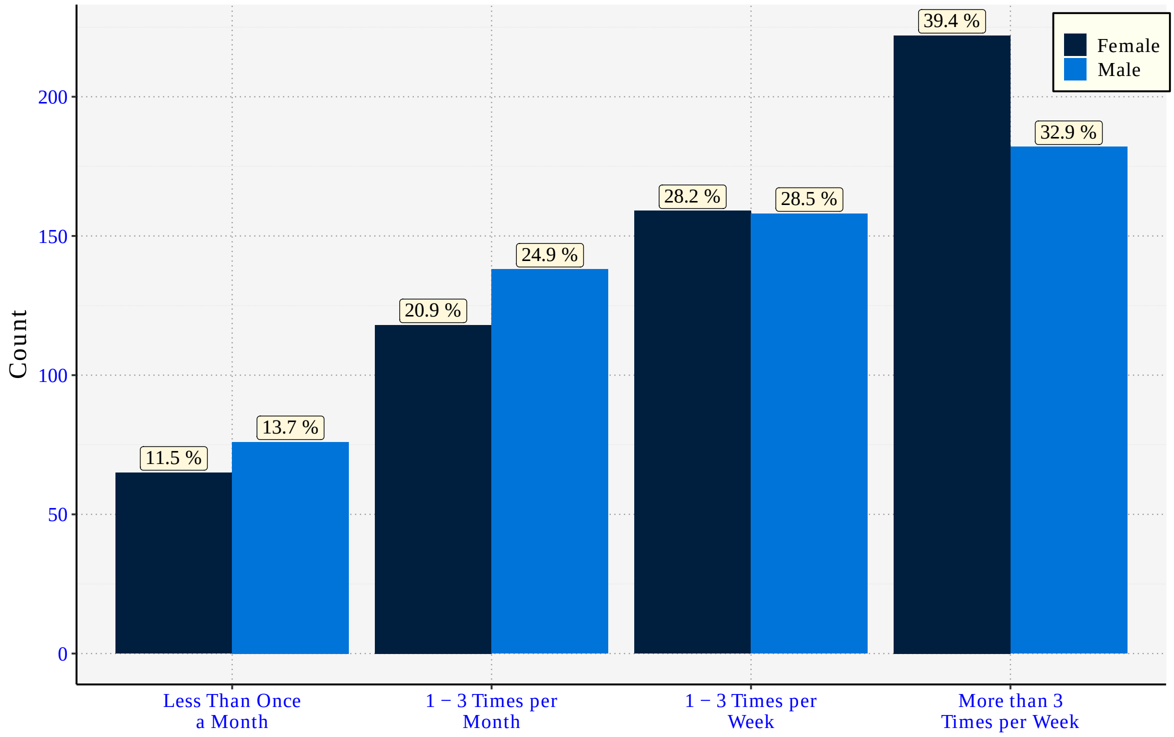

Figure 4.28 – Relative Jetty Use Frequency Distribution per Gender

Trips Temporal Distribution

Jetty trips' departure times follow the expected travel pattern with two peaks, one in the mornings and the other in the evenings. Figure-4.29 shows that the morning peak hour is at 6:00 am, and the evening peak is between 17:00 and 19:00. There is a significant difference between the percentage of morning and evening trips, with the number of morning trips being double the number of evening trips. Jetty temporal demand follows CDMX traffic patterns, where morning peak demand is more sever than evening peak demand [190]. The increasing demand for Jetty in the mornings could be due to Jetty's saving in travel time, which is more valued by the people in the mornings. People generally value the time before reaching the workplace more than the time after leaving work due to the time constraints [191]. Table-4.14 shows the summary of the percentage of morning trips from the total Jetty use per user.

There is no significant difference between the total sample and the partial sample regarding the trip hourly and daily distribution. Figure-4.29 shows that the partial sample is following the same distribution of the total sample. Therefore, the analysis carried out below will only consider the partial sample. This sample will be used for the modeling as it has full information about the home-work locations and Jetty trips' details. There is no difference in Jetty use patterns during the weekdays, with Friday has a slightly lower use percentage compared to other days, as shown in Figure-A.1.

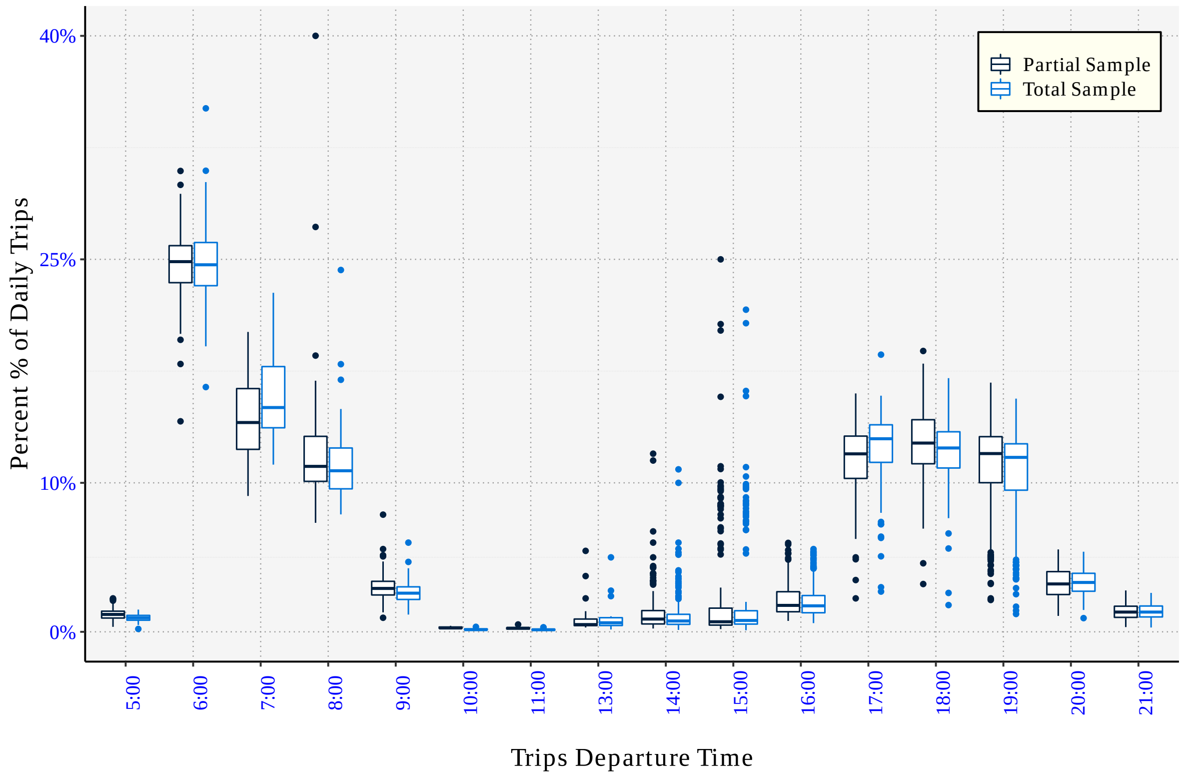

Figure 4.29 – Jetty Trips Departure Time

Table 4.14 – Percent of Morning Trips Summary Statistics

Statistic	Mean	SD	Min	$1^{st}Q.$	Median	$3^{rd}Q.$	Max	N
Pct.% of Morning Trips	49.7	34.8	0.00	15.8	51.7	78.3	100.0	1118

Partial Sample

Jetty Trips characteristics

Trip Distance and Route Distance: The average Jetty trip distance is 25.1 km, with a ±6.35 km standard deviation. The maximum trip distance is 52.4 km, and the minimum distance is 4.9 km. Every trip is a part of a route, and the average Jetty route distance is 27.8 km, with a ±7.97 km standard deviation. The maximum route distance is 55.1 km, and the minimum distance is 6.5 km. Figure-A.2 shows the route and the distance distribution. Figure-4.30 shows that 75% of the trips utilize at least 90% of total route length, which is a good indicator for the directness of the routes and their potential in time and transfer saving.

Trip in Vehicle Time: The average trip duration is 46.06 minutes with ±11.27 minute standard deviations, 75% of the trips are performed in less than 49 minutes.

Number of Booked Tickets: One person can book more than one ticket per trip or can use promotion codes from previous trips. The maximum number of tickets bought by one person during a trip was four tickets, with 97% of the passengers booking one ticket on average, and 2.3% buying two tickets.

Vehicle Capacity: Jetty offers trips in four-vehicle sizes. 1) The taxi, which is a three-seat car, 2) Caddy, which is a six-seat minivan, 3) The van, which capacity varies from 13 to 19 seats. 4) The bus where capacity varies from 30 - 45 seats. More than 50% of the trips are performed in buses, and this could be because buses are relatively cheaper than other vehicles.

Figure-4.31 shows the distribution of the trips done by the users of the partial sample per each vehicle size, (1118 users performed 54,174 trips). The shown percentage is the ration of the trips in the assigned vehicle capacity to the total number of trips. It is worth mentioning that Jetty starts the operations for the new routes in small-sized vehicles, increasing the vehicle sizes according to the demand increase.

Trip Fare: The average trip fare is 74.7 MXN for Caddy, 58.8 MXN for Taxi, 62.8 MXN for Van, and 40.2 MXN for the bus. The Fares are inversely proportional to the vehicle's capacity.

Table 4.15 – Trip Characteristic Summary Statistics

Variables (Unit)	Mean	SD	Min.	1^{st}Q.	Median	3^{rd}Q.	Max.
Trip Distance, (Km)	25.10	6.35	4.87	23.91	24.23	26.65	52.37
Route Distance, (Km)	27.80	7.97	6.50	24.20	26.50	26.60	55.12
(Trip Dist./ Route Dist.) (%)	91.56	11.6	17.33	91.08	93.69	100.00	100.00
Duration, (Minutes)	46.06	11.27	9.00	44.00	45.00	49.00	97.00
Number of Ticket	1.02	0.20	0.00	1.00	1.00	1.00	4.00
Capacity, (Pax*)	34.36	13.67	3.00	17.00	42.00	45.00	45.00
Fare, (MXN)	47.40	18.49	0.00	38.00	40.00	59.00	267.00

*Pax = Passenger, N =1118

To further investigate the individual use pattern and the preference for the different Jetty vehicle types, the percentage of each user's trips in a specific vehicle type was calculated, and Figure-4.32 shows on the X-axis the percent of trips in a specific mode for each user. The figure's Y-axis is the percentage of users in reference to the total number of users. Figure-4.32 presents an interesting fact that almost 82% of all users use only one type of vehicle for their travel, 30% of those users use only vans, and the other 52% use buses and these users never used any other type of services.

Jetty Demand and Occupancy

The received demand and occupancy data for May show that for buses, 40% of the vehicle's capacity is utilized, and the rate increases in other vehicles to 60% of its capacity. The indicated capacity is the number of seats in the different vehicle types. Also, Jetty demand is almost stable during May for the different vehicle types with a slight increase in the middle of the weekdays and a slight decrease on Fridays, as shown in Figure-4.33.

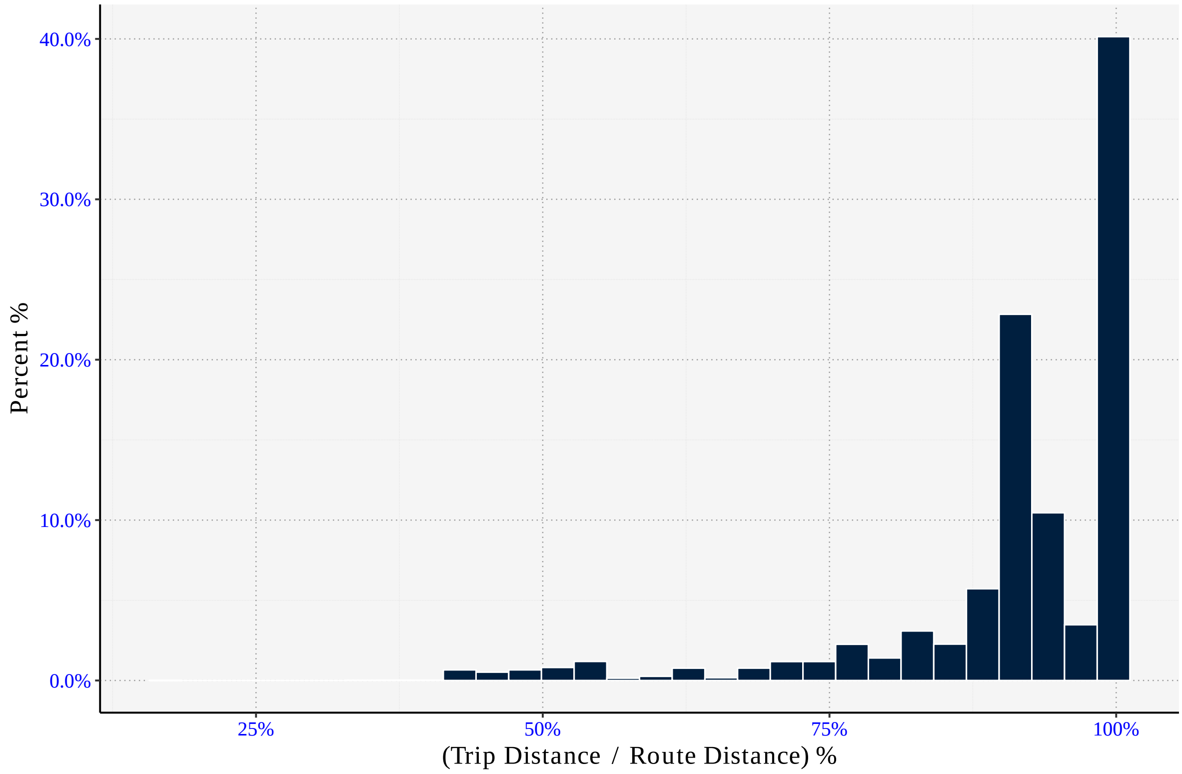

Figure 4.30 – Ratio of Jetty Trip Distance to Jetty Total Route Distance

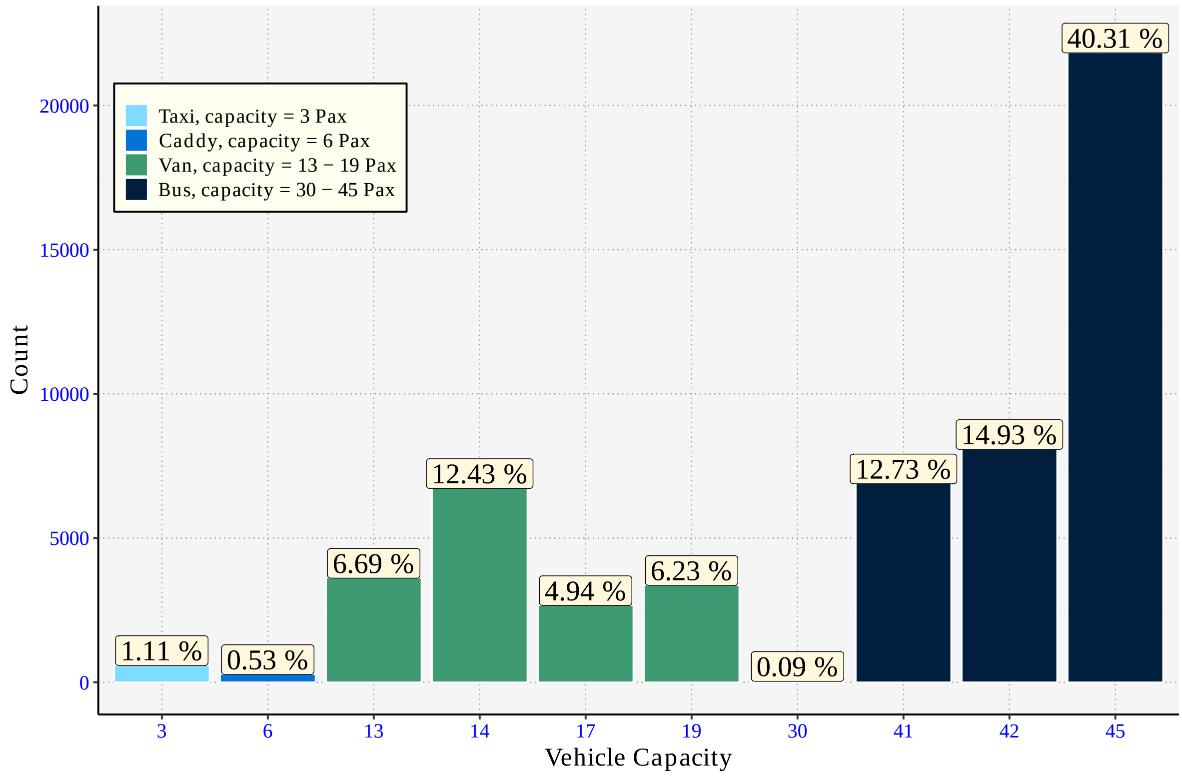

Figure 4.31 – Percentage of Total Jetty Trips per Vehicle Type

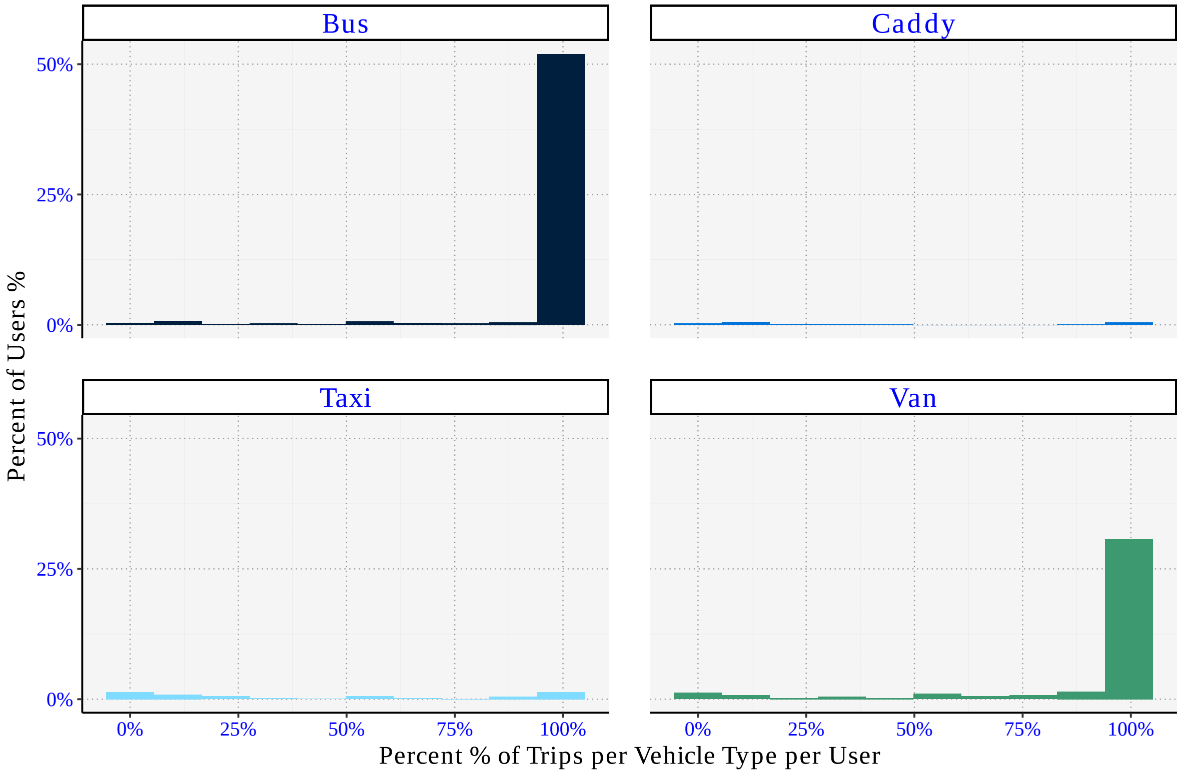

Figure 4.32 – Percentage of User's Trips per Vehicle Type

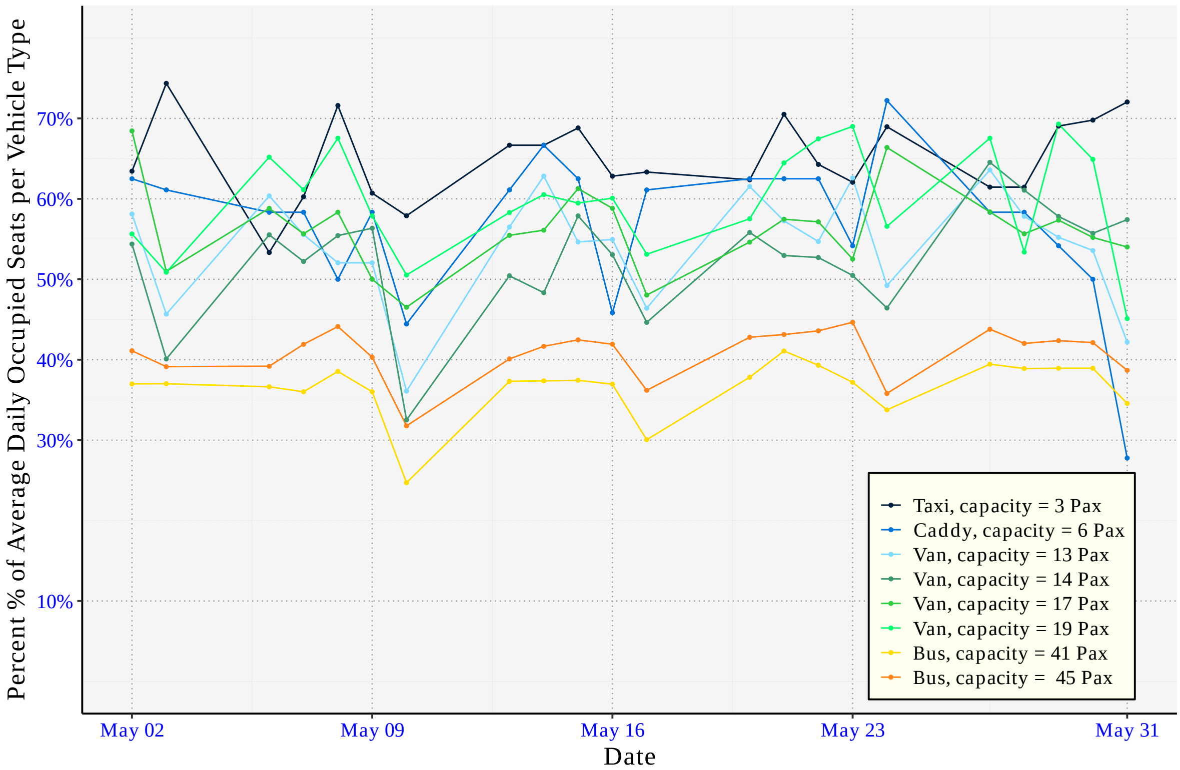

Figure 4.33 – Jetty Vehicles Occupancy Rate, May 2019

4.3.3. GTFS Files

The third source of information is the GTFS files. GTFS files were included in the analysis to study the synergy between users' home locations, available public transport modes, and the use of Jetty.

GTFS files were retrieved from the open mobility data platform[11]. The Nearest neighbor search algorithm was implemented using *nngeo* package [73] in R [199] to identify the closest station of each of the public transport modes available in the GTFS files to each home location. Afterward, the headway in the nearest station for each mode is assigned for the corresponding users, and the direct distance to the nearest station is calculated.

Table-4.16 shows the summary statistics for the headway and distance for the different modes for the partial sample users. The analysis of the GTFS files conforms to the properties of the public transport network in CDMX. For example, the average mean distance to the nearest suburban train is significantly larger than the other modes since the suburban train line is limited to the north of the city. Moreover, the broad coverage of the RTP network (Figure-4.2) is evident, where the average access distance is 1.94 km, which is the smallest distance compared to all other modes.

Table 4.16 – GTFS Files Data Summary Statistics

variable	Abbreviation	Mean	SD	Min	Max	Unit
Distance to the nearest Bus(MCDX) station	dist_km_Bus_MCDX	9.88	6.05	0.02	37.83	Km
Distance to the nearest Subway station	dist_km_METRO	5.07	4.86	0.09	36.05	Km
Distance to the nearest Metrobus station	dist_km_Metrobus	5.17	4.36	0.12	33.55	Km
Distance to the nearest Night Bus station	dist_km_Night Bus	8.91	5.67	0.05	38.40	Km
Distance to the nearest RTP station	dist_km_RTP	1.94	3.80	0.02	29.21	Km
Distance to the nearest RTP-ESP station	dist_km_RTP_ESP	3.26	4.12	0.10	31.52	Km
Distance to the nearest Suburban Train station	dist_km_Suburban Train	13.02	6.74	0.48	32.12	Km
Distance to the nearest Trolleybus station	dist_km_Trolleybus	3.65	4.85	0.01	35.35	Km
Headway at the nearest CDMX Bus station	Hw_m_Bus_MCDX	12.11	2.46	4.20	15.60	Min.
Headway at the nearest Subway Station	Hw_m_METRO	3.73	1.32	2.00	5.83	Min.
Headway at the nearest Metrobus station	Hw_m_Metrobus	5.36	3.08	3.00	30.00	Min.
Headway at the nearest Light Rail station	Hw_m_Night Bus	9.24	3.59	7.00	15.00	Min.
Headway at the nearest RTP station	Hw_m_RTP	31.86	16.70	4.00	85.00	Min.
Headway at the nearest RTP-ESP station	Hw_m_RTP_ESP	5.00	0.00	5.00	5.00	Min.
Headway at the nearest Suburban Train station	Hw_m_Suburban Train	10.00	0.00	10.00	10.00	Min.
Headway at the nearest Trolleybus station	Hw_m_Trolleybus	4.05	0.73	2.00	6.00	Min.

[11] transitfeeds.com

4.4. Data Preparation

The next step after performing the descriptive data analysis was to prepare the data for the modeling process. This step was done by checking for zero and near-zero variance variables, checking for multicollinearity between the different variables, scale the numerical variables to zero mean, one-hot encoding the categorical variables, and performing explanatory factor analysis (EFA).

4.4.1. Zero and Near-Zero Variance

The headway at Suburban Train station and RTP-ESP station has a zero variance; both variables were excluded from the data used for modeling.

4.4.2. Collinearity Detection

Collinearity detection was done on each group of variables separately.

Sociodemographic Variables

The first group contains the Sociodemographic variables that can be treated as ordinal categorical variables: age, household size, personal income, household income, number of cars in the household, and education level. The variables are ordered in ascending order, and the polychoric correlations, using *polycor* package [88], between the variables, were computed.

Figure-4.34 shows that there is a high correlation between personal income and household income, which is intuitive. The modeling strategy, discussed in the next chapter, was built on the individual level; the personal income variable was used for the modeling process, and the household income was omitted from the model. Also, there is a weak positive correlation between the number of cars in the household and household income. But since the correlation is weak, all of the three variables will be used in the model building process.

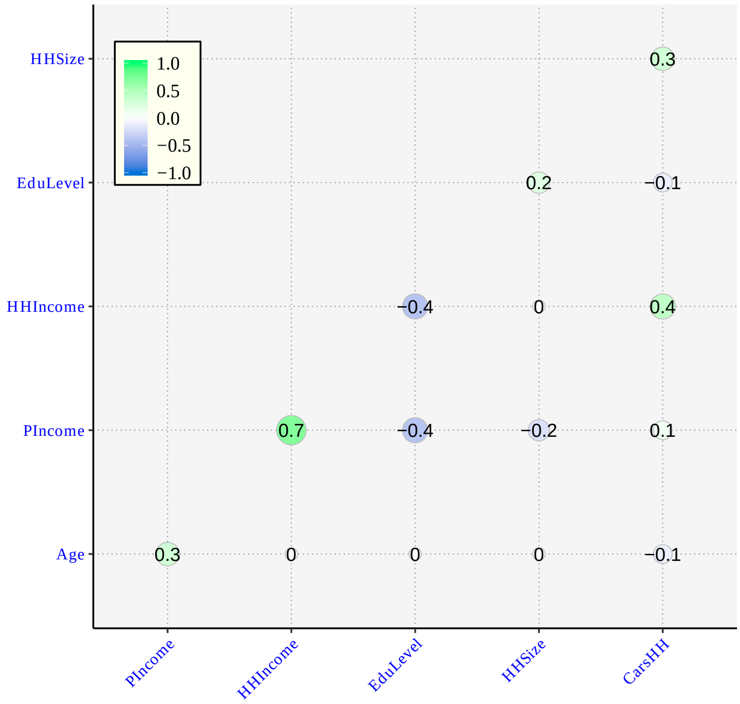

Figure 4.34 – Sociodemographic Variables Correlation Coefficients

Jetty Trip Related Variables

The second group of variables is the Jetty trip related variables: the most used vehicle type, vehicle average capacity, Jetty mean access and egress distances, average fare, and average Jetty trip distance. There is a high negative correlation between the vehicle used and average vehicle capacity and the average fare, as shown in Figure-4.35. The average fare was removed from the modeling data.

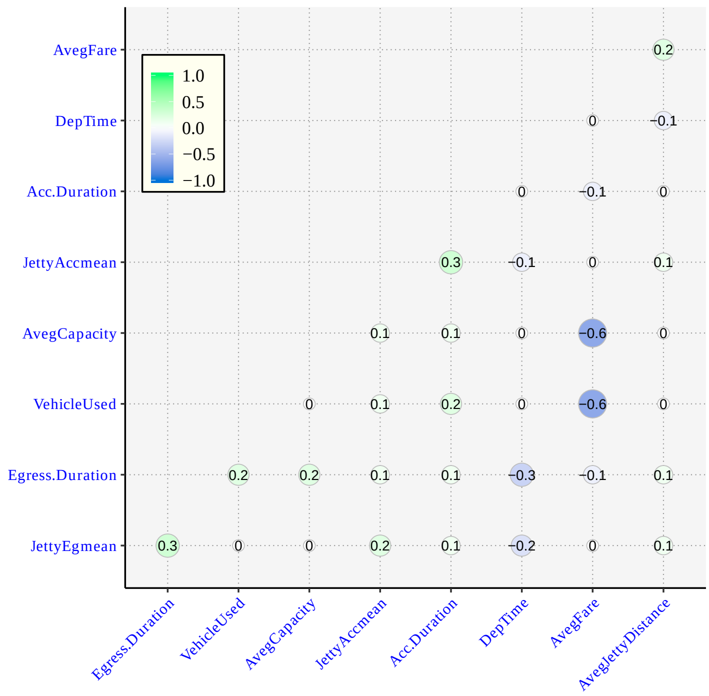

Figure 4.35 – Jetty Use Related Variables Correlation Coefficients

GTFS Data

The third group of variables is the distance to the nearest public transit station in meters and their corresponding headway in seconds retrieved from GTFS files. Figure- 4.36 shows the correlation coefficient between the different variables of this group. The distance to the nearest Trolley bus station is highly correlated with the distances to the nearest RTP-ESp, RTP, Night buses, Metro, Metrobus, and MCDX-buses stations. Only distances to the nearest Metro station and trolleybus station will be kept in the model.

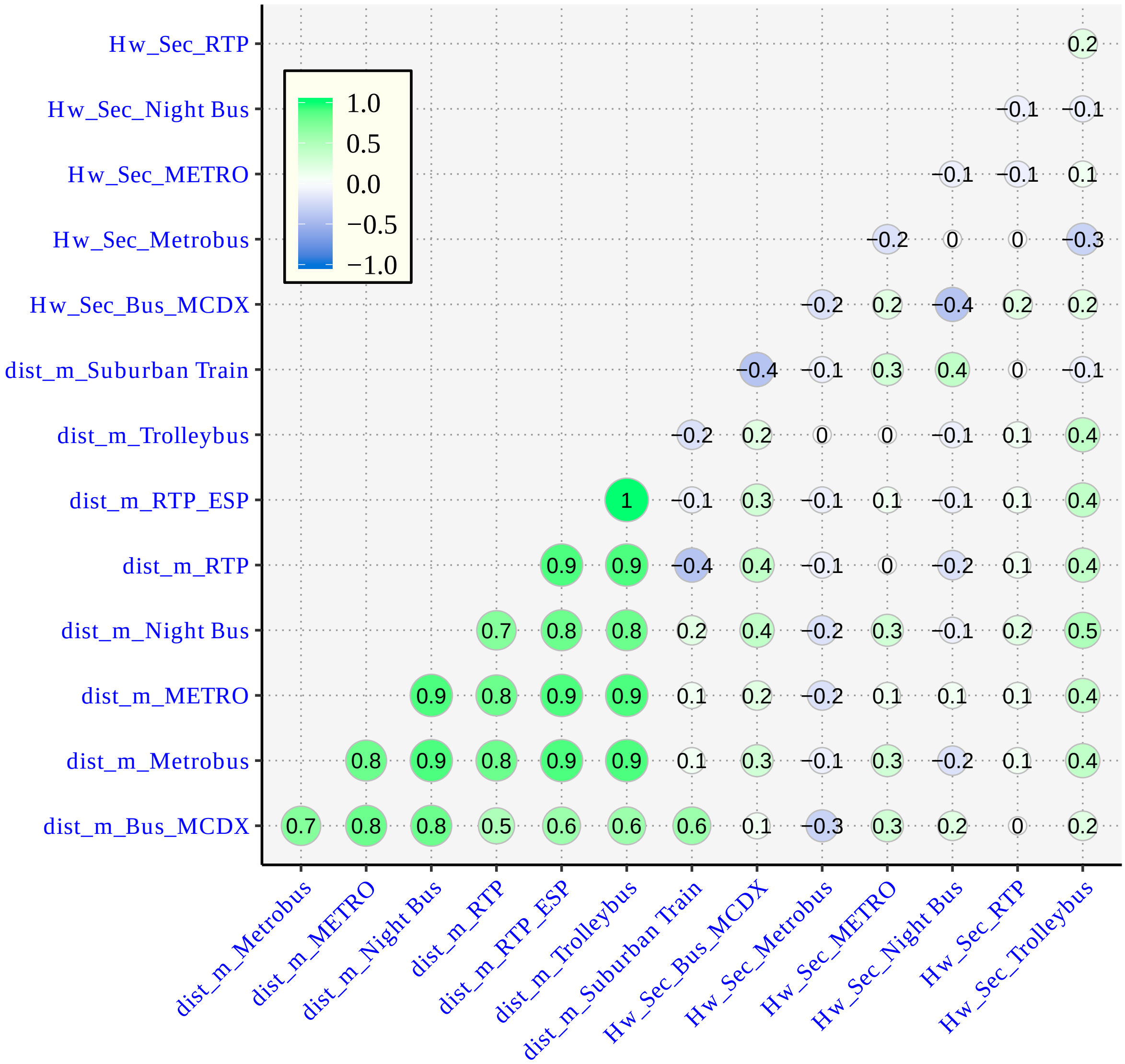

Figure 4.36 – GTFS Files Data Correlation Coefficients

4.4.3. Numerical Variables Scaling

All the numerical variables were scaled to have a mean value of zero. Scaling was done in two steps. The first step was subtracting the mean of each column from all the corresponding observations. The next step was dividing the variables by their standard deviation. The core function (*Scale*) in R [199] was used for this transformation.

4.4.4. One-Hot Encoding

The next step was to perform the one-hot encoding for the categorical variable or to transform them into binary variables. Function (*one_hot*) from package *mltools* [100] in R [199] was used for this task. This step was performed to make the categorical variables usable by the different packages used in the modeling process.

4.4.5. Exploratory Factor Analysis

EFA was performed on the frequency of use Likert data to infer the latent construct between the different variables. Before running the EFA, the initial hypotheses were built. The EFA results expected to reveal three factors that indicate the use of public transportation (PUT) and paratransit as the first factor, the use of taxi and private car as the second factor, and the third factor is the use of micro-mobility. The initial factors number was estimated using a scree test from (*psych*) [203] package in R [199], and considering the initial Hypotheses. An iterative process obtained the EFA by removing the variables that caused noise during the EFA estimation process, and by trying different factor numbers. A polychoric correlation was used to calculate the EFA as it was preferred over the commonly used Pearson correlation in the case of ordered nominal data [115]. Starting from 20 variables, only factors that explain at least ten percent of the data variability were kept. Twelve variables and two factors capturing 39% of the data variance were estimated.

The estimated EFA revealed two factors represent two user groups with two distinct travel patterns. These factors can be explained as I)The first factor is the frequent public transportation and paratransit users (referred to as PUT-Users). II) The second factor is micro-mobility and shared-micro-mobility users (refereed to MA-users); Table-4.17 shows the EFA analysis results.

Table 4.17 – EFA Results

	Factor 1	**Factor 2**
Frequency of Metro use	0.64	
Frequency of Metrobus use	0.42	
Frequency of Light-Rail use	0.58	
Frequency of Trolleybus use	0.60	
Frequency of RTP use	0.60	
Frequency of Bus use	0.50	
Frequency of Minibus use	0.66	
Frequency of Combi use	0.58	
Frequency of Bicycle use		0.52
Frequency of Shared-Bicycle use		0.87
Frequency of Shared-Scooter use		0.70
Frequency of Walk		0.45
	Put-Users	**MM-Users**
SS loadings	2.79	1.94
Proportion Var	0.23	0.16
Cumulative Var	0.23	0.39

Chapter 5

Modeling Process

This chapter shows the models' estimation process, and the estimated models' coefficient interpretation following the research methodology shown in Chapter 3. Four choice models and two ICLV models were developed to answer the different research questions.

5.1. Shift to Jetty

The first developed model was the model that investigates which factors affect the user's choice to shift from the different modes to Jetty. This model answers the third research question. A choice model and an HCM were estimated to investigate the questioned factors, and their results were compared.

Dependent Variable

The first step in developing the model was to define the dependent variable. For the subject model, the answer to the question of which modes would have been used to replace the last Jetty trip was used as the dependent variable. Modes were grouped into four groups that have common operational and usage attributes:

- Group **A** contains: Motorcycle, Car as a driver or passenger (**Private Modes**)
- Group **B** contains: Ride hailing, and E-ride hailing (**Taxi**)
- Group **C** contains: Shared taxi, Minibus, Combi, and Camion (**Paratransit**)
- Group **D** contains: Metro, Metrobus, Ecobus, and Suburban train (**PUT**)

The reported modes to replace Jetty trip, 13 modes, were coded to one of the respective four categories, and if two modes were in the same group, they were coded only once. Eighty-one percent (81%) of the trips were performed in one or two of the main modes categories. The rest of the trips (17.5%) were completed in three different modes noting that one of the three modes belongs to group A or B. Table-5.1 shows the trips summary details, and Table-5.2 shows the details of the trips that were done in three different mode categories.

Table 5.1 – Combined Modes Replacing Last Jetty Trip Summary Statistics

Category	Count (Pct.%)	Category	Count (Pct%)
A (Private Modes)	140 (12.5%)	A+C	20 (1.8%)
B (Taxi)	62 (5.5%)	A+D	37 (3.3%)
C (Shared taxi, Collective Services)	67 (6%)	B+C	33 (3%)
D (PUT)	96 (8.6%)	B+D	46 (4.1%)
C+D	350 (31.3%)	3 Different Modes	196 (17.5%)
A+B	57 (5.1%)	No Trip	14 (1.3%)
Total			1118 (100%)

Table 5.2 – Trips in Three Different Modes Summary

Category	Count (Pct.%)	Combination	Count (Pct.%)
A	112 (57.1%)	A+B+C	13 (6.6%)
B	129 (65.8%)	A+B+D	32 (16.3%)
C	164 (83.7%)	A+C+D	67 (34.2%)
D	183 (93.4%)	B+C+D	84 (42.9%)
Total			196 (100%)

The dependent variable was coded as a binary variable that was set to be equal to (zero) if a trip was completed in groups (C, D, and C+D), and it was set to (one) otherwise. The dependent variable defines the choice between the PUT and paratransit trips, and on the other side, the trips that were made, or partially made in private modes and taxi, which will be referred to it in the next sections as car trips.

Hybrid Choice Model

Following the procedures of the modeling process mentioned in Section-3.3 and starting from the EFA analysis results, which revealed two latent variables indicating the frequent use of public transportation and the frequent use of micro-mobility were investigated with several specifications for the structure equation part. Only one latent variable, the frequent use of PUT, was significantly different from zero. Figure-5.1 shows the full path diagram of the final model.

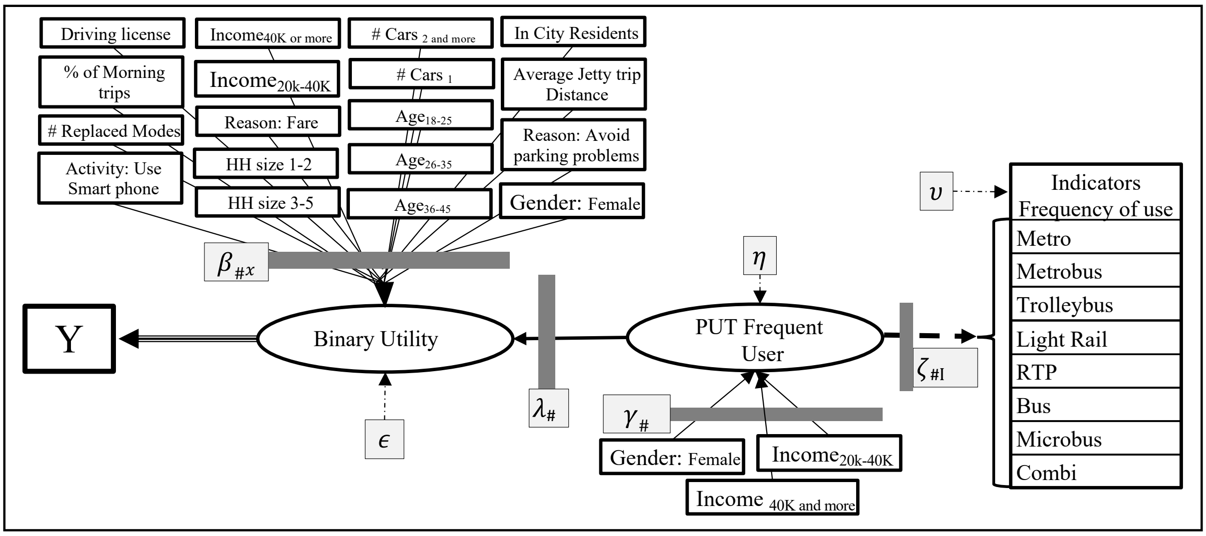

Figure 5.1 – Shift to Jetty HCM Full Path Diagram (*own illustration*)

The model in Figure-5.1 can is generally formulated as follows:

Structure Model :

$$\alpha_n = \gamma_n X + \eta_n \quad , \quad n = 1 \quad , \quad \eta \sim N(0, \textstyle\sum_\eta); \quad \textit{(1 equation)}$$

$$U = X\beta_1 + \alpha\lambda + \epsilon \quad , \quad \epsilon \sim N(0, 1); \quad \textit{(1 equation)}$$

Measurement Model :

$$I_r = \alpha\zeta_r + \upsilon_r \quad , \quad r = 1, ..., 8 \quad , \quad \upsilon \sim N(0, \textstyle\sum_\upsilon); \quad \textit{(8 equations)}$$

From the different tested specifications for the structure model part, only two personal characteristics, gender, and income were appropriate for use based on their significance, their coefficient interpretation in the context of the model results, and their impact on the overall model's goodness of fit. The latent attitude (l), for person number (n) was defined as:

$$\begin{aligned}\alpha_{n,1} =& \gamma_{1,Female} \cdot Z_{n,Female} + \gamma_{1,Income_{20K-40K}} \cdot Z_{n,Income_{20K-40K}} + \\ & \gamma_{1,Income_{40Kormore}} \cdot Z_{n,Income_{40Kormore}} + \eta_{n,1}\end{aligned}$$

The structure model consists of two parts one is deterministic, and the other is random. ($\eta_{n,l}$) is a standard normally distributed term with a mean of zero and a standard deviation of one distributed across individuals representing the random element. In the choice model, (λ) represents the impact of the latent variable (α) on the choice model.

Table-5.3 left side shows the estimation results for the HCM, and the right side shows the reduced choice model, Table-5.4 shows the estimation of the latent model part of the HCM. The HCM was estimated using classical integration estimation.

Model Coefficients Interpretation

Choice Model

The interpretation for all the variables could be explained as follows, noting that the description below for each coefficient is in reference to hold and control other variables:

Gender: The coefficient's positive sign indicates that female users are more likely to shift to Jetty from car trips compared to male users.

Age: Age group 46 and older was set as a reference category for the age variable. Parameters number (three, four, and five) can be interpreted as young people are more likely to shift to Jetty from car trips compared to older age groups.

Household Size: Household size six or more was set as the reference category. Only households with a one-to-two-person level is significantly different from zero at a 90% level of significance. Coefficient numbers six and seven can be interpreted as small households containing one-to-two-person households more likely to shift from car trips towards Jetty compared to larger-sized households.

Personal Income: Personal income group with income level equal to 20 thousand pesos or less was set as a reference category. All the income groups are statistically different from zero at least a significant level of 95%. Parameters number (eight and nine) can be interpreted as users with higher income groups are more likely to shift from car trips to Jetty compared to the lower-income groups. This finding matches with the literature regarding the profile of shared mobility users being wealthier than the average population.

Driving License Availability: Driving License Availability is a binary variable that suggests a person has a driving license. The estimated positive sign coefficient indicates that People with a driving license are more likely to shift from car trips to Jetty compared to people without driving license.

In City Residents: "In city" is a binary factor that indicates if someone resides within the limits of CDMX or not. The estimated negative coefficient shows that residents of the city are less likely to shift from car trips to Jetty.

Number of Cars in Household: The group of people with no cars available in the household was set as the reference group. All the categories of the number of vehicles in the household are statistically different from zero. Parameters number (12 and 13) can be interpreted as users with cars in the household are more likely to shift to Jetty compared to people with no cars in the households.

Number of Modes Replaced by Jetty: The estimated coefficient shows that the increase in the number of modes that were to be used to replace Jetty, the more likely the users to shift from car trips to Jetty.

Average Jetty Trips Distance: The estimated coefficient negative sign shows that the longer the average Jetty trips per user, the less likely the user to shift from car trips.

Percentage of Morning Trips: The estimated positive coefficient shows that users who make more trips using Jetty in the morning (before noon) are more likely to shift from car trips to Jetty.

Use Smart Phone: Use a smartphone is a binary variable that indicates if someone specified that he uses a smartphone during the Jetty trip. The estimated positive coefficient shows that people who use a smartphone during Jetty trips are more likely to shift from car trips to Jetty.

Fare: Fare is a binary variable that indicates the reason why people use Jetty. The estimated positive coefficient shows that people shift to Jetty from car trips due to the Jetty fare.

Avoid Parking Problems: Avoid parking problems is a binary variable that indicates the reason people use Jetty. The estimated positive coefficient shows that people shift to Jetty from car trips to avoid parking problems.

Latent Variable Model

Table-5.4 lower part shows the measurement model part of the latent variable model. The measurement model estimated positive coefficients (ζ) show that the higher the levels of the answer (the more frequent the use), the more the use of the PUT in general, which is intuitive.

Table-5.4 upper part shows the structure equation part of the latent variable model. Coefficients of the structure model (γ) need to be explained in company with the measurement model. For gender, females' negative sign coefficients have a negative impact on the latent variable compared to males; in other words, females are less frequent users for PUT. For the income, the base group is the income group with 20k or less income, and the estimated coefficient negative sign shows that high-income groups are less frequent PUT users, and the income level 40k or more have the most impact on the LV.

It is worth mentioning that the threshold between the different levels of the indicators (τ_p) are only reflecting the orders of the threshold, and this is why their estimation results are not shown.

For the choice model part, the latent variable impact on the model can be interpreted as the latent variable (frequent PUT users) that are less likely to shift from car trips to Jetty. The coefficient of the latent variable in the choice model is the second-highest coefficient, which shows the LV impact on the choice.

It is noticed that the inclusion of the attitudinal factors, LV, shrinks the magnitude of the sociodemographic variables estimated coefficients, which confirms that sociodemographic variables act as representative for latent attitudes, the same phenomenon was noticed in similar studies using similar modeling techniques [8].

Statistical Evaluation

As explained before in the methodology modeling section-3.3, and to check the adequacy of using HCM over the reduced choice model, the rho-squared-adjusted ($\rho^2_{Adjusted}$) for the choice part of the HCM and the binary model were calculated considering a degree of freedom of one extra variable for the HCM. The estimated ($\rho^2_{Adjusted}$) for the HCM is lower (0.10) than the reduced choice model (0.12), which is an indicator that the reduced choice model slightly fits the data better than the HCM. However, the extra insights given by the HCM compensates for the reduced fit as the primary use of this model is to investigate the variable affecting the shifting process, and the model will not be used in any prediction. Ben-Akiva et al. used the previous evaluation methodology for comparing the fit of the HCM and the reduced choice model [29].

The estimated Gender coefficient was not statistically different from zero in the HCM ($P - value = 0.7$); however, it was statistically significant in the reduced choice model. To further investigate the impact of removing the coefficient from the HCM an additional model without the gender coefficient was estimated and LRT between the choice part of the two HCM was performed with the following Hypothesis:

$$H_0 : \text{restricted model, (no female), is the true model}$$

$$-2[-567.621 - (-567.445)] \sim \chi_{1,0.05}$$

$$0.3516 \sim \chi_{1,0.05} = 3.84$$

The LRT result (0.35) is smaller than $\chi_{1,0.05} = 3.84$ with one degree of freedom (the difference between the two models number of variables), and (95%) confidence interval; therefore, the restricted model cannot be rejected for all the reasonable significance levels.

Table-B.1 shows the estimation results for the restricted HCM, and Table-B.2 shows the estimation of the HCM with two latent variables, frequency of PUT use, and frequency of micro-mobility use, noting that the HCM with two latent variables did not yield significant results.

Table-5.5 shows the summary of the model's significant variables against both categories of the model-dependent variable.

Table 5.3 – Shift to Jetty from Car-Based Modes HCM and Binary Logit Model Results

		HCM Model		Choice Model	
No	Variable	β (P-value)	Rob.Std. Error	β (P-value)	Rob.Std. Error
1	Intercept	-4.45 (0.00)	0.69	-3.74 (0.00)	0.59
2	Gender: Female (vs Male)	0.07 (0.71)	0.18	0.50 (0.00)	0.16
3	Age between 18 and 25 (vs 46 and older)	1.20 (0.00)	0.35	1.12 (0.00)	0.31
4	Age between 26 and 35 (vs 46 and older)	0.57 (0.04)	0.28	0.47 (0.06)	0.25
5	Age between 36 and 45 (vs 46 and older)	0.57 (0.05)	0.29	0.46 (0.07)	0.26
6	Household Size between 1-2 (vs 6 and more)	0.64 (0.07)	0.35	0.78 (0.01)	0.29
7	Household Size between 3-5 (vs 6 and more)	0.33 (0.29)	0.31	0.40 (0.13)	0.26
8	Personal Income between 20K- 40K (vs 20K or less)	0.39 (0.05)	0.20	0.71 (0.00)	0.17
9	Personal Income 40K or more (vs 20K or less)	0.76 (0.01)	0.31	1.47 (0.00)	0.28
10	Driving License Availability (vs no)	0.76 (0.00)	0.22	0.70 (0.00)	0.20
11	In City Resident (vs no)	-0.46 (0.06)	0.24	-0.56 (0.01)	0.22
12	#No of Cars in Household = 1 (vs zero cars)	0.72 (0.00)	0.26	0.91 (0.00)	0.22
13	#No of Cars in Household = 2 or more (vs zero cars)	0.64 (0.02)	0.28	1.02 (0.00)	0.24
14	#No of modes replaced by Jetty	0.55 (0.00)	0.12	0.22 (0.02)	0.09
15	Average Jetty trips distance	-0.33 (0.00)	0.11	-0.31 (0.00)	0.10
16	Pct (%) of morning trips	0.54 (0.02)	0.24	0.66 (0.00)	0.22
17	Activity: Use Smart Phone	0.32 (0.09)	0.19	0.37 (0.02)	0.16
18	Reason: Fare	0.58 (0.00)	0.18	0.55 (0.00)	0.16
19	Reason: Avoid Parking Problem	0.54 (0.04)	0.26	0.52 (0.02)	0.23
20	LV: Frequent Put User (λ)	-1.13 (0.00)	0.14	—	—
	$\rho^2_{Adjusted}$	**0.10**		**0.12**	

N = 941, P-values are reported in parentheses are based on the robust standard errors, used to control for heteroscedasticity that might exist

Table 5.4 – Shift to Jetty from Car-Based Modes Latent Variable Model

Latent Variable Model			
	Structure Model (Frequency of PUT Use)	ζ (P-value)	Rob.Std. Error
1	Gender:Female	-0.51 (0.00)	0.08
2	Personal Income between 20K-40K (vs 20K or less)	-0.52 (0.00)	0.09
3	Personal Income 40K or more (vs 20K or less)	-1.07 (0.00)	0.13
	Measurement Model (Frequency of PUT Use)	γ (P-value)	Rob.Std. Error
	Indicators		
1	Frequency of Metro use	1.40 (0.00)	0.12
2	Frequency of Metrobus use	0.71 (0.00)	0.09
3	Frequency of Light-Rail use	1.27 (0.00)	0.14
4	Frequency of Trolleybus use	1.37 (0.00)	0.20
5	Frequency of RTP use	1.34 (0.00)	0.13
6	Frequency of Bus use	0.99 (0.00)	0.10
7	Frequency of Microbus use	1.33 (0.00)	0.11
8	Frequency of Combi use	1.26 (0.00)	0.12

P-values are reported in parentheses are based on the robust standard errors, used to control for heteroscedasticity that might exist

Table 5.5 – Shift to Jetty from Car-Based Modes Significant variables Summary

	Car Trips	No-Car trips
Variable	Count (Pct.%)	Count (Pct.%)
Age		
46 or More	66 (10.91 %)	77 (15.01 %)
Between 18 and 25	87 (14.38 %)	57 (11.11 %)
Between 26 and 35	282 (46.61 %)	244 (47.56 %)
Between 36 and 45	165 (27.27 %)	131 (25.54 %)
Missing	5 (00.83 %)	4 (00.78 %)
Gender		
Female	311 (51.4 %)	253 (49.32 %)
Male	294 (48.6 %)	260 (50.68 %)
Household Size		
Between 1 and 2	189 (31.24 %)	125 (24.37 %)
Between 3 and 5	353 (58.35 %)	317 (61.79 %)
6 and More	38 (6.28 %)	56 (10.92 %)
Missing	25 (4.13 %)	15 (2.92 %)
Personal Income		
20K or less	218 (36.03 %)	281 (54.78 %)
20K - 40K	214 (35.37 %)	141 (27.49 %)
40K or More	81 (13.39 %)	25 (4.87 %)
Missing	92 (15.21 %)	66 (12.87 %)
Driving License		
No	83 (13.72 %)	138 (26.9 %)
Yes	522 (86.28 %)	375 (73.1 %)
Number of Cars in the Household		
No cars	73 (12.07 %)	137 (26.71 %)
1	288 (47.6 %)	232 (45.22 %)
2 or more	244 (40.33 %)	144 (28.07 %)
Education		
Bachelor or higher	564 (93.22 %)	442 (86.16 %)
Other	38 (6.28 %)	69 (13.45 %)
Missing	3 (0.50 %)	2 (0.39 %)
Employment		
Employ_Full time	538 (88.93 %)	465 (90.64 %)
Employ_Other	67 (11.07 %)	48 (9.36 %)
In City Residents		
No	135 (22.31 %)	90 (17.54 %)
Yes	470 (77.69 %)	423 (82.46 %)
Activity Use of Smart phone		
No	150 (24.79 %)	158 (30.8 %)
Yes	455 (75.21 %)	355 (69.2 %)
Reason Fare		
No	345 (57.02 %)	368 (71.73 %)
Yes	260 (42.98 %)	145 (28.27 %)
Reason Avoid Parking Problems		
No	521 (86.12 %)	469 (91.42 %)
Yes	84 (13.88 %)	44 (8.58 %)
	Mean± (SD)	Mean± (SD)
Percent of Morning Trip %	54 ± (33.5)	45 ± (36)
Jetty trip Distance (km)	25.3 ± (6.88)	23.9 ± (6.77)
Number of modes replaced by Jetty	2.16 ± (0.71)	2.11 ± (0.92)

N = 1118

5.2. Services Choice

The next research question to answer was investigating which factors motivate the users to choose between the different vehicle types. Jetty is available in four vehicular categories: taxi, caddy, van, and bus. The main differences between these categories are the capacity and price, were taxi being the smallest and most expensive, and the bus being the largest and the cheapest service among the four services. In some cases and after filling all the booked seats, extra users can stand in the corridor of the buses, which could stimulate the felling of using PUT.

Dependent Variable

The first step was to assign the most frequently used vehicular type to each person based on their use pattern estimated from the received Jetty database. When a user performs 50% or more of his trips in a specific vehicle type, it was assigned to him as the most frequently used vehicle. The majority of users (97%) performed their trips in buses and vans, as shown in Table-5.6. Based on the previous analysis, the factors affecting the choice between the different services will be investigated between the frequent bus users and the frequent van users, and other users will be removed. There are two main reasons behind comparing the factors impacting the choice between buses and vans. The first reason, there is no enough users in the sample that use other vehicles category. The second reason is the distinctive differences between the two services where van been smaller in size, almost half the bus capacity on average, increasing the level of service use convenience, and the van ticket is more expensive than the bus tick. A binary logit and probit models were developed to investigate the factors affecting the choice between the two services.

The dependent variable was a binary variable that was equal to (zero) when the most frequently used vehicle is a bus, and (one) when the most frequently used vehicle is a van

Table 5.6 – Most Frequently Used vehicle Type per User Summary Statistics

Vehicle Type	Capacity (Pax)	Count (Pct.%)
Bus	30 - 45	663 (59.3%)
Van	13 - 19	417 (37.3%)
Caddy	6	6 (0.5%)
Taxi	3	32 (2.9%)
Total		1118 (100%)

Model Coefficients Interpretation

Table-5.7 shows the estimated coefficient of the final model. The left side of the table shows the point estimate for the model. The right side of the table shows the 95% confidence interval (CI) limits for each coefficient estimated using bootstrapping resampling technique. The number of samples used for bootstrapping was 1000 samples. The interpretation for all the significant variables could be explained as:

Gender: The female coefficient's positive sign indicates that female users are more likely to use the van compared to male users.

Household Size: Household size is a categorical variable, where the household size six or more was set as the reference level. The two parameters (three and four) show that small size households prefer the use of the bus over the use of van compared to the larger household sizes.

Personal Income: The category of income 40k or more was set as the reference to the personal income categorical variable. The estimated coefficient (five and six) can be interpreted as users with higher incomes are more likely to use the van over the bus compared with the lower-income groups. This finding complies with the fact that buses are cheaper than vans.

Employed: The employment status is a categorical variable with the category other than full-time employment is set as the reference category. The estimated coefficient of full-time employees shows that full time employed people are more likely to choose van compared to other categories.

Access and Egress modes: The modes used to access and egress Jetty are categorical variables with two main categories. The first category is the walk and bike modes, and all the other modes were set to the other category, which was also set as the reference category for both the variables. Parameters (eight and nine) can be interpreted as users who access or egress the service by walk or bike are more likely to use the van over the bus. This could be due to the fact that the access distances to the van is on average shorter than the access distances to the bus.

Percentage of Morning Trips: The estimated coefficient shows that users who make more trips using Jetty in the morning (before noon) are more likely to use a van, this can be due to the fact that access distances to vans are shorter than buses on average.

Headway in the Nearest Metro Station: The estimated coefficient negative sign shows that metro higher headway in the nearest station to the users increase to use of buses over the vans.

Jetty Trip Distance: The estimated coefficient shows that longer Jetty trip distances increase the use of vans over buses. This finding complies with the fact that the level of convenience of using a van is higher than using a bus.

Willingness to Walk: The categorical variable of the willingness to walk to the access point is a dual-level categorical variable. The category the willingness to walk ten minutes or less was set as the reference category. The estimated coefficient shows that the people who are willing to walk a long time opt for using the bus over the van. This finding complies with the fact that access distanced to buses are on average longer than average access distances to vans. Also, this finding matches with the findings of the willingness to walk model as shown in Table-4.11

Activities: All the activities during the trip are used in the modeling process as binary variables. Only two activities are significant at a 90% level of significance, the working, and talk on the phone during the trip. The estimated coefficients show that users who work during the trip are more likely to use the van, and the people who talk on the phone during the trip are more likely to use the bus. This findings also complies with the nature of the two services.

Reasons: All the reasons to use Jetty are used in the modeling process as binary variables. Only four reasons are significantly different from zero, the booking of the seats, the ease of payment, security against theft, and fare. The estimated coefficients show that users who use Jetty for the seat-booking and the ease of payment are more likely to use the bus over the van, and users how to use Jetty for security against theft and fare are more likely to use vans over buses.

Frequency of Use The relative Jetty frequency of use is a categorical variable with four levels represent the average use rate for each user during his actual use period in the last seven months before the survey start date. The use rates are I)Less than once a month, II)one to three times a month, III)one to three times a week, and IV)More than three times a week (reference level). The estimated coefficients show that the increase in the Jetty use rate increase the use of buses over vans. This finding complies with the fact that buses are cheaper than vans.

Note: EFA results were used to investigate the impact of the latent variables (travel attitudes) on the service choice preference; however, no acceptable results were obtained trying multiple structure and choice model specifications and their combinations.

Additional Analytic

Additional analysis was performed to assess and verify the impacts of the significant variables on the choice model. Table-5.8 shows the summary statistics for the questioned parameters, and it shows that the average trip distance is almost the same for buses and vans users with the vans users have more trip distances distribution. Moreover, access and egress distance for van users are shorter. The metro headway is shorter by half a minute for the van users. Moreover, Table-5.8 shows that bus users are more willing to walk a long time to access Jetty.

Table 5.7 – Service Choice Binary Logit Model Results

No	Variable	β (P-value)	Rob.std. Error	95% CI	
				LL	UL
1	Intercept	-2.74 (0.00)	0.57	-2.88	-2.80
2	Gender: Female (vs Male)	0.36 (0.03)	0.17	0.36	0.38
3	Household Size between 1-2 (vs 6 or more)	1.14 (0.00)	0.33	1.17	1.21
4	Household Size between 3-5 (vs 6 or more)	0.71 (0.02)	0.31	0.71	0.75
5	Personal Income Less than 20K (vs more than 40K)	-0.94 (0.00)	0.29	-0.98	-0.94
6	Personal Income 20-40K (vs more than 40K)	-0.71 (0.01)	0.29	-0.75	-0.72
7	Employed Full time (vs other)	0.74 (0.01)	0.30	0.76	0.80
8	Access by Walk or Bike (vs other)	1.13 (0.00)	0.17	1.15	1.18
9	Egress by Walk or Bike (vs other)	0.65 (0.00)	0.18	0.64	0.67
10	Egress Duration	-0.34 (0.01)	0.13	-0.37	-0.35
11	Pct (%) of morning trips	0.46 (0.06)	0.25	0.46	0.49
12	Headway in the nearest Metro station	-0.34 (0.00)	0.08	-0.36	-0.35
13	Average Jetty Trip Distance	0.33 (0.00)	0.09	0.34	0.35
14	Welling to Walk More than 10 min. (Vs 10 or less)	-0.42 (0.01)	0.16	-0.44	-0.42
15	Activity Working	0.43 (0.06)	0.22	0.42	0.45
16	Activity Talk on Phone	-0.51 (0.04)	0.25	-0.56	-0.53
17	Reason Booking of Seats	-0.43 (0.02)	0.18	-0.47	-0.44
18	Reason Ease of Payment	-0.38 (0.04)	0.19	-0.41	-0.39
19	Reason Security Against Theft	0.44 (0.01)	0.18	0.44	0.47
20	Reason Fare	0.31 (0.07)	0.17	0.31	0.33
21	Use Frequency: Less than Once a Month (vs more than 3 times a week)	0.98 (0.00)	0.27	0.99	1.02
22	Use Frequency: 1-3 Times a Month (vs more than 3 times a week)	0.69 (0.00)	0.23	0.70	0.73
23	Use Frequency: 1-3 Times a Week (vs more than 3 times a week)	0.60 (0.00)	0.21	0.62	0.65
Model Diagnostics					
Number of observations		1118			
Number of excluded observations		212			
Number of estimated Parameters		21			
$\mathcal{L}(\beta_0)$		-627.99			
$\mathcal{L}(\hat{\beta})$		-483.49			
$-2[\mathcal{L}(\beta_0) - \mathcal{L}(\hat{\beta})]$		289.00			
ρ^2		0.230			
$\rho^2_{Adjusted}$		0.19			
AIC		1012.97			
BIC		1123.58			

P-values are reported in parentheses are based on the robust standard errors, used to control for heteroscedasticity that might exist

Table 5.8 – Service Choice Model Significant Variables Summary

	Bus	Van
Variable	Count (Pct.%)	Count (Pct.%)
Age		
Between 18 and 25	98(14.78%)	44(10.55%)
Between 26 and 35	280(42.23%)	220(52.76%)
Between 36 and 45	184(27.75%)	106(25.42%)
46 or More	97(14.63%)	42(10.07%)
Missing	4(0.6%)	5(1.2%)
Gender		
Female	331(49.92%)	218(52.28%)
Male	332(50.08%)	199(47.72%)
Household Size		
Between 1 and 2	148(22.32%)	149(35.73%)
Between 3 and 5	424(63.95%)	228(54.68%)
6 and More	68(10.26%)	23(5.52%)
Missing	23(3.47%)	17(4.08%)
Personal Income		
20K or less	326(49.17%)	158(37.89%)
20K - 40K	204(30.77%)	138(33.09%)
40K or More	42(6.33%)	57(13.67%)
Missing	91(13.73%)	64(15.35%)
Driving License		
No	133(20.06%)	81(19.42%)
Yes	530(79.94%)	336(80.58%)
Cars in the Household		
zero	124(18.7%)	72(17.27%)
1	305(46%)	202(48.44%)
2 or more	234(35.29%)	143(34.29%)
Education		
Bachelor or higher	579(87.33%)	389(93.29%)
Other	84(12.67%)	23(5.52%)
Missing	00(00.00%)	5(1.2%)
Employment		
Full time	579(87.33%)	387(92.81%)
Other	84(12.67%)	30(7.19%)
Access to Jetty Modes		
Walk or Bike	188(28.36%)	204(48.92%)
Other	475(71.64%)	213(51.08%)
Egress from Jetty Modes		
Walk or Bike	312(47.06%)	295(70.74%)
Other	351(52.94%)	122(29.26%)

	Bus	Van
Variable	Count (Pct.%)	Count (Pct.%)
In City Residence		
No	19(2.87%)	203(48.68%)
Yes	644(97.13%)	214(51.32%)
Willingness to Walk to access point		
10 minutes or less	322(48.57%)	273(65.47%)
More than 10 minutes	341(51.43%)	144(34.53%)
Activity Working		
No	587(88.54%)	345(82.73%)
Yes	76(11.46%)	72(17.27%)
Activity Talk on Phone		
No	555(83.71%)	368(88.25%)
Yes	108(16.29%)	49(11.75%)
Reason Booking of Seat		
No	159(23.98%)	135(32.37%)
Yes	504(76.02%)	282(67.63%)
Reasons Security against theft		
No	210(31.67%)	123(29.5%)
Yes	453(68.33%)	294(70.5%)
Reasons Ease of Payment		
No	451(68.02%)	313(75.06%)
Yes	212(31.98%)	104(24.94%)
Jetty Use rate		
Less Than Once a Month	72(10.86%)	63(15.11%)
1-3 Times per Month	138(20.81%)	104(24.94%)
1-3 Times per Week	190(28.66%)	120(28.78%)
3 Times or More per Week	263(39.67%)	130(31.18%)
	Mean ± SD	Mean ± SD
Average fare (MXN)	40.90 ± (11.75)	64.93 ± (17.06)
Average Trip Distance (km)	24.38 ± (3.44)	25.18 ± (10.20)
Pct of Morning Trips (%)	48 ± (33)	0.54 ± (36)
Access Duration (min)	19.3±(14.6)	15.6±(14.4)
Egress Duration (min)	17.20 ± (13.65)	11.94 ±(12.59)
Access Distance (Km)	3.18 ± (2.59)	2.70 ± (2.83)
Egress Distance (Km)	3.70 ±(3.16)	3.57 ±(3.64)
Metro Headway (sec)	239 ± (87.3)	203±(57.4)
N = 1080		

5.3. Use Frequency

The third research question to answer was: what are the factors affecting the frequency of Jetty use? An ordinal logistic HCM and multinomial HCM were developed to investigate the factors affecting the Jetty use frequency.

Dependent Variable

The dependent variable was set as the relative use frequency for each user. The variable was calculated by dividing the number of Jetty trips over the period between the first and last rides for each user. Table-5.9 shows the summary of the dependent variable.

Table 5.9 – Jetty Use Frequency Summary Statistics

Use rate	# No (Pct%)
Less than Once a Month	141 (12.6%)
1 - 3 Times per Month	256 (22.9%)
1 - 3 Times per Week	317 (28.4%)
More than 3 Times per Week	404 (36.1%)
Total	1118 (100%)

5.3.1. Ordinal Logit Hybrid Choice Model

Starting from the EFA analysis results, two latent variables indicating the frequent use of public transportation and the frequent use of micro-mobility were investigated with several specifications for the structure equation part. Both the latent variables were proven to be significantly different from zero on reasonable significance levels. Figure-5.2 shows the full path diagram for the HCM. The measurement model considered the ordered logit model for the measurement equation. The model in Figure-5.1 is formulated as follows:

Structure Model :

$$\alpha_n = \gamma_n X + \eta_n \quad , \quad n = 2 \quad , \quad \eta \sim N(0, \textstyle\sum_\eta); \quad \textit{(2 equations)}$$
$$U = X\beta_1 + \alpha\lambda + \epsilon \quad , \quad \epsilon \sim N(0, 1); \quad \textit{(1 equation)}$$

Measurement Model :

$$I_r = \alpha\zeta_r + \upsilon_r \quad , \quad r = 1, ..., 12 \quad , \quad \upsilon \sim N(0, \textstyle\sum_\upsilon); \quad \textit{(12 equations)}$$

Trying different specifications for the structure model, three personal characteristics, gender, income, and the number of cars in the household were appropriate to use based on their significance and their coefficient interpretation in the context of the model results.

The latent attitude (l), $(l = 1, 2)$, for person number (n) can be defined as:

$$\begin{aligned}\alpha_{n,1} =&\gamma_{1,Female} \cdot Z_{n,Female} + \gamma_{1,Income_{lessthan20K}} \cdot Z_{n,Income_{lessthan20K}} +\\ &\gamma_{1,Income_{20K-40K}} \cdot Z_{n,Income_{20K-40K}} + \eta_{n,1}\\ \alpha_{n,2} =&\gamma_{2,Female} \cdot Z_{n,Female} + \gamma_{2,Income_{lessthan20K}} \cdot Z_{n,Income_{lessthan20K}} +\\ &\gamma_{2,Income_{20K-40K}} \cdot Z_{n,Income_{20K-40K}} + \gamma_{2,Income_{nocars}} \cdot Z_{n,Income_{nocars}} +\\ &\gamma_{2,Income_{cars=1}} \cdot Z_{n,Income_{cars=1}} + \eta_{n,2}\end{aligned}$$

$(\eta_{n,l})$ the random part of the structure model is a standard normally distributed term with a mean of zero and a standard deviation of 1 distributed across the individual.

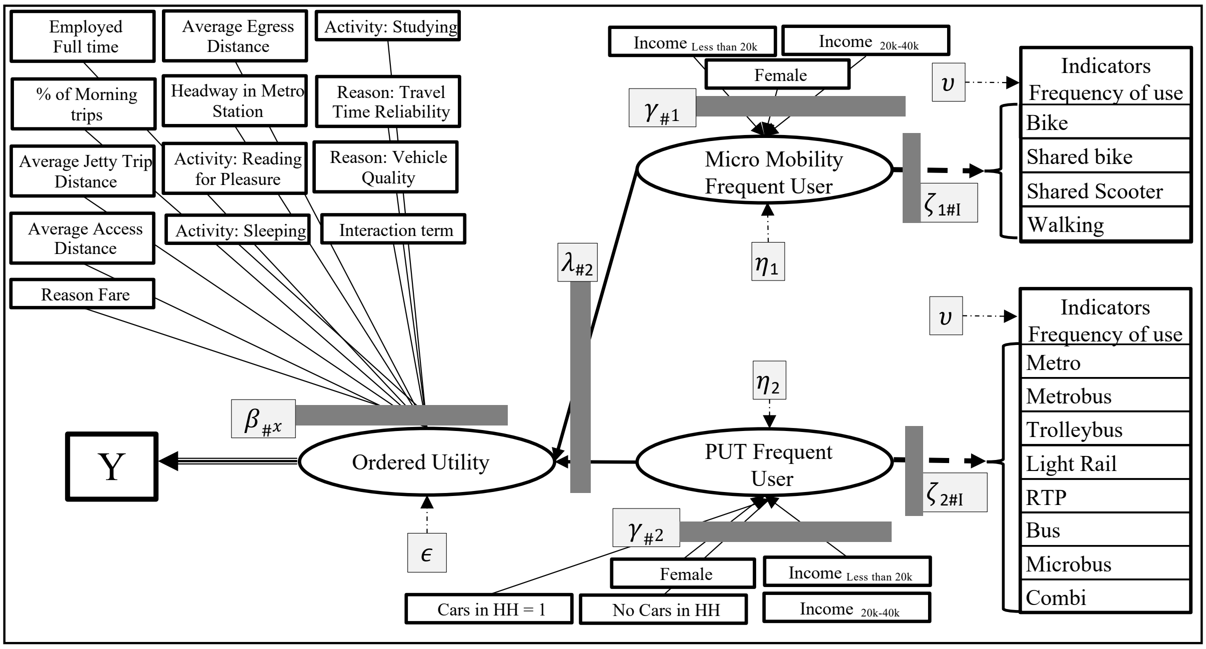

Figure 5.2 – Jetty Use Frequency Ordered-HCM Full Path Diagram (*own illustration*)

Table-5.10 left side shows the estimated coefficient of the HCM and the right side shows the reduced choice model, the lower part of the table shows the latent model part of the HCM.

Model Coefficients Interpretation

The interpretation for all the significant variables could be explained in control of other variables as:

Employed Status: The full-time employed positive estimated coefficient shows that users that are employed full-time are more likely to use Jetty more frequently compared to other employed category.

Percentage of Morning Trips: The estimated coefficient shows that the higher the percentage of Jetty trips done in the morning, the more likely the person uses Jetty more frequent.

Average Jetty Trips Distance: The estimated positive coefficient shows that the longer the average Jetty trips, the more frequent the people to use Jetty. This coefficient indicates the convenience introduced by Jetty, as more extended trips could increase transfers and longer travel time if done in modes rather than Jetty.

Access and Egress Distances: Parameters number (four and five) can be interpreted as the longer the access or the egress distance to or from Jetty; the less frequency the person would use Jetty.

Headway in the Nearest Metro Station: The estimated positive coefficient shows that the longer the headway of the nearest metro station, which indicates longer waiting times for Metro, the more likely the person to use Jetty.

Activity During Trips: Four activities are proven to be significantly different from zero. These activities are: working, reading for pleasure, sleeping, and studying. The interpretation for the estimated coefficient is that people who perform these activities during the Jetty trip increases the use of Jetty.

Reasons to Use Jetty: Three reasons are proven to be significantly different from zero. These reasons are I) travel-time reliability, II) Quality of vehicle, and III) Fare. The estimated coefficient shows that reasons I, and II, increase the use of Jetty, and the third reason, fare, decreases the use of Jetty.

Interaction term (Female x Reason: Security against Harassment): The interaction term represents the female users that use the service because of its security against harassment, and the estimated coefficient shows that this user group is more likely to use the service compared to other people.

Thresholds: The three thresholds between the four levels of the use rate are significantly different from zero, indicating that there is a noticeable difference between these levels.

It is worth mentioning that for the OLM the estimated coefficients have different values for each of the thresholds between the different levels, and the previous interpretation is the overall magnitude of each of the estimated coefficients

Latent Variable Model

Table-5.10 lower part shows the measurement model part of the latent variable model. The measurement model estimated positive coefficients (ζ) shows that the higher the levels of the answer (the more frequent the use), the more the use of the PUT and micro-mobility in general.

The structure equation part of the latent variable model shows that the estimated coefficients of the structure model (γ) need to be explained alongside the measurement model.

LV1 (MM use): Gender's negative sign indicates that females have a negative impact on the latent variable compared to males, or in other words, females are less frequent users of MM. For the income, the base group is the income group with 40k or more income, and the estimated coefficient negative sign shows that high-income groups are more likely to use MM compared to the low-income groups.

LV2 (PUT use): Gender's negative sign indicates that females have a negative impact on the latent variable compared to males, or in other words, females are less frequent users of PUT. For the income, the reference group is the income group with 40k or more income, and the estimated coefficient positive sign shows that the lower-income groups are more likely to use the PUT compared with the higher income group. For the number of cars in the household, the reference category is the two or more cars in the household. The positive coefficients show that households with no cars are more likely to use PUT compared to the other categories. It is to be noticed that the income level less than 20k and no cars in the household have the most substantial impact on this latent variable.

For the choice model part, the latent variables impact on the model, can be interpreted as the latent variable (frequent PUT users) reduces Jetty use, and the frequent micro-mobility latent variable increases Jetty use.

It is worth mentioning that the threshold between the different levels of the indicators (τ_p) are only reflecting the orders of the threshold, and this is why their estimation results are not shown

Statistical Evaluation

The adjusted-rho-squared- ($\rho^2_{Adjusted}$) for both models was calculated considering a degree of freedom with two more variables (LV) for the HCM was calculated to compare HCM and the reduced choice model fit. The calculated $\rho^2_{Adjusted}$ for both models are equal (0.05), which is an indicator that both models fit the data equally; however, the HCM has the advantage of the extra insight it gives from the structure model part.

Table 5.10 – Jetty Use Frequency Model-1 Results

		HCM		Choice Model	
No	Variable	β (P-value)	Rob.Std. Error	β (P-value)	Rob.Std. Error
1	Employed Full time (vs others)	0.67 (0.00)	0.21	0.69 (0.00)	0.22
2	Percent % of Trips in Mornings	0.69 (0.00)	0.20	0.76 (0.00)	0.20
3	Average Jetty Trip Distance	0.21 (0.00)	0.07	0.2 (0.00)	0.07
4	Average Access Distance to Jetty	-0.13 (0.08)	0.07	-0.15 (0.03)	0.07
5	Average Egress Distance from Jetty	-0.29 (0.00)	0.06	-0.32 (0.00)	0.06
6	Headway in the nearest Metro Station	0.16 (0.01)	0.06	0.15 (0.02)	0.06
7	Activity Reading For Pleasure	0.37 (0.03)	0.17	0.34 (0.04)	0.16
8	Activity Sleeping	0.76 (0.00)	0.16	0.70 (0.00)	0.15
9	Activity Studying	0.91 (0.00)	0.31	0.85 (0.00)	0.26
10	Reason Travel Time Reliability	0.35 (0.01)	0.13	0.34 (0.01)	0.12
11	Reason Quality of Vehicle	0.37 (0.00)	0.13	0.33 (0.01)	0.13
12	Reason Fare	-0.38 (0.00)	0.13	-0.32 (0.01)	0.13
13	Interaction Female X Reason Security Against Harassment	0.36 (0.06)	0.19	0.34 (0.05)	0.17
14	LV1 Frequent MA user (λ_1)	0.14 (0.14)	0.10	—	—
15	LV2 Frequent PUT user (λ_2)	-0.24 (0.00)	0.07	—	—
Threshold 1	Less Than Once a Month|1 - 3 Times per Month	-1.89 (0.00)	0.17	-2.01 (0.00)	0.17
Threshold 2	1 - 3 Times per Month|1 - 3 Times per Week	-0.37 (0.02)	0.15	-0.51 (0.00)	0.15
Threshold 3	1 - 3 Times per Week| More than 3 Times per Week	1.01 (0.00)	0.15	0.85 (0.00)	0.15
	$\rho^2_{Adjusted}$	**0.05**		**0.05**	
Latent Variable Model					
	Structure Model	Frequent MA Users		Frequent Put Users	
		γ (P-value)	Rob.Std. Error	γ (P-value)	Rob.Std. Error
1	Gender: Female (vs male)	-0.48 (0.00)	0.11	-0.5 (0.00)	0.09
2	Personal Income Less than 20K (vs more than 40k)	-0.47 (0.00)	0.16	0.99 (0.00)	0.14
3	Personal Income 20K-04K (vs more than 40k)	-0.33 (0.03)	0.15	0.5 (0.00)	0.15
4	No Cars in Household (vs 2 or more cars)	–	–	0.93 (0.00)	0.10
5	Cars in Household = 1 (vs 2 or more cars)	–	–	0.43 (0.00)	0.09
	Measurement Model	Frequent MA		Frequent PUT	
	Indicators	ζ (P-value)	Rob.Std. Error	ζ (P-value)	Rob.Std. Error
1	Frequency of Bike use	1.06 (0.00)	0.17	—	—
2	Frequency of Shared bike use	2.4 (0.00)	0.63	—	—
3	Frequency of Shared Scooter use	1.83 (0.00)	0.35	—	—
4	Frequency of Walk use	0.82 (0.00)	0.16	—	—
5	Frequency of Metro use	—	—	1.34 (0.00)	0.12
6	Frequency of Metrobus use	—	—	0.72 (0.00)	0.09
7	Frequency of Light-Rail use	—	—	1.19 (0.00)	0.17
8	Frequency of Trolleybus use	—	—	1.35 (0.00)	0.22
9	Frequency of RTP use	—	—	1.26 (0.00)	0.14
10	Frequency of Bus use	—	—	0.94 (0.00)	0.11
11	Frequency of Microbus use	—	—	1.17 (0.00)	0.13
12	Frequency of Combi use	—	—	1.15 (0.00)	0.12

N =941, P-values are reported in parentheses are based on the robust standard errors, used to control for heteroscedasticity that might exist

5.3.2. Multinomial Logit Hybrid Choice Model

The ordered nature of the dependent variable, frequency of use, was ignored to investigate the different variables impact the different Jetty use rates. A Multinomial Logit (MNL) was estimated using the same dependent variable for the previous model consisted of the four use levels, as shown in Table-5.9.

Table-5.11 shows the estimated coefficient of the reduced choice model and the HCM, and Figure-5.3 shows the full path diagram for the same model.

The interpretation for all the significant variables could be explained in control of other variables as:

Level 1: Less than once a Month

Egress distance: People who need extended access distances to Jetty are more likely to use the service less than once a month, or in other words, the increase of egress distance decreases Jetty use rate.

Activity Sleeping: Sleeping is the only statistically significant activity on this level of use. Sleeping during the trip reduces the probability to use Jetty with less than once a month, or people who sleep during Jetty trip are more likely to use Jetty at a higher rate.

Level 2: One to Three Times per Month

Personal Income: Income group level 40k or more was set as the reference category for this variable, and the estimated coefficients number three and four can be interpreted as a lower-income group are more likely to use Jetty on this rate more than the higher income groups.
Security Against Theft: The negative estimated coefficient for security against theft variables indicates that the variable reduces Jetty use rate on this level.

Level 3: One to Three Times per Week

Income: Income group level 40k or more was set as the reference category for this variable, and the estimated coefficients number (six and seven) can be interpreted as lower-income groups are more likely to use Jetty on this rate more than the higher income groups.

Level 4: More than three times a week

Employment: Full time employed users are more likely to use Jetty on this level than other people.

Average trip Distance: The increase of the trip distance increases the odds of using Jetty on this level.

Access and Egress Distances: The increase of access or egress distances reduce the use of Jetty on this level.

Percentage of Morning Trips: Users with a high percentage of Jetty trips done in the morning are more likely to use the service more than other users on this level of use.

Metro Headway: The increase in Metro headway increases Jetty use on this level.

Activities During the Trip: Two activities are significant on this level of use: sleeping and reading for pleasure; both the activities increase the use of the service on this level.

Reasons to Use Jetty: Four reasons are significantly different from zero on this level of use; security against harassment, quality of the vehicle, travel time reliability, and user of time while traveling. All four reasons increase the use of Jetty on this level.

Latent Variable Model

The estimated latent variable model with both of its components, measurement model, and structure model, yields the same result and the same interpretation for the latent variable model estimated for the ordered logit model in the previous section. Their impact on the choice model is the same as in the case of the ordered logit model.

Statistical Evaluation

The adjusted-rho-squared ($\rho^2_{Adjusted}$) for both the reduced choice model and the choice part of the HCM were calculated considering a degree of freedom with two more variables (LV) for the HCM. The calculated $\rho^2_{Adjusted}$ for both models equals (0.10), which is an indicator that both models fit the data equally; however, the HCM has the advantage for the extra insight it gives from the structure model part. Also, ($\rho^2_{Adjusted}$) shows that the multinomial HCM fits the data better than the ordered HCM.

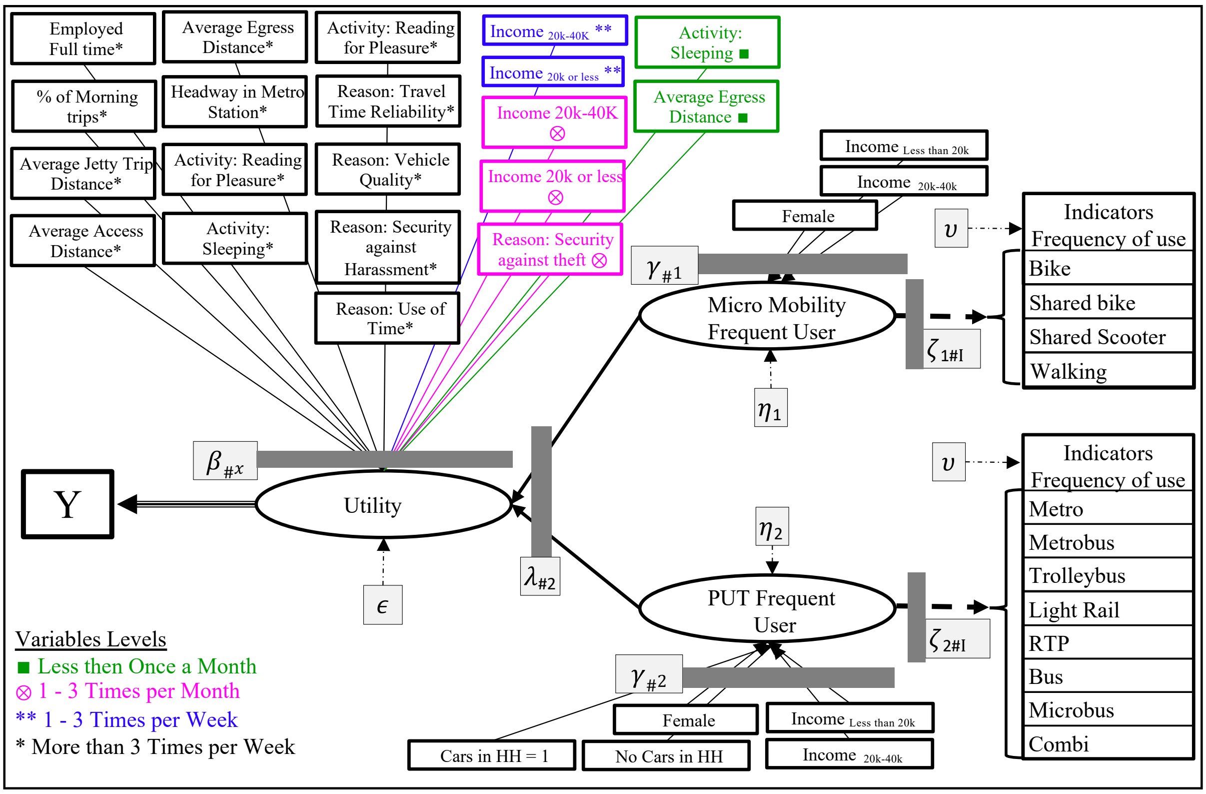

Figure 5.3 – Jetty Use Frequency MNL-HCM Full Path Diagram (*own illustration*)

Note: All the significant variables estimated in the hybrid ordered logit model (Table-5.10) were also proven significant in the hybrid logit multinomial model, except for reason fare, activity sleeping, and the interaction term of female and reason security against harassment. However, the hybrid multinomial logit model added some insight into the factor affecting the frequency of use represented by the introduction of the income impact on the different use levels.

Table 5.11 – Jetty Use Frequency Model-2 Results

No	Variable	ICLV Model β (P-value)	ICLV Model Rob.Std. Error	Choice Model β (P-value)	Choice Model Rob.Std. Error
Level Less Than Once a Month					
1	Activity Sleep	-0.72 (0.00)	0.23	-0.64 (0.00)	0.20
2	Jetty Mean Egress Distance	0.17 (0.06)	0.09	0.18 (0.03)	0.08
1 - 3 Times per Month					
3	Personal Income between 20K- 40K (vs 40K or more)	0.80 (0.01)	0.32	0.82 (0.00)	0.21
4	Personal Income 20K or Less (vs 40K or more)	0.39 (0.24)	0.33	0.52 (0.01)	0.21
5	Reason: Security Against Theft	-0.51 (0.02)	0.22	-0.51 (0.00)	0.16
1 - 3 Times per Week					
6	Personal Income between 20K- 40K (vs 40K or more)	0.60 (0.02)	0.25	0.61 (0.00)	0.19
7	Personal Income 20K or Less (vs 40K or more)	0.31 (0.18)	0.24	0.44 (0.02)	0.18
4 or More Times per Week					
8	ASC	-1.22 (0.06)	0.66	-1.45 (0.00)	0.36
9	Employed Full time	0.59 (0.11)	0.37	0.59 (0.03)	0.27
10	Average Jetty Trip Distance	0.23 (0.01)	0.08	0.22 (0.00)	0.08
11	Jetty Mean Access Distance	-0.23 (0.03)	0.10	-0.24 (0.01)	0.09
12	Jetty Mean Egress Distance	-0.37 (0.00)	0.11	-0.41 (0.00)	0.10
13	Pct (%) of Morning Trips	0.89 (0.00)	0.23	0.94 (0.00)	0.21
14	Headway in the nearest Metro station	0.31 (0.00)	0.08	0.28 (0.00)	0.08
15	Activity Reading for Pleasure	0.42 (0.06)	0.23	0.37 (0.05)	0.19
16	Activity Sleep	0.53 (0.04)	0.25	0.49 (0.02)	0.20
17	Reason Security against Harassment	0.76 (0.02)	0.33	0.74 (0.00)	0.26
18	Reason Quality of Vehicle	0.49 (0.01)	0.19	0.44 (0.00)	0.15
19	Reason Travel Time Reliability	0.36 (0.04)	0.17	0.34 (0.02)	0.15
20	Reason Use of Time while Travelling	0.49 (0.01)	0.18	0.51 (0.00)	0.17
21	LV1: Frequent MM user (λ_1)	0.18 (0.18)	0.14	—	—
22	LV2: Frequent PUT user (λ_2)	-0.27 (0.00)	0.09	—	—
	$\rho^2_{Adjusted}$	**0.10**		**0.10**	
Latent Variable Model					
	Structure Model	Frequent MA Users γ (P-value)	Frequent MA Users Rob.Std. Error	Frequent Put Users γ (P-value)	Frequent Put Users Rob.Std. Error
1	Gender: Female (vs male)	-0.49 (0.00)	0.11	-0.49 (0.00)	0.09
2	Personal Income Less than 20K (vs more than 40k)	-0.52 (0.02)	0.22	1.04 (0.00)	0.18
3	Personal Income 20K-04K (vs more than 40k)	-0.29 (0.21)	0.23	0.55 (0.00)	0.18
4	No Cars in Household (vs 2 or more cars)	–	–	0.88 (0.00)	0.14
5	Cars in Household = 1 (vs 2 or more cars)	–	–	0.39 (0.00)	0.14
	Measurement Model Indicators	Frequent MA ζ (P-value)	Frequent MA Rob.Std. Error	Frequent PUT ζ (P-value)	Frequent PUT Rob.Std. Error
1	Frequency of Bike use	1.06 (0.00)	0.17	—	—
2	Frequency of Shared bike use	2.24 (0.00)	0.47	—	—
3	Frequency of Shared Scooter use	1.8 (0.00)	0.38	—	—
4	Frequency of Walk use	0.78 (0.00)	0.15	—	—
5	Frequency of Metro use	—	—	1.36 (0.00)	0.16
6	Frequency of Metrobus use	—	—	0.69 (0.00)	0.10
7	Frequency of Light-Rail use	—	—	1.11 (0.00)	0.16
8	Frequency of Trolleybus use	—	—	1.25 (0.00)	0.17
9	Frequency of RTP use	—	—	1.25 (0.00)	0.16
10	Frequency of Bus use	—	—	0.97 (0.00)	0.10
11	Frequency of Microbus use	—	—	1.22 (0.00)	0.11
12	Frequency of Combi use	—	—	1.23 (0.00)	0.15

N=941, P-values are reported in parentheses are based on the robust standard errors, used to control for heteroscedasticity that might exist

Table 5.12 – Jetty Use Frequency Model Significant Variables Summary Statistics

Variable	Less Than Once a Month Count(pct%)	1 - 3 Times per Month Count(pct%)	1 - 3 Times per Week Count(pct%)	More than 3 Times per Week Count(pct%)
Age				
Between 18 and 25	16 (11.35%)	37 (14.45%)	31 (9.78%)	60 (14.85%)
Between 26 and 35	69 (48.94%)	116 (45.31%)	160 (50.47%)	181 (44.8%)
Between 36 and 45	34 (24.11%)	62 (24.22%)	82 (25.87%)	118 (29.21%)
46 or More	19 (13.48%)	39 (15.23%)	41 (12.93%)	44 (10.89%)
Missing	3 (2.13%)	2 (0.78%)	3 (0.95%)	1 (0.25%)
Gender				
Female	65 (46.1%)	118 (46.09%)	159 (50.16%)	222 (54.95%)
Male	76 (53.9%)	138 (53.91%)	158 (49.84%)	182 (45.05%)
Household size				
Between 1 and 2	42 (29.79%)	80 (31.25%)	79 (24.92%)	113 (27.97%)
Between 3 and 5	79 (56.03%)	145 (56.64%)	194 (61.2%)	252 (62.38%)
6 and More	12 (8.51%)	24 (9.38%)	32 (10.09%)	26 (6.44%)
Missing	8 (5.67%)	7 (2.73%)	12 (3.79%)	13 (3.22%)
Personal Income				
40K or More	19 (13.48%)	18 (7.03%)	28 (8.83%)	41 (10.15%)
20K - 40K	32 (22.7%)	92 (35.94%)	104 (32.81%)	127 (31.44%)
20K or less	67 (47.52%)	110 (42.97%)	142 (44.79%)	180 (44.55%)
Missing	23 (16.31%)	36 (14.06%)	43 (13.56%)	56 (13.86%)
Driving License				
No	25 (17.73%)	41 (16.02%)	66 (20.82%)	89 (22.03%)
Yes	116 (82.27%)	215 (83.98%)	251 (79.18%)	315 (77.97%)
Number of Cars in Household				
No cars	23 (16.31%)	55 (21.48%)	63 (19.87%)	69 (17.08%)
1	73 (51.77%)	97 (37.89%)	157 (49.53%)	193 (47.77%)
2 or more	45 (31.91%)	104 (40.62%)	97 (30.6%)	142 (35.15%)
Education Level				
Bachelor or higher	125 (88.65%)	229 (89.45%)	284 (89.59%)	368 (91.09%)
Other	16 (11.35%)	25 (9.77%)	33 (10.41%)	33 (8.17%)
Missing	00(00.0%)	2 (0.78%)	00(00.0%)	3 (0.74%)
Employment				
Full time	118 (83.69%)	223 (87.11%)	289 (91.17%)	373 (92.33%)
Other	23 (16.31%)	33 (12.89%)	28 (8.83%)	31 (7.67%)
In City Residence				
No	33 (23.4%)	44 (17.19%)	65 (20.5%)	83 (20.54%)
Yes	108 (76.6%)	212 (82.81%)	252 (79.5%)	321 (79.46%)
Activity Reading For Pleasure				
No	111 (78.72%)	207 (80.86%)	255 (80.44%)	299 (74.01%)
Yes	30 (21.28%)	49 (19.14%)	62 (19.56%)	105 (25.99%)
Activity Sleeping				
NO	54 (38.3%)	73 (28.52%)	80 (25.24%)	67 (16.58%)
Yes	87 (61.7%)	183 (71.48%)	237 (74.76%)	337 (83.42%)
Activity Studying				
No	132 (93.62%)	242 (94.53%)	293 (92.43%)	370 (91.58%)
Yes	9 (6.38%)	14 (5.47%)	24 (7.57%)	34 (8.42%)
Reasons Travel Time Reliability				
No	79 (56.03%)	132 (51.56%)	150 (47.32%)	173 (42.82%)
Yes	62 (43.97%)	124 (48.44%)	167 (52.68%)	231 (57.18%)
Reasons Quality of Vehicle				
No	91 (64.54%)	173 (67.58%)	205 (64.67%)	233 (57.67%)
Yes	50 (35.46%)	83 (32.42%)	112 (35.33%)	171 (42.33%)
Reason Fare				
No	75 (53.19%)	147 (57.42%)	213 (67.19%)	278 (68.81%)
Yes	66 (46.81%)	109 (42.58%)	104 (32.81%)	126 (31.19%)
Reason Security Against Theft				
No	44 (31.21%)	102 (39.84%)	96 (30.28%)	103 (25.5%)
Yes	97 (68.79%)	154 (60.16%)	221 (69.72%)	301 (74.5%)
Reason Security Against Harassment				
No	130 (92.2%)	220 (85.94%)	273 (86.12%)	323 (79.95%)
Yes	11 (7.8%)	36 (14.06%)	44 (13.88%)	81 (20.05%)
	Mean ± (SD)	Mean ±(SD)	Mean ±(SD)	Mean ±(SD)
Average Trip Distance (km)	24.00 ± (7.58)	23.91 ± (6.92)	24.34 ±(6.68)	25.24 ± (6.63)
Percent % of Morning Trips	45% ± (48%)	42%± (37%)	47% ± (36%)	58% ± (25%)
Access Mean Distance (km)	3.70 ± (3.81)	2.97 ± (2.45)	3.07 ± (2.95)	2.65 ±(2.00)
Egress Mean Distance (km)	4.87 ± (4.44)	4.11 ± (3.53)	3.78 ±(3.61)	2.80 ± (2.15)
Metro Headway (secs)	221.35 ± (76.20)	214.67 ±(79.61)	215.58 ±(80.00)	236.31 ± (78.58)

N = 1118

Chapter 6

Discussions

This chapter discusses the findings of the data analysis and the estimated models and how they answer the research questions and fulfill the research objectives. The implication of the findings on transportation planning and policy-making are discussed, in addition to the results of the tested hypotheses mentioned in the methodology, chapter-3. Lastly, the lessons learned from the survey are summarized.

6.1. Findings

6.1.1. Data Analysis Findings

Jetty users are young, full-time employees, highly educated with high income, high availability of a driving license, and their car ownership rate is high compared to the average population of CDMX. The gender of the users is balanced between males and females (Table-4.3). Jetty user's profile is similar to other shared services user's profile in other developed countries discussed in section-2.2.3. The only difference in the Jetty case is that women are frequent Jetty users, which was also similar to London, where the majority of ATS services were females [266], considering Jetty have similar characteristics for ATS, this could be an interesting finding that women are more entitled for the use of ATS compared to male counterparts. However, more studies for similar systems are required to verify this finding. Female use presence in the case of Jetty could have resulted from the deteriorated security condition of PUT in CDMX, discussed in detail in section-4.1. The survey analysis supports the fact that female Jetty use is motivated by the security condition of PUT, where (27%) of female respondents reported using the service for its security against harassment, only (4%) of males reported the same reason. The estimated models support the same finding as discussed in the next sections.

Ninety-five percent (95%) of the users specified that their last Jetty trip purpose is commuting to work(Figure-4.8). The analysis shows that the general purpose of using Jetty is commuting to work and two observations support this fact: I) the trips geographic routes, where 50% of the users used only one OD of Jetty to perform the majority of their Jetty trips (Figure-4.7); II)Jetty users are solo travelers, 97% of users booked one ticket on their trips (Table-4.15), this observation complies with the fact that is commuting to work trips are generally solo trips [142]

The adequacy of Jetty pick-up and drop-off points locations is reflected in the ways users access and egress the service, mainly done by walking. The fact that users access and egress by walking is a good indicator to avoid the potential last-mile problem that might occur if the stations' location is not coordinated with the users' OD (Figure-4.9and Figure-4.10). The adequacy of the pick-up and drop-off locations comes from the fact that Jetty dynamically relocate their stops based on actual user data and users' requests [188].

Jetty trips replace multi-modal trips for three-quarters of the users (on average, 2.14 modes per trip (Table-4.6). The reported disaggregated modes to replace Jetty trip 65% of the users would replace a link in their trips that would have been done in a car or a taxi (Table-4.7 and Table-4.6). Jetty replacing car trips could be an opportunity to reduce VKT; however, this opportunity depends on the replaced modes and the modes used to access and egress to Jetty. Regardless of the VKT and other externalities that could result in form Jetty use, the direct connection Jetty provides between the different OD indicated by the majority of the users utilizing almost the entire route between the start station and the end station (Figure-4.30), saves travel time and saves the number of transfers for majority of the users.

Jetty users are frequent users of e-hailing service (Figure-4.18). They prefer the use of e-hailing over taxi use, which is a different preference compared to the general CDMX population, where e-hailing trips represent only 30% of daily hailing trips [190, 188]. This finding is supported in the literature for the shared mobility user's profile, that shared mobility users are wealthier, more educated than the average population, refer to section-2.2.3. Also, this finding can indicate that shared mobility users are more willing to adopt several forms of service compared to other users.

The reported users' activities during the trip are equally distributed between both genders. Sleeping and using smart-phone are the highest reported activities by 70% of the users, which are double the third-ranked activity, looking out the windows (Figure-4.19). Sleeping during the trip indicates the level of safety users feel while using Jetty, while they might not be able do the same activity while using PUT

The distribution of reasons to use Jetty is almost equal between the two genders except for security against harassment were females were six times more than their male counterparts, and using Jetty to avoid parking problems, where males were double the female in choosing this reason because they relay more in cars as drivers to perform the same Jetty trip (Figure-4.20). The deteriorated condition of the PUT in CDMX is also reflected in the users' reasons to use Jetty. The top four reasons specified to use Jetty are: booking of the seat, security against theft, trip duration, and travel time reliability. These reasons were chosen by at least 50% of the users. Fare is only chosen by 35% of the users, might be because Jetty trip cost is more expensive than PUT, but still cheaper than e-hailing. Details of the PUT conditions in CDMX is discussed in details in section-4.1.

The temporal distribution of CDMX traffic is reflected on Jetty Demand, where Jetty's morning peak hour demand volume is double the evening peak hour demand on average, as revealed by analyzing the Jetty use database (Figure-4.29), following the same traffic demand pattern of the city. This difference in Jetty use by the time of the day could be due to the difference in the VOT between the different times of the day [191], where people in the morning have more time constraints compared to the evening peak hour; therefore, they are willing to use the more expensive service to save the travel time and have more reliable travel times. Also, the higher rate of Jetty use is combined with a higher percentage of the trips are performed in the morning; people who use Jetty more than three times per week are performing 58% of their trips in the morning hours (before noon).

The habitually of the travel pattern is evident on the users' choice of Jetty vehicles, where 82% of the users performed their Jetty trips in one vehicle type only (Figure-4.32).

Jetty occupancy rate is utilizing, on average, 50% of the vehicle's seating capacity (Figure-4.33). This fact is an essential advantage for services similar to Jetty over pooled ride-hailing services with fixable routes provided by Uber and Lyft, where using the pooled service of these companies does not grant the matching with other riders; however, those companies do not share their data to verify such debate [66].

The comments of the survey show that users' are satisfied with the service, where people indicated that the service is a safe, reliable, and convenient transport option (Figure-4.24). These virtues are not available in the city's PUT system, and pointing them shows why people immigrate from PUT to Jetty. Also, users requested the expansion of routes that also indicates users' satisfaction with the service and the low accessibility of PUT. On the other hand, people even asked to lower the trip cost, which suggests that Jetty is not affordable to all the users.

6.1.2. EFA Findings

EFA results revealed two factors explaining two underlying travel patterns: frequent PUT and paratransit use, frequent micro-mobility, and shared micro-mobility use (Table-4.17). These factors demonstrate two distinctive user's travel attitudes driven by the characteristics and taste of the users that they formed over an extended time, and it affects their general travel decisions and adoption of new services. User travel attitudes were hypothesized as; frequent PUT and paratransit users will be less likely to use Jetty or even shift from PUT and paratransit to Jetty. Also, frequent shared micro-mobility users will be more likely to use shared service, Jetty, in our case. Both hypotheses were proven correct, as discussed in the modeling finding section.

6.1.3. Models Findings

Females are more likely to shift to Jetty from car trips (trips that were done in cars, e-hailing, and ride-hailing or combination of these modes) compared to males. Also, females are using more convenient, more safe service (vans) that stimulate the use of shared mobility over the use of PUT represented by buses in the Jetty case. However, gender does not impact the service use frequency except for females who use the service to avoid harassment in the other available modes.

The age impact only appeared when moving from car trips to Jetty, where young users are more likely to shift to Jetty from car trips, which follows the data analysis finding and coincides with the general profile of shared mobility user. However, age did not impact use-frequency.

Personal income is the most significant sociodemographic attribute in deciding on the shift from car trips to Jetty, choice of service type, and use frequency. The high-income group is more likely to shift to Jetty, use van (the more expensive service), and use the service more frequently. This finding makes it necessary to question the equity of the service and how it should be addressed on a broader scale to include all the income groups under the coverage of using such services.

Small household sizes (1-2 persons) are more likely to shift from car trips to Jetty, use the more expensive service (van); however, the household size does not impact the use frequency. This finding matches with the shared mobility users profile for other services in other cities, where single users are adopting the service more than the other population.

Full-time employee is more likely to use vans and use the service more frequently than other users; this also matches with the profile of shared mobility users in other cities.

Number of cars in the household impact shift to Jetty; users with cars in the household are more likely to shift to Jetty from car trips than other users; however, car ownership rate does not impact Jetty use frequency in a direct way.

Trip distance is a significant factor that affects Jetty use in different aspects. The longer the distance, the less likely the users to shift from car trips, which could be because car trips are more convenient than Jetty. Also, the longer trip distance increases the odds of using van over the bus, which is also because vans are more convenient than buses. Lastly, the longer the user's average trip, the more likely the user to use Jetty more frequently, mainly because the long trip distance will save more time than public transport and will save the number of required transfers of the users. The main indication of trip distance impact on Jetty uses the convince provided by the service in term of trip time and cost-saving.

Access and egress times and modes impact the service use. The proximity of the access and egress locations to the trip's origin and destination increases the use frequency of the service. This point is directly related to service planning, which should consider the different land uses.

The percentage of the morning trips of Jetty use was a significant variable in all the estimated models, which is a clear indication of Jetty's importance in the mornings. Percentage of morning trips increases the odds of shifting from car trips to Jetty, it increases the odds of using van over the bus, increasing Jetty use frequency. This finding is related to the city traffic pattern, where the morning peak is more severe than the evening peak hour. Also, this finding confirms the fact that people value the saving in trip times in the mornings more than the evening due to the time constraints to reach workplaces at a certain time [191].

The synergy between Jetty use and other modes use is evident in the model's findings, where the low accessibility to the Metro (representing PUT) increases the use of Jetty. Also, the deteriorating condition of PUT vehicles and frequent break downs impact PUT travel time [230, 188], and subsequently increase Jetty use where the user-specified reasons of vehicle quality and travel time reliability increase Jetty use as per the estimated models.

Also, the relation between Jetty use and private vehicle use is represented by people shifting from car trips to Jetty to avoid parking problems, and performing multi-tasking activities such as reading for pleasure, sleeping, studying, and use of time during the trip increase Jetty use.

Fare significantly influences Jetty's use in different aspects; fare increases the odds to shift from car trips to Jetty (mainly due to the savings in travel cost between the two options). Also, increased fare boosts the odds to choose bus over van (buses are almost 50% cheaper than vans), and fare, reduces Jetty use frequency.

Also the deteriorated security situation in CDMX [188, 206, 164] increase Jetty use; where people use van over the bus (the van is smaller in size, and it promotes the security felling compared to the bus) for its security against theft, and Jetty security against harassment increases Jetty use rate.

6.2. Tested Hypotheses

The following section shows the results for the proposed hypothesis that was tested by the data analysis and the estimated models:

Hypothesis 1: Females are more likely to shift to Jetty from car trips compared to males (Table-5.3); however, the introduction of the LV reduces the statistical significance of the gender impact. Also, the interaction of female and security against harassment is more likely to increase the use as estimated in the hybrid ordered choice model (Table-5.10), but the multinomial hybrid (Table-5.11) choice model did not yield any significance coefficient for gender; therefore the hypothesis is partially accepted.

Hypothesis 2: Jetty users are younger than the average CDMX population as per the results of the data analysis (Table-4.3) .Young age group users are more likely to shift from car trips to Jetty compared to older groups (Table-5.3); however, the Age did not prove to impact Jetty frequency of use (Table-5.10 and Table-5.11); therefore, this hypothesis is partially validated

Hypothesis 3: The data analysis results show that Jetty users are wealthier than the average city population (Table-4.3); moreover, high-income users are more likely to shift from car trips to Jetty compared to lower-income groups. However, the middle-income groups (less than 40k MXN per month) are moderately using Jetty, between 1-3 times per month to 1-3 times per week, and the income level does not directly increase Jetty use rate. This hypothesis is partially validated.

Hypothesis 4: Jetty users car ownership rate is higher than the average city population car ownership rate (Table-4.3). Also, households with cars are more likely to shift from car trips to Jetty compared to households with no cars; however, the ownership of cars does not directly impact Jetty use frequency. High car ownership rate has a negative impact on the frequent PUT use attitude (LV), or in other words, the high car ownership rate reduces the attitude of being a frequent PUT user, which indirectly reduces Jetty use frequency; therefore, this hypothesis is partially valid.

Hypothesis 5: The household size distribution of Jetty users almost matching with CDMX household size distribution (Table-4.3), also small size household are more likely to shift from car trips to Jetty; but household size does not impact the usage frequency; therefore, this hypothesis is rejected.

Hypothesis 6: The data analysis shows that the majority of Jetty users are full-time employed (Table-4.3), also full-time employee are more likely to use Jetty compared to non-full-time employee (Table-5.10 and Table-5.11). Thus, this hypothesis is accepted.

Hypothesis 7: Although Jetty users are more educated compared to the average CDMX population (Table-4.3), education did not prove to impact Jetty use at any level; therefore, this hypothesis is rejected.

Hypothesis 8: The increase in access and egress distances are proven to reduce Jetty use rate (Table-5.10 and Table-5.11); therefore, this hypothesis is accepted.

Hypothesis 9: The increase in Jetty trip distance is proven to increase Jetty use rate (Table-5.10 and Table-5.11); therefore, this hypothesis is accepted.

Hypothesis 10: The impact of Jetty's fare on the service use rate is significant in the ordered hybrid choice model, but not statistically significant in the multinomial hybrid choice mode; therefore, this hypothesis is neither rejected nor accepted (Table-5.10 and Table-5.11).

Hypothesis 11: The multitasking offered by Jetty use (in terms of sleeping, reading for pleasure) is increasing Jetty use rate; therefore, this hypothesis is accepted (Table-5.10 and Table-5.11).

Hypothesis 12: In Jetty use frequency ordered hybrid model, the interaction of security against harassment and gender (females) increase the use of Jetty, also in the multinomial hybrid model shows that security against harassment increase Jetty use; therefore, this hypothesis is accepted.

Hypothesis 13: The increase in the Metro headway increases Jetty use rate; however, the rest of accessibility measures to the other PUT did not prove to be significant, so this hypothesis is neither accepted nor rejected.

Hypothesis 14: The increase in Jetty use in the mornings is proven to increase Jetty use rate (Table-5.10 and Table-5.11); therefore, this hypothesis is accepted.

Hypothesis 15: The inclusion of LV reduced the fitting of the binary model investigating the shift to Jetty, and did not enhance the models of use frequency; therefore, this hypothesis is neither accepted or rejected (Table-5.3, Table-5.10, and Table-5.11).

Table 6.1 – Tested Hypotheses Summary

Hypothesis	Validated	Neither Validated nor Rejected	Rejected
1	✓(partially)		
2	✓(partially)		
3	✓(partially)		
4	✓(partially)		
5			✓
6	✓		
7			✓
8	✓		
9	✓		
10		✓	
11	✓		
12	✓		
13		✓	
14	✓		
15		✓	

6.3. Policy implication

In the previous sections, Jetty users' travel behavior and factors impacting their travel decisions were analyzed and discussed. The analysis involved developing different models investigating the shift to the service from other modes, the choice between different service categories, and the service use frequency as functions of users' sociodemographics, attitudes, accessibility to PUT, and trip and service characteristics. Considering that Jetty is a platform for organized pooled ride service, the analysis outcomes are used to more extensive discussions with implication on similar services planning and transportation policymaking.

Occupancy and VKT

The majority of Jetty trips are commuting trips. Ninety-five percent (95%) of the users declared that their last Jetty trip's purpose was commuting to work. Also, the comparison of users' last trip route to their total Jetty use considering the provided home, and work locations indicate that 50% of the users use only one OD (Home to Work or vice versa) to perform more than 50% of Jetty trips, which confirms that Jetty is mainly used for commuting trips. Commuting trips are an essential part of travel demand as they represent a substantial share of the yearly VKT [65, 161]. Also, commuting trips are generally solo trips [142], which is reflected in Jetty's case by 97% of the users booking only one ticket per trip. Jetty users reported that 66% of the disaggregated trips replacing Jetty would have been done in small-size vehicles (passenger vehicles or taxi), which are low occupancy vehicles, and they are expected to have one user per vehicle [142].

On the other hand, the analysis shows that the Jetty occupancy rate is 40% on average for buses and around 60% on average for other vehicle types. This occupancy rate is mainly due to the nature of the service, which utilizes fixed routes and fixed schedules compared to the other pooled services operated by the TNC companies. TNC pooled services do not grant matching between users due to the flexible nature of the service routes; nevertheless, verifying this point is not easy as TNC companies do not share detailed data about their occupancy and ridership. Moreover, Jetty operation scheme of fixed routes minimizes the impact of the empty VKT as vehicles does not search for users away from their original routes, which is not the case of TNC pooled services, where a significant amount of empty VKT could be generated to relocate and search for the different users.

Moreover, there are two land-use-related factors support the potential reduction of VKT. These factors are specific to Jetty, and might not be considered in other services: I)The accessibility to formal jobs in the north of the city, where Jetty pickup locations are concentrated, is reduced due to the offered limited number of jobs and the low coverage of public transportation network as indicated by the city urban marginalization index[1], which is an accessibility index to jobs and essential services. II) The car ownership rate in the north-west of the city, where the pick locations are located, is higher than the rest of the city [102], which is evident in the study sample, where 80% of Jetty users has at least one car in the household (Table-4.3).

The model investigated the factor impacting the shift to Jetty, revealed that households with cars are willing to shift to Jetty from car trips, and the convenience offered by the service in terms of multi-tasking and avoiding associated driving problems attract car users to the services. Moreover, the cost saving of Jetty compared to e-hailing and ride-hailing attracts the users to Jetty.

However, the reduction of VKT is debatable on a citywide scale. The core of people movement in CDMX is the public transport not private vehicles, 53% of the daily trips are done in PUT compared to 25.2% done in private transport [190] (Table-4.1), so attracting the users from public transportation would not always yield optimum results and reduce VKT.

All the previous points show pooled service potential in large-sized vehicles with fixed routes and schedules (Jetty-similar services) in reducing the overall VKT. Policies instruments such as service subsidize to cater to the broader population, subsidize on the tolled routes, and to give priority to the vehicle in park and ride location, especially in the suburban areas, will encourage such services over TNC pooled services.

[1] www.conapo.gob.mx/en/CONAPO/Indice_de_marginacion_urbana_2010, Last accessed on 15 March, 2020

Service planning and Multimodality

The analysis shows that Jetty planning is essential in attracting new users, and considering such a planning approach for new services will be a good practice that should promote positive results in many aspects, especially for pooled services.

The pick-up and drop-off locations are well located considering the different land use. The good location is reflected by the majority of the users accessing and egressing the service by walk. Jetty also modifies its pick-up and drop-off locations based on users' actual trip data and based on users' requests [188]. Such an approach to service planning could result in an overall increase in active mobility use and the avoidance of the first and last-mile dilemma. Also, users with remote access and egress distances to Jetty stations are accessing the service by motorized modes, which promotes the idea of multimodality, especially under the poor condition of PUT. The planning of similar services should follow the same concept of Jetty considering the land use of both origins and destinations.

City policies should support multimodality. For example, the first mile for the different taxi and TNC services should be increased to support the use of active mobility and PUT. Also, free parking locations should be dedicated to the people who come from outside the city by driving and access the city by pooled services.

Jetty routes are providing a direct connection between the residential location in the north of the city and the job centers in Santa Fe and Polanco, offering additional accessibility that could be extended to the broader population if integrated with the city's transit system. Such integration of shared services with the broader PUT network could help in covering low-density areas that PUT cannot cover under the limited resources.

The change of Jetty's station locations based on consumer's request is a good practice and an example for citizen participation in service planning that needs to be adopted by the city on the broader transit network scale. Also, periodical surveys of the user's evaluation for the different transit services should be included as a part of the general operation scheme to measure the user's actual needs and to guide the different stakeholders towards user-oriented services.

Equity

Jetty users belong to the high-income segment of the population, which raises the question of the equity of using Jetty, especially for low-income groups, considering that some users asked to lower Jetty trip cost as indicated in the survey comments part.

The equity of using Jetty should alert policymakers and city planners, noting that Jetty service areas with low PUT accessibility and a low number of jobs[1], which does not qualify Jetty service to be lavish product inclusive to a specific segment of users. Solving the low PUT accessibility and the limited job opportunities in the north of CDMX will entitle a significant investment in infrastructure and land use, which will require a long time to materialize. Shared mobility offers a low-cost intermediate solution for such problems.

The integration of Jetty with other PUT services could be hindered by the fact that shared mobility use needs smartphones and credit cards, which are not always available to everyone, especially marginalized groups such as low-income old users, and women. Such points need to be addressed by the authorities in the form of establishing shared mobility access points for people with no smartphones and subsidize the service for low-income groups and households with no cars [220].

6.4. Lessons Learned from the Survey

The survey had some drawbacks in its design that were discovered during the data processing stage; however, those drawbacks did not affect the quality of the collected data, or the final results and conclusions. Avoiding these drawbacks when conducting similar studies is advisable. The details of the survey were discussed in section-3.1.1. The main points that are recommended to be amended in case of similar surveys are:

The survey did not investigate the family status and the number of children in the household, which could be important factors impacting Jetty use. Number of children in the household was proven to be a significant factor in similar shared mobility studies [72], and it impacts the general users travel behavior [52].

The choice set for the access and egress modes from or to Jetty in question number (Q1.2 and Q1.4) did not match with the choice set for Jetty's replacement modes in question (Q1.6). The access egress choice set also had fewer options compared to Jetty's replacement modes. For example, "car as a driver" and "car as a passenger" are two modes in Jetty's replacement modes, but only "car" was available as a mode to access and egress from Jetty. The difference between the three modes is mainly in their capacity. So when users specified "car" as their used mode to access or egress Jetty, it was not clear if "car" as a driver or "car" as a passenger.

In question number (Q1.6), there is no option to specify if the same mode is used more than once on the same trip. For example, if a person would replace the latest Jetty trip using two buses, it is not possible to indicate this option, a bus will be specified only once. Although this problem did not affect the replaced trip's distance calculation, it did not make it possible to calculate the exact number of transfers in the replaced trip or the exact number of trip's links Jetty replaced.

Not all the modes available for travel within Mexico City were included in the different related questions' choice sets, (Q1.2, Q1.4, Q1.6, and Q2.1). Moreover, the survey asked the users about their frequency of using Light rail, Trolley bus, shared scooter and shared Bicycle in (Q.2.1), and did not include these options in other questions (Q1.2, Q1.4, Q1.6). However, users reported the use of unmentioned modes under the other options choice for question (Q1.6, modes to replace Jetty); however, the use of these modes was not significant.

The specified levels of the sociodemographic attributes in the survey part three questions differ from the available published levels for Mexico city. This mismatch made the comparison between the survey representation of the total population not clear for some attributes like age, and income level.

The question of estimating the car's parking costs for the last trip, if replaced by car (Q1.7), was a mandatory question to answer. While users might not use a car as a mode of transport because they do not own one or have a driving license, it is misleading for these users to estimate parking costs. A suitable replacement for this question could be asking about the availability of parking at the trip's final destination to study the parking effect on mode choice. In the analysis process to overcome this problem, the users' group who own at least one car, and they have driving licenses was isolated, and their parking cost estimation was compared to the general population. Interestingly, no difference between this group parking cost estimation and the general population parking cost estimation was observed.

Chapter 7
Conclusion

This chapter presents the conclusion of the thesis, the limitation of the research, and recommendations for future work.

7.1. Conclusion

Developing cities face immerse challenges by the overgrowing stress on its infrastructure imposed by the growing population and the resulting travel demand; with limited resources, the gap between the travel demand and travel supply in such cities is growing with no expected short term improvement unless a distributive solution is adopted. Traditional long term solutions such as the change in land use and extension of infrastructure need high capital investment that might not always be available. Shared mobility presents the opportunity for an immediate solution to cater for the growing travel demand. Specifically, under the scope of this study, organized pooled services similar to Jetty have positive potential impacts with their high capacity and their ability to attract car users, as shown in the analysis. Jetty operation scheme does not offer door-to-door service; it has a fixed schedule, restrictive time supply (peak hour only), and fixed routes. Integrating such service in the broader PUT network could even increase the positive impacts; however, this integration is conditioned by some rules to produce successful results. These rules can be summarized as: I) Planning the service based on the actual spatial and temporal travel demand fulfilling people's needs. II) Considering the land use of potential origins and destination locations. III) Synchronize the service operations spatially and temporally with other PUT services. IV) Subsidize the service for marginalized groups such as; low-income people, elders, people with no smartphone, or internet access.

In this work, the factors affecting the adoption of pooled rides (Jetty) is investigated using online survey data collected in June 2019. The estimated models can be integrated into the broader travel demand models to increase their quality by incorporating shared services into the travel models. Also, the estimated models can be informative in term of strategic planning as sociodemographic variables representing the demographics of the population are included.

In addition to the previous points, the answers of the leading research questions can be summarized as

1)***Is*** the shared mobility users' profile in Mexico city, an example of a city in a developing country, the same as in Europe and in North America?

Shared mobility user's profile in Mexico City, Mexico, is similar to other shared mobility users in developing countries. Users are high income, well educated, young, with high car ownership rate compared to the average city population.

2)***What*** are the synergies between public and private transportation and shared mobility services?

There is a direct relation between Jetty use and PUT, where the reduced accessibility to PUT increase the use of Jetty. The relation between Jetty use and car ownership is complex, where Jetty users with cars in the household are more likely to shift to Jetty compared to other users. On the other hand, users attitude of using PUT is increased by the unavailability of a car in the household, and this attitude reduces Jetty use in general.

3) ***What*** factors affect a users' decision to adopt shared mobility services?

Gender, age, household size, personal income, availability of a driving license, number of cars, percent of morning use, fare, driving associated problems, and general travel behavior are proven to be the main factors affecting individual choices on moving from different travel modes to use Jetty, noting that age and general travel pattern have the highest impact on users decision.

4) ***What*** factors affect a users' decision to choose between different shared mobility services?

Choice between the different service categories, distinct by fare and size, is mainly governed by three main groups of variables; I) Sociodemographic represented by gender, household size, personal income, employment status. II) Service accessibility in terms of modes used to access and egress the service, and accessibility to PUT. III) Travel behavior represented by percent of morning use. IV) Service convenience in terms of trip distance, multitasking, security, and fare.

5) ***What*** are the factors that affect the frequency of use of shared mobility?

The service use rate is mainly affected by employment status, percent of morning trips, personal income, distance to access and egress the service and trip distance, accessibility to the PUT, convenience and security, general travel attitude where frequent PUT users are less likely to be frequent users, and active and micro-mobility users are more likely to use the service

7.2. Study Limitation

This thesis presents a framework for investigating the factors affecting the adoption and use of shared mobility services. A case study of Jetty was used as an example. Notwithstanding, some limitations were observed that should be addressed in similar studies. First, all the limitations of the survey discussed in detail in the lesson learned section-6.4 need to be addressed. Second, the survey was an online survey that did not cover not-Jetty-commuter increasing the potential of non-coverage bias [8]. Third, the collected sample is not representative of the city average population, mainly in terms of low-income groups, low or no-education groups, and households with no cars who need to be included in future work. Fourth, Jetty is a relatively new service (the application beta version was launched in July 2017), and the service user is growing exponentially [188] so that the observed behavior could be due to the users' early adoption behavior, and not actual user behavior. However, this study gave an interesting insight into the use of commercially organized pooled services, which is not widely discussed in the literature.

7.3. Recommendations and Future work

The finding of this work gives an interesting insight into the adaptation and use of pooled services; however, there are some few recommendations for future work.

Future similar studies should consider interviewing non-users, and also it should consider using face-to-face interviews to avoid coverage-bias and investigate the behavior of none users and why they do not use it.
In this research two attitudinal latent variables were used to investigate the impact of users travel behavior on using pooled rides, other attitudinal and latent behavioral factors such as evaluation of safety and security, the value of time, technology savviness, and adoption of other shared economy services could be included in the survey design to investigate the impact of those factor on adapting pooled rides.

Considering the novelty of Jetty service a follow-up survey investigating the users' adoption and use rate and how it changes over time will be informative of which direction is the use of such services is going.

A stated preference survey investigating how users of pooled rides value their time compared to non-users will guide the operators and the authorities for a fair pricing scheme for such services.

Bibliography

[1] Aasebø, Haakon Konrad William. «Shared Micro Mobility: What is happening in our streets?» MA thesis. 2019.

[2] Abrahamse, Wokje and Keall, Michael. «Effectiveness of a web-based intervention to encourage carpooling to work: A case study of Wellington, New Zealand». In: *Transport policy* 21 (2012), pp. 45–51.

[3] Adjei, Isaac Akpor et al. «An Application of Bootstrapping in Logistic Regression Model». In: *Open Access Library Journal* 3.09 (2016), pp. 1–9.

[4] Agyeman, Julian, McLaren, Duncan, and Schaefer-Borrego, Adrianne. «Sharing cities». In: *Friends of the Earth Briefing* (2013), pp. 1–32.

[5] Akaike, H. «A new look at the statistical identification model». In: *IEEE Transactions on Automatic Control* 19 (1974), p. 716.

[6] Akanwa, Angela Oyilieze and Joe-Ikechebelu, Ngozi. «The Developing World's Contribution to Global Warming and the Resulting Consequences of Climate Change in These Regions: A Nigerian Case Study». In: *Global Warming and Climate Change*. IntechOpen, 2019.

[7] Alam, Bhuiyan, Nixon, Hilary, Zhang, Qiong, et al. *Investigating the Determining Factors for Transit Travel Demand by Bus Mode in US Metropolitan Statistical Areas*. Tech. rep. Mineta Transportation Institute, 2015.

[8] Alemi, Farzad, Circella, Giovanni, Handy, Susan, and Mokhtarian, Patricia. «What influences travelers to use Uber? Exploring the factors affecting the adoption of on-demand ride services in California». en. In: *Travel Behaviour and Society* 13 (Oct. 2018), pp. 88–104. ISSN: 2214-367X. DOI: `10.1016/j.tbs.2018.06.002`. URL: `http://www.sciencedirect.com/science/article/pii/S2214367X17300947` (visited on 11/29/2019).

[9] Alemi, Farzad, Circella, Giovanni, Mokhtarian, Patricia, and Handy, Susan. *On-demand ride services in California: Investigating the factors affecting the frequency of use of Uber/-Lyft*. Tech. rep. 2018.

[10] Allen, Darcy et al. «The sharing economy». In: *Institute of Public Affairs Review: A Quarterly Review of Politics and Public Affairs, The* 67.3 (2015), p. 24.

[11] American Public Transportation Association. *Potential impact of gasoline price increases on US public transportation ridership, 2011-2012*. Tech. rep. 2011.

[12] Anable, Jillian, Chatterton, Dr Tim, Doughty, Dr Karolina, and Marsden, Greg. «Rethinking habitual travel patterns what might fleximobility mean for sustainable transport policies?» In: *. MOBILITY* (2015), p. 10.

[13] Anaya-Sánchez, Rafael, Aguilar-Illescas, Rocıo, Nasiff-Seiffert, Milagros, and Molinillo, Sebastian. «Why Rideshare?: An Analysis of Factors Influencing Intention to Use». In: *Sharing Economy and the Impact of Collaborative Consumption*. IGI Global, 2020, pp. 185–203.

[14] Ardra, M and Rejikumar, G. «Examining the Adoption Intentions of Women in Kochi regarding Uber Services». In: *International Journal of Pure and Applied Mathematics* 117.20 (2017), pp. 937–943.

[15] Arteaga-Sánchez, Rocıo, Belda-Ruiz, Maria, Ros-Galvez, Alejandro, and Rosa-Garcia, Alfonso. «Why continue sharing: Determinants of behavior in ridesharing services». In: *International Journal of Market Research* (2018), p. 1470785318805300.

[16] El-Assi, Wafic, Mahmoud, Mohamed Salah, and Habib, Khandker Nurul. «Effects of built environment and weather on bike sharing demand: a station level analysis of commercial bike sharing in Toronto». In: *Transportation* 44.3 (2017), pp. 589–613.

[17] Atkinson-Palombo, Carol, Varone, Lorenzo, and Garrick, Norman W. «Understanding the surprising and oversized use of ridesourcing services in poor neighborhoods in New York City». In: *Transportation Research Record* 2673.11 (2019), pp. 185–194.

[18] Audirac, Ivonne. «Stated preference for pedestrian proximity: an assessment of new urbanist sense of community». In: *Journal of Planning Education and Research* 19.1 (1999), pp. 53–66.

[19] Aziz, HM Abdul, Nagle, Nicholas N, Morton, April M, Hilliard, Michael R, White, Devin A, and Stewart, Robert N. «Exploring the impact of walk–bike infrastructure, safety perception, and built-environment on active transportation mode choice: a random parameter model using New York City commuter data». In: *Transportation* 45.5 (2018), pp. 1207–1229.

[20] Bachand-Marleau, Julie, Lee, Brian HY, and El-Geneidy, Ahmed M. «Better understanding of factors influencing likelihood of using shared bicycle systems and frequency of use». In: *Transportation Research Record* 2314.1 (2012), pp. 66–71.

[21] Balac, Milos, Ciari, Francesco, and Axhausen, Kay W. «Carsharing demand estimation: Zurich, switzerland, area case study». In: *Transportation Research Record* 2563.1 (2015), pp. 10–18.

[22] Banzhaf, H Spencer and Kasim, M Taha. «Fuel consumption and gasoline prices: The role of assortative matching between households and automobiles». In: *Journal of Environmental Economics and Management* 95 (2019), pp. 1–25.

[23] Bao, Jie, Xu, Chengcheng, Liu, Pan, and Wang, Wei. «Exploring bikesharing travel patterns and trip purposes using smart card data and online point of interests». In: *Networks and Spatial Economics* 17.4 (2017), pp. 1231–1253.

[24] Barr, Lawrence C. «Testing for the significance of induced highway travel demand in metropolitan areas». In: *Transportation Research Record* 1706.1 (2000), pp. 1–8.

[25] Becker, Henrik, Balac, Milos, Ciari, Francesco, and Axhausen, Kay W. «Assessing the welfare impacts of shared mobility and Mobility as a Service (MaaS)». In: *Transportation Research Part A: Policy and Practice* 131 (2020), pp. 228–243.

[26] Belsley, David A. «Conditioning diagnostics». In: *Wiley StatsRef: Statistics Reference Online* (2014).

[27] Ben-Akiva, Moshe E, Lerman, Steven R, and Lerman, Steven R. *Discrete choice analysis: theory and application to travel demand.* Vol. 9. MIT press, 1985.

[28] Ben-Akiva, Moshe, McFadden, Daniel, Gärling, Tommy, Gopinath, Dinesh, Walker, Joan, Bolduc, Denis, Börsch-Supan, Axel, Delquié, Philippe, Larichev, Oleg, Morikawa, Taka, et al. «Extended framework for modeling choice behavior». In: *Marketing letters* 10.3 (1999), pp. 187–203.

[29] Ben-Akiva, Moshe, Walker, Joan, Bernardino, Adriana T, Gopinath, Dinesh A, Morikawa, Taka, and Polydoropoulou, Amalia. «Integration of choice and latent variable models». In: *Perpetual motion: Travel behaviour research opportunities and application challenges* (2002), pp. 431–470.

[30] Berry, William D, Feldman, Stanley, and Stanley Feldman, Dr. *Multiple regression in practice*. 50. Sage, 1985.

[31] Bhat, Chandra R, Dubey, Subodh K, and Nagel, Kai. «Introducing non-normality of latent psychological constructs in choice modeling with an application to bicyclist route choice». In: *Transportation Research Part B: Methodological* 78 (2015), pp. 341–363.

[32] Bhat, Chandra R and Lawton, T Keith. «Passenger travel demand forecasting». In: *Transportation Research Board* (2000).

[33] Bhattacharya, Amar, Romani, Mattia, and Stern, Nicholas. «Infrastructure for development: meeting the challenge». In: *CCCEP, Grantham Research Institute on Climate Change and the Environment and G* 24 (2012).

[34] Bierlaire, Michel. *PandasBiogeme: a short introduction*. Tech. rep. EPFL, 2018.

[35] Bilgin, Berker, Magne, Pierre, Malysz, Pawel, Yang, Yinye, Pantelic, Vera, Preindl, Matthias, Korobkine, Alexandre, Jiang, Weisheng, Lawford, Mark, and Emadi, Ali. «Making the case for electrified transportation». In: *IEEE Transactions on Transportation Electrification* 1.1 (2015), pp. 4–17.

[36] Blumenberg, Evelyn and Smart, Michael. «Brother can you spare a ride? Carpooling in immigrant neighbourhoods». In: *Urban Studies* 51.9 (2014), pp. 1871–1890.

[37] Boarnet, Marlon and Crane, Randall. «The influence of land use on travel behavior: specification and estimation strategies». In: *Transportation Research Part A: Policy and Practice* 35.9 (2001), pp. 823–845.

[38] Böcker, Lars and Meelen, Toon. «Sharing for people, planet or profit? Analysing motivations for intended sharing economy participation». en. In: *Environmental Innovation and Societal Transitions*. Sustainability Perspectives on the Sharing Economy 23 (June 2017), pp. 28–39. ISSN: 2210-4224. DOI: 10.1016/j.eist.2016.09.004. URL: http://www.sciencedirect.com/science/article/pii/S2210422416300892.

[39] Botsman, Rachel and Rogers, Roo. «What's mine is yours». In: *The rise of collaborative consumption* (2010).

[40] Bryant, Katie. «Research From the Global South: The important role of context in international research activities». In: *McGill Journal of Education/Revue des sciences de l'éducation de McGill* 49.1 (2014), pp. 227–235.

[41] Buck, Darren, Buehler, Ralph, Happ, Patricia, Rawls, Bradley, Chung, Payton, and Borecki, Natalie. «Are bikeshare users different from regular cyclists? A first look at short-term users, annual members, and area cyclists in the Washington, DC, region». In: *Transportation Research Record* 2387.1 (2013), pp. 112–119.

[42] Buliung, Ron N, Soltys, Kalina, Bui, Randy, Habel, Catherine, and Lanyon, Ryan. «Catching a ride on the information super-highway: toward an understanding of internet-based carpool formation and use». In: *Transportation* 37.6 (2010), pp. 849–873.

[43] Burris, Mark W and Winn, Justin R. «Slugging in Houston—Casual carpool passenger characteristics». In: *Journal of Public Transportation* 9.5 (2006), p. 2.

[44] Cascetta, Ennio. *Transportation Systems Analysis: Models and Applications*. 2nd ed. Springer Optimization and Its Applications. Springer US, 2009. ISBN: 9780387758565. DOI: 10.1007/978-0-387-75857-2. URL: https://www.springer.com/gp/book/9780387758565.

[45] Caulfield, Brian, O'Mahony, Margaret, Brazil, William, and Weldon, Peter. «Examining usage patterns of a bike-sharing scheme in a medium sized city». In: *Transportation Research Part A: Policy and Practice* 100 (2017), pp. 152–161.

[46] *CDMX*. es. = https://cablebus.cdmx.gob.mx, last accessed on 1/1/20.

[47] Cerutti, Priscila Sardi, Martins, Rosiane Dutra, Macke, Janaina, and Sarate, João Alberto Rubim. «Green, but not as green as that: An analysis of a Brazilian bike-sharing system». In: *Journal of Cleaner Production* 217 (2019), pp. 185–193.

[48] Cervero, Robert. «City CarShare: First-year travel demand impacts». In: *Transportation Research Record* 1839.1 (2003), pp. 159–166.

[49] Cervero, Robert. «Linking urban transport and land use in developing countries». In: *Journal of Transport and Land Use* 6.1 (2013), pp. 7–24.

[50] Cervero, Robert. «Transport infrastructure and the environment in the global South: sustainable mobility and urbanism». In: *Journal of Regional and City Planning* 25.3 (2014), pp. 174–191.

[51] Cervero, Robert and Kockelman, Kara. «Travel demand and the 3Ds: Density, diversity, and design». In: *Transportation Research Part D: Transport and Environment* 2.3 (1997), pp. 199–219.

[52] Chakrabarti, Sandip and Joh, Kenneth. «The effect of parenthood on travel behavior: Evidence from the California Household Travel Survey». In: *Transportation Research Part A: Policy and Practice* 120 (2019), pp. 101–115.

[53] Chan, Nelson D and Shaheen, Susan A. «Ridesharing in North America: Past, present, and future». In: *Transport Reviews* 32.1 (2012), pp. 93–112.

[54] Chang, Annie YJ, Miranda-Moreno, Luis, Clewlow, Regina, and Sun, Lijun. *"Trend or Fad? Deciphering the Enablers of Micromobility in the U.S."* SAE International, 2019.

[55] Chaniotakis, Emmanouil, Efthymiou, Dimitrios, and Antoniou, Constantinos. «Data aspects of the evaluation of demand for emerging transportation systems». In: *Demand for Emerging Transportation Systems*. Elsevier, 2020, pp. 77–99.

[56] Chen, Xiaohong, Cheng, Jiaqi, Ye, Jianhong, Jin, Yong, Li, Xi, and Zhang, Fei. «Locating station of one-way carsharing based on spatial demand characteristics». In: *Journal of Advanced Transportation* 2018 (2018).

[57] Cheng, Long, Chen, Xuewu, Yang, Shuo, Wu, Jingxian, and Yang, Min. «Structural equation models to analyze activity participation, trip generation, and mode choice of low-income commuters». In: *Transportation Letters* 11.6 (2019), pp. 341–349.

[58] Circella, Giovanni, Alemi, Farzad, Tiedeman, Kate, Handy, Susan, and Mokhtarian, Patricia. *The adoption of shared mobility in California and its relationship with other components of travel behavior*. UC Davis: National Center for Sustainable Transportation., 2018. URL: `https://escholarship.org/uc/item/1kq5d07p`.

[59] Clewlow, Regina and Mishra, Gouri Shankar. *Shared mobility: Current adoption, use, and potential impacts on travel behavior*. Tech. rep. 2017.

[60] Cohen, Adam and Shaheen, Susan. *Planning for Shared Mobility*. 2018, p. 111. ISBN: 9781611901863.

[61] Cohen, Boyd and Kietzmann, Jan. «Ride on! Mobility business models for the sharing economy». In: *Organization & Environment* 27.3 (2014), pp. 279–296.

[62] Collins, Craig, Hasan, Samiul, and Ukkusuri, Satish V. «A novel transit rider satisfaction metric: Rider sentiments measured from online social media data». In: *Journal of Public Transportation* 16.2 (2013), p. 2.

[63] *Company Information | Uber Newsroom Germany*. `https://www.uber.com/en-DE/newsroom/company-info/`, last accessed on 11/29/19.

[64] Contrino, Heather and McGuckin, Nancy. «Demographics matter: travel demand, options, and characteristics among minority populations». In: *Public Works Management & Policy* 13.4 (2009), pp. 361–368.

[65] Convery, Sheila and Williams, Brendan. «Determinants of Transport Mode Choice for Non-Commuting Trips: The Roles of Transport, Land Use and Socio-Demographic Characteristics». In: *Urban Science* 3.3 (2019), p. 82.

[66] Conway, Matthew Wigginton, Salon, Deborah, and King, David A. «Trends in taxi use and the advent of ridehailing, 1995–2017: Evidence from the US National Household Travel Survey». In: *Urban Science* 2.3 (2018), p. 79.

[67] Costain, Cindy, Ardron, Carolyn, and Habib, Khandker Nurul. «Synopsis of users' behaviour of a carsharing program: A case study in Toronto». In: *Transportation Research Part A: Policy and Practice* 46.3 (2012), pp. 421–434.

[68] Costello, Anna B and Osborne, Jason. «Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis». In: *Practical Assessment, Research, and Evaluation* 10.1 (2005), p. 7.

[69] De Lorimier, Alexandre and El-Geneidy, Ahmed M. «Understanding the factors affecting vehicle usage and availability in carsharing networks: A case study of Communauto carsharing system from Montréal, Canada». In: *International Journal of Sustainable Transportation* 7.1 (2013), pp. 35–51.

[70] DeCoster, Jamie. *Overview of factor analysis*. 1998.

[71] Degele, Jutta, Gorr, Anna, Haas, Katja, Kormann, Dimitri, Krauss, Sascha, Lipinski, Paulina, Tenbih, Muhammet, Koppenhoefer, Christine, Fauser, Jan, and Hertweck, Dieter. «Identifying E-scooter sharing customer segments using clustering». In: *2018 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC)*. IEEE. 2018, pp. 1–8.

[72] Dias, Felipe F, Lavieri, Patrıcia S, Garikapati, Venu M, Astroza, Sebastian, Pendyala, Ram M, and Bhat, Chandra R. «A behavioral choice model of the use of car-sharing and ride-sourcing services». In: *Transportation* 44.6 (2017), pp. 1307–1323.

[73] Dorman, Michael. *nngeo: k-Nearest Neighbor Join for Spatial Data*. R package version 0.3.4. 2020. URL: `https://CRAN.R-project.org/package=nngeo`.

[74] Duffy, Peter, Fitzpatrick, Colin, Conway, Thomas, and Lynch, Robert P. «Energy Sources and Supply Grids–The Growing Need for Storage». In: *Energy Storage Options and Their Environmental Impact*. Royal Society of Chemistry, 2018. Chap. 1, pp. 1–41.

[75] Duran-Rodas, David, Chaniotakis, Emmanouil, and Antoniou, Constantinos. «Built Environment Factors Affecting Bike Sharing Ridership: Data-Driven Approach for Multiple Cities». In: *Transportation Research Record* (2019), p. 0361198119849908.

[76] Efron, Bradley. «Estimating the error rate of a prediction rule: improvement on cross-validation». In: *Journal of the American statistical association* 78.382 (1983), pp. 316–331.

[77] Eisenmeier, S RJ., ed. *Ride Sharing Platforms in Developing Countries: Effects and Implications for Mexico City*. 3. New York: Oxford University Press, 2019.

[78] Ekström, Joakim. «The phi-coefficient, the tetrachoric correlation coefficient, and the Pearson-Yule Debate». In: *UCLA: Department of Statistics, UCLA.* (2011). URL: `https://escholarship.org/uc/item/7qp4604r`.

[79] Eros, Emily, Mehndiratta, Shomik, Zegras, Chris, Webb, Kevin, and Ochoa, Maria C. «Applying the General Transit Feed Specification to the Global South: Experiences in Mexico City, Mexico—and Beyond». In: *Transportation Research Record* 2442.1 (2014), pp. 44–52.

[80] Estache, Antonio. «Infrastructure finance in developing countries: An overview». In: *EIB Papers* 15.2 (2010), pp. 60–88.

[81] European Commission. *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions–A European Agenda for the collaborative economy*. 2016.

[82] European Commission. Directorate-General Transport and Statistical Office of the European Communities. *EU Transport in Figures: Statistical Pocket book*. Office for Official Publications of the European Communities, 2016.

[83] *Ferrocarriles Suburbanos*. = http://fsuburbanos.com/, last accessed on 1/1/20.

[84] Fishman, Elliot, Washington, Simon, and Haworth, Narelle. «Bike share: a synthesis of the literature». In: *Transport Reviews* 33.2 (2013), pp. 148–165.

[85] Fishman, Elliot, Washington, Simon, and Haworth, Narelle. «Bike share's impact on car use: Evidence from the United States, Great Britain, and Australia». In: *Transportation Research Part D: Transport and Environment* 31 (2014), pp. 13–20.

[86] Fishman, Elliot, Washington, Simon, Haworth, Narelle, and Watson, Angela. «Factors influencing bike share membership: An analysis of Melbourne and Brisbane». In: *Transportation Research Part A: Policy and Practice* 71 (2015), pp. 17–30.

[87] Fleischer, Aliza, Tchetchik, Anat, and Toledo, Tomer. «The impact of fear of flying on travelers' flight choice: Choice model with latent variables». In: *Journal of Travel Research* 51.5 (2012), pp. 653–663.

[88] Fox, John. *polycor: Polychoric and Polyserial Correlations*. R package version 0.7-10. 2019. URL: `https://CRAN.R-project.org/package=polycor`.

[89] Frenken, Koen and Schor, Juliet. «Putting the sharing economy into perspective». In: *A Research Agenda for Sustainable Consumption Governance*. Edward Elgar Publishing, 2019.

[90] Fukuda, Daisuke and Morichi, Shigeru. «Incorporating aggregate behavior in an individual's discrete choice: An application to analyzing illegal bicycle parking behavior». In: *Transportation Research Part A: Policy and Practice* 41.4 (2007), pp. 313–325.

[91] Furuhata, Masabumi, Dessouky, Maged, Ordóñez, Fernando, Brunet, Marc-Etienne, Wang, Xiaoqing, and Koenig, Sven. «Ridesharing: The state-of-the-art and future directions». In: *Transportation Research Part B: Methodological* 57 (2013), pp. 28–46.

[92] Galdames, Camila, Tudela, Alejandro, and Carrasco, Juan-Antonio. «Exploring the role of psychological factors in mode choice models by a latent variables approach». In: *Transportation Research Record* 2230.1 (2011), pp. 68–74.

[93] Gehrke, Steven. *A Survey of Ride-Hailing Passengers*. TREC Friday Seminar Series 152. May 2018.

[94] Gehrke, Steven R, Felix, Alison, and Reardon, Timothy G. «Substitution of ride-hailing services for more sustainable travel options in the greater Boston region». In: *Transportation Research Record* 2673.1 (2019), pp. 438–446.

[95] Gillingham, Kenneth and Munk-Nielsen, Anders. «A tale of two tails: Commuting and the fuel price response in driving». In: *Journal of Urban Economics* 109 (2019), pp. 27–40.

[96] Giuliano, Genevieve, Levine, Douglas W, and Teal, Roger F. «Impact of high occupancy vehicle lanes on carpooling behavior». In: *Transportation* 17.2 (1990), pp. 159–177.

[97] *Glossary Listing - Glossary - Planning - FHWA*. `https://www.fhwa.dot.gov/Planning/glossary/glossary_listing.cfm`, last accessed on 11/12/19.

[98] Goodman, Anna, Green, Judith, and Woodcock, James. «The role of bicycle sharing systems in normalising the image of cycling: An observational study of London cyclists». In: *Journal of Transport & Health* 1.1 (2014), pp. 5–8.

[99] Goodspeed, Robert, Xie, Tian, Dillahunt, Tawanna R, and Lustig, Josh. «An alternative to slow transit, drunk driving, and walking in bad weather: An exploratory study of ridesourcing mode choice and demand». In: *Journal of Transport Geography* 79 (2019), p. 102481.

[100] Gorman, Ben. *mltools: Machine Learning Tools*. R package version 0.3.5. 2018. URL: `https://CRAN.R-project.org/package=mltools`.

[101] Grahn, Rick, Harper, Corey D, Hendrickson, Chris, Qian, Zhen, and Matthews, H Scott. «Socioeconomic and usage characteristics of transportation network company (TNC) riders». In: *Transportation* (2019), pp. 1–21.

[102] Guerra, Erick. «The geography of car ownership in Mexico City: a joint model of households' residential location and car ownership decisions». In: *Journal of Transport Geography* 43 (2015), pp. 171–180.

[103] Gunn, Holly. «Web-based surveys: Changing the survey process». In: *First Monday* 7.12 (2002).

[104] Gwilliam, KM, Kojima, Masami, and Johnson, Todd. *Reducing air pollution from urban transport*. World Bank Washington, DC, 2004.

[105] Haghighi, N Nima, Liu, Xiaoyue Cathy, Wei, Ran, Li, Wenwen, and Shao, Hu. «Using Twitter data for transit performance assessment: a framework for evaluating transit riders' opinions about quality of service». In: *Public Transport* 10.2 (2018), pp. 363–377.

[106] Hair, JF, Anderson, RE, Tatham, RL, and Black, WC. *Multivariate Data Analysis, 5th international ed*. 1998.

[107] Hampshire, Robert C and Gaites, Craig. «Peer-to-peer carsharing: Market analysis and potential growth». In: *Transportation Research Record* 2217.1 (2011), pp. 119–126.

[108] Handy, Susan. «Methodologies for exploring the link between urban form and travel behavior». In: *Transportation Research Part D: Transport and Environment* 1.2 (1996), pp. 151–165.

[109] Harrell Jr, Frank E. *Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis*. Springer, 2015.

[110] Heinrichs, Harald et al. «Sharing economy: a potential new pathway to sustainability». In: *GAIA-Ecological Perspectives for Science and Society* 22.4 (2013), pp. 228–231.

[111] Heinze, Georg and Schemper, Michael. «A solution to the problem of separation in logistic regression». In: *Statistics in medicine* 21.16 (2002), pp. 2409–2419.

[112] Henao, Alejandro. «Impacts of Ridesourcing-Lyft and Uber-on Transportation Including VMT, Mode Replacement, Parking, and Travel Behavior». PhD thesis. 2017.

[113] Hess, Stephane and Palma, David. «Apollo: A flexible, powerful and customisable freeware package for choice model estimation and application». In: *Journal of choice modelling* 32 (2019), p. 100170.

[114] Högel, Markus and Schröder, Axel. «Differences in Sharee Motivations to Participate in Car Sharing with Regard to Ownership Allocation». MA thesis. 2018.

[115] Holgado–Tello, Francisco Pablo, Chacón–Moscoso, Salvador, Barbero–Garcıa, Isabel, and Vila–Abad, Enrique. «Polychoric versus Pearson correlations in exploratory and confirmatory factor analysis of ordinal variables». In: *Quality & Quantity* 44.1 (2010), p. 153.

[116] Howe, Enrico and Bock, B. *Global Scootersharing Market Report 2018*. 2018.

[117] Hu, Songhua, Chen, Peng, Lin, Hangfei, Xie, Chi, and Chen, Xiaohong. «Promoting carsharing attractiveness and efficiency: An exploratory analysis». In: *Transportation Research Part D: Transport and Environment* 65 (2018), pp. 229–243.

[118] Hupp, R Craig. «Vanpool Travel Characteristics in Southeast Michigan». In: *Transportation Research Record* 823 (1981), pp. 15–17.

[119] Ilavarasan, P Vigneswara, Verma, Ravinder Kumar, and Kar, Arpan Kumar. «Sharing economy platforms as enablers of urban transport in the global south: Case of digital taxi aggregators in New Delhi, India». In: *Urban Transport in the Sharing Economy Era* (2018), p. 128.

[120] INEGI. *México en Cifras*. es. `https://en.www.inegi.org.mx/app/areasgeograficas/?ag=09#tabMCcollapse-Indicadores`, last accessed on 1/1/20. Jan. 1998.

[121] *Inicio | Sistema de Transporte Masivo y Teleférico*. = http://sitramytem.edomex.gob.mx/, last accessed on 1/1/20.

[122] James, Gareth, Witten, Daniela, Hastie, Trevor, and Tibshirani, Robert. *An introduction to statistical learning*. Vol. 112. Springer, 2013.

[123] Jirón, Paola. «Sustainable urban mobility in Latin America and the Caribbean». In: *Global Report on Human Settlements* (2013).

[124] Joo, Jae-Hun. «Motives for participating in sharing economy: Intentions to use car sharing services». In: *Journal of Distribution Science* 15.2 (2017), pp. 21–26.

[125] Jorge, Diana and Correia, Gonçalo. «Carsharing systems demand estimation and defined operations: a literature review». In: *European Journal of Transport and Infrastructure Research* 13.3 (2013).

[126] Kahle, David and Wickham, Hadley. «ggmap: Spatial Visualization with ggplot2». In: *The R Journal* 5.1 (2013), pp. 144–161. URL: `https://journal.r-project.org/archive/2013-1/kahle-wickham.pdf`.

[127] Kamargianni, Maria, Polydoropoulou, Amalia, and Goulias, Konstadinos G. «Teenager's Travel Patterns for School and After-School Activities.» In: *Procedia-Social and Behavioral Sciences* 48 (2012), pp. 3635–3650.

[128] Kang, Alexandra S, Jayaraman, K, Soh, Keng-Lin, and Wong, Wai Peng. «Convenience, flexible service, and commute impedance as the predictors of drivers' intention to switch and behavioral readiness to use public transport». In: *Transportation Research Part F: Traffic Psychology and Behaviour* 62 (2019), pp. 505–519.

[129] Kang, Junhee, Hwang, Keeyeon, and Park, Sungjin. «Finding factors that influence carsharing usage: Case study in seoul». In: *Sustainability* 8.8 (2016), p. 709.

[130] Karlaftis, Matthew G and McCarthy, Patrick S. «Subsidy and public transit performance: A factor analytic approach». In: *Transportation* 24.3 (1997), pp. 253–270.

[131] Khan, Mubassira and Machemehl, Randy. «The impact of land-use variables on free-floating carsharing vehicle rental choice and parking duration». In: *Seeing Cities Through Big Data*. Springer, 2017, pp. 331–347.

[132] Kim, Anna Joo, Brown, Anne, Nelson, Marla, Ehrenfeucht, Renia, Holman, Nancy, Gurran, Nicole, Sadowski, Jathan, Ferreri, Mara, Sanyal, Romola, Bastos, Marta, and Kresse, Klaas. «Planning and the So-Called 'Sharing' Economy / Can Shared Mobility Deliver Equity?/ The Sharing Economy and the Ongoing Dilemma about How to Plan for Informality/ Regulating Platform Economies in Cities – Disrupting the Disruption?/ Regulatory Combat? How the 'Sharing Economy' Is Disrupting Planning Practice/ Corporatised Enforcement: Challenges of Regulating AirBnB andOther Platform Economies/ Nurturing a Generative Sharing Economy for Local Public Goods and Service Provision». en. In: *Planning Theory & Practice* 20.2 (Mar. 2019), pp. 261–287. ISSN: 1464-9357, 1470-000X. DOI: 10.1080/14649357.2019.1599612.

[133] Kim, Daejin, Ko, Joonho, and Park, Yujin. «Factors affecting electric vehicle sharing program participants' attitudes about car ownership and program participation». In: *Transportation Research Part D: Transport and Environment* 36 (2015), pp. 96–106.

[134] Kim, DongJun, Shin, HeeCheol, Im, Hyeongjun, and Park, Junsik. «Factors influencing travel behaviors in bikesharing». In: *Transportation Research Board 91st Annual Meeting.* 2012.

[135] Kim, Jinhee, Rasouli, Soora, and Timmermans, Harry. «Hybrid choice models: principles and recent progress incorporating social influence and nonlinear utility functions». In: *Procedia Environmental Sciences* 22 (2014), pp. 20–34.

[136] «Pearson's Correlation Coefficient». In: *Encyclopedia of Public Health.* Ed. by Kirch, Wilhelm. Dordrecht: Springer Netherlands, 2008, pp. 1090–1091. ISBN: 978-1-4020-5614-7. DOI: 10.1007/978-1-4020-5614-7_2569. URL: https://doi.org/10.1007/978-1-4020-5614-7_2569.

[137] Le-Klähn, Diem-Trinh, Hall, C Michael, and Gerike, Regine. «Analysis of visitor satisfaction with public transport in Munich». In: *Journal of Public Transportation* 17.3 (2014), p. 5.

[138] Kohavi, Ron et al. «A study of cross-validation and bootstrap for accuracy estimation and model selection». In: *Ijcai.* Vol. 14. 2. Montreal, Canada. 1995, pp. 1137–1145.

[139] Koppelman, Frank S. «Predicting transit ridership in response to transit service changes». In: *Journal of Transportation Engineering* 109.4 (1983), pp. 548–564.

[140] Kuhn, Max. «Building predictive models in R using the caret package». In: *Journal of Statistical Software* 28.5 (2008), pp. 1–26.

[141] Kuhn, Max. *caret: Classification and Regression Training.* R package version 6.0-85. 2020. URL: https://CRAN.R-project.org/package=caret.

[142] Lavieri, Patrıcia S and Bhat, Chandra R. «Investigating objective and subjective factors influencing the adoption, frequency, and characteristics of ride-hailing trips». In: *Transportation Research Part C: Emerging Technologies* 105 (2019), pp. 100–125.

[143] Lempert, Rainer, Zhao, Jiaying, and Dowlatabadi, Hadi. «Convenience, savings, or lifestyle? Distinct motivations and travel patterns of one-way and two-way carsharing members in Vancouver, Canada». PhD thesis. University of British Columbia, Dec. 2018. DOI: 10.14288/1.0377753. URL: https://open.library.ubc.ca/cIRcle/collections/graduateresearch/42591/items/1.0377753.

[144] Li, Zuhe, Fan, Yangyu, Jiang, Bin, Lei, Tao, and Liu, Weihua. «A survey on sentiment analysis and opinion mining for social multimedia». In: *Multimedia Tools and Applications* 78.6 (2019), pp. 6939–6967.

[145] Likert, Rensis. «The likert-type scale». In: *Archives of Psychology* 140.55 (1932), pp. 1–55.

[146] Lin, Pengfei, Weng, Jiancheng, Liang, Quan, Alivanistos, Dimitrios, and Ma, Siyong. «Impact of Weather Conditions and Built Environment on Public Bikesharing Trips in Beijing». In: *Networks and Spatial Economics* (2018), pp. 1–17.

[147] Litman, Todd. «Valuing transit service quality improvements». In: *Journal of Public transportation* 11.2 (2008), p. 3.

[148] Litman, Todd. *You Can Get There From Here: Evaluating Transportation System Diversity*. en. Tech. rep. 2015, p. 42. URL: `https://www.vtpi.org/choice.pdf`.

[149] Liu, Xiaoyue Cathy, Zuo, Qian, Na, Shenruoyang, Wei, Ran, Golub, Aaron, Wang, Liming, and Davis, Jake. *Leveraging Twitter and Machine Learning for Real-Time Transit Network Evaluation*. Tech. rep. Portland, OR: Transportation Research and Education Center (TREC), 2019.

[150] Lu, Xuedong and Pas, Eric I. «Socio-demographics, activity participation and travel behavior». In: *Transportation Research Part A: Policy and Practice* 33.1 (1999), pp. 1–18.

[151] Ma, Liang, Seydel, John, Zhang, Xin, and Ding, Xiao Yan. «Users' recommendation intentions for shared bike services: A social exchange theory perspective». In: *International Journal of Sustainable Transportation* (2019), pp. 1–10.

[152] Ma, Liang, Zhang, Xin, Ding, Xiaoyan, and Wang, Gaoshan. «Bike sharing and users' subjective well-being: An empirical study in China». In: *Transportation Research Part A: Policy and Practice* 118 (2018), pp. 14–24.

[153] Malhotra, Arvind and Van Alstyne, Marshall. «The dark side of the sharing economy... and how to lighten it». In: *Communications of the ACM* 57.11 (2014), pp. 24–27.

[154] Mannucci, Pier Mannuccio and Franchini, Massimo. «Health effects of ambient air pollution in developing countries». In: *International Journal of Environmental Research and Public Health* 14.9 (2017), p. 1048.

[155] *Mapa Red de Rutas RTP*. es. `https://www.rtp.cdmx.gob.mx/red-de-rutas`, last accessed on 1/1/20.

[156] Marcińczak, Szymon and Bartosiewicz, Bartosz. «Commuting patterns and urban form: Evidence from Poland». In: *Journal of Transport Geography* 70 (2018).

[157] Marschak, J. «Binary choice constraints on random utility indications». In: *Stanford Symposium on Mathematical Methods in the Social Sciences*. Stanford University Press, 1960.

[158] Mattson, Jeremy and Godavarthy, Ranjit. «Bike share in Fargo, North Dakota: Keys to success and factors affecting ridership». In: *Sustainable cities and society* 34 (2017), pp. 174–182.

[159] McFadden, Daniel. *Conditional logit analysis of qualitative choice behavior*. 1973.

[160] McFadden, Daniel. «The choice theory approach to market research». In: *Marketing science* 5.4 (1986), pp. 275–297.

[161] McGuckin, Nancy and Srinivasan, Nandu. «The journey-to-work in the context of daily travel». In: *Census Data for Transportation Planning Conference*. 2005.

[162] McKenzie, Grant. «Spatiotemporal comparative analysis of scooter-share and bike-share usage patterns in Washington, DC». In: *Journal of transport geography* 78 (2019), pp. 19–28.

[163] Meenar, Mahbubur, Flamm, Bradley, and Keenan, Kevin. «Mapping the emotional experience of travel to understand cycle-transit user behavior». In: *Sustainability* 11.17 (2019), p. 4743.

[164] Mejía-Dorantes, Lucía and Soto Villagrán, Paula. «A review on the influence of barriers on gender equality to access the city: A synthesis approach of Mexico City and its Metropolitan Area». In: *Cities* 96 (Jan. 2020), p. 102439. ISSN: 02642751. DOI: 10.1016/j.cities.2019.102439. URL: https://linkinghub.elsevier.com/retrieve/pii/S0264275118309983 (visited on 01/30/2020).

[165] Mela, Carl F and Kopalle, Praveen K. «The impact of collinearity on regression analysis: the asymmetric effect of negative and positive correlations». In: *Applied Economics* 34.6 (2002), pp. 667–677.

[166] *Metro CDMX*. es. https://www.metro.cdmx.gob.mx, last accessed on 1/1/20.

[167] *Metrobús*. es. https://www.metrobus.cdmx.gob.mx, last accessed on 1/1/20.

[168] *Mexico City*. es. https://brtdata.org/location/latin_america/mexico/mexico_city/, last accessed on 1/1/20.

[169] Midi, Habshah, Sarkar, Saroje Kumar, and Rana, Sohel. «Collinearity diagnostics of binary logistic regression model». In: *Journal of Interdisciplinary Mathematics* 13.3 (2010), pp. 253–267.

[170] Mohammad, Saif M. and Turney, Peter D. «Crowdsourcing a Word-Emotion Association Lexicon». In: 29.3 (2013), pp. 436–465.

[171] Mohan, Dinesh and Tiwari, Geetam. «Mobility, environment and safety in megacities: dealing with a complex future». In: *IATSS research* 24.1 (2000), pp. 39–46.

[172] Morikawa, Taka, Ben-Akiva, Moshe, and McFadden, Daniel. «Discrete choice models incorporating revealed preferences and psychometric data». In: *Advances in Econometrics* 16 (2002), pp. 29–56.

[173] Motoaki, Yutaka and Daziano, Ricardo A. «Assessing Goodness of Fit of Hybrid Choice Models: An Open Research Question». In: *Transportation Research Record* 2495.1 (2015), pp. 131–141.

[174] Mourad, Abood, Puchinger, Jakob, and Chu, Chengbin. «A survey of models and algorithms for optimizing shared mobility». In: *Transportation Research Part B: Methodological* (2019).

[175] Mueller, Johannes, Schmoeller, Stefan, and Giesel, Flemming. «Identifying users and use of (electric-) free-floating carsharing in Berlin and Munich». In: *18th International Conference on Intelligent Transportation Systems*. IEEE. 2015, pp. 2568–2573.

[176] Müller, Johannes, Correia, Gonçalo Homem de Almeida, and Bogenberger, Klaus. «An explanatory model approach for the spatial distribution of free-floating carsharing bookings: A case-study of German cities». In: *Sustainability* 9.7 (2017), p. 1290.

[177] Murphy, Enda and Usher, Joe. «The role of bicycle-sharing in the city: Analysis of the Irish experience». In: *International Journal of Sustainable Transportation* 9.2 (2015), pp. 116–125.

[178] Nakov, Preslav, Ritter, Alan, Rosenthal, Sara, Sebastiani, Fabrizio, and Stoyanov, Veselin. «SemEval-2016 task 4: Sentiment analysis in Twitter». In: *arXiv preprint arXiv:1912.01973* (2019).

[179] Nazari, Fatemeh, Seyedabrishami, Seyedehsan, and Mamdoohi, Amir Reza. «A Direct Demand Model of Departure Time and Mode for Intercity Passenger Trips». en. In: *International Journal of Transportation Engineering* 3.2 (2015), p. 17.

[180] Neter, John, Kutner, Michael H, Nachtsheim, Christopher J, and Wasserman, William. *Applied linear statistical models.* Irwin Chicago, 1996.

[181] NEZU, Yuta and MIURA, Takao. «Extracting stopwords on social network service». In: *Information Modelling and Knowledge Bases XXXI* 321 (2020), p. 59.

[182] Nielsen, F. Å. *AFINN.* Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby: Informatics and Mathematical Modelling, Technical University of Denmark, Mar. 2011.

[183] Nikitas, Alexandros. «Understanding bike-sharing acceptability and expected usage patterns in the context of a small city novel to the concept: A story of 'Greek Drama'». In: *Transportation Research Part F: Traffic Psychology and Behaviour* 56 (2018), pp. 306–321.

[184] Nikitas, Alexandros, Wallgren, Pontus, and Rexfelt, Oskar. «The paradox of public acceptance of bike sharing in Gothenburg». In: *Proceedings of the Institution of Civil Engineers-Engineering Sustainability.* Vol. 169. 3. Thomas Telford Ltd. 2015, pp. 101–113.

[185] Noland, Robert B. «Trip patterns and revenue of shared e-scooters in Louisville, Kentucky». In: *Transport Findings* (Apr. 2019), p. 7747. DOI: `10.32866/7747`.

[186] Ogilvie, Flora and Goodman, Anna. «Inequalities in usage of a public bicycle sharing scheme: socio-demographic predictors of uptake and usage of the London (UK) cycle hire scheme». In: *Preventive Medicine* 55.1 (2012), pp. 40–45.

[187] Olaru, Doina, Smith, Brett, and Taplin, John HE. «Residential location and transit-oriented development in a new rail corridor». In: *Transportation Research Part A: Policy and Practice* 45.3 (2011), pp. 219–237.

[188] Onésimo Flores Dewey. «App-Based Collective Transport Service in Mexico City: A Start-Up Case Study». In: *International Transport Forum Discussion Papers* (2019).

[189] *OpenMobilityData.* `https://transitfeeds.com/search?q=mexico+city`, last accessed on 1/1/20.

[190] *Origin-Destination Survey in Households of the Metropolitan Zone of the Valley of Mexico (EOD) 2017.* http://en.www.inegi.org.mx/programas/eod/2017/, last accessed on 1/1/20.

[191] Paleti, Rajesh, Vovsha, Peter, Givon, Danny, and Birotker, Yehoshua. «Impact of individual daily travel pattern on value of time». In: *Transportation* 42.6 (2015), pp. 1003–1017.

[192] *Panorama profesional por estados | OLA*. `http://www.observatoriolaboral.gob.mx/static/estudios-publicaciones/Panorama_profesional_estados.html`, last accessed on 1/1/20.

[193] Paulssen, Marcel, Temme, Dirk, Vij, Akshay, and Walker, Joan L. «Values, attitudes and travel behavior: a hierarchical latent variable mixed logit model of travel mode choice». In: *Transportation* 41.4 (2014), pp. 873–888.

[194] Pedersen, Thomas Lin. *ggraph: An Implementation of Grammar of Graphics for Graphs and Networks*. R package version 2.0.0. 2019. URL: `https://CRAN.R-project.org/package=ggraph`.

[195] Prato, Carlo Giacomo, Bekhor, Shlomo, and Pronello, Cristina. «Latent variables and route choice behavior». In: *Transportation* 39.2 (2012), pp. 299–319.

[196] Provin, Daniel, Angerer, Peter, and Zimmermann, Steffen. «Economics of B2C Sharing Platforms». 2016.

[197] Pucher, John, Korattyswaroopam, Nisha, and Ittyerah, Neenu. «The crisis of public transport in India: overwhelming needs but limited resources». In: *Journal of Public Transportation* 7.4 (2004), p. 1.

[198] Puth, Marie-Therese, Neuhäuser, Markus, and Ruxton, Graeme D. «On the variety of methods for calculating confidence intervals by bootstrapping». In: *Journal of Animal Ecology* 84.4 (2015), pp. 892–897.

[199] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria, 2020. URL: `http://www.R-project.org/`.

[200] Raux, Charles, Zoubir, Ayman, and Geyik, Mirkan. «Who are bike sharing schemes members and do they travel differently? The case of Lyon's "Velo'v" scheme». In: *Transportation Research Part A: Policy and Practice* 106 (2017), pp. 350–363.

[201] Rayle, Lisa, Dai, Danielle, Chan, Nelson, Cervero, Robert, and Shaheen, Susan. «Just a better taxi? A survey-based comparison of taxis, transit, and ridesourcing services in San Francisco». In: *Transport Policy* 45 (2016), pp. 168–178.

[202] Rayle, Lisa, Shaheen, Susan A, Chan, Nelson, Dai, Danielle, and Cervero, Robert. *App-based, on-demand ride services: comparing taxi and ridesourcing trips and user characteristics in San Francisco*. Tech. rep. Citeseer, 2014.

[203] Revelle, William. *psych: Procedures for Psychological, Psychometric, and Personality Research*. R package version 1.9.12. Northwestern University. Evanston, Illinois, 2019. URL: `https://CRAN.R-project.org/package=psych`.

[204] Ribeiro, Filipe N, Araújo, Matheus, Gonçalves, Pollyanna, Gonçalves, Marcos André, and Benevenuto, Fabrıcio. «Sentibench-a benchmark comparison of state-of-the-practice sentiment analysis methods». In: *EPJ Data Science* 5.1 (2016), pp. 1–29.

[205] Ricci, Miriam. «Bike sharing: A review of evidence on impacts and processes of implementation and operation». In: *Research in Transportation Business & Management* 15 (2015), pp. 28–38.

[206] Rivadeneyra, Aldo Tudela, Dodero, Abel Lopez, Mehndiratta, Shomik Raj, Alves, Bianca Bianchi, and Deakin, Elizabeth. «Reducing Gender-Based Violence in Public Transportation: Strategy Design for Mexico City, Mexico». In: *Transportation Research Record* 2531.1 (2015), pp. 187–194.

[207] Roszak, Julie and Marechal, Florine. *The Importance of Environmental Sustainability in the Decision to Participate in the Sharing Economy.* 2017.

[208] Saif, Hassan, Fernández, Miriam, He, Yulan, and Alani, Harith. «On stopwords, filtering and data sparsity for sentiment analysis of twitter». In: *LREC 2014, Ninth International Conference on Language Resources and Evaluation. Proceedings* (2014), pp. 810–817.

[209] Santos, Georgina. «Sustainability and shared mobility models». In: *Sustainability* 10.9 (2018), p. 3194.

[210] Schaefers, Tobias. «Exploring carsharing usage motives: A hierarchical means-end chain analysis». In: *Transportation Research Part A: Policy and Practice* 47 (Jan. 2013), pp. 69–77. ISSN: 0965-8564. DOI: `10.1016/j.tra.2012.10.024`.

[211] Schlich, Robert and Axhausen, Kay W. «Habitual travel behaviour: evidence from a six-week travel diary». In: *Transportation* 30.1 (2003), pp. 13–36.

[212] Schor, Juliet et al. «Debating the sharing economy». In: *Journal of Self-Governance and Management Economics* 4.3 (2016), pp. 7–22.

[213] Schwarz, Gideon et al. «Estimating the dimension of a model». In: *The annals of statistics* 6.2 (1978), pp. 461–464.

[214] Shaheen, Susan. «Shared Mobility: The Potential of Ridehailing and Pooling». In: *Three Revolutions.* Springer, 2018. Chap. 3, pp. 55–76.

[215] Shaheen, Susan A., Chan, Nelson D., and Gaynor, Teresa. «Casual carpooling in the San Francisco Bay Area: Understanding user characteristics, behaviors, and motivations». In: *Transport Policy.* Transit Investment and Land Development. Edited by Xinyu (Jason) Cao and Qisheng Pan Shared Use Mobility Innovations. Edited by Susan Shaheen 51 (Oct. 2016), pp. 165–173. ISSN: 0967-070X. DOI: `10.1016/j.tranpol.2016.01.003`.

[216] Shaheen, Susan A, Cohen, Adam P, Zohdy, Ismail H, Kock, Beaudry, et al. *Smartphone applications to influence travel choices: practices and policies.* Tech. rep. United States. Federal Highway Administration, 2016.

[217] Shaheen, Susan A, Guzman, Stacey, and Zhang, Hua. «Bikesharing in Europe, the Americas, and Asia: past, present, and future». In: *Transportation Research Record* 2143.1 (2010), pp. 159–167.

[218] Shaheen, Susan A, Martin, Elliot W, Cohen, Adam P, Chan, Nelson D, and Pogodzinski, Mike. *Public Bikesharing in North America During a Period of Rapid Expansion: Understanding Business Models, Industry Trends & User Impacts, MTI Report 12-29.* Tech. rep. Mineta Transportation Institute, 2014.

[219] Shaheen, Susan A, Zhang, Hua, Martin, Elliot, and Guzman, Stacey. «China's Hangzhou public bicycle: understanding early adoption and behavioral response to bikesharing». In: *Transportation Research Record* 2247.1 (2011), pp. 33–41.

[220] Shaheen, Susan, Bell, Corwin, Cohen, Adam, and Yelchuru, Balaji. *Travel behavior: Shared mobility and transportation equity.* Tech. rep. 2017.

[221] Shaheen, Susan, Chan, Nelson, Bansal, Apaar, and Cohen, Adam. *Shared mobility: A sustainability & technologies workshop: definitions, industry developments, and early understanding.* Tech. rep. UC Berkeley: Transportation Sustainability Research Center, 2015.

[222] Shaheen, Susan and Cohen, Adam. *Shared Micromoblity Policy Toolkit: Docked and Dockless Bike and Scooter Sharing.* Tech. rep. UC Berkeley: Transportation Sustainability Research Center, 2019.

[223] Shaheen, Susan and Cohen, Adam. *Shared Mobility Policy Briefs: Definitions, Impacts, and Recommendations.* Tech. rep. University of California Institute of Transportation Studies, 2018.

[224] Shaheen, Susan, Cohen, Adam, and Bayen, Alexandre. *The Benefits of Carpooling.* Tech. rep. Oct. 22, 2018. DOI: `10.7922/G2DZ06GF`. URL: `https://escholarship.org/uc/item/7jx6z631` (visited on 04/14/2020).

[225] Shaheen, Susan, Cohen, Adam, Zohdy, Ismail, et al. *Shared mobility: current practices and guiding principles.* Tech. rep. United States. Federal Highway Administration, 2016.

[226] Shaheen, Susan and Martin, Elliot. «Unraveling the Modal Impacts of Bikesharing». en. In: *ACCESS Magazine* 47 (2015), p. 9.

[227] Shaheen, Susan, Stocker, Adam, Lazarus, Jessica, and Bhattacharyya, Abhinav. *RideKC: Bridj Pilot evaluation: impact, operational, and institutional analysis.* Tech. rep. 2016.

[228] Shaheen, Susan, Stocker, Adam, and Mundler, Marie. *Online and app-based carpooling in France: Analyzing users and practices—A study of BlaBlaCar.* Tech. rep. University of California Institute of Transportation Studies, 2017.

[229] Shared and Digital Mobility Committee. *Taxonomy and Definitions for Terms Related to Shared Mobility and Enabling Technologies.* en. Tech. rep. SAE International, Sept. 2018. DOI: `10.4271/J3163_201809`.

[230] Sheinbaum, Claudia. *Plan estratégico de movilidadde la Ciudad de México 2019 Una ciudad, un sistema. [Strategic mobility plan of Mexico City 2019 One city, one system.]* 2018. URL: `http://claudiasheinbaum.org/files/Plan_de_movilidad.pdf`.

[231] Shen, Yu, Zhang, Xiaohu, and Zhao, Jinhua. «Understanding the usage of dockless bike sharing in Singapore». In: *International Journal of Sustainable Transportation* 12.9 (2018), pp. 686–700.

[232] Silge, Julia and Robinson, David. «tidytext: Text mining and analysis using tidy data principles in R». In: *Journal of Open Source Software* 1.3 (2016), p. 37.

[233] *Slug-Lines.Com - Slugging and Slug Lines Information For Washington DC.* `http://www.slug-lines.com/`, last accessed on 11/12/19.

[234] Souza Silva, Laize Andréa de, Andrade, Maurıcio Oliveira de, and Maia, Maria Leonor Alves. «How does the ride-hailing systems demand affect individual transport regulation?» In: *Research in Transportation Economics* 69 (2018), pp. 600–606.

[235] Spearman, Charles. «" General Intelligence" Objectively Determined and Measured.» In: *American Journal of Psychology* 15.2 (1904), pp. 201–293.

[236] Sperling, Daniel and Clausen, Eileen. «The developing world's motorization challenge». In: *Issues in Science and Technology* 19.1 (2002), pp. 59–66.

[237] Stiglic, Mitja, Agatz, Niels, Savelsbergh, Martin, and Gradisar, Mirko. «The benefits of meeting points in ride-sharing systems». In: *Transportation Research Part B: Methodological* 82 (2015), pp. 36–53.

[238] Stillwater, Tai, Mokhtarian, Patricia L, and Shaheen, Susan A. «Carsharing and the built environment: Geographic information system–based study of one US operator». In: *Transportation Research Record* 2110.1 (2009), pp. 27–34.

[239] Su, Qing and Zhou, Liren. «Parking management, financial subsidies to alternatives to drive alone and commute mode choices in Seattle». In: *Regional Science and Urban Economics* 42.1-2 (2012), pp. 88–97.

[240] Sun, Yeran, Mobasheri, Amin, Hu, Xuke, and Wang, Weikai. «Investigating impacts of environmental factors on the cycling behavior of bicycle-sharing users». In: *Sustainability* 9.6 (2017), p. 1060.

[241] Susan Shaheen, Chan, Nelson, Apaar Bansal, and Adam Cohen. *Shared Mobility A Sustainbility & Technologies Workshop, Definitions, Industry Development and Early Understanding.* Nov. 2015.

[242] Syed, Sharfuddin I and Khan, Ata M. «Factor analysis for the study of determinants of public transit ridership». In: *Journal of Public Transportation* 3.3 (2000), p. 1.

[243] Tahmasseby, Shahram, Kattan, Lina, and Barbour, Brian. «Propensity to participate in a peer-to-peer social-network-based carpooling system». In: *Journal of Advanced Transportation* 50.2 (2016), pp. 240–254.

[244] Tang, Bao-Jun, Li, Xiao-Yi, Yu, Biying, and Wei, Yi-Ming. «How app-based ride-hailing services influence travel behavior: An empirical study from China». In: *International Journal of Sustainable Transportation* (2019), pp. 1–15.

[245] Temme, Dirk, Paulssen, Marcel, and Dannewald, Till. *Integrating latent variables in discrete choice models–How higher-order values and attitudes determine consumer choice.* Tech. rep. Berlin : Humboldt-Universität, 2007.

[246] Tenenboim, Einat and Shiftan, Yoram. «Accuracy and bias of subjective travel time estimates». In: *Transportation* 45.3 (2018), pp. 945–969.

[247] Teubner, Timm and Flath, Christoph M. «The economics of multi-hop ride sharing». In: *Business & Information Systems Engineering* 57.5 (2015), pp. 311–324.

[248] Tirachini, Alejandro. «Ride-hailing, travel behaviour and sustainable mobility: an international review». In: *Transportation* (Nov. 2019). ISSN: 0049-4488, 1572-9435. DOI: `10.1007/s11116-019-10070-2`. URL: `http://link.springer.com/10.1007/s11116-019-10070-2`.

[249] Tirachini, Alejandro and Gomez-Lobo, Andres. «Does ride-hailing increase or decrease vehicle kilometers traveled (VKT)? A simulation approach for Santiago de Chile». In: *International Journal of Sustainable Transportation* 14.3 (2020), pp. 187–204.

[250] Tirachini, Alejandro and Río, Mariana del. «Ride-hailing in Santiago de Chile: Users' characterisation and effects on travel behaviour». In: *Transport Policy* 82 (2019), pp. 46–57.

[251] Train, Kenneth E. *Discrete choice methods with simulation*. Cambridge university press, 2009.

[252] Train, Kenneth E, McFadden, Daniel L, and Goett, Andrew A. «Consumer attitudes and voluntary rate schedules for public utilities». In: *The review of economics and statistics* (1987), pp. 383–391.

[253] Transportation Research Board and National Academies of Sciences, Engineering, and Medicine. *Operational Experiences with Flexible Transit Services*. en. Washington, D.C.: Transportation Research Board, Aug. 2004. ISBN: 978-0-309-42780-7. DOI: 10.17226/23364.

[254] *Tren Ligero*. https://www.ste.cdmx.gob.mx/tren-ligero, last accessed on 1/1/20.

[255] *Trolebús*. https://www.ste.cdmx.gob.mx/red-de-servicio/lineas-de-trolebus, last accessed on 1/1/20.

[256] Tsekeris, Theodore and Tsekeris, Charalambos. «Demand forecasting in transport: Overview and modeling advances». In: *Economic research-Ekonomska istraživanja* 24.1 (2011), pp. 82–94.

[257] Tsirimpa, Athena, Polydoropoulou, Amalia, and Antoniou, Constantinos. «Development of a latent variable model to capture the impact of risk aversion on travelers' switching behavior». In: *Journal of Choice Modelling* 3.1 (2010), pp. 127–148.

[258] United Nations Department of Economic and Social Affairs. *The World's Cities in 2016*. en. Statistical Papers - United Nations (Ser. A), Population and Vital Statistics Report. UN, Sept. 2016. DOI: 10.18356/8519891f-en. URL: https://www.un-ilibrary.org/population-and-demography/the-world-s-cities-in-2016_8519891f-en (visited on 12/27/2019).

[259] Vanoutrive, Thomas, Van De Vijver, Elien, Van Malderen, Laurent, Jourquin, Bart, Thomas, Isabelle, Verhetsel, Ann, and Witlox, Frank. «What determines carpooling to workplaces in Belgium: location, organisation, or promotion?» In: *Journal of Transport Geography* 22 (2012), pp. 77–86.

[260] Vatcheva, Kristina P, Lee, MinJae, McCormick, Joseph B, and Rahbar, Mohammad H. «Multicollinearity in regression analyses conducted in epidemiologic studies». In: *Epidemiology (Sunnyvale, Calif.)* 6.2 (2016).

[261] Velicer, Wayne F and Jackson, Douglas N. «Component analysis versus common factor analysis: Some further observations». In: *Multivariate Behavioral Research* 25.1 (1990), pp. 97–114.

[262] Venables, W. N. and Ripley, B. D. *Modern Applied Statistics with S*. Fourth. ISBN 0-387-95457-0. New York: Springer, 2002. URL: http://www.stats.ox.ac.uk/pub/MASS4.

[263] Verboven, Hans and Vanherck, Lise. «The sustainability paradox of the sharing economy». en. In: *uwf- UmweltWirtschaftsForum* 24.4 (Dec. 2016), pp. 303–314. ISSN: 1432-2293. DOI: 10.1007/s00550-016-0410-y. URL: https://doi.org/10.1007/s00550-016-0410-y.

[264] Vij, Akshay and Walker, Joan L. «How, when and why integrated choice and latent variable models are latently useful». In: *Transportation Research Part B: Methodological* 90 (2016), pp. 192–217.

[265] Wan, Wan Nor Ainun Al Baiyah, Mohamad, Anis Fatihah Mahamad Fuad, Shahib, Nor Syafiqah, Azmi, Azila, Kamal, Saiful Bahari Mohd, and Abdullah, Dahlan. «A framework of customer's intention to use Uber service in tourism destination». In: *International Academic Research Journal of Business and Technology* 2.2 (2016), pp. 102–106.

[266] Wang, Chao, Quddus, Mohammed, Enoch, Marcus, Ryley, Tim, and Davison, Lisa. «The impact of socio-economic factors on the demand for demand responsive transport (DRT) in Greater Manchester: 2 a multilevel modelling approach 3». In: *Transport* 8 (2012), p. 9.

[267] Wang, Yu, Wang, Shanyong, Wang, Jing, Wei, Jiuchang, and Wang, Chenglin. «An empirical study of consumers' intention to use ride-sharing services: using an extended technology acceptance model». In: *Transportation* (2018), pp. 1–19.

[268] Washbrook, Kevin, Haider, Wolfgang, and Jaccard, Mark. «Estimating commuter mode choice: A discrete choice analysis of the impact of road pricing and parking charges». In: *Transportation* 33.6 (2006), pp. 621–639.

[269] Washington, Simon, Karlaftis, Matthew G, Mannering, Fred, and Anastasopoulos, Panagiotis. *Statistical and econometric methods for transportation data analysis*. CRC press, 2020.

[270] Wickham, Hadley, Averick, Mara, Bryan, Jennifer, Chang, Winston, McGowan, Lucy D'Agostino, François, Romain, Grolemund, Garrett, Hayes, Alex, Henry, Lionel, Hester, Jim, Kuhn, Max, Pedersen, Thomas Lin, Miller, Evan, Bache, Stephan Milton, Müller, Kirill, Ooms, Jeroen, Robinson, David, Seidel, Dana Paige, Spinu, Vitalie, Takahashi, Kohske, Vaughan, Davis, Wilke, Claus, Woo, Kara, and Yutani, Hiroaki. «Welcome to the tidyverse». In: *Journal of Open Source Software* 4.43 (2019), p. 1686. DOI: 10.21105/joss.01686.

[271] Wilhelms, Mark-Philipp, Henkel, Sven, and Falk, Tomas. «To earn is not enough: A means-end analysis to uncover peer-providers' participation motives in peer-to-peer carsharing». In: *Technological Forecasting and Social Change* 125 (2017), pp. 38–47.

[272] Willing, Christoph, Klemmer, Konstantin, Brandt, Tobias, and Neumann, Dirk. «Moving in time and space–Location intelligence for carsharing decision support». In: *Decision Support Systems* 99 (2017), pp. 75–85.

[273] Wu, Xuemei and Zhi, Qiang. «Impact of Shared Economy on Urban Sustainability: From the Perspective of Social, Economic, and Environmental Sustainability». en. In: *Energy Procedia*. Clean Energy for Clean City: CUE 2016–Applied Energy Symposium and Forum: Low-Carbon Cities and Urban Energy Systems 104 (Dec. 2016), pp. 191–196. ISSN: 1876-6102. DOI: 10.1016/j.egypro.2016.12.033. URL: http://www.sciencedirect.com/science/article/pii/S1876610216315892 (visited on 11/23/2019).

[274] Yáñez, Marıa F, Raveau, Sebastián, and Ortúzar, J de D. «Inclusion of latent variables in mixed logit models: modelling and forecasting». In: *Transportation Research Part A: Policy and Practice* 44.9 (2010), pp. 744–753.

[275] Yoon, Taekwan, Cherry, Christopher R, and Jones, Luke R. «One-way and round-trip carsharing: A stated preference experiment in Beijing». In: *Transportation Research Part D: transport and environment* 53 (2017), pp. 102–114.

[276] Young, Mischa, Allen, Jeff, and Farber, Steven. «Measuring when Uber behaves as a substitute or supplement to transit: An examination of travel-time differences in Toronto». In: *Journal of Transport Geography* 82 (2020), p. 102629.

[277] Young, Mischa and Farber, Steven. «The who, why, and when of Uber and other ride-hailing trips: An examination of a large sample household travel survey». In: *Transportation Research Part A: Policy and Practice* 119 (2019), pp. 383–392.

[278] Zervas, Georgios, Proserpio, Davide, and Byers, John W. «The rise of the sharing economy: Estimating the impact of Airbnb on the hotel industry». In: *Journal of marketing research* 54.5 (2017), pp. 687–705.

[279] Zhang, Zhongheng. «Model building strategy for logistic regression: purposeful selection». In: *Annals of translational medicine* 4.6 (2016).

[280] Zhen, Chen. «Impact of ride-sourcing services on travel habits and transportation planning». PhD thesis. University of Pittsburgh, 2015.

[281] Zuo, Ting, Wei, Heng, Liu, Hao, and Yang, Y Jeffrey. «Bi-level optimization approach for configuring population and employment distributions with minimized vehicle travel demand». In: *Journal of Transport Geography* 74 (2019), pp. 161–172.

Additional Analysis

Table A.1 – Combined Modes to Replace Last Jetty Trip

Rank	Replacement.total	Count	Pct.%	Rank	Replacement.total	Count	Pct.%
1	Metro+Microbus	482	19.4%	28	Bus+Taxi	16	0.6%
2	Car	283	11.4%	29	Shared-taxi+Taxi	16	0.6%
3	Microbus	130	5.2%	30	Bus+Shared-taxi+Car	14	0.6%
4	Taxi	130	5.2%	31	Bus+Microbus+Car	13	0.5%
5	Metro+Microbus+Taxi	128	5.2%	32	Metro+Bus+Shared-taxi	12	0.5%
6	Taxi+Car	118	4.8%	33	Shared-taxi+Taxi+Car	12	0.5%
7	Metro+Bus+Microbus	102	4.1%	34	Metro+Shared-taxi+Taxi	11	0.4%
8	Metro	88	3.5%	35	Microbus+Shared-taxi+Taxi	10	0.4%
9	Metro+Microbus+Shared-taxi	83	3.3%	36	Shared-taxi	10	0.4%
10	Metro+Microbus+Car	81	3.3%	37	Bus+Microbus+Shared-taxi	7	0.3%
11	Metro+Bus	66	2.7%	38	Bus+Shared-taxi+Taxi	7	0.3%
12	Bus+Microbus	65	2.6%	39	Bus+Shared-taxi	6	0.2%
13	Metro+Taxi	57	2.3%	40	Car+Motorcycle	6	0.2%
14	Bus	50	2.0%	41	Motorcycle	5	0.2%
15	Microbus+Car	46	1.9%	42	Car+Taxi+Motorcycle	4	0.2%
16	Metro+Bus+Taxi	41	1.7%	43	Shared-taxi+Car	4	0.2%
17	Metro+Car	41	1.7%	44	Microbus+Shared-taxi+Car	3	0.1%
18	Metro+Shared-taxi	40	1.6%	45	Taxi+Motorcycle	3	0.1%
19	Microbus+Taxi	37	1.5%	46	Metro+Microbus+Motorcycle	2	0.1%
20	Microbus+Taxi+Car	37	1.5%	47	Bus+Microbus+Motorcycle	1	0.0%
21	Metro+Taxi+Car	36	1.4%	48	Metro+Car+Motorcycle	1	0.0%
22	Bus+Taxi+Car	27	1.1%	49	Metro+Shared-taxi+Motorcycle	1	0.0%
23	Bus+Microbus+Taxi	23	0.9%	50	Microbus+Car+Motorcycle	1	0.0%
24	Metro+Bus+Car	22	0.9%	51	Microbus+Car+Shared-taxi	1	0.0%
25	Bus+Car	20	0.8%	52	Microbus+Metro+Motorcycle	1	0.0%
26	Metro+Shared-taxi+Car	20	0.8%	53	Microbus+Motorcycle	1	0.0%
27	Microbus+Shared-taxi	17	0.7%	54	Shared-taxi+Bus+Motorcycle	1	0.0%

Table A.2 – Access and Egress Time by Mode, Top Five Modes Summary Statistics

Mode	Access Time						Egress Time					
	Count	Pct.%	Avg	SD	Min	Max	Count	Pct.%	Avg	SD	Min	Max
Walking	866	34.9%	13.0	13.4	1.0	120.0	1,418	57.1%	9.8	8.6	0.0	100.0
Car	384	15.5%	16.0	12.1	2.0	120.0	88	3.5%	21.0	25.1	1.0	202.0
Microbus	315	12.7%	24.9	14.7	5.0	80.0	327	13.2%	22.4	15.8	2.0	130.0
Taxi	309	12.4%	17.1	11.7	0.0	120.0	220	8.9%	16.2	8.2	1.0	60.0
Combi	128	5.2%	24.6	19.8	1.0	120.0	73	2.9%	25.5	19.4	2.0	120.0
Other	482	19.4%	23.1	14.8	1.0	120.0	358	14.4%	22.3	19.4	1.0	202.0

Taxi includes taxi and ride-hailing combined

Table A.3 – Disaggregated Activities During Jetty Trip Summary Statistics

Rank	Activity	Pct%
1	Sleeping	75.50%
2	Use of Smart Phone	72.50%
3	Look out of the Window	33.90%
4	Reading for Pleasure	22.00%
5	Talk on Phone	14.50%
6	Working	13.70%
7	Studying	7.30%
8	Talk to Other Passengers	6.20%
9	Others	0.90%

Table A.4 – Disaggregated Reasons to Use Jetty Summary Statistics

Rank	Reason	Pct.%
1	Booking of seat	69.30%
2	Security against theft	67.10%
3	Travel time (trip duration)	65.80%
4	Travel time reliability	52.70%
5	Access/Egress time	42.40%
6	Quality of vehicle	36.10%
7	Fare	35.6%
8	Ease of payment	27.50%
9	Use of time while travelling	25.30%
10	Fare transparency	20.70%
11	Traffic safety	20.10%
12	Security against harassment	15.50%
13	Avoid parking problems	11.40%
14	Socializing with other passengers	0.70%

Table A.5 – Modes Use Frequency Summary Statistics

Mode	Level	Count	Pct. %	Mode	Level	Count	Pct. %
Car as a Driver				**RTP**			
	Never	1043	41.99		Never	1681	67.67
	Less than once a month	271	10.91		Less than once a month	326	13.12
	1-3 times a month	378	15.22		1-3 times a month	230	9.26
	1-3 times a week	394	15.86		1-3 times a week	103	4.15
	4 or more times a week	398	16.02		4 or more times a week	144	5.8
Car as a Passenger				**Bus**			
	Never	710	28.58		Never	1430	57.57
	Less than once a month	385	15.5		Less than once a month	333	13.41
	1-3 times a month	537	21.62		1-3 times a month	266	10.71
	1-3 times a week	375	15.1		1-3 times a week	150	6.04
	4 or more times a week	477	19.20		4 or more times a week	305	12.28
Taxi				**Microbus**			
	Never	1153	46.42		Never	1079	43.44
	Less than once a month	537	21.62		Less than once a month	308	12.4
	1-3 times a month	395	15.9		1-3 times a month	327	13.16
	1-3 times a week	212	8.53		1-3 times a week	244	9.82
	4 or more times a week	187	7.53		4 or more times a week	526	21.18
E-Hailing				**Combi**			
	Never	299	12.04		Never	1500	60.39
	Less than once a month	516	20.77		Less than once a month	289	11.63
	1-3 times a month	801	32.25		1-3 times a month	204	8.21
	1-3 times a week	495	19.93		1-3 times a week	151	6.08
	4 or more times a week	373	15.02		4 or more times a week	340	13.69
Motorcycle				**Jetty like Service**			
	Never	2321	93.44		Never	457	18.4
	Less than once a month	69	2.78		Less than once a month	383	15.42
	1-3 times a month	32	1.29		1-3 times a month	443	17.83
	1-3 times a week	31	1.25		1-3 times a week	374	15.06
	4 or more times a week	31	1.25		4 or more times a week	827	33.29
Metro				**Shared App Vehicle**			
	Never	659	26.53		Never	1413	56.88
	Less than once a month	450	18.12		Less than once a month	371	14.94
	1-3 times a month	448	18.04		1-3 times a month	325	13.08
	1-3 times a week	254	10.23		1-3 times a week	189	7.61
	4 or more times a week	673	27.09		4 or more times a week	186	7.49
Metrobus				**Bike**			
	Never	1113	44.81		Never	2070	83.33
	Less than once a month	532	21.42		Less than once a month	166	6.68
	1-3 times a month	379	15.26		1-3 times a month	99	3.99
	1-3 times a week	164	6.6		1-3 times a week	68	2.74
	4 or more times a week	296	11.92		4 or more times a week	81	3.26
Light Rail				**Shared Bike**			
	Never	1935	77.9		Never	2335	94
	Less than once a month	281	11.31		Less than once a month	46	1.85
	1-3 times a month	120	4.83		1-3 times a month	37	1.49
	1-3 times a week	48	1.93		1-3 times a week	22	0.89
	4 or more times a week	100	4.03		4 or more times a week	44	1.77
Suburban Train				**Shared Scooter**			
	Never	2174	87.52		Never	2305	92.79
	Less than once a month	151	6.08		Less than once a month	78	3.14
	1-3 times a month	69	2.78		1-3 times a month	58	2.33
	1-3 times a week	30	1.21		1-3 times a week	23	0.93
	4 or more times a week	60	2.42		4 or more times a week	20	0.81
Trolleybus				**Walking**			
	Never	2201	88.61		Never	737	29.67
	Less than once a month	187	7.53		Less than once a month	176	7.09
	1-3 times a month	60	2.42		1-3 times a month	286	11.51
	1-3 times a week	21	0.85		1-3 times a week	288	11.59
	4 or more times a week	15	0.60		4 or more times a week	997	40.14

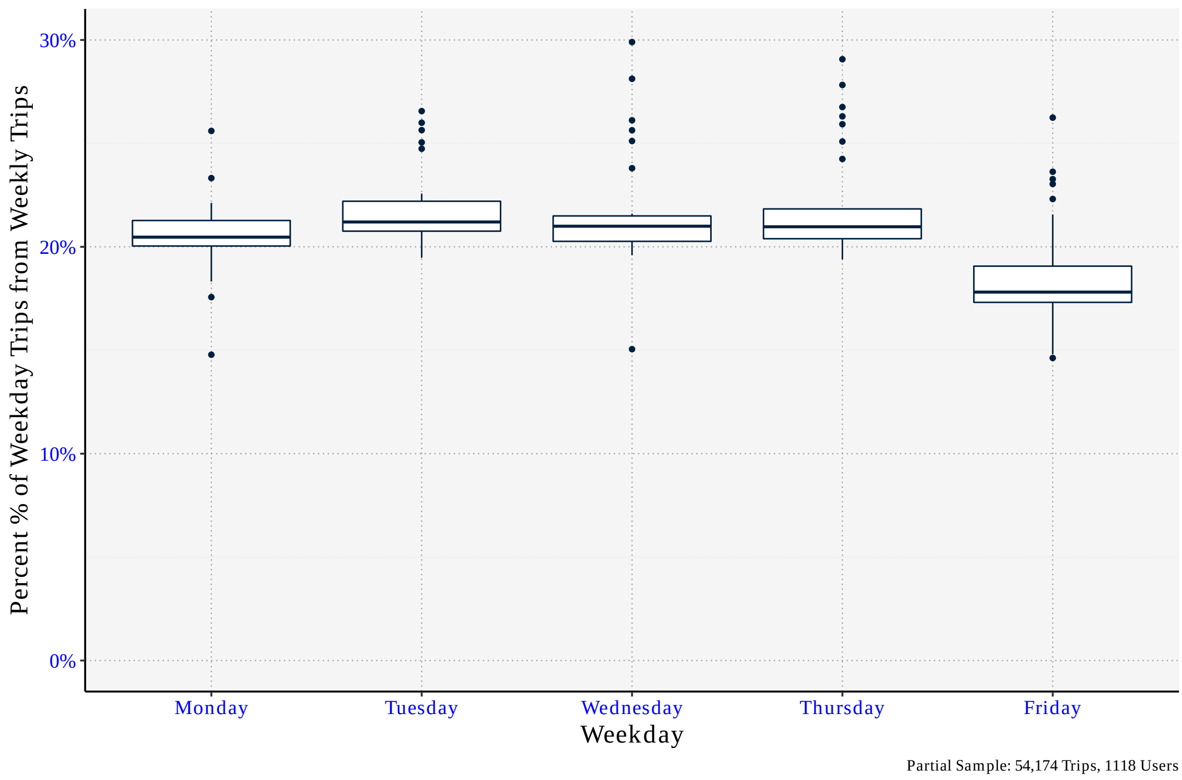

Figure A.1 – Percentage of Jetty Trips per Workday

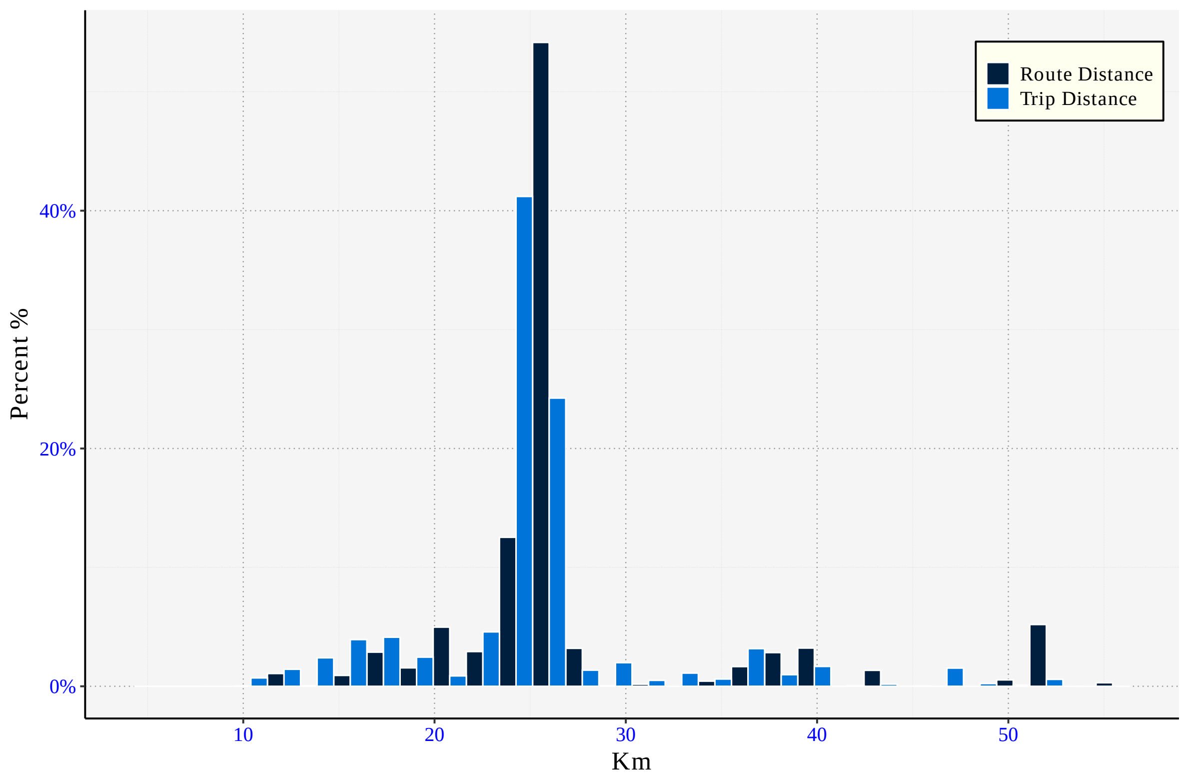

Figure A.2 – Jetty Route Distance and Trip Distance Distribution

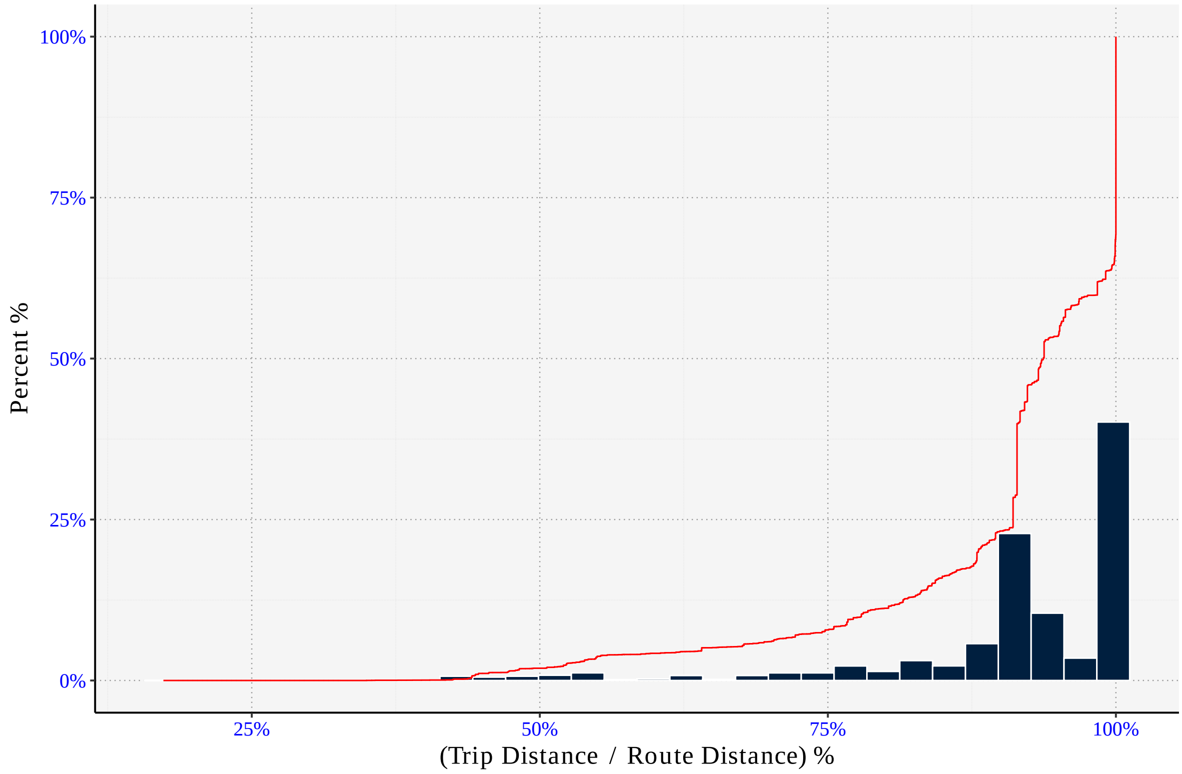

Figure A.3 – Ratio of Jetty Trip Distance to Route Distance, CDF

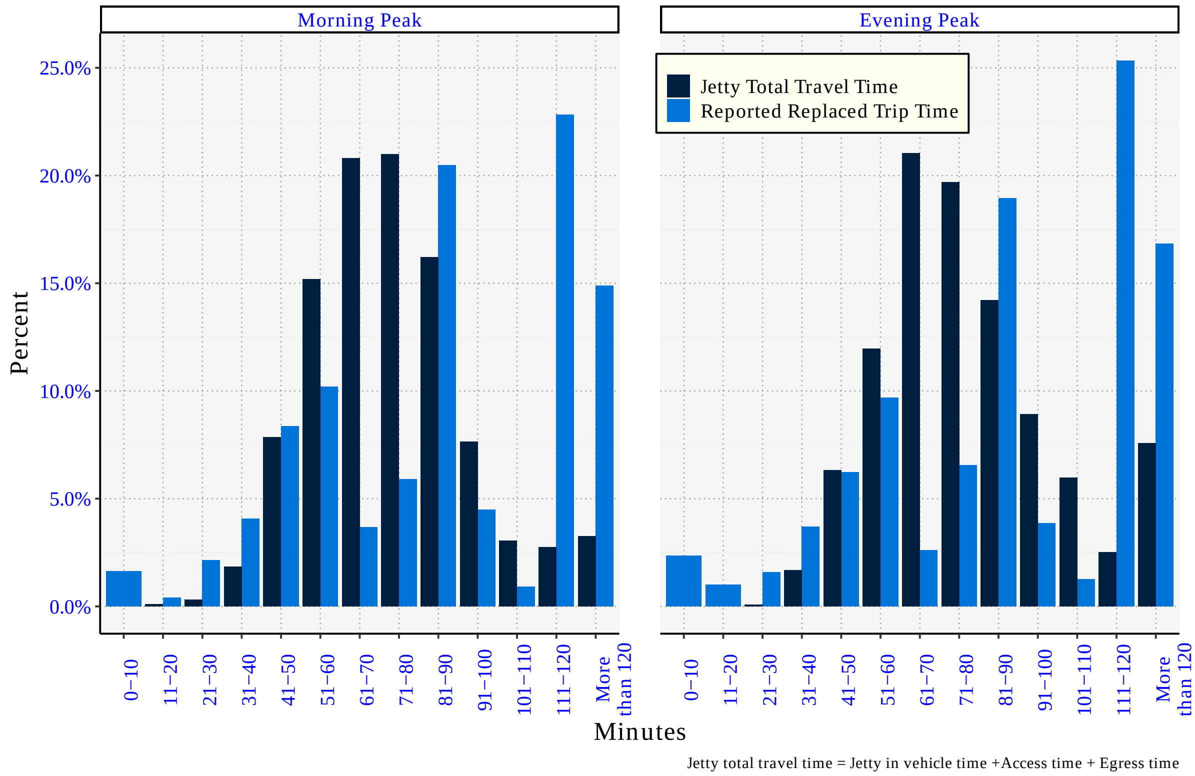

Figure A.4 – Jetty and Replaced Trip Travel Time Distribution by Peak Hour

Table A.6 – Shift to Jetty Users' Sentiment Score Summary Statistics

Group	Trip	Mean	SD	Min	Q1	Median	Q3	Max	Word list
All	Car	3.43	4.24	-5	1	3	5	39	AFINN
Femal	Car	3.72	3.64	-5	2	3	5	18	AFINN
Male	Car	3.03	4.95	-3	1	2.5	4	39	AFINN
All	No-Car	3.2	3.59	-3	1	3	5	20	AFINN
Femal	No-Car	3.93	4.13	-3	1	3	6	20	AFINN
Male	No-Car	2.53	2.86	-3	0.5	3	5	10	AFINN
All	Car	1.55	1.89	-4	1	1	2	8	Emolex
Femal	Car	1.62	1.78	-4	1	1	2	6	Emolex
Male	Car	1.46	2.05	-3	1	1	2	8	Emolex
All	No-Car	1.3	1.74	-2	1	1	2	9	Emolex
Femal	No-Car	1.66	1.96	-2	1	1	2	9	Emolex
Male	No-Car	0.97	1.44	-2	0	1	1	6	Emolex

Table A.7 – Jetty Service Choice Users' Sentiment Score Summary Statistics

Users' Group	Service	Mean	Std.Dev	Min	Q1	Median	Q3	Max	Word List
All	Van	3.03	4.62	-5	1	3	5	39	AFINN
Female	Van	3.12	3.21	-5	1	3	5	12	AFINN
Male	Van	2.93	5.89	-3	0	2	4	39	AFINN
All	Bus	3.52	3.55	-5	1	3	5	20	AFINN
Female	Bus	4.16	4.13	-5	2	3	6	20	AFINN
Male	Bus	2.76	2.51	-3	1	3	5	10	AFINN
All	Van	1.43	1.86	-2	1	1	2	8	Emolex
Female	Van	1.52	1.65	-1	1	1	2	6	Emolex
Male	Van	1.33	2.09	-2	0	1	2	8	Emolex
All	Bus	1.44	1.81	-4	1	1	2	9	Emolex
Female	Bus	1.68	1.95	-4	1	2	3	9	Emolex
Male	Bus	1.15	1.58	-3	1	1	2	6	Emolex

Table A.8 – Jetty Use Frequency Users' Sentiment Score Summary Statistics

Frequency of Use	Mean	Std.Dev	Min	Q1	Median	Q3	Max	Word list
Less than once a month	3.85	5.39	-3	0	3	6	20	AFINN
1-3 times a month	2.65	3.14	-2	1	2	4	18	AFINN
1-3 times per week	3.07	4.8	-5	1	3	5	39	AFINN
more than 3 times per week	3.77	3.29	-5	2	3	5	15	AFINN
Less than once a month	1.75	2.72	-2	0.5	1	2	9	Emolex
1-3 times a month	1.13	1.71	-3	0	1	2	6	Emolex
1-3 times per week	1.2	1.55	-2	1	1	2	8	Emolex
More than 3 times per week	1.74	1.84	-4	1	1	2	7	Emolex

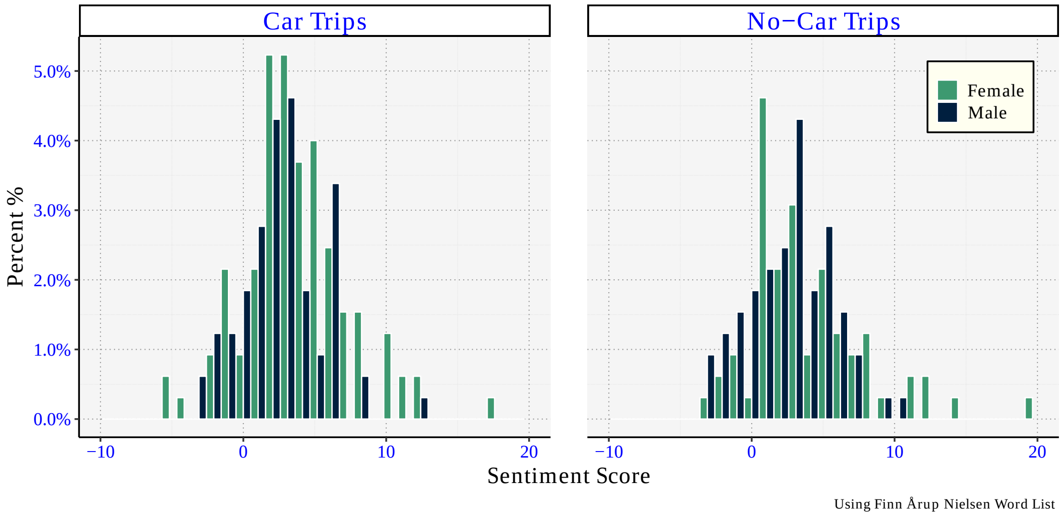

Figure A.5 – Shift to Jetty Users' Sentiment Score

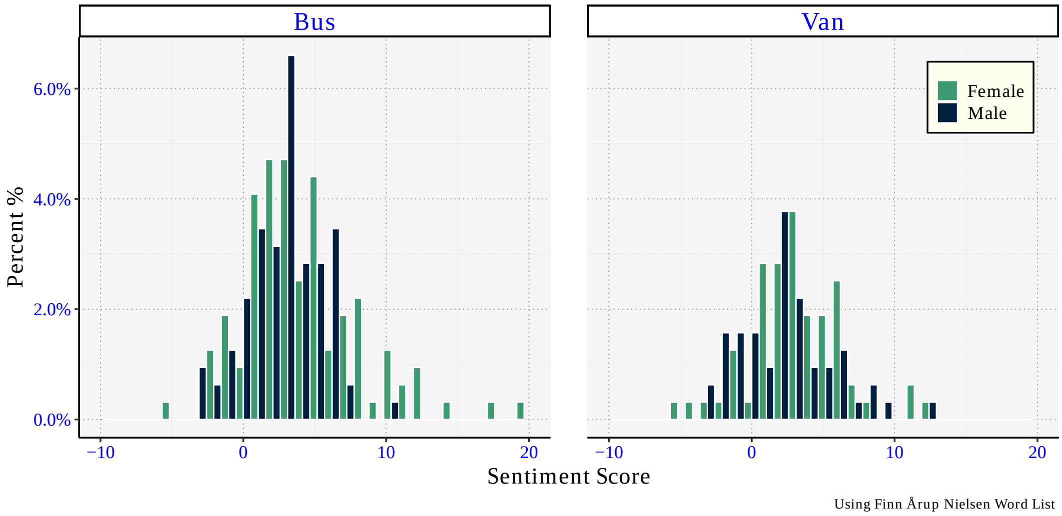

Figure A.6 – Jetty Service Choice Users' Sentiment Score

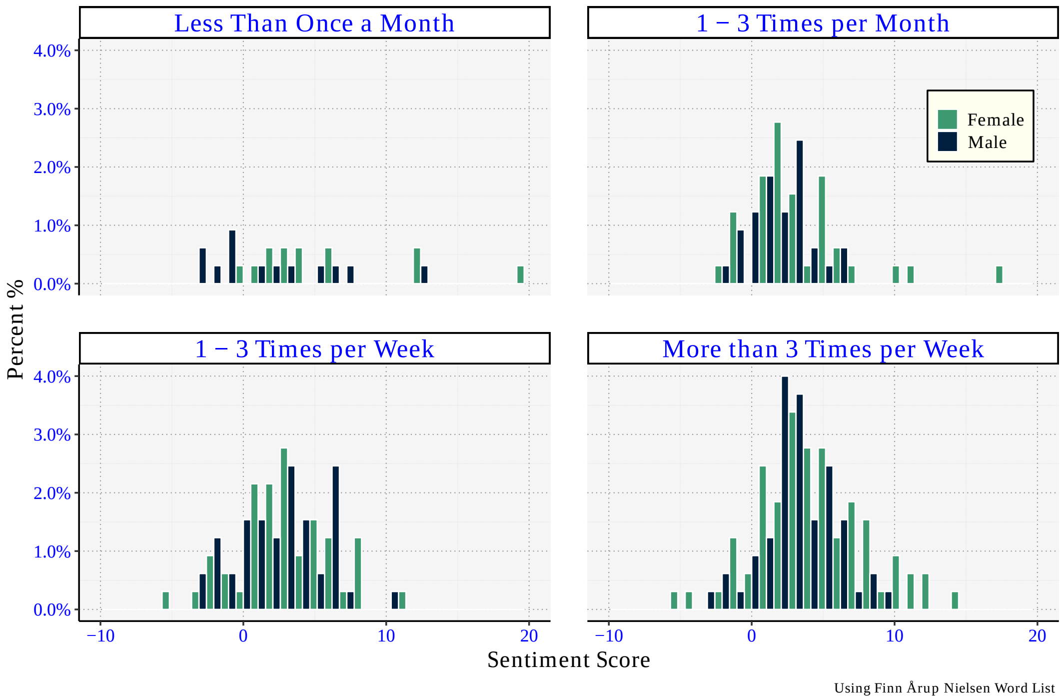

Figure A.7 – Jetty Use Frequency Users' Sentiment Score

Additional Models

This appendix contains all the models that were not presented in the modeling chapter

Shift to Jetty

Table-B.1 shows the restricted binary logit HCM, the female coefficient is removed.

Table B.1 – Shift to Jetty HCM, Model-2

No	Variable	β (P-value)	Rob.Std. Error
1	Intercept	-4.41 (0.00)	0.68
2	Age between 18 and 25 (vs 46 and older)	1.2 (0.00)	0.35
3	Age between 26 and 35 (vs 46 and older)	0.57 (0.04)	0.28
4	Age between 36 and 45 (vs 46 and older)	0.57 (0.05)	0.29
5	Household Size between 1-2 (vs 6 and more)	0.64 (0.06)	0.35
6	Household Size between 3-5 (vs 6 and more)	0.33 (0.29)	0.31
7	Personal Income between 20K- 40K (vs 20K or less)	0.38 (0.05)	0.19
8	Personal Income 40K or more (vs 20K or less)	0.73 (0.01)	0.30
9	Driving License Availability	0.74 (0.00)	0.22
10	In City Resident	-0.46 (0.06)	0.24
11	#No of Cars in Household = 1 (vs zero cars)	0.73 (0.00)	0.26
12	#No of Cars in Household = 2 or more (vs zero cars)	0.65 (0.02)	0.28
13	#No of modes replaced by Jetty	0.55 (0.00)	0.12
14	Average Jetty trips distance	-0.33 (0.00)	0.11
15	Pct (%) of morning trips	0.55 (0.02)	0.24
16	Activity: Use Smart Phone	0.32 (0.09)	0.19
17	Reason: Fare	0.57 (0.00)	0.18
18	Reason: Avoid Parking Problem	0.53 (0.04)	0.26
19	LV: Frequent Put User	-1.14 (0.00)	0.13
Latent Variable Model			
	Structure Model (Frequent Put Users)	ζ (P-value)	Rob.Std. Error
1	Gender: Female	-0.51 (0.00)	0.08
	Personal Income between 20K- 40K (vs 20K or less)	-0.52 (0.00)	0.09
2	Personal Income 40K or more (vs 20K or less)	-1.08 (0.00)	0.13
	Measurement Model (Frequent Put Users)	γ(P-value)	Rob.Std. Error
1	Frequency of Metro use	1.4 (0.00)	0.12
2	Frequency of Metrobus use	0.71 (0.00)	0.09
3	Frequency of Light rail use	1.27 (0.00)	0.13
4	Frequency of Trolleybus use	1.37 (0.00)	0.20
5	Frequency of RTP use	1.34 (0.00)	0.13
6	Frequency of Bus use	0.99 (0.00)	0.10
7	Frequency of Microbus use	1.32 (0.00)	0.11
8	Frequency of Combi use	1.26 (0.00)	0.12

N=941, P-values are reported in parentheses are based on the robust standard errors, used to control for heteroscedasticity that might exist

$\rho^2_{Adjusted} = 0.10$

Table-B.2 shows the shift to Jetty binary logit HCM with two latent variables. The impact of the micro-mobility frequent user latent variable is noticeable, where it reduces the statistical significance of the other coefficients compared to Model-2 shown in Table-B.1.

Table B.2 – Shift to Jetty HCM, Model-3

		ICLV Model			
No	Variable	β (P-value)	Rob. Std. Error		
1	Intercept	-4.37 (0.00)	0.74		
2	Gender: Female (vs Male)	0.16 (0.58)	0.29		
3	Age between 18 and 25 (vs 46 and older)	1.13 (0.23)	0.94		
4	Age between 26 and 35 (vs 46 and older)	0.51 (0.47)	0.71		
5	Age between 36 and 45 (vs 46 and older)	0.53 (0.47)	0.74		
6	Household Size between 1-2 (vs 6 and more)	0.59 (0.18)	0.44		
7	Household Size between 3-5 (vs 6 and more)	0.31 (0.43)	0.39		
8	Personal Income between 20K- 40K (vs 20K or less)	0.33 (0.25)	0.29		
9	Personal Income 40K or more (vs 20K or less)	0.63 (0.28)	0.58		
10	Driving License Availability	0.75 (0.01)	0.28		
11	In City Resident	-0.46 (0.34)	0.48		
12	#No of Cars in Household = 1 (vs zero cars)	0.72 (0.22)	0.59		
13	#No of Cars in Household = 2 or more (vs zero cars)	0.65 (0.31)	0.64		
14	#No of modes replaced by Jetty	0.56 (0.00)	0.12		
15	Average Jetty trips distance	-0.31 (0.02)	0.13		
16	Pct. (%) of morning trips	0.51 (0.12)	0.33		
17	Activity: Use Smart Phone	0.33 (0.15)	0.23		
18	Reason: Fare	0.6 (0.05)	0.30		
19	Reason: Avoid Parking Problem	0.54 (0.38)	0.61		
20	LV: Frequent MM User (λ)	0.22 (0.28)	0.20		
21	LV: Frequent PUT User (λ)	-1.19 (0.00)	0.23		
Latent Variable Model					
	Structure Model	Frequent MM		Frequent PUT	
	Variable	ζ (P-value)	Rob.Std. Error	ζ (P-value)	Rob.Std. Error
1	Gender: Female (vs male)	-0.5 (0.00)	0.12	-0.48 (0.00)	0.07
2	Personal Income 20K-04K (vs 20k or Less)	0.17 (0.22)	0.14	-0.52 (0.00)	0.09
3	Personal Income 40K or more (vs 20k or Less)	0.47 (0.01)	0.18	-1.06 (0.00)	0.21
	Measurement Model	Frequent MM		Frequent PUT	
	Indicators	γ (P-value)	Rob.Std. Error	γ (P-value)	Rob.Std. Error
1	Frequency of Bike use	1.2 (0.00)	0.32	—	—
2	Frequency of Shared bike use	2.4 (0.09)	1.43	—	—
3	Frequency of Shared Scooter use	1.6 (0.18)	1.20	—	—
4	Frequency of Walk use	0.83 (0.00)	0.15	—	—
5	Frequency of Metro use	—	—	1.42 (0.00)	0.12
6	Frequency of Metrobus use	—	—	0.74 (0.00)	0.10
7	Frequency of Light-Rail use	—	—	1.27 (0.00)	0.16
8	Frequency of Trolleybus use	—	—	1.39 (0.00)	0.32
9	Frequency of RTP use	—	—	1.35 (0.00)	0.13
10	Frequency of Bus use	—	—	1.02 (0.00)	0.11
11	Frequency of Microbus use	—	—	1.35 (0.00)	0.12
12	Frequency of Combi use	—	—	1.22 (0.00)	0.14

N=941, P-values are reported in parentheses are based on the robust standard errors, used to control for heteroscedasticity that might exist

$\rho^2_{Adjusted} = 0.09$

Service Choice

Table B.3 – Service Choice Binary Probit Model Results

No	Variable	β (P-value)	Rob.Std. Error	95% CI	
				LL	UL
1	Intercept	-1.65 (0.00)	0.34	-1.72	-1.68
2	Gender: Female (vs Male)	0.21 (0.03)	0.10	0.21	0.22
3	Household Size between 1-2 (vs 6 or more)	0.69 (0.00)	0.19	0.70	0.73
4	Household Size between 3-5 (vs 6 or more)	0.43 (0.02)	0.18	0.42	0.45
5	Personal Income Less than 20K (vs more than 40K)	-0.53 (0.00)	0.16	-0.43	-0.41
6	Personal Income 20-40K (vs more than 40K)	-0.4 (0.01)	0.16	-0.56	-0.54
7	Employed Full time (vs other)	0.44 (0.01)	0.16	0.45	0.47
8	Access by Walk or Bike (vs other)	0.66 (0.00)	0.10	0.68	0.69
9	Egress by Walk or Bike (vs other)	0.39 (0.00)	0.11	0.39	0.40
10	Egress Duration	-0.17 (0.00)	0.06	-0.19	-0.18
11	Pct (%) of morning trips	0.19 (0.00)	0.05	0.28	0.30
12	Headway in the nearest Metro station	-0.25 (0.01)	0.10	-0.22	-0.22
13	Average Jetty Trip Distance	0.28 (0.05)	0.15	0.19	0.20
14	Welling to Walk More than 10 min. (Vs 10 or less)	0.24 (0.08)	0.14	-0.26	-0.25
15	Activity Working	-0.3 (0.04)	0.14	0.24	0.25
16	Activity Talk on Phone	-0.22 (0.00)	0.05	-0.33	-0.31
17	Reason Booking of Seats	-0.26 (0.01)	0.10	-0.28	-0.27
18	Reason Ease of Payment	-0.22 (0.04)	0.11	-0.23	-0.22
19	Reason Security Against Theft	0.26 (0.01)	0.10	0.26	0.27
20	Reason Fare	0.19 (0.05)	0.10	0.19	0.20
21	Use Frequency: Less than Once a Month (vs more than 3 times a week)	0.57 (0.00)	0.16	0.57	0.59
22	Use Frequency: 1-3 Times a Month (vs more than 3 times a week)	0.39 (0.00)	0.13	0.40	0.41
23	Use Frequency: 1-3 Times a Week (vs more than 3 times a week)	0.35 (0.00)	0.12	0.36	0.38

Model Diagnostics	
Number of observations	1118
Number of excluded observations	212
Number of estimated Parameters	21
$\mathcal{L}(\beta_0)$	-601.01
$\mathcal{L}(\hat{\beta})$	-483.70
$-2[\mathcal{L}(\beta_0) - \mathcal{L}(\hat{\beta})]$	234.62
ρ^2	0.200
$\rho^2_{Adjusted}$	0.16
AIC	1013.40
BIC	1124.00

Jetty Use Frequency

Table B.4 – Jetty Use Frequency Ordered Probit Model

No	Variable	beta (P-value)	Rob. Std. Error	95% CI	
				LL	UL
1	Employed Full time (vs others)	0.38 (0.00)	0.12	0.37	0.39
2	Pct. % of Morning Trips	0.4 (0.00)	0.10	0.40	0.41
3	Average Jetty Trip Distance	0.11 (0.00)	0.04	0.11	0.12
4	Average Access Distance to Jetty	-0.1 (0.01)	0.04	-0.11	-0.10
5	Average Egress Distance from Jetty	-0.19 (0.00)	0.04	-0.20	-0.19
6	Headway in the nearest Metro Station	0.08 (0.03)	0.04	0.08	0.09
7	Activity Reading For Pleasure	0.18 (0.05)	0.09	0.18	0.19
8	Activity Sleeping	0.41 (0.00)	0.09	0.41	0.42
9	Activity Studying	0.48 (0.00)	0.15	0.47	0.49
10	Reason Travel Time Reliability	0.2 (0.01)	0.07	0.20	0.21
11	Reason Quality of Vehicle	0.18 (0.02)	0.08	0.18	0.19
12	Reason Fare	-0.2 (0.01)	0.08	-0.20	-0.19
13	Interaction (Female X Reason Security Against Harassment)	0.22 (0.05)	0.11	0.21	0.22
Threshold 1	Less Than Once a Month|1 - 3 Times per Month	-0.24 (0.13)	0.16	-0.25	-0.23
Threshold 2	1 - 3 Times per Month|1 - 3 Times per Week	0.62 (0.00)	0.15	0.62	0.64
Threshold 3	1 - 3 Times per Week| More than 3 Times per Week	1.44 (0.00)	0.16	1.44	1.46
Model Diagnostics					
Number of observations		1118			
Number of excluded observations		212			
Number of estimated Parameters		21			
$\mathcal{L}(\beta_0)$		-1242.4			
$\mathcal{L}(\hat{\beta})$		-1165.0			
$-2[\mathcal{L}(\beta_0) - \mathcal{L}(\hat{\beta})]$		154.8			
ρ^2		0.06			
$\rho^2_{Adjusted}$		0.05			
AIC		2362			
BIC		2439.6			