The Dissertation Committee for Patricia Sauri Lavieri Certifies that this is the approved version of the following Dissertation:

Is the Future of Urban Mobility Shared? Modeling Ride-Hailing Adoption and Preferences for Ownership and Sharing of Autonomous Vehicles

Committee:

Chandra R. Bhat, Supervisor

Randy B. Machemehl

Stephen D. Boyles

Chandler Stolp

Johanna Zmud

Is the Future of Urban Mobility Shared? Modeling Ride-Hailing Adoption and Preferences for Ownership and Sharing of Autonomous Vehicles

by

Patricia Sauri Lavieri

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August, 2018

Dedication

To my parents, Maribel and Durval.

Acknowledgements

I am grateful for having the opportunity to be mentored by Dr. Chandra Bhat and I am honored to be a part of his legacy. Dr. Bhat, I thank you for the support and guidance, for encouraging me to pursue my own ideas, and for your friendship. You lead by example and you have inspired me to raise my personal standards. I would also like to thank the members of my dissertation committee, Dr. Randy Machemehl, Dr. Stephen Boyles, Dr. Chandler Stolp, and Dr. Johanna Zmud, for their constructive comments and suggestions to improve this dissertation. I also owe sincere thanks to my graduate student mentor, Subodh Dubey, for always having an open door to help, even when that door was on the other side of the world. A special thanks goes to Lisa Macias for her friendly assistance and for bringing Macy to the office to provide us with a more than welcomed 4-year-old moral support and art-work. My student colleagues and co-workers also added many important elements into my journey. I would like to thank all researchers in our group, especially Alice Chu, Felipe Dias, and Dr. Sebastian Astroza for their team-work and friendship.

Friends, family, and professors from Brazil also played a very important role in my doctoral studies. I would like to voice my sincere appreciation to Dr. Eduardo Campello for inspiring me to pursue a Ph.D. in a top institution in the U.S., and Dr. Orlando Strambi for all the advice and encouragement. I would also like to thank all my friends who were not physically with me during this process, but who gave me long-distance encouragement. Patricia Aguiar and Gregory Valente, your visits were memorable highlights of the past four years. I would like to acknowledge my greatest inspiration, my Family. My parents and my brother, for countless reasons, have motivated me to realize my potential and to take on unprecedented challenges and never give up. They have also provided unconditional support and love to ensure my success and happiness. I want to extend my thanks to my new family, the “turning points” in my Ph.D. life, James and Kaia. You fill my life with cheer and joy.

Finally, I would like to acknowledge and thank the funding I received from the Brazilian Government through the program “Ciencia Sem Fronteiras”. Without the funding from CAPES, I would not have had the opportunity to pursue my doctoral studies and develop this dissertation.

Abstract

Is the Future of Urban Mobility Shared? Modeling Ride-Hailing Adoption and Preferences for Ownership and Sharing of Autonomous Vehicles

Patricia Sauri Lavieri, Ph.D.

The University of Texas at Austin, 2018

Supervisor: Chandra R. Bhat

Society is experiencing the initial stages of a technological revolution that promises to disrupt urban transportation as known today and induce behavioral and social changes. The main factors guiding the transformation of urban mobility are the growth of Information and Communication Technology (ICT)-enabled transportation services and the development of autonomous vehicle (AV) technologies. While the use of ICTs and vehicular automation are expected to provide direct road capacity improvements due to the real-time provision of traffic information, crash reductions, and platooning capabilities, these gains may be offset by latent demand effects. That is, the increase in level of service may actually result in the generation of more trips and escalation of vehicle miles traveled. In this sense, proactive planning and policy guided towards promoting the use of shared vehicles and pooled rides are important to minimize possible negative externalities of automation. The current dissertation provides initial guidance to such planning by examining individuals' preferences toward the adoption of current and future mobility services and technologies. A research framework containing four independent but related analysis components is developed to allow a comprehensive investigation of travelers' characteristics and behaviors associated with ride-hailing use and preferences regarding AVs. Empirical analyses are conducted using advanced econometric techniques applied to different types of data from three different cities. The results of the empirical analyses are compared and implications to transportation planning and policy are discussed.

The results from the analyses undertaken in the dissertation show that, from a behavioral perspective, a service-based transportation future where people predominantly travel using shared vehicles and pooled rides instead of their own vehicles is on its way but still distant. A complex combination of actions is required to promote the use of shared services both today and in an AV future. Among these actions, we identify the need for campaigns to (a) increase technology awareness among older individuals and individuals from lower income households, and (b) reduce privacy-sensitivity among non-Hispanic Whites and millennials. Such efforts should also be complemented by a decrease in service fares and urban densification. The results also suggest the need to promote policies and integrated multi-modal systems that discourage individuals from substituting the use of active and public transit modes by ride-hailing and AV-based services. Finally, we observe that individuals seem to be less sensitive to the presence of strangers in a commute trip than in a leisure trip, but the sensitivity to time is the opposite. The implications of these results are that pooled services may have a large market penetration potential for commute trips as long as operated efficiently with minimal detour and pick-up/drop-off delays.

Table of Contents

List of Tables

List of Figures

CHAPTER 1. Introduction

1.1 Introduction

Society and technology evolve together in a synergistic relationship that is occasionally impacted by revolutionary inventions. In the past century, the large scale production of automobiles and the advent of the internet were groundbreaking technological advancements that drastically influenced the world's economic and spatial organization, as well as social behaviors and norms. Automobiles changed the shape of cities and people's lifestyles, transforming the meaning of flexibility and freedom, and increasing the geographic area that humans could reach to pursue daily activities. This physical solution to geographic access enabled by automobiles was then augmented and renovated by virtual access opportunities brought by the internet and other information and communication technologies (ICTs). Instantaneous and ubiquitous remote access to people, information, goods, services, and activities became a natural and expected aspect of individuals' lives, changing how and when work, consumption, and social activities take place. ICTs have also promoted direct impacts on transportation by improving network efficiency, facilitating innovative transportation services, and providing real-time information to guide user's choices and enhance travel experiences.

Society is now about to experience another technological revolution that promises to disrupt urban transportation as known today and also induce behavioral and social shifts. Two of the key components of this radical change are: the growth of ICT-enabled transportation services that challenge the need for private vehicle ownership and the development of self-driving automotive technology. In the next sections, we discuss each of these two components and their implications to urban transportation and travel behavior.

1.2 New transportation services and the concept of Mobility-as-a-Service

Mobility-as-a-Service (MaaS) systems refer to the vision of shifting transportation from an ownership-based perspective to an access-based perspective. This paradigm switch is occurring after a decade of worldwide development and popularization of new transportation services such as bicycle-sharing, car-sharing, and ride-hailing[1]. The key idea behind MaaS is to use ICT to offer users with tailored mobility packages that facilitate multimodal door-to-door travel (USDOT, 2016; Jittrapirom et al., 2017). That is, using an interconnected network of public and private transportation services (such as transit, bicycle-sharing, car-sharing, and ride-hailing or taxi) and an online platform that provides users with multiple options of personalized trip plans and offers an integrated payment system (per distance and/or time traveled), MaaS systems are designed to enable convenient, cost-effective, and environmentally sustainable alternatives to the use of private cars. Most of the currently existing MaaS schemes are established in Europe and have transit as a main structuring component, while other modes, such as bicycle and car-sharing, are used as first and last mile connectors (Jittrapirom et al., 2017). For cities with deficient transit systems and medium/low-density land use patterns (common characteristics among U.S cities), microtransit and ride-hailing, especially pooled ride-hailing, can play important roles as MaaS facilitators (see Enoch, 2015 and Frei et al., 2017) [2,3].

Ride-hailing services have experienced exponential growth in the past years. For instance, it took Uber six years to reach the one billion-trip milestone in 2015, but only six additional months to reach the two billion milestone. One year after that, the company exceeded 5 billion trips (Uber, 2017). Indeed, among all new mobility services, ride-hailing has the highest penetration rate in the U.S. In 2017, Uber alone (a single ride-hailing company) had more than ten times the number of active subscribers of all North American car-sharing programs together, and more than four times the number of bicycle-sharing frequent users (20 million [U.S] : 1.8 million [North America] : 4 million [U.S], according to Statista, 2018a, 2018b, 2018c). This substantial growth reflects the convenience that ride-hailing offers to users by being a reliable, lower cost (compared to traditional taxi services), on-demand and door-to-door transportation service that does not require subscription fees and does not involve cognitive or physical efforts

[1] Ride-hailing services, also referred to as transportation network companies (TNCs), use a smartphone or web application to pair passengers with drivers who offer paid rides in their non-commercial vehicles. The service is analogous to a taxi, but offers scheduling and pricing advantages. The largest and most well-known ride-hailing company in the U.S. is called Uber.

[2] Microtransit refers to private multi-passenger transportation services (using SUVs, mini-vans or shuttle buses), that serve passengers using dynamically generated routes, and may expect passengers to make their way to and from common pick-up or drop-off points (USDOT-FTA, 2018).

[3] Ride-hailing services can be hired in a pooled mode, in which the user accepts to share a ride with strangers in exchange for a cheaper fare.

from the traveler (compared to car-sharing that requires the traveler to drive, and bicycle-sharing that the traveler needs to pedal).

Even as ride-hailing has gained considerable traction and is widely prevalent today in most urban areas, its impacts on individual travel are unclear and have not been adequately examined. This includes limited knowledge of which travel modes are being substituted, what its potential impacts on private vehicle ownership are, how it may affect peak and off-peak travel patterns, and whether its convenience induces more or less travel. A main reason for the lack of studies on ride-hailing impacts is the scarcity of publicly available data. To fill this void, some researchers have resorted to specialized user surveys or large-scale household travel surveys that collect limited information on ride-hailing preferences. These studies suggest that users replace trips by modes other than taxi, including public transit and driving (Rayle et al., 2016), while effects on vehicle ownership are less clear. A small proportion of the population may be willing to dispose one or more household vehicles because of ride-hailing availability (Clewlow and Mishra, 2017), but still the majority of the users own personal vehicles (Smith, 2016; Dias et al., 2017). In summary, ride-hailing plays a significant role in MaaS systems and it is critical to understand its demand. Therefore, a deeper understanding of the variables that contribute to its adoption, the factors that incentivize the use of pooled rides, and the potential competition with transit ridership and vehicle ownership is required.

1.3 New self-driving technologies: autonomous vehicles

Alongside the innovation on mobility services, as discussed in the previous section, automotive technology is also passing through a period of significant transformation. Autonomous vehicles (AVs) utilize a set of sensing equipment (such as video cameras, radars, LIDAR, GPS and, in the case of connected vehicles, communication devices) and computational power to identify and predict the environment in their surroundings in order to take automatic action. The tasks that the vehicle can accomplish independently may vary in the degree of sophistication, which translates into different levels of automation. From no automation at Level 0 to high and full automation at Level 4 and Level 5, respectively (SAE, 2014). High automation means that in most environments (particular areas) the automated driving system can operate independently of human action (that is, autonomous driving is mode-specific), and full automation means that steering wheels are no longer necessary and the automated system can manage every situation.

Announcements by car manufacturers and technology companies promise Level 4 vehicles to reach the roads by 2020-21, while Level 5 vehicles should take a lot longer (CTR, 2017; NVIDIA, 2017; Gibbs, 2017; Ford Motor Company, 2018; GM, 2018). For AVs applied to ride-hailing services, which are also known as shared autonomous vehicles (SAVs), levels 4 and 5 of automation overlap. This is because SAVs can be programed to serve a very specific area (under specific weather conditions) where the software controlling the vehicle has been fully trained to react to the environment without human supervision. Although the vehicle is constrained to a certain area of coverage, it operates similarly to a Level 5 vehicle in this area. Since it may not be practical for a car owner to have a vehicle that does not require driving only in a subset of situations, it is becoming common sense that Level 4 AVs will enter the market through ride-hailing services years before they reach dealerships. For instance, Waymo, prior Google, already launched a program in Phoenix, Arizona, where volunteers can subscribe to be early riders and use Level 4-5 cars for daily trips in specific areas (Waymo, 2018). In that sense, developing a deeper understanding of demand characteristics and travel behavior associated with the use of ride-hailing today may also contribute to the understanding of the early stages impacts of AVs on transportation.

1.4 Potential impacts of automation on transportation and the use of AVs as MaaS providers

Automation will bring significant traffic safety enhancements (Fernandes and Nunes, 2012; Winkle, 2016); however, vehicles with Level 4-5 of automation (which will be called AVs from now on) may also engender substantial changes in urban transportation and land-use. We can classify the potential effects of automation on transportation into two categories: (1) direct technological effects on supply and operations, and (2) indirect effects due to changes in demand behavior. The first category of effects includes the increase of network capacity and efficiency due to platooning capabilities, better traffic coordination and reduced accidents (as identified by Fernandes and Nunes, 2012, and Tientakool et al., 2015, for example). Additionally, transit systems may be expanded by utilizing smaller vehicles to provide first and last mile services. Transit costs of operation should also reduce when drivers are no longer required, and park and ride areas may be retrofitted for other land use purposes. Similarly, parking spaces in central areas may be repurposed because vehicles will be able to self-park in less dense areas, and return to the owner's home or reposition to serve other trips. For instance, simulation studies conducted by Zhang et al. (2015) and Zhang and Guhathakurta (2017) observe that even low penetration rates of AVs may allow proportionally high reductions of parking needs.

The second category of effects is associated with the impacts that automation has on transportation consumers. Segments of the population that previously could not use cars because of the inability to drive (such as children, elders, and physically or mentally challenged individuals) will have their level of accessibility increased leading to an increase in social inclusion. At the same time, current car users may experience increased comfort due to both changes in vehicle design and elimination of the need to drive, which should allow for a meaningful use of the time traveling (socializing, working or sleeping, for example) and multitasking. Such factors may also reduce the disutility commonly attributed to traveling (especially driving) and, thus, decrease individual's value of travel time (VTT). The consequences may be the increase of number of activities and/or distance between activities resulting in the growth of vehicle miles traveled. In the long term, certain segments of the population could also choose to relocate to more affordable or isolated areas, resulting in urban sprawl. The chauffeuring capabilities of AVs may also generate empty vehicle miles of travel, especially if households opt to own fewer vehicles that are frequently moving back and forth to pick-up and drop-off household members. Together, these indirect effects may offset network efficiency gains generated by the direct technological effects of automation, and congestion levels and energy consumption could actually increase[4].

The extent to which AVs would produce the positive and negative externalities discussed above may vary depending on their long-term adoption paradigm. Figure 1-1 contrasts two hypothetical AV adoption patterns that could lead to different outcomes. The figure shows that the prevalence of privately owned AVs (by individuals and households) could lead to high rates of empty vehicle travel, significant decrease in value of travel time, and increased congestion and energy consumption, as discussed above. On the other hand, if the common practice becomes using MaaS systems and hiring SAVs by time and distance traveled (similarly to today's ride-hailing services), then significant drops in value of travel time, as well as increases in empty vehicle travel, could be avoided. As a consequence, lower congestion levels and energy

[4] Individuals may become less sensitive to congestion once they are not driving and can have a productive use of travel time.

consumption (compared to the other scenario and potentially compared to today) could be achieved. Indeed, supply perspective, SAVs and pooled SAVs (PSAVs) are receiving significant attention from researchers (for some recent studies, see Frei et al., 2017, Levin et al., 2017, and Wang et al., 2018). The studies suggest that PSAVs have good potential to quite substantially reduce overall VMT relative to the case of privately owned AVs or solo-rider SAVs, and also that additional travel times due to pick-up and drop-off of multiple passengers could be compensated by reductions in congestion if shared rides are massively adopted by users. Therefore, identifying factors associated with the use of MaaS, particularly ride-hailing and pooled ride-hailing, today that could contribute to future preferences for using SAVs and PSAVs instead of owning AVs is critical to guide transportation planning and policy in the years to come.

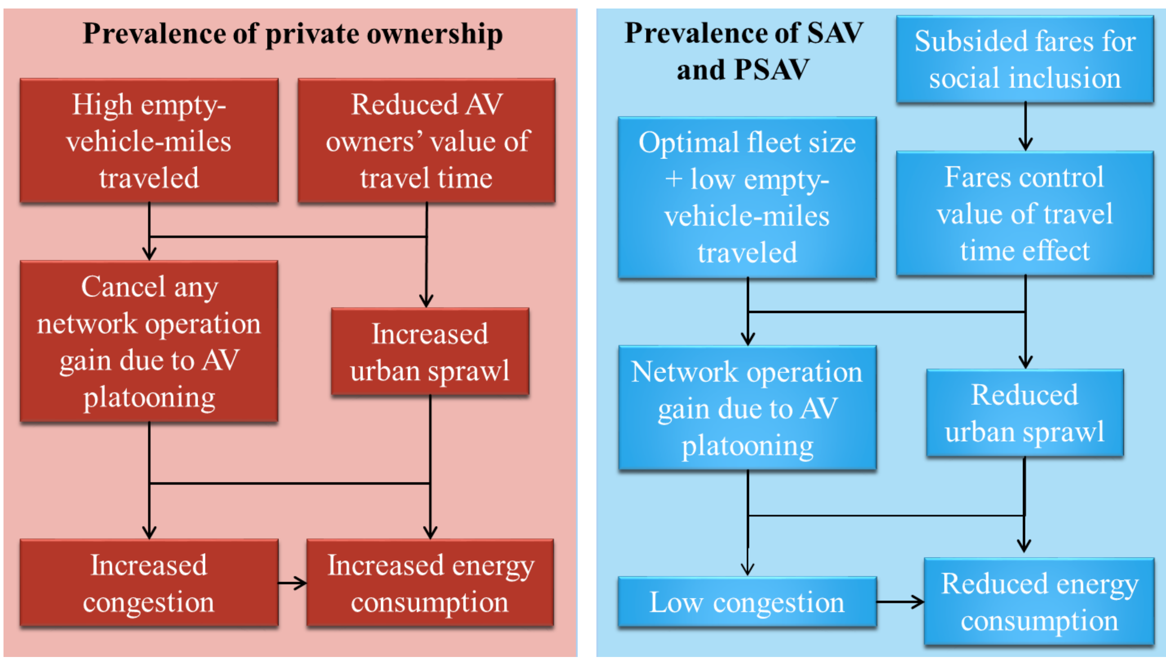


Figure 1-1 Comparison of hypothetical effects of automation based on an ownership adoption paradigm and in a shared adoption paradigm

1.5 Research questions

The earlier discussion may be summarized in five main points: (1) ICTs are allowing the integration between different modes of transportation and the creation of MaaS systems; (2) Ride-hailing services are growing exponentially and they may play a key role in MaaS systems, especially in cities where transit is limited; (3) Automation has become a reality and vehicles with Level 4 of automation will likely be reaching the streets by 2020; (4) Automation can provide direct road capacity improvements but it can also generate externalities depending on the adoption paradigm; (5) Proactive planning and policy guided towards promoting the use of shared vehicles and pooled rides are important factors to both minimize negative and maximize positive externalities of automation. To inform such planning, a deep understanding of the current use of ride-hailing services, together with an examination of individual's preferences regarding AV adoption, is critical.

Despite a growing literature focusing on travel behavior associated with ride-hailing and AV preferences, the current efforts are mostly descriptive and involve limited analysis of the travel behavior dimensions impacted and their determinant factors (Rayle et al., 2016; Clewlow and Mishra, 2017; Becker and Axhausen, 2017). Considering that decisions regarding the use of ride-hailing and automation are interweaved with other transportation and lifestyle decisions, comprehensive modeling efforts encompassing user's multiple dimensions of behavior are required. The objective of this dissertation is to develop such models. In that sense, we propose an analytic framework that facilitates the investigation of the following research questions:

(1) What segments of the population already use ride-hailing services? Who is sharing rides? Who are the frequent users?
(2) What land use and transportation aspects contribute to the use of ride-hailing?
(3) Is there evidence of positive and negative externalities of ride-hailing adoption?
(4) What segments of the population have the intention to adopt AVs? Who wants to share vehicles? Who wants to own? And who wants both?
(5) How much individuals would be willing to pay to not share rides in a SAV scenario? How does the willingness-to-pay to not share relate with the value of travel time?
(6) What are the impacts of current ride-hailing experience on the intentions to adopt AVs, SAVs and PSAVs?

Four independent but related analysis components are developed to address the questions above. The first two analyses focus on users' current ride-hailing behavior, and the other two simultaneously investigate current travel behaviors and future intentions to use AVs. The first analysis applies a two-step aggregate modeling approach to investigate the generation and distribution of daily ride-hailing trips in the city of Austin, Texas. The second analysis complements the first by modeling the multiple choices associated with the use of ride-hailing at the individual level (instead of trip counts per TAZ), while the third analysis jointly models individuals' ride-hailing experience and their preferences towards ownership and sharing of AVs. The final analysis focuses on individuals' perceptions toward sharing rides with strangers in an AV future as well as their willingness to pay to ride alone (or to avoid sharing rides with strangers). These analyses are conducted using advanced econometric techniques applied to different types of data from three different cities. The results are compared based on the research questions.

1.6 Dissertation structure

The dissertation is organized as follows. Chapter 2 introduces the overall analytic framework of the dissertation, positioning each of the analysis components of the dissertation in relation to each other. The chapter also discusses the data, and presents the two main modeling methodologies that are used in the four analysis components. Each of the following chapters presents one of the analysis components. Chapter 3 and Chapter 4 analyze current ride-hailing adoption and use based on two distinct types of data and modeling approaches (they constitute the first two analysis components). Chapter 5 uses survey data to investigate travelers' interest in adopting AV technology, and determines the extent to which individuals are inclined to acquire such vehicles for private ownership or use them in a shared mobility service configuration (the third analysis component). Chapter 6 examines pooling behavior in an AV context. Each of the four main chapters contains its own discussion of policy implications, conclusions, and recommendations for future research. The final dissertation chapter, Chapter 7, presents a discussion of the main results of each chapter under the scope of the seven exploratory research questions described above.

CHAPTER 2. Analytic Framework

2.1 Introduction

In this dissertation, the examination of current ride-hailing demand behavior and preferences for the adoption of AVs is undertaken through four modeling efforts. The first two analyses focus on users' current ride-hailing behavior, while the other two simultaneously investigate current behavior and future preferences. All four models have the current ride-hailing behavior as a key endogenous variable. The first analysis (Chapter 3) applies a two-step aggregate modeling approach to investigate the generation and distribution of daily ride-hailing trips. A spatially lagged multivariate count model is used to describe how many trips are generated in a specific traffic analysis zone (TAZ) on both weekdays and weekend days. A fractional split model is applied to identify the characteristics of zones that attract ride-hailing trips.

The second, third and fourth analyses investigate individuals' choices and are based on the same modeling methodology –the Generalized Heterogeneous Data Model (GHDM). The GHDM is ideal for the multivariate behavioral frameworks proposed in each of these analyses as it allows for a simultaneous estimation of multiple types of dependent variables (including multiple nominal outcomes, multiple ordinal variables, and multiple count variables, as well as multiple continuous variables) by representing the covariance relationships among them through a reduced number of stochastic latent factors. Specifically, the second analysis (Chapter 4) complements the first by modeling the multiple choices associated with use of ride-hailing at the individual level (instead of trip counts per TAZ). The multiple outcomes in this second analysis component include the choice to use ride-hailing, the frequency of both solo and pooled rides, and the characteristics (purpose, time of the day, companion, and mode substituted) of the latest ride-hailing trip of survey respondents. These multiple outcomes are jointly modeled as functions of socio-demographic characteristics, latent constructs representing attitudes and lifestyles, and endogenous variables representing residential location and vehicle ownership. The third analysis (Chapter 5) models preferences regarding the adoption of AVs. Based on the person's lifecycle, lifestyle (represented by latent constructs), and current transportation related behavior, the model explains whether an individual is inclined to purchase an AV or use only SAVs (or both or none) in the future. In addition to the AV preferences, the main endogenous variables considered are residential location density, vehicle ownership, and experience with car-sharing and ride-hailing services. The final analysis (Chapter 6) focuses on individuals' perceptions toward pooling (or sharing) rides. The current experience with ride-hailing services is modeled together with stated choices between hiring a solo and a pooled ride for commute and leisure trips in a SAV future. Again, latent constructs representing attitudes and socio-demographic characteristics are used to explain the current behavior and stated intentions.

In the remainder of this chapter, we provide further details on the data and modeling methodologies used in the four analyses. Different types of data from multiple sources and regions are used in this dissertation, thus we start with an overview of the common types of data used in travel demand models, presenting their advantages and limitations. Then we explain which data type is employed in each of the models developed. The formulations of the two modeling approaches are presented together with a discussion of the importance of certain features of these models.

2.2 Data

2.2.1 Aggregate and Disaggregate

The most basic unit of analysis in passenger travel demand modeling is the individual or the household, while a common form of aggregate data relates to examining one or more travel dimensions at the spatial level of a traffic analysis zone (TAZ) or some other geographic space unit. The disadvantages of the use of aggregate data compared to disaggregate data in travel demand modeling have been extensively discussed in the transportation literature and practice (see Ortuzar and Willumsen, 2001, for example). First, aggregation of any kind means the replacement of multiple observations by summary statistics of the group created. Thus, by definition, information is being lost and so is variability, resulting in the reduction of the explanatory power of the constructed model. Second, when data is aggregated spatially, a number of unobserved variables related to the area corresponding to the unit of analysis become confounding factors. An additional problem comes from the definition of the boundaries and size of each spatial unit. Historically, the delimitation of TAZs, census tracts, or other spatial units has been based on the roadway network or other geographic characteristic, meaning that no spatial homogeneity in terms of demand characteristics and land use is guaranteed within a zone. As an illustration of spatial heterogeneity, one may observe that a zone generates daily 4 trips per habitant (the U.S. average). However, the high income individuals of that zone may be responsible for the generation of 80% of the trips, but represent only 40% of the population. An analysis based on the spatially aggregate data (that involves only the number of trips per zone and the income distribution of the zone) will fail to observe that low income individuals may be excluded from the opportunity to travel (pursue out-of-home activities). It is also possible that individuals in a boundary area of a zone behave more similarly to individuals in the adjacent zone than individuals in the other extreme of their same zone.

Spatially aggregated data has been extensively used in association with the classic transportation model (the four-step model), which is a trip-based model comprised of four sub-models[5] of trips per spatial unit (TAZ, for example). In the past 20 years, improvements in computational power together with methodological advancements has led to the transition from the classic transportation model to activity-based models, which are not only more theoretically grounded but also use disaggregate individual and household level data. Disaggregate data does not present limitations discussed above but require the administration of more complex surveys, which are consequently more costly both to collect and to process and prepare the data.

2.2.2 Active and Passive

Travel demand data, whether at a disaggregate level or an aggregate level, is usually obtained through active surveys, in which individuals are directly asked about their characteristics, activities and travel. Besides the cost, this type of data collection is also limited by the burden imposed to the respondent and hence is limited in the amount of information and especially timespan. Individuals are asked about usual behavior or to describe a single day of activities and travel. For over a decade, passive data collection methods have also been discussed and incorporated to travel surveys in order to expand the amount of information gathered without increasing the respondent burden. Passive data collection is unobtrusive and does not require the direct questioning of participants, individuals are just recorded. In-vehicle or individual GPS trackers have been used to passively collect time and location information that can help infer activity location, travel distances, time of the day and even travel mode. Currently, this type of data is also being collected through the respondent's smartphone (see Zmud et al., 2013).

[5] The four sub-models are: tip generation, trip distribution, model split and assignment.

A new type of passive data that is becoming increasingly prevalent, is obtained in a more indirect manner (not as part of a travel surveys), and is often labeled as *big data*[6]. It encompasses all information that is stored in computers about transactions, actions, interactions and movements that occur digitally based on ICT technology. Examples in transportation are data based on smart cards (transit) and the large number of applications for smartphones, such as basic navigation apps, other GPS based apps, and ride-hailing and bicycle-sharing apps. These data allow the observation of where individuals start and end trips, at what times, what types of transportation services are being used and how frequently, among other information. Unfortunately, these data sets are typically proprietary and the few that are open to the public usually have all information related to the users removed because of privacy issues. Thus, they allow monitoring of what is happening but, generally, do not provide much information on who is doing it or why. In that sense, other sources of information need to be used to complement big data use for travel behavior analysis. A potential source of user information can be the spatially aggregated survey data mentioned in the previous section.

2.2.3 Revealed and Stated Choice

As discussed earlier, travel demand data can be collected through active or passive participation of users and can be aggregate or disaggregate. Disaggregate level and actively collected data are typically used by travel behavior studies because they are concerned with the analysis of individual's choices. Choice behavior is explained based on two types of variables, characteristics of the individuals making choices and the attributes of the choice alternatives. The relative importance of these attributes can be measured through revealed or stated choice data. Revealed data refers to situations in which choices are actually made in a real market, while stated data is based on choices made under hypothetical scenarios. Typically stated choice experiments are used to understand users' acceptability of alternatives that are not yet available in the real world or to test their sensitivity to changes in different choice attributes. Although commonly obtained through active surveys, revealed choice may be inferred from passive data collection methods, while stated choice must be collected through active methods. An extensive discussion on the advantages and disadvantages of revealed and stated choice data is provided by Hensher et al. (2005). Here we summarize a few aspects that are relevant to this research.

[6] Passive data is one component of *big data*. Also to be cited is that the definition of what constitutes *big data* includes not only the data type but also size, storage, and analysis components (Ward and Baker, 2013).

One key limitation of revealed data is that it is restricted to currently existing alternatives within a stable existing market. Therefore, this type of data cannot be used to predict market changes a priori to the introduction of new alternatives, such as new transportation technologies and services. Additionally, revealed choice data involves real life environments that are not fully observed by the analyst, and the actual choice set (alternatives, attributes and levels) perceived by the consumer may not be the same considered during the analysis. Stated choice data has the advantage of allowing the measurement of situations that are not yet real. However, it is up to the analyst to determine which attributes and levels should be considered in the choice and possibly these will not reflect what the individual would consider in real life due to personal constraints. Decisions are bounded by real world constraints, so when we use revealed choice these constraints are necessarily influencing the choice but this may not be true in a hypothetical stated choice setting.

2.2.4 The Data in the Current Study

The data classification taxonomy presented above is important in the context of this dissertation because we analyze ride-hailing usage through multiple perspectives based on different types of data. In the first analysis, we use a combination of passive data (*big data*) released by a ride-hailing company and several publicly available data sources that provide socio-demographic and land-use distributions across TAZs. Recently, Ride-Austin, a non-profit Austin-based ride-hailing company, released a large dataset containing trip-level information. Using six months of this trip data, including trip origin and destination location, distance traveled and corresponding dates, we develop a two-step modeling framework to investigate the generation and distribution of ride-hailing trips on an average weekday and weekend day. Although the RideAustin data set contain trip information at a disaggregate level, it does not provide user information and corresponding socio-demographic characteristics. Therefore, our analysis is undertaken at the zonal level (trips are spatially aggregated according to TAZs) and relies on zonal demographics to infer ridesourcing demand characteristics. Using this combination of data sets, we are able to explore the impacts of zonal distributions of income, gender, race/ethnicity, age, population and employment density, as well as transit supply characteristics and land use information regarding presence of parks and universities, on the generation and distribution of ride-hailing trips. Despite the limitations of aggregate data discussed earlier, the possibility of analyzing six months of daily trips provides an opportunity that would not be possible through survey data.

Additionally, instead of working with a sample, we are able to use the entire population of trips of a company that had approximately one third of the ride-hailing trip market share in Austin (during the period analyzed).

Considering that the aggregate trip-based model allows us to identify general patterns but is limited in terms of individual behavior inference, we perform a second analysis based on disaggregate revealed choice survey data. Ideally, this data set would also be from Austin, Texas, to allow for a direct comparison of results. However, the data available for this analysis is from the Metropolitan Area of Dallas-Fort Worth, Texas (DFW). The data was obtained through a web-based survey conducted with commuters in the second semester of 2017. The survey was administered using the online tool “Qualtrics” (Qualtrics, 2017) and the distribution was performed through mailing lists from multiple sources (local transportation planning organizations, universities, private transportation sector companies, non-profit organizations, and social media) reaching a final clean sample of 1,607 respondents. Comparisons between the sample distribution and DFW population distribution will be presented in Chapter 4.

The third and fourth analyses utilize individual-level revealed and stated choice survey data. Both analyses have two sets of endogenous variables, one representing current (revealed) behavior, and the other representing future intended (stated) behavior. The third analysis relies on data from the Puget Sound Region Household Travel Survey (PSRC, 2016), which used an online tool and telephone calls for the survey administration but had its recruitment performed through regular mail, ensuring the desired geographic coverage. Further details about the sample are provided in Chapter 5. The final analysis utilizes data from the same online survey used in Chapter 4 but it includes an additional section of stated choice variables.

2.3 Trip-based analysis: bivariate spatially lagged count model and fractional split model

In this section, we describe the modeling framework of our first analysis component. We develop a two-step procedure to analyze ride-hailing trip generation and distribution between TAZs on weekdays and weekends. In the trip generation analysis, the average number of trips generated at each TAZ on an average weekday is modeled jointly with the average number of trips generated on an average weekend day. We utilize a spatial bivariate count model that takes into consideration the spatial dependence between TAZs as well as the correlation between the two types of days.

Accounting for spatial dependency in trip generation models is important because one can expect neighboring zones to present similar travel demand patterns, especially when considering that the delimitation of zone borders is usually made based on the transportation network and does not necessarily reflect differences in demand patterns, as discussed earlier. For example, it is possible that individuals in close proximity (say in neighboring TAZs) will be influenced by each other's ride-hailing propensity through social interactions, leading to a lagged endogenous variable effect. Therefore, we consider this type of spatial dependence through the use of a spatial lag model. The spatial lag specification, in reduced form, allows spatial dependence through both spatial spillover effects (observed exogenous variables at one location having an influence on the dependent variable at that location and neighboring locations) as well as spatial error correlation effects (unobserved exogenous variables at one location having an influence on the dependent variable at that location and neighboring locations).

Of equal importance is to recognize that ride-hailing trip generation rates may vary between weekdays and weekends, but that common TAZ-level unobserved factors may influence the counts on both types of days. For instance, a zone with very limited parking (an unobserved variable in our analysis) is likely to be associated with high ride-hailing trip generation rates, both on weekdays and weekend days.

For the trip distribution analysis, we develop a fractional split distribution model that analyzes the fractions of trips originating from a zone that terminate in each destination zone. This analysis provides an understanding of factors that "pull" ride-hailing trips toward a zone. The next two sections provide an overview of both the generation and distribution models.

2.3.1 Spatial Multivariate Count Model

The spatial multivariate count model is based on Bhat et al. (2014). There are two components to this model. The first part is the recasting of the basic count model and the second part is the spatial dependency formulation.

2.3.1.1 Count Model Recasting

The framework used here is based on a recasting of the basic count model as a special case of a generalized ordered-response (GOR) model, as proposed by Castro, Paleti, and Bhat, 2012. In this approach, the count is viewed as a result of a latent demand generation propensity that gets mapped into the observed trip counts through thresholds that are themselves functions of exogenous variables. This approach offers the advantage of accommodating over-dispersion and excess zeros, which is useful when modeling zones that do not produce any trips (for example, open areas) and zones that produce unusually high numbers of trips (for example, zones that have high density of bars and active night life).

Let q (q = 1, 2, …, Q) be the index for the territorial unit of analysis (a "TAZ" in the current analysis) and let s (s = 1, 2, …, S) be the index for day-type (weekday or weekend day in the current analysis). Let y_{qs} be the index for the count of trips generated in a day-type s in a TAZ q, and let m_{qs} be the actual observed count of trips in the day-type s in the TAZ q. Next, consider that there is a TAZ-specific demand function that represents the propensity for trip generation on day-type s. This propensity is not directly observed, and so may be represented by a latent (unobserved to the analyst) variable y^*_{qs}. Then, in the generalized ordered response (GOR) notation, the latent propensity y^*_{qs} is written as a function of a $(K \times 1)$-vector of observed covariates $\boldsymbol{x}_q$ (excluding the constant) as:

$$y^*_{qs} = \boldsymbol{\beta}'_s \boldsymbol{x}_q + \eta_{qs}, \quad y_{qs} = m_{qs} \Rightarrow \psi_{qs,m_{qs}-1} < y^*_{qs} < \psi_{qs,m_{qs}}. \tag{1}$$

In the above specification, $\boldsymbol{\beta}_s$ is a $(K \times 1)$-vector whose elements capture the effects of the $\boldsymbol{x}_q$ variable vector on latent demand propensity y^*_{qs}. Finally, η_{qs} captures TAZ-specific unobserved factors that increase or decrease the latent propensity for generating trips in a week or weekend day. The thresholds in Equation (1) take the following form:

$$\psi_{qs,m_{qs}} = \Phi^{-1}\left(e^{-\lambda_{qs}} \sum_{l=0}^{m_{qs}} \frac{\lambda_{qs}^{l}}{l!}\right) + \alpha_{s,m_{qs}}, \ \lambda_{qs} = e^{\boldsymbol{\gamma}'_s \boldsymbol{h}_q}, \alpha_{s,0} = 0, \ \alpha_{s,m_{qs}} = \alpha_{s,L_s} \text{ if } m_{qs} > L_s, \tag{2}$$

where Φ^{-1} is the inverse function of the univariate cumulative standard normal, $\psi_{qs,-1} = -\infty \ \forall q, \ s$, $\alpha_{s,0} = 0 \ \forall s$ (this restriction is needed for identification, given the parameterization of the thresholds), $\boldsymbol{h}_q$ is a vector of exogenous variables (including a constant) associated with TAZ q (there can be common variables in $\boldsymbol{h}_q$ and $\boldsymbol{x}_q$), $\boldsymbol{\gamma}_s$ is a corresponding coefficient vector to be estimated for day-type s, and L_s is a pre-defined count level that is determined based on empirical testing and on the context under consideration. Note that thresholds are impacted by the TAZ characteristics so that the translation from the trip generation propensity into an actual number of trips generated may vary between two zones (based on the exogenous variables $\boldsymbol{h}_q$) even if they have exactly the same $\boldsymbol{x}_q$. As in the typical ordered-response framework, the values of $\alpha_{s,m_{qs}}$ should be such that the ordering condition on the thresholds $(-\infty < \psi_{qs,0} < \psi_{qs,1} < \psi_{qs,2} < ....)$ is satisfied. The presence of the $\alpha_{s,m_{qs}}$ terms provides flexibility to accommodate high or low probability masses for specific count outcomes without the need for cumbersome treatment using hurdle or zero-inflated models. If these terms are set to zero, and all elements of the vector $\boldsymbol{\beta}_s$ are also set to zero, the result is the traditional Poisson count model mechanisms (see Castro, Paleti, and Bhat, 2012).

2.3.1.2 Spatial Component

The model adopted in this study considers spatial dependence across TAZs in observed covariates $\boldsymbol{x}_q$ vector as well as in the unmeasurable terms η_{qs}. To conserve on space, we refer the reader to Bhat et al. (2014) for a complete explanation and formulation of the spatial structure of the model. In general terms, we have the usual distance-based spatial weight matrix (**W**), which indicates whether a pair of TAZ should be considered spatially dependent $(w_{qq'} > 0)$.

The trip generation propensity from a zone is influenced by exogenous variables specific to that zone, and the trip generation propensity from neighboring zones based on an autoregressive coefficient represented by δ_s. δ_s may be positive or negative $(-1 < \delta_s < 1 \ \forall \ s)$. In our model we adopt weight matrices based on functions of the distance between the center points of two TAZs. Since we are analyzing the central area of Austin, we consider that spatial dependence may occur between zones that are up to 3 kilometers apart. The final equation of the multivariate count model incorporating spatial dependence for a specific zone for a specific type of day is:

$$y_{qs}^* = \delta_s \sum_{q'=1}^{Q} w_{qq'} y_{q's}^* + \boldsymbol{b}_s' \, \boldsymbol{x}_q + \eta_{qs}, \quad y_{qs} = m_{qs} \ \Rightarrow \ \psi_{qs,m_{qs}-1} < y_{qs}^* < \psi_{qs,m_{qs}} . \tag{3}$$

We consider the joint nature of the demand propensities across day-types for each TAZ *q* by allowing the elements η_{qs} to be correlated across the two day types (*s*=1, 2) for each TAZ *q.* A final important point to be noted here is that the spatial dependency in counts is generated through spatial "spillover" effects and spatial error correlation effects in the latent ride-hailing demand propensity, not through the localized TAZ-specific characteristics that impact the thresholds in the count model.

2.3.2 *Fractional Split Distribution Model*

To estimate the fractional split model, we use a quasi-likelihood estimation approach (see Sivakumar and Bhat, 2002; Gourieroux et al., 1984). Let f_{qj} be the fraction of the total trips that originate in zone *q* that terminate in zone *j*, such that $\sum_{j=1}^{J} f_{qj} = 1$. We can write the log-likelihood for the qj^{th} zone pairing as follows:

$$LL_{qj} = f_{qj} \log\left[G\left(\boldsymbol{\mu}, \boldsymbol{z}_{qj}\right)\right] + \left(1 - f_{qj}\right)\log\left[1 - G\left(\boldsymbol{\mu}, \boldsymbol{z}_{qj}\right)\right] \tag{4}$$

where $\boldsymbol{\mu}$ is a vector of coefficients to be estimated and $\boldsymbol{z}_{qj}$ is a vector of exogenous variables with characteristics of zones *q* and *j*. The function $G(.)$ takes the following logit functional form:

$$G\left(\boldsymbol{\mu}, \boldsymbol{z}_{qj}\right) = \frac{1}{1 + e^{-\boldsymbol{\mu}'\boldsymbol{z}_{qj}}} \tag{5}$$

2.4 Individual-based Analysis: the Generalized Heterogeneous Data Model (GHDM) System

A second modeling methodology is adopted in this dissertation to examine individual behavior. It consists of a very comprehensive approach that allows for the consideration that transportation decisions are made as a bundle. With this approach we are able to investigate the relationship between ride-hailing adoption and many other transportation decisions, as well as individual's future intentions, while controlling for observed and unobserved factors that simultaneously influence these multiple decisions and intentions.

The GHDM (Bhat, 2015) is an evolution of a class of models known as Integrated Choice and Latent Variable (ICLV) models (see Ben-Akiva et al., 2002; Bhat and Dubey, 2014), which was inspired by structural equation modeling (SEM) techniques used in psychology and social sciences. In these approaches, unobserved psychological constructs serve as latent factors that provide a structure to the dependence among the many indicator variables (dependent variables), while the constructs themselves are explained by exogenous variables and may be correlated with one another in a structural relationship. The equations that explain the regression of the indicators onto latent constructs are called measurement equations. In traditional SEM, all indicators are usually continuous variables, while in ICLV models they are usually continuous or ordinal with one single outcome being nominal. In the GHDM, a mix of continuous, ordinal, count, and nominal outcomes is allowed without any restriction on the number. While all three approaches can be seen as parsimonious attempts to explain covariance relationships among multiple outcomes, the GHDM approach represents a powerful dimension-reduction technique that allows for the representation of the covariance relationship of high-dimensional heterogeneous outcome data.

In discrete choice transportation analysis, the use of latent constructs representing psychological factors is motivated by a need to represent choice behaviors more realistically. It is argued that choices are shaped by individuals' perceptions, which do not reflect the objective reality. Therefore, either perceptions or other psychological factors should be taken into account when modeling choice. In that sense, the inclusion of factors representing attitudes and lifestyles has become a common practice. Initial methods would enter psychometric data, such as attitudinal indicators, directly in the choice utilities. However, this procedure ignores possible measurement errors of the items (which are especially significant when trying to measure very subjective factors). Additionally, the attitudinal indicators may be correlated with other unobserved individual-specific factors that influence choice, potentially generating estimation inconsistency (Bhat and Dubey, 2014). An evolution of this method, which is still frequently used, but is also econometrically inconsistent, is the use of a two-step procedure. First, the latent factor is estimated based on multiple continuous indicators (factor analysis) and subsequently it added to the alternatives utilities as an exogenous variable. The ICLV and GHDM were solutions proposed to allow the use of conceptually sound psychological factors in choice models in an econometrically appropriate manner.

The ability to estimate jointly multiple choice outcomes and to use latent variables representing attitudes and lifestyles has other positive implications for transportation modeling. A common problem when using choice models to evaluate potential policy impacts is controlling for self-selection effects. For example, if a traditional multinomial choice model is used to investigate mode choice and residential density is entered as exogenous variable, the positive effect of this variable on the use of non-motorized modes will be likely overestimated. This is because, by considering residential density as an exogenous factor, the analyst ignores that unobserved factors that influence the individual's choice of residential location may also impact his/her mode choice. By incorporating a latent variable representing green-lifestyle, for example, and including both residential density and mode choice as dependent variables (residential density also impacting mode choice directly), it is possible to control for such unobserved factors and "isolate" the true effect of residential density on mode choice.

A final comment on the representation of taste heterogeneity with latent variables is also pertinent. The magnitude of each latent variable is different for individuals depending on the values of their socio-demographic characteristics that are considered in the structural equation component. Therefore, when the latent variable is added to the utility, it incorporates taste heterogeneity by increasing or decreasing utility depending on the individual's characteristics. However, besides this additive taste heterogeneity, latent variables can also be interacted with other explanatory variables, especially alternative specific variables, and act as moderators. For example, in a modeling investigating the impacts of sharing a ride with strangers, a latent variable representing privacy-sensitivity attitude can be interacted with the exogenous variable representing the number of additional passengers in a ride. Resulting in the identification that additional passengers affect more privacy-sensitive individuals.

2.4.1 The GHDM Formulation

In this section, we present an overview of GHDM formulation proposed by Bhat (2015a). We refer the reader to the original paper for additional details on the formulation, estimation, and identification conditions.

There are two components to the model: (1) the latent variable SEM, and (2) the latent variable measurement equation model (MEM). As illustrated in Figure 2-1, the SEM component defines latent variables as functions of exogeneous variables. In the MEM component, the endogeneous variables are described as functions of both latent variables and exogeneous variables. The error terms of the structural equations (which define the latent variables) permeate into the measurement equations (which describe the outcomes variables) creating a parsimoious dependence structure among all dependent variables. The measurement equations have different characteristics depending on the type of dependent variable (continuous, ordinal, count, or nominal), however all have continuous underlying functions, as described in detail in the next sections.

In the following presentation, we will consider a cross-sectional model, and we will suppress the index q for decision-makers (q=1,2,…,Q) in parts of the presentation, and assume that all error terms are independent and identically distributed across decision-makers.

2.4.1.1 <u>Latent Variable Structural Equation Model</u>

Let l be an index for latent variables (l=1,2,…,L). Consider the latent variable z_l^* and write it as a linear function of covariates:

$$z_l^* = \boldsymbol{\alpha}_l' \boldsymbol{w} + \eta_l, \tag{1}$$

where $\boldsymbol{w}$ is a $(\tilde{D} \times 1)$ vector of observed covariates (excluding a constant), $\boldsymbol{\alpha}_l$ is a corresponding $(\tilde{D} \times 1)$ vector of coefficients, and η_l is a random error term assumed to be standard normally distributed for identification purposes (See Bhat, 2015a). Next, define the $(L \times \tilde{D})$ matrix $\boldsymbol{\alpha} = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, ..., \boldsymbol{\alpha}_L)'$, and the $(L \times 1)$ vectors $\boldsymbol{z}^* = (z_1^*, z_2^*, ..., z_L^*)'$ and $\boldsymbol{\eta} = (\eta_1, \eta_2, \eta_3, \ldots, \eta_L)'$. Unlike much of the earlier research in ICLV modeling, we allow an MVN correlation structure for $\boldsymbol{\eta}$ to accommodate interactions among the unobserved latent variables: $\boldsymbol{\eta} \sim MVN_L[\boldsymbol{0}_L, \boldsymbol{\Gamma}]$, where $\boldsymbol{0}_L$ is an $(L \times 1)$ column vector of zeros, and $\boldsymbol{\Gamma}$ is $(L \times L)$ correlation matrix. In matrix form, we may write Equation (1) as:

$$\boldsymbol{z}^* = \boldsymbol{\alpha} \boldsymbol{w} + \boldsymbol{\eta}. \tag{2}$$

A general covariance structure for the latent variables as in Equation (2) is adopted, therefore, no causal relationship between latent variables is allowed. Bhat (2015) discusses the identification considerations for both cases.

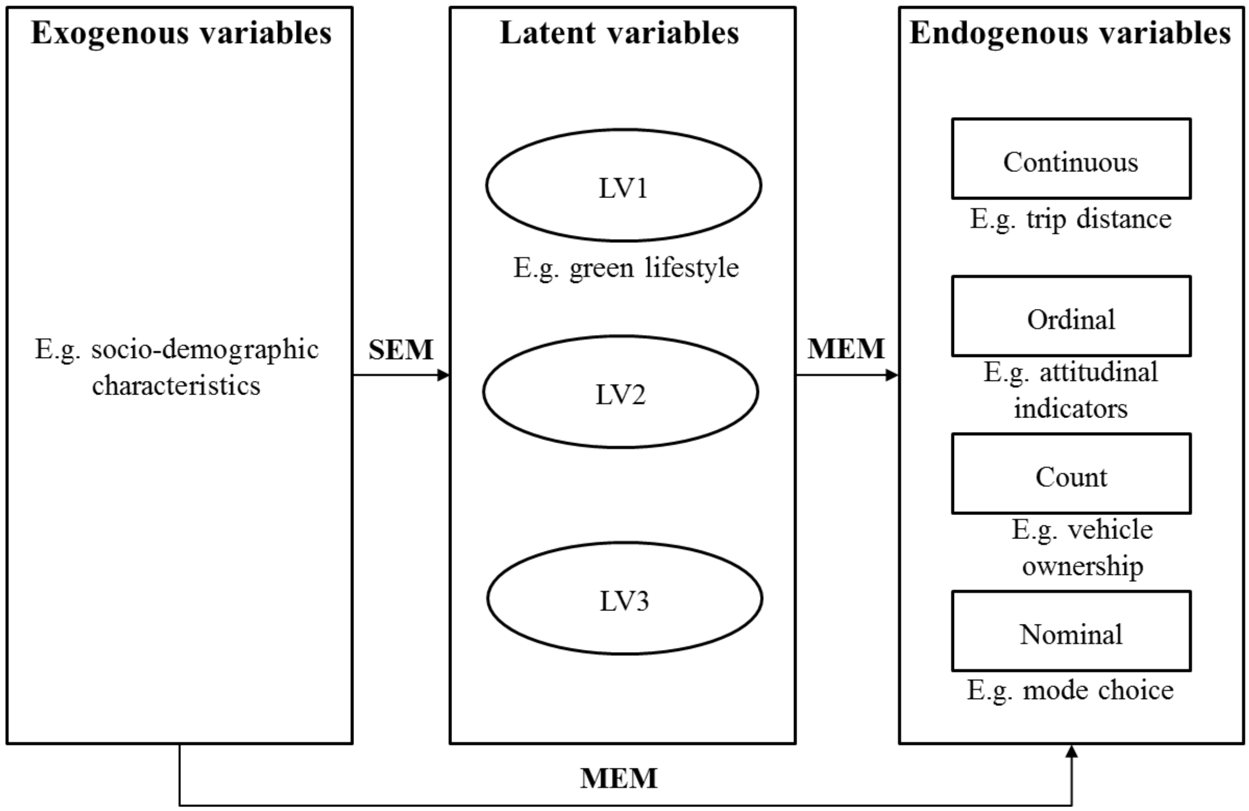


Figure 2-1 Simplified diagram of the GHDM framework

2.4.1.2 Latent Variable Measurement Equation Model Components

We will consider a combination of continuous, ordinal, count, and nominal outcomes (indicators) of the underlying latent variable vector $\boldsymbol{z}^*$. However, these outcomes may be a function of a set of exogenous variables too.

Let there be H continuous outcomes $(y_1, y_2, ..., y_H)$ with an associated index h $(h=1,2,...,H)$. Let $y_h = \boldsymbol{\gamma}_h'\boldsymbol{x} + \boldsymbol{d}_h'\boldsymbol{z}^* + \varepsilon_h$ in the usual linear regression fashion, where $\boldsymbol{x}$ is an $(A\times 1)$ vector of exogenous variables (including a constant) as well as possibly the observed values of other endogenous continuous variables, other endogenous ordinal variables, other endogenous count variables, and other endogenous nominal variables (introduced as dummy variables). $\boldsymbol{\gamma}_h$ is a corresponding compatible coefficient vector.[7] $\boldsymbol{d}_h$ is an $(L\times 1)$ vector of latent

[7] In joint limited-dependent variable systems in which one or more dependent variables are not observed on a continuous scale, such as the joint system considered in the current paper that has discrete dependent and count variables (which we will more generally refer to as limited-dependent variables), the structural effects of one limited-dependent variable on another can only be in a single direction. That is, it is not possible to have correlated unobserved effects underlying the propensities determining two limited-dependent variables, as well as have the observed limited-dependent variables themselves structurally affect each other in a bi-directional fashion.

variable loadings on the h^{th} continuous outcome, and ε_h is a normally distributed measurement error term. Stack the H continuous outcomes into an $(H \times 1)$ vector $\boldsymbol{y}$, and the H error terms into another $(H \times 1)$ vector $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, ..., \varepsilon_H)'$. Also, let $\boldsymbol{\Sigma}$ be the covariance matrix of ε, which is restricted to be diagonal. This helps identification because there is already an unobserved latent variable vector $\boldsymbol{z}^*$ that serves as a vehicle to generate covariance between the outcome variables (as we discuss in the next section). Define the $(H \times A)$ matrix $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, ...\gamma_H)'$ and the $(H \times L)$ matrix of latent variable loadings $\boldsymbol{d} = \left(\boldsymbol{d}_1, \boldsymbol{d}_2, ..., \boldsymbol{d}_H\right)'$. Then, one may write, in matrix form, the following measurement equation for the continuous outcomes:

$$\boldsymbol{y} = \boldsymbol{\gamma x} + \boldsymbol{dz}^* + \boldsymbol{\varepsilon}. \tag{3}$$

Next, consider N ordinal outcomes (indicator variables) for the individual, and let n be the index for the ordinal outcomes $(n = 1, 2, ..., N)$. Also, let J_n be the number of categories for the n^{th} ordinal outcome $(J_n \geq 2)$ and let the corresponding index be j_n $(j_n = 1, 2, ..., J_n)$. Let $\tilde{y}_n^*$ be the latent underlying variable whose horizontal partitioning leads to the observed outcome for the n^{th} ordinal variable. Assume that the individual under consideration chooses the a_n^{th} ordinal category. Then, in the usual ordered response formulation, for the individual, we may write:

$$\tilde{y}_n^* = \tilde{\boldsymbol{\gamma}}_n' \boldsymbol{x} + \tilde{\boldsymbol{d}}_n' \boldsymbol{z}^* + \tilde{\varepsilon}_n, \text{ and } \tilde{\psi}_{n,a_n-1} < \tilde{y}_n^* < \tilde{\psi}_{n,a_n}, \tag{4}$$

where $\boldsymbol{x}$ is a vector of exogenous and possibly endogenous variables as defined earlier, $\tilde{\boldsymbol{\gamma}}_n$ is a corresponding vector of coefficients to be estimated, $\tilde{\boldsymbol{d}}_n$ is an $(L \times 1)$ vector of latent variable loadings on the n^{th} continuous outcome, the $\tilde{\psi}$ terms represent thresholds, and $\tilde{\varepsilon}_n$ is the standard normal random error for the n^{th} ordinal outcome. For each ordinal outcome, $\tilde{\psi}_{n,0} < \tilde{\psi}_{n,1} < \tilde{\psi}_{n,2} ... < \tilde{\psi}_{n,J_n-1} < \tilde{\psi}_{n,J_n}$; $\tilde{\psi}_{n,0} = -\infty$, $\tilde{\psi}_{n,1} = 0$, and $\tilde{\psi}_{n,J_n} = +\infty$. For later use, let $\tilde{\boldsymbol{\psi}}_n = (\tilde{\psi}_{n,2}, \tilde{\psi}_{n,3} ..., \tilde{\psi}_{n,J_n-1})'$ and $\tilde{\boldsymbol{\psi}} = (\tilde{\boldsymbol{\psi}}_1', \tilde{\boldsymbol{\psi}}_2', ..., \tilde{\boldsymbol{\psi}}_N)'$. Stack the N underlying continuous variables $\tilde{y}_n^*$ into an $(N \times 1)$ vector $\tilde{\boldsymbol{y}}^*$, and the N error terms $\tilde{\varepsilon}_n$ into another $(N \times 1)$ vector $\tilde{\boldsymbol{\varepsilon}}$. Define $\tilde{\boldsymbol{\gamma}} = (\tilde{\gamma}_1, \tilde{\gamma}_2, ..., \tilde{\gamma}_H)'$ [$(N \times A)$ matrix] and $\tilde{\boldsymbol{d}} = \left(\tilde{\boldsymbol{d}}_1, \tilde{\boldsymbol{d}}_2, ..., \tilde{\boldsymbol{d}}_N\right)$ [$(N \times L)$ matrix], and let $\mathbf{IDEN_N}$ be the identity matrix of dimension N representing the correlation matrix of $\tilde{\boldsymbol{\varepsilon}}$ (so,

$\tilde{\boldsymbol{\varepsilon}} \sim MVN_N(\mathbf{0}_N, \mathbf{IDEN_N})$; again, this is for identification purposes, given the presence of the unobserved $\boldsymbol{z}^*$ vector to generate covariance. Finally, stack the lower thresholds for the decision-maker $\tilde{\psi}_{n,a_n-1}(n=1,2,\ldots,N)$ into an $(N \times 1)$ vector $\tilde{\boldsymbol{\psi}}_{\boldsymbol{low}}$ and the upper thresholds $\tilde{\psi}_{n,a_n}(n=1,2,\ldots,N)$ into another vector $\tilde{\boldsymbol{\psi}}_{up}$. Then, in matrix form, the measurement equation for the ordinal outcomes (indicators) for the decision-maker may be written as:

$$\tilde{\boldsymbol{y}}^* = \tilde{\boldsymbol{\gamma}}\boldsymbol{x} + \tilde{\boldsymbol{d}}\boldsymbol{z}^* + \tilde{\boldsymbol{\varepsilon}}, \quad \tilde{\boldsymbol{\psi}}_{\boldsymbol{low}} < \tilde{\boldsymbol{y}}^* < \tilde{\boldsymbol{\psi}}_{\boldsymbol{up}}. \tag{5}$$

Let there be C count variables for a household, and let c be the index for the count variables $(c=1,2,\ldots,C)$. Let the count index be $k_c\ (k_c = 0,1,2,\ldots,\infty)$ and let r_c be the actual observed count value for the household. Then, following the recasting of a count model in a generalized ordered-response probit formulation (see Castro, Paleti, and Bhat, 2012 and Bhat et al., 2014b), a generalized version of the negative binomial count model may be written as:

$$\breve{y}_c^* = \breve{\boldsymbol{d}}_{\boldsymbol{c}}'\boldsymbol{z}^* + \breve{\varepsilon}_c,\ \breve{\psi}_{c,r_c-1} < \breve{y}_c^* < \breve{\psi}_{c,r_c}, \tag{6}$$

$$\breve{\psi}_{c,r_c} = \Phi^{-1}\left[\frac{(1-\upsilon_c)^{\theta_l}}{\Gamma(\theta_c)}\sum_{t=0}^{r_c}\left(\frac{\Gamma(\theta_c+t)}{t!}(\upsilon_c)^t\right)\right] + \varphi_{c,r_c},\ \upsilon_c = \frac{\lambda_c}{\lambda_c+\theta_c},\ \text{and}\ \lambda_c = e^{\breve{\gamma}_c'\boldsymbol{x}}. \tag{7}$$

In the above equation, $\breve{y}_c^*$ is a latent continuous stochastic propensity variable associated with the count variable c that maps into the observed count r_c through the $\breve{\boldsymbol{\psi}}_{\boldsymbol{c}}$ vector (which is a vertically stacked column vector of thresholds $(\breve{\psi}_{c,-1}, \breve{\psi}_{c,0}, \breve{\psi}_{c,1}, \breve{\psi}_{c,2}, \ldots)'$. $\breve{\boldsymbol{d}}_{\boldsymbol{c}}$ is an $(L \times 1)$ vector of latent variable loadings on the c^{th} count outcome, and $\breve{\varepsilon}_c$ is a standard normal random error term. $\breve{\boldsymbol{\gamma}}_{\boldsymbol{c}}$ is a column vector corresponding to the vector $\boldsymbol{x}$. Φ^{-1} in the threshold function of Equation (7) is the inverse function of the univariate cumulative standard normal. θ_c is a parameter that provides flexibility to the count formulation, and is related to the dispersion parameter in a traditional negative binomial model $(\theta_c > 0\ \forall\, c)$. $\Gamma(\theta_c)$ is the traditional gamma function; $\Gamma(\theta_c) = \int_{\tilde{t}=0}^{\infty} \tilde{t}^{\,\theta_c-1} e^{-\tilde{t}}\, d\tilde{t}$. The threshold terms in the $\breve{\boldsymbol{\psi}}_{\boldsymbol{c}}$ vector satisfy the ordering condition (*i.e.*, $\breve{\psi}_{c,-1} < \breve{\psi}_{c,0} < \breve{\psi}_{c,1} < \breve{\psi}_{c,2} \ldots. < \infty\ \forall\ c$) as long as $\varphi_{c,-1} < \varphi_{c,0} < \varphi_{c,1} < \varphi_{c,2} \ldots. < \infty$. The presence of the φ_c terms in the thresholds provides substantial flexibility to accommodate high or low probability masses for specific count outcomes without the need for cumbersome traditional treatments using zero-inflated or related mechanisms in multi-dimensional model systems (see Castro, Paleti and Bhat, 2012, for a detailed discussion). For identification, we set $\varphi_{c,-1} = -\infty$ and $\varphi_{c,0} = 0$ for all count variables *c*. In addition, we identify a count value e_c^* $(e_c^* \in \{0,1,2,......\})$ above which φ_{c,k_c} $(k_c \in \{1,2,......\})$ is held fixed at φ_{c,e_c^*}; that is, $\varphi_{c,k_c} = \varphi_{c,e_c^*}$ if $k_c > e_c^*$, where the value of e_c^* can be based on empirical testing. Doing so is the key to allowing the count model to predict beyond the range available in the estimation sample. For later use, let $\boldsymbol{\varphi}_c = (\varphi_{c,1}, \varphi_{c,2}, \dots \varphi_{c,e_c^*})'$ $(e_c^* \times 1$ vector) (assuming $e_c^* > 0$), $\boldsymbol{\varphi} = (\boldsymbol{\varphi}_1', \boldsymbol{\varphi}_2', \dots, \boldsymbol{\varphi}_C')'$ $\left[\left(\sum_c e_c^*\right) \times 1 \text{ vector}\right]$, and $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots \theta_C)'$ $[\text{C} \times 1 \text{ vector}]$. Also, stack the *C* latent variables $\breve{y}_c^*$ into a $(C \times 1)$ vector $\breve{\boldsymbol{y}}^*$, and the C error terms $\breve{\varepsilon}_c$ into another $(C \times 1)$ vector $\breve{\varepsilon}$. Let $\breve{\boldsymbol{\varepsilon}} \sim MVN_C(\mathbf{0}_C, \mathbf{IDEN_C})$ from identification considerations, and stack the lower thresholds of the individual $\breve{\psi}_{c,r_c-1}(c = 1, 2, ..., C)$ into a $(C \times 1)$ vector $\breve{\boldsymbol{\psi}}_{\boldsymbol{low}}$, and the upper thresholds $\breve{\psi}_{c,r_c}(c = 1, 2, ..., C)$ into another $(C \times 1)$ vector $\breve{\boldsymbol{\psi}}_{\boldsymbol{up}}$. Define $\breve{\boldsymbol{\gamma}} = (\breve{\gamma}_1, \breve{\gamma}_2, ..., \breve{\gamma}_C)'$ $[(C \times A)$ matrix] and $\breve{\boldsymbol{d}} = (\breve{\boldsymbol{d}}_1, \breve{\boldsymbol{d}}_2, ..., \breve{\boldsymbol{d}}_C)'$ $[(C \times L)$ matrix]. With these definitions, the latent propensity underlying the count outcomes may be written in matrix form as:

$$\breve{\boldsymbol{y}}^* = \breve{\boldsymbol{d}}\boldsymbol{z}^* + \breve{\varepsilon}, \;\; \breve{\boldsymbol{\psi}}_{\boldsymbol{low}} < \breve{\boldsymbol{y}}^* < \breve{\boldsymbol{\psi}}_{\boldsymbol{up}}. \tag{8}$$

Note also that the interpretation of the generalized ordered-response recasting is that consumers have a latent "long-term" propensity $\breve{y}_c^*$ associated with the demand for each product/service represented by the count *c*, which is a linear function of the latent variable vector $\boldsymbol{z}^*$ (see Castro, Paleti and Bhat, 2012, for a discussion of the interpretation of the generalized ordered-response recasting of count models). Such a specification enables covariance across the count outcomes (through the propensity variables $\breve{y}_c^*$) and between the count outcomes and other mixed outcomes. On the other hand, there may be some specific consumer contexts and characteristics (embedded in $\boldsymbol{x}$) that may dictate how the long-term propensity is manifested in a count demand at any given *instant of time*. Our implicit assumption is that the latent variable vector $\boldsymbol{z}^*$ affects the "long-term" latent demand propensity $\breve{y}_c^*$, but does not play a role in the instantaneous translation of propensity to actual manifested count demand. This allows us to easily incorporate count outcomes within a mixed outcome model, and estimate the resulting model using Bhat (2011) MACML approach. Similarly, an implicit assumption in Equation (8) is that the factors/constraints that are responsible for the instantaneous translation of propensity to manifested count demand (that is, the elements of the $\boldsymbol{x}$ vector) do not affect the "long-term" demand propensity, though this is being imposed purely for parsimony purposes. Relaxing this assumption does not complicate the model system or the estimation process in any way.

Finally, let there be G nominal (unordered-response) variables for an individual, and let g be the index for the nominal variables $(g = 1, 2, 3, ..., G)$. Also, let I_g be the number of alternatives corresponding to the g^{th} nominal variable ($I_g \geq 3$) and let i_g be the corresponding index $(i_g = 1, 2, 3, ..., I_g)$. Consider the g^{th} nominal variable and assume that the individual under consideration chooses the alternative m_g. Also, assume the usual random utility structure for each alternative i_g.

$$U_{gi_g} = \boldsymbol{b}'_{gi_g}\boldsymbol{x} + \boldsymbol{\vartheta}'_{gi_g}(\boldsymbol{\beta}_{gi_g}\boldsymbol{z}^*) + \varsigma_{gi_g}, \tag{9}$$

where $\boldsymbol{x}$ is as defined earlier, $\boldsymbol{b}_{gi_g}$ is an $(A \times 1)$ column vector of corresponding coefficients, and ς_{gi_g} is a normal error term. $\boldsymbol{\beta}_{gi_g}$ is an $(N_{gi_g} \times L)$-matrix of variables interacting with latent variables to influence the utility of alternative i_g, and $\boldsymbol{\vartheta}_{gi_g}$ is an $(N_{gi_g} \times 1)$-column vector of coefficients capturing the effects of latent variables and their interaction effects with other exogenous variables. If each of the latent variables impacts the utility of the alternatives for each nominal variable purely through a constant shift in the utility function, $\boldsymbol{\beta}_{gi_g}$ will be an identity matrix of size L, and each element of $\boldsymbol{\vartheta}_{gi_g}$ will capture the effect of a latent variable on the constant specific to alternative i_g of nominal variable g. Let $\boldsymbol{\varsigma_g} = (\varsigma_{g1}, \varsigma_{g2}, ...\varsigma_{gI_g})'$ $(I_g \times 1$ vector), and $\boldsymbol{\varsigma_g} \sim MVN_{I_g}(\mathbf{0}, \mathbf{\Lambda_g})$. Taking the difference with respect to the first alternative, the only estimable elements are found in the covariance matrix $\breve{\mathbf{\Lambda}}_g$ of the error differences, $\boldsymbol{\breve{\varsigma}_g} = (\breve{\varsigma}_{g2}, \breve{\varsigma}_{g3}, ..., \breve{\varsigma}_{gI_g})$ (where $\breve{\varsigma}_{gi} = \varsigma_{gi} - \varsigma_{g1}, i \neq 1)$. Further, the variance term at the top left diagonal of $\breve{\mathbf{\Lambda}}_g$ $(g = 1, 2, ..., G)$ is set to 1 to account for scale invariance. $\mathbf{\Lambda_g}$ is constructed from $\breve{\mathbf{\Lambda}}_\mathbf{g}$ by adding a row on top and a column to the left. All elements of this additional row and column are filled with values of zero. In addition, the usual identification restriction is imposed such that one of the alternatives serves as the base when introducing alternative-specific constants and variables that do not vary across alternatives (that is, whenever an element of $\boldsymbol{x}$ is individual-specific and not alternative-specific, the corresponding element in $\boldsymbol{b}_{gi_g}$ is set to zero for at least one alternative i_g). To proceed, define $\boldsymbol{U}_g = (U_{g1}, U_{g2}, ..., U_{gI_g})'$ $(I_g \times 1$ vector), $\boldsymbol{b}_g = (\boldsymbol{b}_{\boldsymbol{g}1}, \boldsymbol{b}_{\boldsymbol{g}2}, \boldsymbol{b}_{\boldsymbol{g}3}, ..., \boldsymbol{b}_{\boldsymbol{g}I_g})'$ $(I_g \times A$ matrix), and $\boldsymbol{\beta}_g = (\boldsymbol{\beta}'_{g1}, \boldsymbol{\beta}'_{g2}, ..., \boldsymbol{\beta}'_{gI_g})'$ $\left(\sum_{i_g=1}^{I_g} N_{gi_g} \times L\right)$ matrix.

Also, define the $\left(I_g \times \sum_{i_g=1}^{I_g} N_{gi_g}\right)$ matrix $\boldsymbol{\vartheta_g}$, which is initially filled with all zero values. Then, position the $(1 \times N_{g1})$ row vector $\boldsymbol{\vartheta}'_{g1}$ in the first row to occupy columns 1 to N_{g1}, position the $(1 \times N_{g2})$ row vector $\boldsymbol{\vartheta}'_{g2}$ in the second row to occupy columns N_{g1}+1 to $N_{g1} + N_{g2}$, and so on until the $(1 \times N_{gI_g})$ row vector $\boldsymbol{\vartheta}'_{gI_g}$ is appropriately positioned. Further, define $\boldsymbol{\varpi}_g = (\boldsymbol{\vartheta}_g \boldsymbol{\beta}_g)$ $(I_g \times L$ matrix), $\ddot{G} = \sum_{g=1}^{G} I_g$, $\tilde{G} = \sum_{g=1}^{G} (I_g - 1)$, $\boldsymbol{U} = \left(\boldsymbol{U}'_1, \boldsymbol{U}'_2, \dots, \boldsymbol{U}'_G\right)'$ $(\ddot{G} \times 1$ vector), $\boldsymbol{\varsigma} = (\varsigma_1, \varsigma_2, ... \varsigma_G)'$ $(\ddot{G} \times 1$vector), $\boldsymbol{b} = (\boldsymbol{b}'_1, \boldsymbol{b}'_2, ..., \boldsymbol{b}'_G)'$ $(\ddot{G} \times A$ matrix), $\boldsymbol{\varpi} = (\boldsymbol{\varpi}'_1, \boldsymbol{\varpi}'_2, ..., \boldsymbol{\varpi}'_G)'$ $(\ddot{G} \times L$ matrix), and $\boldsymbol{\vartheta} = \text{Vech}(\boldsymbol{\vartheta}_1, \boldsymbol{\vartheta}_2, ..., \boldsymbol{\vartheta}_G)$ (that is, $\boldsymbol{\vartheta}$ is a column vector that includes all elements of the matrices $\boldsymbol{\vartheta}_1, \boldsymbol{\vartheta}_2, ..., \boldsymbol{\vartheta}_G$). Then, in matrix form, we may write Equation (9) as:

$$\boldsymbol{U} = \boldsymbol{bx} + \boldsymbol{\varpi}\, \boldsymbol{z}^* + \boldsymbol{\varsigma}, \tag{10}$$

where $\boldsymbol{\varsigma} \sim MVN_{\ddot{G}}(\boldsymbol{0}_{\ddot{G}}, \boldsymbol{\Lambda})$. As earlier, to ensure identification, we specify $\boldsymbol{\Lambda}$ as follows:

$$\boldsymbol{\Lambda} = \begin{bmatrix} \boldsymbol{\Lambda}_1 & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Lambda}_2 & \boldsymbol{0} & \boldsymbol{0} \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{\Lambda}_3 & \boldsymbol{0} \cdots & \boldsymbol{0} \\ \vdots & \vdots & \vdots & \vdots \dots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \dots & \boldsymbol{\Lambda}_G \end{bmatrix} (\ddot{G} \times \ddot{G} \text{ matrix}). \tag{11}$$

In the general case, this allows the estimation of $\sum_{g=1}^{G}\left(\frac{I_g*(I_g-1)}{2}-1\right)$ terms across all the G nominal variables, as originating from $\left(\frac{I_g*(I_g-1)}{2}-1\right)$ terms embedded in each $\breve{\mathbf{\Lambda}}_g$ matrix; (g=1,2,…,G) .

Now, we can organize the above as $E=(H+N+C)$. Define $\ddot{\boldsymbol{y}}=\left(\boldsymbol{y}',\left[\tilde{\boldsymbol{y}}^*\right]',\left[\breve{\boldsymbol{y}}^*\right]'\right)'$ [$E\times 1$ vector], $\ddot{\boldsymbol{\gamma}}=(\boldsymbol{\gamma}',\tilde{\boldsymbol{\gamma}}',\mathbf{0}_{AC})'$ [$E \times A$ matrix], $\ddot{\boldsymbol{d}}=(\boldsymbol{d}',\tilde{\boldsymbol{d}}',\breve{\boldsymbol{d}}')'$ [$E\times L$ matrix], and $\ddot{\boldsymbol{\varepsilon}}=(\boldsymbol{\varepsilon}',\tilde{\boldsymbol{\varepsilon}}',\breve{\boldsymbol{\varepsilon}}')'$ ($E\times 1$ vector), where $\mathbf{0}_{AC}$ is a matrix of zeros of dimension $A\times C$. Let $\boldsymbol{\delta}$ be the collection of parameters to be estimated: $\boldsymbol{\delta}=[\text{Vech}(\boldsymbol{\alpha}),\text{Vech}(\boldsymbol{\Sigma}),\text{Vech}(\ddot{\boldsymbol{\gamma}}),\text{Vech}(\ddot{\boldsymbol{d}}),\text{Vech}(\breve{\boldsymbol{\gamma}}),\ \boldsymbol{\varphi},\ \boldsymbol{\theta},\text{Vech}(\boldsymbol{b}),\boldsymbol{\vartheta},\text{Vech}(\boldsymbol{\Lambda})]$, where the operator "$\text{Vech}(\mathbf{.})$" vectorizes all the non-zero elements of the matrix/vector on which it operates. We will assume that the error vectors $\boldsymbol{\tau}$, ε, $\boldsymbol{\xi}$, and $\boldsymbol{\varsigma}$ are independent of each other. While this assumption is not strictly necessary (and can be relaxed in a very straightforward manner within the estimation framework of our model system as long as the resulting model is identified), the assumption aids in developing general sufficiency conditions for identification of parameters in a mixed model when the latent variable vector $\boldsymbol{z}^*$ already provides a mechanism to generate covariance among the mixed outcomes.

With the matrix definitions above, the continuous components of the model system may be written compactly as:

$$\boldsymbol{z}^*=\boldsymbol{\alpha}\boldsymbol{w}+\boldsymbol{\eta}, \tag{12}$$

$$\ddot{\boldsymbol{y}}=\ddot{\boldsymbol{\gamma}}\boldsymbol{x}+\ddot{\boldsymbol{d}}\boldsymbol{z}^*+\ddot{\boldsymbol{\varepsilon}},\ \text{with Var}(\ddot{\boldsymbol{\varepsilon}})=\ddot{\boldsymbol{\Sigma}}=\begin{bmatrix}\boldsymbol{\Sigma} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{IDEN_N} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{IDEN_C}\end{bmatrix}\ (E\times E\ \text{matrix})\ , \tag{13}$$

$$\boldsymbol{U}=\boldsymbol{b}\boldsymbol{x}+\boldsymbol{\varpi}\boldsymbol{z}^*+\boldsymbol{\varsigma}. \tag{14}$$

To develop the reduced form equations, replace the right side of Equation (12) for $\boldsymbol{z}^*$ in Equations (13) and (14) to obtain the following system:

$$\ddot{\boldsymbol{y}} = \ddot{\boldsymbol{\gamma}}\boldsymbol{x} + \ddot{\boldsymbol{d}}\boldsymbol{z}^* + \ddot{\boldsymbol{\varepsilon}} = \ddot{\boldsymbol{\gamma}}\boldsymbol{x} + \ddot{\boldsymbol{d}}(\boldsymbol{\alpha w} + \boldsymbol{\eta}) + \ddot{\boldsymbol{\varepsilon}} = \ddot{\boldsymbol{\gamma}}\boldsymbol{x} + \ddot{\boldsymbol{d}}\boldsymbol{\alpha w} + \ddot{\boldsymbol{d}}\boldsymbol{\eta} + \ddot{\boldsymbol{\varepsilon}} \;, \tag{15}$$

$$\boldsymbol{U} = \boldsymbol{bx} + \boldsymbol{\varpi}\,\boldsymbol{z}^* + \varsigma = \boldsymbol{bx} + \boldsymbol{\varpi}\,(\boldsymbol{\alpha w} + \boldsymbol{\eta}) + \varsigma = \boldsymbol{bx} + \boldsymbol{\varpi}\,\boldsymbol{\alpha w} + \boldsymbol{\varpi}\,\boldsymbol{\eta} + \varsigma \;.$$

Now, consider the $[(E + \ddot{G}) \times 1)]$ vector $\boldsymbol{yU} = \left[\ddot{\boldsymbol{y}}', \boldsymbol{U}'\right]'$. Define

$$\boldsymbol{B} = \begin{bmatrix} \boldsymbol{B_1} \\ \boldsymbol{B_2} \end{bmatrix} = \begin{bmatrix} \ddot{\boldsymbol{\gamma}}\boldsymbol{x} + \ddot{\boldsymbol{d}}\boldsymbol{\alpha w} \\ \boldsymbol{bx} + \boldsymbol{\varpi}\,\boldsymbol{\alpha w} \end{bmatrix} \text{ and } \boldsymbol{\Omega} = \begin{bmatrix} \boldsymbol{\Omega}_1 & \boldsymbol{\Omega}_{12}' \\ \boldsymbol{\Omega}_{12} & \boldsymbol{\Omega}_2 \end{bmatrix} = \begin{bmatrix} \ddot{\boldsymbol{d}}\boldsymbol{\Gamma}\ddot{\boldsymbol{d}}' + \ddot{\boldsymbol{\Sigma}} & \ddot{\boldsymbol{d}}\boldsymbol{\Gamma}\boldsymbol{\varpi}' \\ \boldsymbol{\varpi}\boldsymbol{\Gamma}\ddot{\boldsymbol{d}}' & \boldsymbol{\varpi}\boldsymbol{\Gamma}\boldsymbol{\varpi}' + \boldsymbol{\Lambda} \end{bmatrix}. \tag{16}$$

Then $\boldsymbol{yU} \sim \boldsymbol{MVN}_{\boldsymbol{E}+\ddot{\boldsymbol{G}}}(\boldsymbol{B}, \boldsymbol{\Omega})$.

The model estimation is performed using Bhat's (2011) MACML. We refer the reader to Bhat (2015a) for the detailed explanation as well as information on model identification criteria.

2.5 Summary

Table 2-1 presents a summary of the four analysis components in this dissertation, listing the main outcomes (endogenous variables), data type and sample size, as well as modeling approach used in each analysis.

Table 2-1: Summary of each analysis component

<table>
<tr><th rowspan="2">Analysis component</th><th rowspan="2">Chapter</th><th colspan="2">Main endogenous variables</th><th rowspan="2">Data type and sample characteristics</th><th rowspan="2">Modeling methodology</th></tr>
<tr><th>Present behavior</th><th>Future intentions</th></tr>
<tr><td rowspan="2">1</td><td rowspan="2">3</td><td rowspan="2">Average number of trips on a weekday and on a weekend day</td><td rowspan="2">--</td><td>Aggregate, passive and revealed</td><td rowspan="2">Spatially lagged multivariate count model and fractional split model</td></tr>
<tr><td>~700 thousand trips averaged across week and weekend days and distributed across 458 TAZ in Austin, TX</td></tr>
<tr><td rowspan="2">2</td><td rowspan="2">4</td><td rowspan="2">Residential density, vehicle ownership, ride-hailing choice and frequency, pooled ride-railing choice and frequency, characteristics of last ride-hailing trip (purpose, time-of-day, companion and mode substituted)</td><td rowspan="2">--</td><td>Disaggregate, active and revealed</td><td rowspan="2">GHDM</td></tr>
<tr><td>1607 commuters of the DFW Metropolitan Area, TX</td></tr>
<tr><td rowspan="2">3</td><td rowspan="2">5</td><td rowspan="2">Residential density, vehicle ownership, ride-hailing choice, and carsharing choice</td><td rowspan="2">Choice between using SAVs, owning an AV, both, and none</td><td>Disaggregate, active, revealed and stated</td><td rowspan="2">GHDM</td></tr>
<tr><td>1832 adults of the Puget Sound Region, WA</td></tr>
<tr><td rowspan="2">4</td><td rowspan="2">6</td><td rowspan="2">Ride-railing choice</td><td rowspan="2">Choice between solo and pooled SAV for commute trips and for leisure trips</td><td>Disaggregate, active, revealed and stated</td><td rowspan="2">GHDM</td></tr>
<tr><td>1607 commuters of the DFW Metropolitan Area, TX</td></tr>
</table>

CHAPTER 3. Ride-Hailing Trip Generation and Distribution

The majority of the content of this chapter is part of a published paper: Lavieri, P.S., Dias, F.F., Juri, N.R., Kuhr, J. and Bhat, C.R., 2017. A model of ridesourcing demand generation and distribution. *Transportation Research Record*. (https://doi.org/10.1177/0361198118756628)

3.1 Introduction

In this chapter, we model and analyze the demand for ride-hailing based on an open source database released by RideAustin, a nonprofit TNC in Austin, Texas. Using six months of detailed trip data, including trip origin and destination location and corresponding time stamps, we develop a two-step modeling framework to investigate the generation and distribution of daily ride-hailing trips at the traffic analysis zone (TAZ) level (the RideAustin data set does not provide user information and corresponding socio-demographic characteristics; therefore, our analysis is undertaken at the zonal level and relies on zonal demographics to infer ride-hailing demand characteristics). As discussed in Chapter 2, we use a spatial bivariate count model to analyze ride-hailing trip generation and inform our understanding of the characteristics of the demand for this service. The use of a spatial analysis technique is important because spatial dependencies in TAZ-level trip generation are likely to exist. Subsequently, we apply a fractional split distribution model to identify zonal characteristics that attract ride-hailing trips and to examine how far individuals are willing to travel by this mode. Examples of explanatory variables used in our analysis are zonal distributions of income, gender, race/ethnicity, age, population and employment density. We also consider transit supply characteristics and land use information regarding presence of parks and universities.

3.2 Data

Several public data sets were compiled to undertake the analysis. The primary data source originated from RideAustin, a TNC operating in Austin, Texas. RideAustin entered the Austin ride-hailing market in 2016, shortly after Uber and Lyft shut down their operations in the city due to disputes over local regulations. The RideAustin data (RideAustin, 2017) provides trip-level information, including the location of trip origins and destinations, total trip length, and corresponding fare. To protect their clients' privacy, RideAustin added noise to the locations of the pickups and drop-offs. The dataset contains a total of 1,494,125 trips that occurred between June 4th, 2016 and April 13th, 2017. Since ridership during the first few months was limited, our analysis only includes data from August 2016 through January 2017. Based on information provided by the Austin Department of Transportation, we estimate that, during that semester, RideAustin was responsible for one third of Austin's ride-hailing market share, suggesting high representativeness of the data. The trip information is supplemented using TAZ-level demographic data obtained from the Capital Area Metropolitan Planning Organization (CAMPO) website and planning toolkit, and the most recent Census estimates. GTFS (General Transit Feed Specification) data is used to estimate the characteristics of the transit system (Texas Government, 2017).

3.2.1 Data Preparation

Raw data, including trip origins and destinations, transit availability, land use, and demographics were mapped to the TAZs defined by CAMPO using GIS software. Given the sparseness of origins and destinations in the outskirts of the city, we chose to focus this study on trips that originate in Central Austin, in the region delimited by Highway 183 to the east and north, Highway 290 to the south, and Texas State Highway Loop 1 (MoPac) to the north. There are 458 TAZs in the area of analysis. The Austin-Bergstrom International Airport is outside the area of interest, but it attracts a large number of trips, so it was modeled as a special external zone in the fractional split model; a second dummy TAZ was used as the destination of all the trips that end outside the study area.

The trip data processing involved calculating the average number of daily trips per origin, and the corresponding average daily split by destination. Separate values were computed for weekdays and weekends. Demographic variables by TAZ were computed using data from the most recent Census, while land-use variables were obtained from the CAMPO planning model. GTFS data was aggregated to generate metrics of transit accessibility, including transit stops per zone, and the average frequency of buses per stop for weekdays and weekends.

3.2.2 Data Description

This research models the average daily count of ride-hailing trips for weekdays and weekends. The spatial unit of analysis is a TAZ. Table 3-1 presents the descriptive statistics for all the variables used in the model, and the year when the corresponding data was collected.

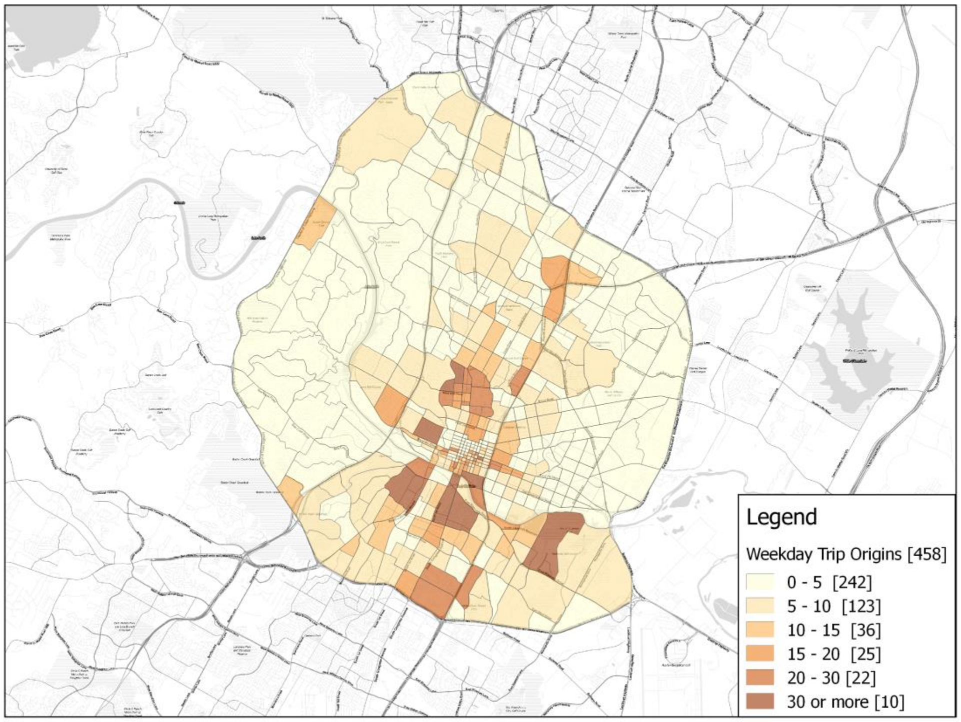
The analysis of descriptive statistics shows a large dispersion in the number of trips generated per zone. Figure 3-1 illustrates the spatial distribution of trips on an average week and weekend day. There is a clear concentration of trips in central and denser areas on both types of days. During weekdays, trips are more concentrated in specific zones that contain universities, parks, or active nightlife. On average there are almost four times more trips generated on a weekend day than on a weekday. These observations are consistent with Rayle et al.'s (2016) results, which suggest, as in San Francisco, that ride-hailing in Austin too is used more for social and leisure activities than work-related activities. Indeed, Hampshire et al. (2017) recently conducted an online survey in Austin and identified the same pattern. On a related note, the average cost of a ride-hailing trip in our sample is US$12.77.

The analysis of transit supply variables suggests that the distribution of transit in Austin is rather heterogeneous. The frequency of bus service averages at 3.12 per hour during weekdays (an average headway of about 20 minutes) and averages 1.64 per hour on weekend days (an average headway of about 37 minutes).

In terms of socio-demographics, there is again a large variation across zones in population density and employment density. The race/ethnicity and education variables indicate a predominantly white and highly educated population. There is a good distribution of individuals in the 18-60 age range. Households are small in size (average of less than 2 individuals), have a mean income of $48,000, and have high vehicle ownership rates (more than half of the sample has at least two vehicles per household). Finally, three variables considered in our model, but not presented in the table, are binary variables representing the presence of parks in a zone, the presence of The University of Texas (UT) territory in a zone, and an indicator of whether a zone is a central business district.

Table 3-1 Sample Descriptive Statistics (458 TAZs)

Variable	Min.	Max.	Mean	Std. Dev.
Outcomes [2016]				
Number of trips in a weekday	0.00	125.00	8.62	11.18
Number of trips in a weekend day	1.00	420.00	31.45	19.89
Transit Supply [2016]				
Number of bus stops	0.00	27.00	3.49	3.62
Frequency of buses in a weekday (bus per hour)	0.00	20.90	3.12	3.25
Frequency of buses in a weekend day (bus per hour)	0.00	13.70	1.64	1.66
Socio-Demographic Variables [2010]				
Population density (population per km^2)	0.00	59,257.67	4,603.24	6,167.92
Employment density (employment per km^2)	0.00	161,932.22	6,813.73	17,545.97
Employment density in retail sector (employment per km^2)	0.00	46,442.92	1,167.67	3,731.79
Race/Ethnicity Variables[2015]				
Proportion of White population	0.00	1.00	0.65	0.34
Proportion of Black and African American population	0.00	0.65	0.05	0.09
Proportion of Asian population	0.00	0.63	0.04	0.06
Proportion of other races/ethnicities	0.00	0.45	0.07	0.08
Educational Attainment Distribution [2015]				
Proportion of population 18 years and above with less than Associate degree	0.00	1.00	0.31	0.26
Proportion of population 18 years and above with Associate or Bachelor's degree or higher	0.00	0.67	0.30	0.19
Proportion of population 18 years and above with Graduate degree	0.00	0.64	0.20	0.15
Age Distribution [2015]				
Proportion of population aged 17 years and below	0.00	0.48	0.13	0.11
Proportion of population aged 18-29 years	0.00	0.99	0.22	0.20
Proportion of population aged 30-39 years	0.00	0.44	0.15	0.10
Proportion of population aged 40-59 years	0.00	0.49	0.20	0.13
Proportion of population aged 60 years and above	0.00	0.44	0.11	0.10
Median Household Size [2010]	0.00	4.00	1.76	1.00
Median Annual Household Income (US$) [2010]	0.00	248,200.00	48,812.00	48,049.00
Household Vehicle Ownership [2010]				
Proportion of household with zero vehicles	0.00	0.08	0.02	0.01
Proportion of household with one vehicle	0.00	0.83	0.37	0.25
Proportion of household with two or more vehicles	0.00	0.95	0.61	0.28
Distance between Centroids of Census Tracts (km)	0.10	19.08	5.75	3.31



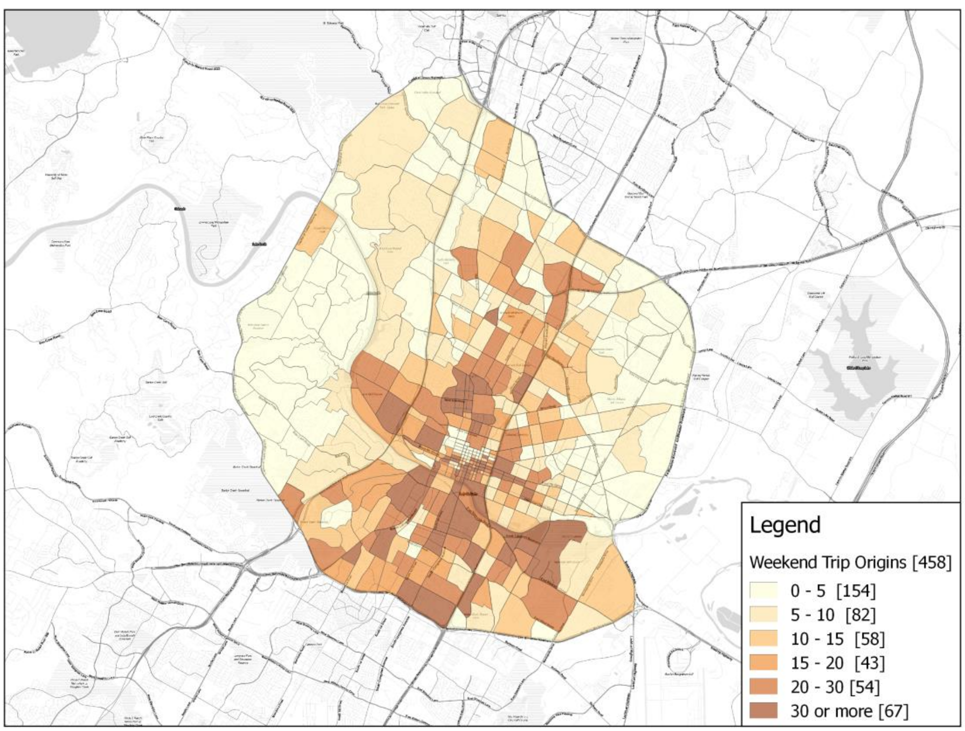


Figure 3-1 Spatial distribution of trips on an average weekday (top) and on an average weekend day (bottom)

3.3 Results

In this section we present and discuss the results for the trip generation and distribution models. We considered all the variables presented in Table 3-1 in our analysis. Several variables and functional forms (including logarithmic transformations) were tested to arrive at the final specification. The model estimation process was guided by prior research, intuitiveness, and parsimony considerations. A few variables that were only marginally statistically significant (i.e., not significant at the 0.05 level of significance) were retained in the final model specification because of their intuitive effects and potential to guide future research efforts.

3.3.1 Trip Generation by Day-Type

As discussed in Chapter 2, the count variable is viewed as a result of a latent demand generation propensity that gets mapped into the observed trip counts through thresholds that are themselves also functions of exogenous variables. The first half of Table 3-2 presents the results for the demand generation propensity. In the first row of results for the weekday trips model, we observe a positive effect of the variable representing the presence of The University of Texas (UT) in the zone. This positive effect indicates that UT is an area with a high intensity of ride-hailing trip origins during a typical weekday, presumably a combination of activity opportunities in the UT area and because students are a segment of the population more likely to use ride-hailing than other population segments. The effect of transit supply in Table 3-2 indicates the expected negative effect, suggesting that ride-hailing decreases as transit service improves. Another perspective is that ride-hailing tends to get used more in areas with relatively poor transit service. Areas with higher residential density and activity intensity, not surprisingly, have more originating ride-hailing trips, a finding supported by earlier studies (Clewlow and Mishra, 2017; Dias et al., 2017). Interestingly, on weekdays, more ride-hailing trips are generated from high activity intensity zones than from high population density zones, while the reverse holds on weekend days. This suggests that ride-hailing is more used after an out-of-home activity on weekdays, and more used from home as an individual leaves home for an out-of-home activity on weekend days.

In terms of population characteristics, zones with higher proportions of white population present a lower propensity to generate ride-hailing trips, both during the weekday and the weekend day. The white population is historically associated with a higher use of the “drive alone” mode (Giuliano, 2003; Smart, 2015) than other segments of the population, which may explain the negative signs on ride-hailing use. The proportion of young adults (18 to 29 years of age) in the zone contributes to an increase in the propensity of ride-hailing trips. This result corroborates findings from previous studies (Rayle et al., 2016; Clewlow and Mishra, 2017). The effect of the median household income of the zone is interesting. It shows that wealthier areas are associated with an increase in the weekday ride-hailing trips, but a decrease in weekend ride-hailing trips. The literature often suggests that ride-hailing users are in the high income segments of the population (Clewlow and Mishra, 2017; Dias et al., 2017). Our results are not inconsistent with the previous literature, but suggest that there is heterogeneity in the income effect based on day of the week. Perhaps high income individuals "buy" time on weekdays through ride-hailing (because they can work or relax rather than drive), while low income individuals gravitate toward ride-hailing on weekend days because of relatively poor transit levels of service. Another possible explanation could be that higher income individuals conduct more social and recreational activities during the week (compared to lower income segments) and use ride-hailing to access these activities. Finally, as expected higher rates of vehicle ownership are associated with a decrease in the generation of ride-hailing trips, a result also observed in Dias et al. (2017) and in Chapter 5.

The second half of Table 3-2 presents the threshold results. The elements of the α vector do not have any substantive interpretations, but play the very important role of accommodating high or low probability masses for specific outcomes The elements in the γ vector are presented next in Table 3-2. The constants within the γ vector do not have any particular interpretation. For the other variables, a positive coefficient shifts all the thresholds toward the left of the demand intensity scale (see Castro, Paleti and Bhat (2012) for a detailed discussion), which has the effect of reducing the probability of the zero trip count. A negative coefficient, on the other hand, shifts all thresholds toward the right of the generation propensity scale, which has the effect of increasing the probability of the zero count. We observe that the proportion of the male population in a zone has opposite effects on weekdays and weekends. Zones that have a higher male population proportion are more likely (than zones with a higher female population proportion) to have non-zero ride-hailing trips during the weekday and zero ride-hailing trips during the weekend days. Also, zones with high vehicle ownership rates are more likely to have zero ride-hailing trips generated in a weekday, while zones that have parks are less likely to have zero trips generated in a weekend. Both results are expected, since having vehicles available in the household reduces the necessity of seeking alternative modes, while parks are associated with recreational activities that are more prevalent on weekends.

Finally, at the bottom of the table we present the cross-correlation between weekdays and weekend days as well as the spatial autoregressive parameter. For the spatial correlation between zones we tested two different weight matrices, one based on the inverse of the distance between the centroid of the zones and another based on the inverse of the squared distance. The best model fit was obtained with the first one. The results confirm ours hypothesis that the number of trips that a zone generates in a weekday is positively associated with the number of trips generated on a weekend day. Additionally, the number of trips in a zone is influenced by observed and unobserved factors of the neighboring zones.

3.3.2 Trip Distribution by Day-Type

The University of Texas has a positive effect on trip attractions during the weekday, suggesting that people might be using ride-hailing to access the campus area at significantly higher rates than other zones. This effect seems to disappear, however, on weekends. This is very likely due to the reduced number of activities on campus during weekends, which results in less people visiting (and traveling to) the area. While zones located in the central business district (CBD) do not attract ride-hailing trips more so than zones in other areas of town, the demand to such CBD zones does decrease on weekends, likely due to the lack of activities during that period. The coefficients related to the airport and external zones are somewhat difficult to interpret directly since zones that fell in these categories had no associated data besides the trip cost. Therefore, many effects are entangled and cannot be immediately interpreted.

The proportion of retail employment, regardless of the day of the week, positively impacts trip attractions. Retail employment may be viewed as a proxy for opportunities for out-of-home activities: people will tend to travel more to places that have activities they want to partake in, such as shopping, and this appears to have a direct positive association with ride-hailing destination points. As expected, there is a positive influence of population density on trip attraction, representing return-home trips. Curiously, though, this effect is not statistically significant during the weekend. This could be a simple reflection of the higher number of out-of-home activities pursued during the weekend days. Thus, even if there are more ride-hailing return-home trips on weekend days than on weekdays, the proportion of such trips (as a fraction of total trips) may be lower on weekend days.

Table 3-2 Estimation Results of the Trip Generation Count Model

Variables	Weekday		Weekend Day	
	Estimate	(t-stat)	Estimate	(t-stat)
Determinants of the latent demand generation propensity (y^*_{qs})				
Special Land-Use				
If all or part of the zone is occupied by The University of Texas	1.271	(2.34)	--	--
Transit Supply				
Average frequency of buses in an average bus stop in the zone (bus per hour)	-0.023	(-1.75)	--	--
Residential Density				
Population density (Natural Logarithm of person per km²)	0.436	(10.17)	0.753	(36.97)
Activity Intensity				
Retail employment density (Natural Logarithm of retail jobs per km²)	0.524	(11.74)	0.492	(14.58)
Population Characteristics				
Proportion of White population	-3.644	(-10.18)	-0.488	(-11.13)
Proportion of population 18-29 years old	2.124	(5.64)	0.298	(1.94)
Proportion of population 30-49 years old	0.225	(1.95)	--	--
Median annual household income (divided by $10,000)	0.056	(3.77)	-0.073	(-3.58)
Proportion of households with 2 or more automobiles	--	--	-2.669	(-25.67)
Demand tipping points (threshold component)				
$\boldsymbol{\alpha}_1$	--	--	--	--
$\boldsymbol{\alpha}_{10}$	-0.240	(-2.96)	-0.350	(-2.66)
$\boldsymbol{\alpha}_{20}$	-0.474	(-2.55)	--	--
$\boldsymbol{\alpha}_{25}$	--	--	-0.661	(-5.48)
Determinants of the thresholds (γ vector elements)				
Constant	2.279	(50.92)	0.182	(6.39)
Population Characteristics				
Proportion of male population	0.366	(14.08)	-0.345	(-6.75)
Households Characteristics				
Proportion of households with 2 or more automobiles	-1.487	(-18.70)	--	--
Special Land-Use				
Presence of parks in the zone	--	--	0.312	(2.94)
Correlation between weekday and weekend	0.394		(9.53)	
Spatial Autoregressive Parameter (δ)	0.561		(29.74)	
Composite Marginal log-likelihood	-250,389.50			

Note: '--' means that the corresponding coefficient was not statistically significantly different from zero at the 90% level of confidence.

As explained previously, the negative effect of the proportion of the white population in a zone on ride-hailing trips generated from the zone may be a reflection of an intrinsic dislike for non-private travel (that is, a generic private auto-inclination). Similarly, the results in Table 3-3 indicate that, as the proportion of households with two or more vehicles in a zone increases, the "attractiveness" of the zone as a terminating point for ride-hailing trips decreases.

The average monetary cost of the trip plays a significant role in the trip distribution process for both weekends and weekdays. Throughout the estimation process, both distance and cost were used, but these two variables were too strongly correlated for both of them to be statistically significant. Therefore, given that the cost variable successfully explained most of the variance of these two variables, the distance variable was omitted from the estimation. The final variable is a pure size effect.

Table 3-3 Estimation Results of the Trip Distribution Split Model

	Weekday		**Weekend Day**	
Variables	Estimate	(t-stat)	Estimate	(t-stat)
Constant	-5.051	(-13.60)	-2.098	(-14.74)
Special Land-Use				
If all or part of the zone is occupied by The University of Texas	0.711	(2.85)	--	--
If the area is in the central business district	--	--	-0.673	(-4.37)
If the area contains the airport	4.142	(14.81)	--	--
If the area is outside the area of interest	3.119	(7.85)	--	--
Residential Density				
Population density (Natural Logarithm of person per km²)	0.110	(2.35)	--	--
Activity Intensity				
Retail employment density (Natural Logarithm of retail jobs per km²)	0.179	(3.33)	0.164	(3.34)
Population Characteristics				
Proportion of White population	--	--	-1.653	(-4.90)
Median annual household income (divided by $10,000)	--	--	0.069	(3.72)
Proportion of households with 2 or more automobiles	-0.851	(-2.36)	-1.723	(-5.03)
Proportion of males	--	--	2.089	(3.94)
Trip Characteristics				
Log of average cost between zones	-0.422	(-2.97)	-0.378	(-4.99)
Other Characteristics				
Area (km²)	0.367	(3.68)	0.218	(1.94)

Note: '--' means that the corresponding coefficient was not statistically significantly different from zero at the 90% level of confidence.

3.4 Conclusions

This chapter has undertaken an analysis of the demand for ride-hailing trips in the city of Austin, Texas. Based on data provided by a non-profit TNC that entered Austin's market after the exit of Uber and Lyft, we develop two models that analyze characteristics of the generation and distribution of ride-hailing trips at a TAZ level. Several public data sets were compiled to complete the analysis, including TAZ-level demographic data obtained from the Capital Area Metropolitan Planning Organization, the most recent Census estimates, and GTFS available from the state of Texas website. The use of open source data is in its early stages and this chapter provides a first glimpse of the potential that these data sources have in informing transportation models.

Our model provides important initial insights on characteristics of ride-hailing demand. Additionally, it identifies interesting heterogeneities between ride-hailing use on weekdays and weekend days. For example, in the context of university campuses, our results suggest that students may be the beneficiaries of the availability of ride-hailing services. This may be either because vehicle ownership rates among university students is lower (compared to working individuals), or because of restricted parking regulations and high parking fees in such areas. Moreover, bus frequencies seem to have a negative impact on the generation of ride-hailing trips during the week, suggesting a substitution effect between ride-hailing and transit use. Another interesting finding is that the effect of the median household income in a zone on trip generation is opposite for weekdays and weekend days, suggesting that different income segments in the population may use ride-hailing for different activity purposes. Overall, the estimated parameters of the multivariate count model can be used to forecast the number of new ride-hailing trips in a TAZ in response to changing TAZ economics and demographics. The trip distribution model indicated that, as expected, the airport is a major ride-hailing trip attraction. This result leads to the question of whether only taxi and carpooling trips to the airport are being substituted, or if travelers who used to park at airports earlier are now opting for ride-hailing instead. A better understanding of this issue can help future parking planning at airports. Finally, the trip distribution model also provides evidence of the substantial use of weekday ride-hailing for returning home, which may suggest that ride-hailing is becoming integrated into the multi-modal use routine of individuals and/or is being used to avoid driving while impaired.

The results and methods used in this study can serve multiple purposes. First, from a travel behavior researcher perspective, we have identified aggregate-level variables that impact ride-hailing, and can guide efforts to better understand the demand for autonomous and connected vehicles in the future. Second, from a planner's perspective, we provide an analytic framework to develop predictive models of ride-hailing movements that can be accommodated in regional and planning network models. Finally, our results may also have relevance to operators in their understanding of travel demand, which can lead to better strategies to allocate drivers to rides, or to estimate optimal fleet size when entering a new market.

CHAPTER 4. Individual Adoption and Use of Ride-Hailing Services

4.1 Introduction

The analysis conducted in this chapter aims to complement that in the previous chapter by developing two disaggregate multi-dimensional models of ride-hailing behavior, one at an individual-level and the second at a trip-level. In the first individual-level model, ride-hailing experience and frequency are jointly modeled as functions of unobserved lifestyle stochastic latent constructs, and observed transportation-related choices and sociodemographic variables. Ride-hailing experience is represented as a nominal dependent variable with three categories: (1) no experience with ride-hailing services, (2) experience only with private services (the individual traveled alone or with people s/he knew), and (3) experience with private and pooled services (the individual has, at least once, traveled with strangers for a cheaper fare). Ride-hailing frequency corresponds to the number of trips made by ride-hailing users within a one-month period prior to the date of the survey. In addition to ride-hailing experience and frequency, we also consider residential location (in three nominal categories, as discussed in the Section 2.1.2) and household vehicle availability (in three ordinal categories, also discussed in Section 2.1.2) as co-endogenous variables in this first individual-level model. These variables are considered in our analysis to account for the possibility that residential location and vehicle availability, along with ride-hailing behavior, are determined as a choice bundle and to accommodate for any self-selection effects in the influence of residential location and vehicle ownership on ride-hailing behavior (our expectation, though, is that these self-selection effects will be rather small, because ride-hailing is a relatively recent mobility option available within the past five years, while residential location and vehicle ownership decisions are typically made at longer time intervals than five years). The modeling methodology adopted is based on the Generalized Heterogeneous Data Model (GHDM) developed by Bhat (2015a), which allows for the joint estimation of multiple outcomes of different types (continuous, ordinal, count and nominal) by establishing a parsimonious dependence structure through stochastic latent variables. In the second trip-level model, four nominal dimensions of the individual's last ride-hailing trip are modeled simultaneously. The first nominal variable is trip purpose, captured in the four categories of airport trips, errand trips (including shopping, personal business, and family errand trips), recreation trips (including leisure, social activities and sports), and work trips (including education trips). The second dimension is time-of-day in the four time windows of morning (6:00 am-10:59 am), mid-day (11:00 am-3:59 pm), evening (4:00 pm-8:59 pm), and night (9:00 pm-5:59 am). The third is companionship (in the two categories of alone or with others). The fourth dimension is the mode substituted by ride-hailing (based on the response to the question "if ride-hailing were not available, which mode would you have used for the trip"), in the four categories of (a) private vehicle, (b) taxi, (c) transit and/or active travel (walk/bicycle), and (d) no trip (that is, the trip would not have been made if ride-hailing were not available). A multivariate multinomial probit (MMNP) modeling approach (see Bhat et al., 2013) is utilized so that common unobserved individual-level factors that affect multiple trip characteristics are captured through error correlations across the choice dimensions of interest. In combination, the results from the two multivariate models developed in this chapter serve as inputs to two broader travel behavior questions: (1) Is pooled ride-hailing a feasible MaaS solution in currently car-dominated cities?, and (2) Is there evidence of the presence of positive and negative externalities of ride-hailing adoption?

4.2 Data and Methodology

The data used for the analysis was obtained through a web-based survey. The distribution was achieved through mailing lists held by multiple entities (local transportation planning organizations, universities, private transportation sector companies, non-profit organizations, and online social media), yielding a final clean sample of 1,607 respondents. To focus on individuals with commute travel, the survey was confined to individuals who had their primary work place outside their homes. Respondents were presented with the definition of ride-hailing as "Ride-hailing services use websites and mobile apps to pair passengers with drivers who provide passengers with transportation in the driver's non-commercial vehicle. Examples are Uber and Lyft.", and then were asked if they had ever used this type of service. The sub-sample that answered positively was further presented with a definition of pooled ride-hailing ("In the carpooling option of ride-sourcing, additional passengers with similar routes get picked and dropped off in the middle of the customer's ride. Customers receive discounted rates when they choose this option") and asked about the use of such a pooled ride-hailing service. Based on the responses to these questions, and as applicable, the respondents were asked to indicate their frequency of use, in the past 30 days, of private and pooled ride-hailing services. Also, all respondents who indicated the use of ride-hailing services at some point in their lives were asked to recall the details of their last ride-hailing trip and provide information on trip purpose, time of day of travel, companionship, and mode substituted. The survey also collected socio-demographic and attitudinal information.

Table 4-1 presents the socio-demographic distribution of the sample. A comparison of the sample with the employed population of DFW (as characterized by the U.S. Census Bureau, 2018d) indicates that the survey has an overrepresentation of males (58.4% in the survey compared to 54.0% from the Census data), individuals between 45 and 64 years of age (53.2% compared to 35.8%), Non-Hispanic Whites (75.0% compared to 51.0%), and individuals with bachelor's or post-graduate degrees (75.6% compared to 33.7%). We also observe that the majority of the sample corresponds to non-students (94.2%) and full time-employees (81.6%). Finally, in terms of household income and household composition, we are unable to compare the statistics from our survey with the Census data, because the latter provides income and household composition data only for all households (while our survey is focused on households with at least one worker with a primary workplace outside home). However, the sample statistics do suggest a skew toward individuals from higher income households and multi-worker households. Overall, there are many possible reasons for the socio-demographic differences between our sample and the Census data. For example, the main topic of the survey was self-driving vehicles, which may be of more interest to highly educated males. Also, the survey was conducted strictly through an online platform and the largest mailing list used in the distribution was of toll-road users, who are likely to be individuals with higher values of time that then correlates with the specific characteristics of our sample. In any case, while the general descriptive statistics of ride-hailing experience and use cannot be generalized to the DFW population, the individual level models still provide important insights on the relationship between ride-hailing travel behavior and socio-demographic/lifestyle characteristics.

4.2.1 Individual-Level Experience and Frequency of Use Model

Figure 4.1 provides the conceptual structure for our ride-hailing experience and frequency model, which are modeled jointly with residential location and household vehicle availability. Exogenous socio-demographic characteristics (left-side box in Figure 4-1) and four endogenous stochastic latent constructs representing attitudinal and lifestyle characteristics of the individual (middle box of Figure 4-1) are used as determinants of the four endogenous variables of interest (residential location density, vehicle availability, ride-hailing experience, and ride-hailing trip frequency in the past 30 days; these are listed in the right-side box of Figure 4-1, along with a host of indicators that enable us to better characterize the four stochastic latent psycho-social constructs in the middle box).

4.2.1.1 Attitudinal and Lifestyle Latent Constructs

Four attitudinal and lifestyle latent constructs are considered in our framework of Figure 4-1: privacy-sensitivity, technology-savviness, variety-seeking lifestyle propensity (VSLP), and green lifestyle propensity (GLP). These are identified based on earlier studies in transportation as well as in the ethnography field that recognize these psycho-social constructs as important determinants of travel-related and technology-use patterns. For instance, the first latent construct, privacy-sensitivity has been acknowledged and included in multiple transportation studies that investigate public transit use (Hunecke et al., 2010; Haustein, 2012; Spears et al., 2013). This is because one of the main aspects of the public transit mode that may discourage use is the presence of strangers in a shared space. Although ride-hailing is a car-based transportation mode, individuals travel with the driver. Hence, understanding how much individuals value being in private environments is a key element to predicting the adoption of ride-hailing, especially the use of pooled ride-hailing. Controlling for privacy-sensitivity is also important because concerns about sharing spaces with strangers influence people's residential location and vehicle availability (through ownership of automobiles) choices as privacy is strongly related to spaciousness and exclusivity considerations, with individuals with a stronger privacy disposition locating in low to medium density neighborhoods and owning many vehicles (see, for example, Bhat et al., 2016 and Bhat, 2015b). Thus, including this construct is important to avoid the overestimation of any positive impacts of dense residential location and low vehicle ownership on ride-hailing use. The second latent construct, tech-savviness, represents the individual's familiarity and affinity with technology, in our case, information and communication technologies (ICTs). This latent construct is relevant because, to hail a ride, the individual needs to use a smartphone app. Indeed, previous studies have found a significant and positive impact of tech-savviness on ride-hailing experience and smart phone use (Alemi et al., 2017; Lavieri et al., 2017; Astroza et al., 2017). The third construct, variety-seeking lifestyle propensity (VSLP) represents the individual's interest in exploration, and his/her openness to new experiences and changes. This construct has also been used in a past ride-hailing study (Alemi et al., 2017) and is important to capture intrinsic heterogeneity in the willingness to deviate from travel habits and mode inertia (Tudela et al., 2011; Rieser-Schüssler and Axhausen, 2012). The construct has been widely used within the theory of basic human values in the cultural-psychology field, and two of the indicators used in our survey to measure this construct are based on Schwartz's core value measures of openness to change (see Schwartz et al., 2001). Finally, the green lifestyle propensity (GLP) construct is used to capture individuals' tendencies toward environmentally friendly behaviors such as reduced use of drive-alone modes, reduced car ownership, and increased preference for dense and walkable neighborhoods. This latent variable is probably the most commonly used lifestyle factor in travel behavior studies (see for example, Van Acker et al., 2014; Bhat, 2015b; Lavieri et al., 2017; Ye and Titheridge, 2017). Similar to privacy-sensitivity, controlling for variety-seeking and green lifestyle is fundamental to capture potential self-selection effects that could bias the impacts of residential density and vehicle ownership on ride-hailing behavior.

The indicators of each construct are presented in Table 4-2, together with their sample distributions. All the indicators are measured on a five-point Likert scale and are modeled as ordinal variables. As may be observed from Table 4-2, the sample shows a general tendency toward being privacy-sensitive, tech-savvy, and having a variety-seeking lifestyle. The concern with privacy during a trip is consistent with the level of car-dominance in DFW, and may possibly impact the adoption of ride-hailing, especially pooled ride-hailing (note that the first indicator for privacy sensitivity is actually a measure of privacy insensitivity as elicited in the survey, and so the response is introduced in a reversed scale in the analysis to capture privacy sensitivity). A clear familiarity with ICTs and a variety-seeking lifestyle in the sample is expected, considering that the sample is skewed toward high levels of education and income. Interestingly, the responses related to the last measure; green lifestyle; show that over 50% of the sample "somewhat" or "strongly" agree that factors other than environmental friendliness dictate their commute mode choices, while just a little over 11% of the sample "somewhat" or "strongly" agree that they do not give much thought to energy saving at home. These descriptive statistics suggest that, while most people are sensitive to energy conservation considerations at home, most people also believe that considerations other than their commute-related environmental footprint dictate their commute mode choices (note again that the two questions pertaining to green lifestyle measure non-green lifestyle in the way they are worded, and so are introduced in a reversed scale in the analysis to capture green lifestyle propensity).

Table 4-1 Sample distribution of socio-demographic characteristics

Variable	Count	%	% Ride-hailing experience
Gender			
Female	668	41.57	54.04
Male	939	58.43	58.04
Age			
18 to 34	261	16.24	75.48
35 to 44	360	22.4	65.28
45 to 54	432	26.88	54.63
55 to 64	423	26.32	45.86
65 or more	131	8.16	33.59
Race/ethnicity			
Non-Hispanic White	1205	74.98	55.19
Non-Hispanic Black	102	6.35	55.88
Hispanic	109	6.78	62.39
Asian/Pacific Islander	101	6.29	65.35
Other	90	5.60	55.55
Education			
Completed high-school	238	14.82	42.44
Completed technical school/associates degree	154	9.58	59.74
Completed undergraduate degree	724	45.05	56.22
Completed graduate degree	491	30.55	62.32
Student (attending institution in person)			
Yes	93	5.79	65.59
No	1514	94.21	55.81
Employment type			
Full-time employee	1312	81.64	57.39
Part-time employee	138	8.59	51.45
Self-employed	157	9.77	52.23
Household income			
Under $49,999	184	11.45	51.09
$50,000-$99,999	443	27.57	46.50
$100,000-$149,999	496	30.86	54.64
$150,000-$199,999	269	16.74	63.94
$200,000 or more	215	13.38	75.81
Household composition			
Single person household	191	11.89	62.30
Single worker multi-person household	265	16.49	44.91
Multi-worker household	1151	71.62	58.04

Table 4-2 Sample distribution of attitudinal and behavioral indicators (n=1607)

Privacy-sensitivity					
	Strongly disagree	Somewhat disagree	Neither	Somewhat agree	Strongly agree
I don’t mind sharing a ride with strangers if it reduces my costs	13.44%	22.15%	20.41%	35.53%	8.46%
Having privacy is important to me when I make a trip	2.80%	10.52%	22.84%	41.19%	22.65%
I feel uncomfortable sitting close to strangers	8.59%	22.53%	27.88%	29.12%	11.89%
Tech-savviness					
	Does not describe me at all	Describes me slightly well	Describes me moderately well	Describes me very well	Describes me extremely well
I frequently use online banking services	2.43%	3.42%	6.41%	18.67%	69.07%
I frequently purchase products online	1.24%	7.28%	14.87%	23.58%	53.02%
Learning how to use new smartphone apps is easy for me	2.49%	5.48%	16.68%	27.13%	48.23%
Variety-seeking lifestyle propensity (VSLP)					
	Does not describe me at all	Describes me slightly well	Describes me moderately well	Describes me very well	Describes me extremely well
I think it is important to have all sorts of new experiences and I am always trying new things.	3.48%	12.62%	29.33%	34.12%	20.45%
Looking for adventures and taking risks is important to me.	13.36%	24.98%	33.25%	21.81%	6.59%
I love to try new products before anyone else	6.90%	15.91%	28.28%	30.33%	18.58%
Green lifestyle propensity (GLP)					
	Strongly disagree	Somewhat disagree	Neither	Somewhat agree	Strongly agree
When choosing my commute mode, there are many things that are more important than being environmentally friendly	4.60%	15.93%	28.38%	34.85%	16.24%
I don’t give much thought to saving energy at home	39.33%	37.59%	11.64%	8.59%	2.86%

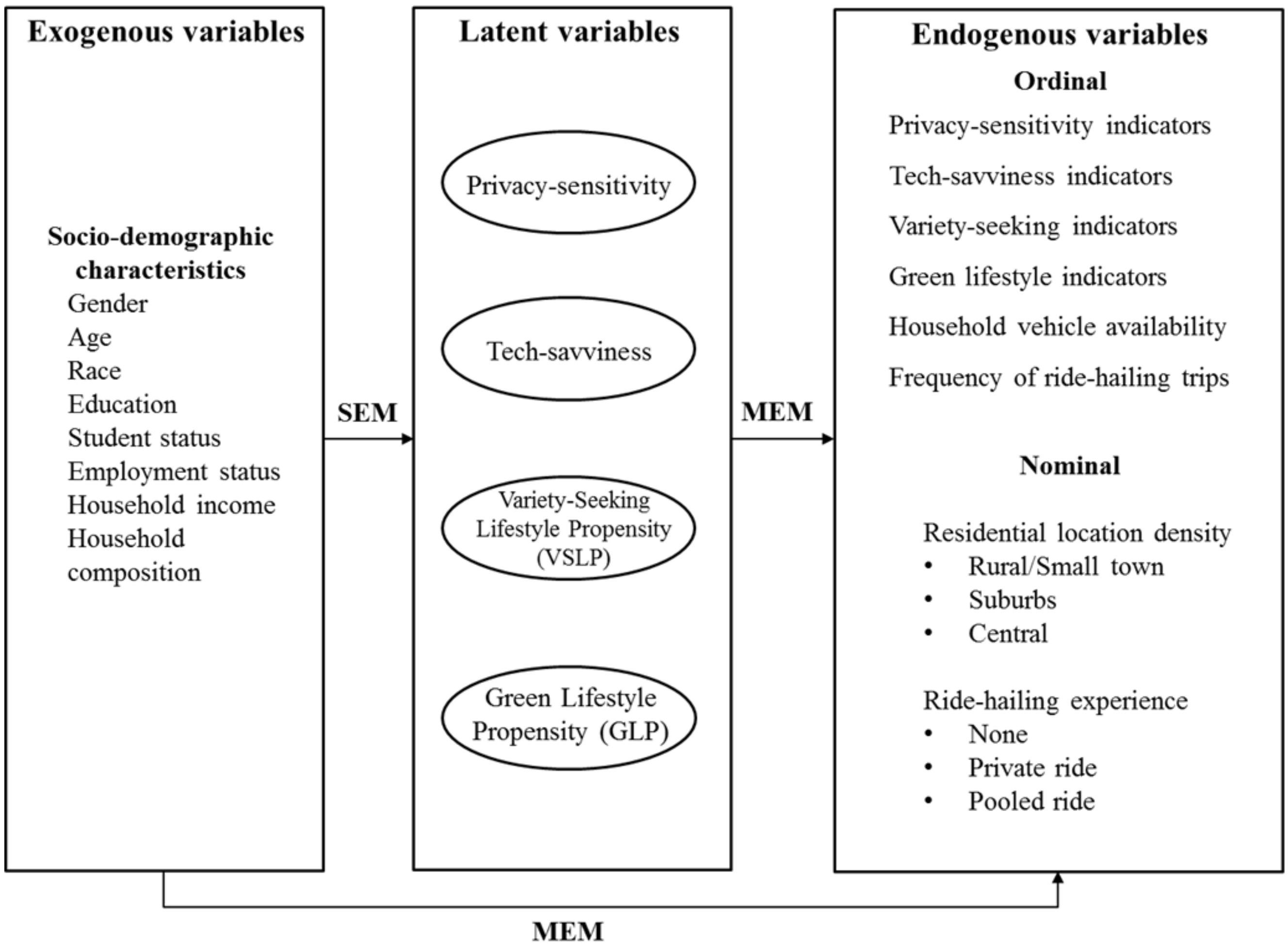


Figure 4-1 Structure of ride-hailing experience and frequency model

4.2.1.2 Main Outcome Variables

As already discussed, there are four endogenous variables of interest (residential location density, vehicle availability, ride-hailing experience, and ride-hailing trip frequency in the past 30 days) in the individual-level model.

Residential location is defined based on a survey item in which the respondents identified the type of neighborhood where they lived: (1) rural area, (2) small town, (3) neighborhood in the suburbs, (4) neighborhood in a central area but not downtown, and (5) downtown. Due to paucity of responses in the “small town” and “downtown” categories, we decided to regroup these five categories into the following three categories of residential location type: rural area or small town (11.6%; n=186), suburban area (65.0%; n=1046), and central area/downtown (23.4%; n=375). For ease in presentation, in the rest of this paper, we will refer to these three residential areas more simply as rural area, suburban area, and urban area, respectively.

Vehicle availability is characterized as the number of vehicles (automobiles) per worker in the household and is categorized in one of three ordinal levels: *less than one vehicle per* *worker, one vehicle per worker, and more than one vehicle per worker*. This definition is widely accepted in the literature as an indicator of vehicle availability or sufficiency for households with workers, because of the role that work schedules and commuting episodes play in shaping household activity schedules and task/vehicle allocation among household members (see, for example, Astroza et al., 2018). The sample distribution in these three ordinal categories is as follows: less than one vehicle per worker (14.7%; n=236), one vehicle per worker (50.8%; n=817), and more than one vehicle per worker (34.5%; n=554).

In terms of ride-hailing experience, about 56.4% of the sample (n=906) reported using ride-hailing services at least once in their lifetimes, although only about 10.0% of the sample (n=157) reported experience with the pooled version of the service. Accordingly, ride-hailing experience is represented in the three nominal categories of *no experience* (43.6%; n=701), *experience with private rides only* (46.6%; n=906-157=749), and *experience with pooled rides* (9.8%; n=157; note that this group may have had experience with private rides too). The column at the far right in Table 1 shows the fraction of individuals with ride-hailing experience by socio-demographic group. We observe that men, young adults (18-44 years of age), individuals of Hispanic and Asian origin, individuals with graduate degrees and students, high income individuals, and individuals living alone and in the central city areas have a higher than average tendency of having used ride-hailing services.

When asked about ride-hailing frequency specifically in the month prior to the survey, 33.7% of all respondents (n=542) reported at least one trip, suggesting that there is a considerable percentage of ride-hailing users (22.7%=56.4%-33.7%) who rely on ride-hailing on a one-off basis rather than on a monthly basis. Coincidentally, data about monthly use of ride-hailing extracted from the 2017 National Household Travel Survey shows that 22.7% of DFW residents used ride-hailing in the month prior to responding to the survey questionnaire (NHTS, 2017), providing an additional level of comfort and veracity to our own data collection effort[8]. It also is important to point out that ride-hailing frequency is relevant only if the individual has had ride-hailing experience (that is, only if the individual is not in the "no experience" category for the ride-hailing experience variable). Within the sub-sample of individuals with some ride-hailing experience (n=906), the frequency of trips in the past 30 days is grouped in one of the following five ordinal levels (the share of each level, as a percentage of 906 individuals with ride-hailing experience, is represented in parentheses: *zero trips* (40.2%; n=364), *1-3 trips*

[8] It is important to mention that the term used in the NHST survey was "ridesharing apps" and there was no specific definition accompanying it. Thus, there may be some differences in the breadth of services considered in the ride-hailing definition of the current study and in the NHTS survey.

(30.9%; n=280), *4-5 trips* (12.6%; n=114), *6-10 trips* (11.0%; n=100), and *more than 10 trips* (5.3%; n=48).[9]

4.2.2 Trip-Level Ride-Hailing Attributes Multivariate Model

The second, trip-level, model we estimate utilizes the subsample of individuals with ride-hailing experience (n=906) and examines the four attributes of trip purpose, time-of-day of trip, trip companionship, and the mode substituted by ride-hailing for the most recent ride-hailing trip undertaken by respondents. This analysis is exploratory in nature, because we are modeling the attributes of an isolated trip outside the broader context of the individual's daily activity-travel schedule. In particular, it is difficult to disentangle whether the choices made for the most recent ride-hailing trip are a reflection of specifically choosing ride-hailing in the last trip or simply a manifestation of the totality of the activity-travel pattern of the individual. For example, if a student is more likely than a non-student to run errands in the last ride-hailing trip relative to traveling to the airport, it is not clear whether this implies that students are more likely than non-students to use ride-hailing to run errands than to go to the airport, or whether this is simply an artifact of students rarely going to the airport in general relative to their non-student counterparts. We will not belabor over this point again when discussing the trip-level results, although all the results there should be viewed through this cautionary interpretive lens. Nonetheless, we use a multivariate modeling approach to study the different trip attributes jointly, allowing us to control for the effects of multiple variables systemically and simultaneously.

4.2.2.1 Exogenous Variables

As in the case of the individual-level model of Section 2.1, the exogenous variables used in the trip-level modeling include the individual level and household level exogenous variables identified in Table 1. However, in addition, we use residential location density, vehicle availability, whether or not the individual has experience with pooled ride-hailing, and ride-hailing frequency as exogenous variables in this exploratory trip-level analysis, assuming that the earlier endogenous variables (in the individual-level model) are higher-level longer-term decisions that impact the more shorter-term trip choice decisions. The last of the endogenous variables from the individual-level model; ride-hailing frequency; is introduced as a binary

[9]Although the frequencies of private and pooled ride-hailing trips were elicited separately in the survey, the number of individuals with at least one pooled ride-hailing trip during the past 30 days was very small (n=48). Thus, we combined the frequencies of private and pooled ride-hailing trips in the modeling of the frequency dimension.

variable in the trip-level analysis, by classifying individuals as either frequent users (at least 4 rides in the past 30 days) or not. Further, we also include the latent constructs as characterized from the individual-level model as exogenous variables by developing an expected value for each latent variable (based on the SEM model estimates from the individual-level model) and each individual.[10]

4.2.2.2 Main Outcome Variables

The alternatives within each of the four trip-level choice dimensions and their sample distributions are presented in Table 4-3. The descriptive statistics corresponding to trip purpose indicate that ride-hailing is mostly being used to access airports and recreational activities (with each of these purposes accounting for about 40% of all ride-hailing trips). The time-of-day shares show a relatively even intensity of trips during the morning and mid-day periods, though there is a definitive spike in the intensity during the evening period (note that all the morning, mid-day, and evening periods are of five hours duration, as we have defined them). The intensity of ride-hailing trips is lower during the nine-hour night period, though this is to be expected given the overall lower intensity of travel during the night relative to the day periods. In terms of trip companionship, about two-fifths of all trips are made alone, while the remaining are with others (co-workers, friends, family, and strangers). The trips with strangers, while having more of a flavor of pooled ride-hailing trips than those with co-workers, friends, and family, amounted to only 13 in number, and so were combined with trips with other accompaniment types. Finally, the dimension of mode substituted from for the ride-hailing trips suggests that much of the draw is from a private vehicle or a taxi. It is also interesting to note that almost 6% of the sample would not have traveled if ride-hailing were not available.

In our exploratory analysis, we adopt an endogeneity hierarchy within our trip-level modeling of different trip attributes by considering trip purpose as a determinant variable in the modeling of the remaining three attributes (time-of-day, companionship, and mode substituted), time-of-day as a determinant variable in modeling the remaining two attributes (companionship

[10] The choice to adopt this approach of treating the latent constructs as exogenous rather than endogenous for our trip-level model (instead of estimating another elaborate GHDM) is based on two considerations. First, the dependence between the trip-level choice dimensions is likely more due to unobserved factors associated with the nature of activities and trips (for example, bars and pubs generally open at night, so recreational trips may be more likely at this time), rather than individuals' psychological and lifestyle factors. Second, we believe that the characterization of the latent attitudinal and life-style constructs would be better based on broad individual-level decisions rather than trip-level decisions. Of course, given the smaller sample available for this trip-level analysis, we also felt a simpler exploratory modeling approach relative to the GHDM would be more appropriate.

and mode substituted), and companionship as a determinant in the modeling of the "mode substituted" trip dimension. The model is not necessarily capturing causal relationships in this exploratory analysis, but only associative relationships. However, note that the model is still a true joint model of all the four attributes simultaneously, because error covariances across the four dimensions are explicitly recognized and modeled, as discussed briefly next.

Table 4-3 Sample distribution of trip characteristics (n=906)

Variable	Count	%
Trip purpose		
Airport	359	39.62
Shopping, personal, or family errands	86	9.49
Recreational and leisure activities	362	39.96
Work or education	99	10.93
Time-of-day		
Moring (6:00am-10:59am)	191	21.08
Mid-day (11:00am-3:59pm	183	20.20
Evening (4:00pm-8:59pm)	305	33.66
Night (9:00pm-5:59am)	227	25.06
Companion		
Alone	370	40.84
With friends, family or co-workers	523	57.73
With a stranger (pooled ride)	13	1.43
Mode substituted		
My own vehicle	419	46.25
Taxi	347	38.30
Transit, bicycle or walk	87	9.60
Would not have traveled	53	5.85

4.2.2.3 Multivariate Multinomial Probit (MMNP) Model

The model adopted for the analysis of trip-level attributes is the MMNP that allows flexible covariances due to unobserved elements within the utilities of each trip dimension's alternatives, and also allows covariances across the utilities of different trip dimensions. The likelihood function for such an MMNP model involves a high-dimensional integral. However, one can use a surrogate likelihood function for estimation in such cases using the composite maximum likelihood inference (CML) approach that preserves the consistency and asymptotically normal properties of the full-information maximum likelihood (FIML) estimator under the same regularity conditions that result in the consistency and asymptotically normal properties of the FIML estimator. The reader is referred to Bhat (2011) and Bhat et al. (2013) for additional details.

4.3 Individual-level Experience and Frequency of Use Model Results

This section presents a detailed discussion of the results of the individual-level ride-hailing experience and frequency model. The final model specification was obtained based on a systematic process of testing alternative combinations of explanatory variables and eliminating statistically insignificant ones. However, some variables that were not statistically significant at a 95% confidence level were still retained due to their intuitive interpretations and important empirical implications. In this regard, the GHDM methodology used involves the estimation of a large number of parameters, so the statistical insignificance of some coefficients may simply be a result of having only 1,607 respondents (and only 906 respondents for the ride-hailing frequency variable). Also, the effects from this analysis, even if not highly statistically significant, can inform specifications in future ride-hailing investigations with larger sample sizes.

In the next section, we discuss the results of the SEM model component of the GHDM, as well as the latent variables' loadings on the attitudinal and lifestyle indicators (which is one part of the MEM). In subsequent sections, we discuss the MEM relationships corresponding to the effects of socio-demographic characteristics and the latent variables on the four main outcomes of interest in the individual-level model (including endogenous effects among these four outcome variables).

4.3.1 Lifestyle and Attitudinal Latent Factors

The structural relationships between socio-demographic variables representing lifecycle stages and the latent constructs are presented in Table 4-4. Gender shows no significant effect on the individual's level of privacy-sensitivity and tech-savviness. Yet, women display lower levels of VSLP and higher levels of GLP. These results are consistent with the social psychology literature. Gender comparisons based on the Theory of Basic Human Values (Schwartz, 1992) identify that men tend to be more open to experiences and changes than women as men generally attribute more value to stimulation, self-direction and hedonism values (Schwartz and Rubel, 2005; Vianello et al., 2013). On the other hand, women are generally more oriented toward prosocial values than men (Liu et al., 2014; Gifford and Nilsson, 2014), which result in more environmentally conscious behaviors (Gilg et al., 2005; Bhat, 2015b).

Age presents generally significant effects on all latent constructs except privacy-sensitivity. In general, younger adults show higher levels of tech-savviness and VSLP than their older counterparts. It is well established that younger generations, through their early exposure to ICT in their formative childhood years, are naturally more familiar and adept with such technologies (Helsper and Eynon, 2010; Twenge, 2013), which contributes to their higher level of tech-savviness. In terms of VSLP, the human values and personality literature identifies that younger individuals are more open to new experiences and more likely to attribute high importance to stimulation values, seeking variety in their daily lives (Gutierrez et al., 2005; Milojev and Sibley, 2017). The marginally significant negative GLP among the youngest group of individuals (18 to 34 years of age) relative to their older peers is interesting, though not inconsistent with findings from recent studies that identify a decrease in the younger generation's environmental consciousness. For example, Liu et al. (2014) and Gifford and Nilsson (2014) suggest that this trend among the youngest generation of adults may be the result of an increase in the importance of material pleasures in the American society as well as with an increased level of optimism that technology will solve environmental problems.

Non-Hispanic White individuals tend to be more privacy-sensitive and exhibit a lower VSLP relative to other races/ethnicities, results that also align with the higher levels of drive-alone travel and vehicle ownership by this ethnic group (Giuliano, 2003; Klein et al., 2018). As expected, individuals who are more highly educated tend to be more green, consistent with results in the social-psychological literature (see, for example, Franzen and Vogl, 2013) that individuals with a higher education are more self-aware of the negative consequences of degrading the environment Usually, education is also an important predictor of tech-savviness (Helsper and Eynon, 2010; Seebauer et al. 2015; Lavieri et al., 2017). However, in our model, such a relationship is not statistically significant, probably because the majority of the sample has at least a bachelor's degree. Part-time employees are less tech-savvy than full-time and self-employed individuals. As Helsper and Eynon (2010) explain, familiarity and ability to use ICTs is largely explained by exposure and experience. In that sense, it is plausible that part-time employees are generally less exposed to technology in the workplace (due to the nature of part-time jobs, and the time spent at work) than full-time and self-employed individuals.

Table 4-4 Determinants of latent constructs

Variables (base category)	Structural Equations Model Component Results							
	Privacy-sensitivity		Tech-savviness		VSLP		GLP	
	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat
Gender (male)								
Female	--	--	--	--	-0.270	-3.24	0.426	1.89
Age (≥55 years)								
18 to 34	--	--	1.144	11.28	0.480	4.89	-1.174	-1.86
35 to 44	--	--	0.899	10.14	0.287	3.51	--	--
45 to 54	--	--	0.441	5.58	--	--	--	--
Race/ethnicity (other)								
Non-Hispanic White	0.187	1.98	--	--	-0.177	-3.34	--	--
Education (≤ undergraduate degree)								
Graduate degree	--	--	--	--	--	--	0.859	2.52
Employment (full-time)								
Part-time employee	--	--	-0.395	-3.29	--	--	--	--
Self-employed	--	--	--	--	--	--	--	--
Household income (< $50,000)								
$50,000-$99,999	--	--	0.283	2.55	--	--	--	--
$100,000-$149,999	--	--	0.446	3.94	--	--	--	--
$150,000-$199,999	--	--	0.668	5.27	--	--	--	--
$200,000 or more	0.259	2.55	0.803	5.98	0.257	2.61	--	--
Household composition (multi-worker and single person)								
Single worker multi-person	--	--	--	--	-0.209	-2.07	--	--
Correlations between latent variables								
Privacy-sensitivity	1.000	n/a						
Tech-savviness	--	--	1.000	n/a				
VSLP	--	--	0.360	2.48	1.000	n/a		
GLP	-0.465	-2.01	--	--	--	--	1.000	n/a

"--" = not statistically significantly different from zero at the 90% level of confidence and removed from the specification.
"n/a" = not applicable

In terms of household demographics, household income contributes to an increase in privacy-sensitivity, tech-savviness and VSLP. The higher privacy-sensitivity among the wealthiest segment of individuals can be a direct result of having more access to private property and/or a need to signal exclusivity through separation and differentiation from others (Chevalier and Gutsatz, 2012; Bhat, 2015b). These individuals may also focus on privacy due to concerns associated with safety and preservation of material assets. Also, higher consumption power allows wealthy individuals early access to new technologies, increasing their exposure and use of technology. Indeed, multiple studies find this positive association between income level and technology use or technology-savviness (see, for example, Astroza et al., 2017; Lavieri et al., 2017; and Liu and Yu, 2017). The higher VSLP in the wealthiest segment of individuals is also reasonable, since this segment has more financial wherewithal to pursue a variety of different types of activities. Finally, compared to multi-worker and single individual (worker) households, individuals living in single-worker multi-person households have lower VSLP.

Two correlations between latent variables are statistically significant (see bottom of Table 4). Privacy-sensitivity is negatively associated with GLP, and tech-savviness is positively associated with VSLP. Both relationships are intuitive. For example, the second positive and reciprocal relationship between tech-savviness and VSLP is to be expected because (a) individuals who seek variety are more likely to experiment with new products and technology, and (b) ICT and internet use expand an individual's awareness and spatial cognition about activity options and opportunities.

The SEM estimation is made possible through the observations on the endogenous variables (far right block of Figure 4-1), which include the latent variable indicators and the four endogenous outcomes of interest. As discussed earlier, the presence of the latent variable indicators is not essential, though they provide stability in the SEM estimation. The loadings of each of the latent constructs on the underlying latent variables characterizing the ordinal indicators of that variable were all as expected (Table 4-5).

Table 4-5 Thresholds and constants of indicators and loadings of latent variables on indicators

Attitudinal and lifestyle indicators	Threshold 2		Threshold 3		Threshold 4		Constant		Loadings	
	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat
Privacy-sensitivity										
I don't mind sharing a ride with strangers if it reduces my costs (inverse scale)	2.523	19.85	3.598	21.06	5.123	19.08	2.504	12.84	1.792	14.09
Having privacy is important to me when I make a trip	0.922	12.13	1.799	22.17	3.076	33.69	2.101	23.01	0.575	16.21
I feel uncomfortable sitting close to strangers	0.954	17.55	1.737	25.04	2.777	25.44	1.409	22.24	0.427	6.19
Tech-savviness										
I frequently use online banking services	1.133	8.67	2.606	18.136	4.099	28.56	2.559	12.83	1.601	55.44
I frequently purchase products online	0.506	6.475	1.017	11.17	1.849	19.27	1.861	14.69	0.681	26.15
Learning how to use new smartphone apps is easy for me	1.138	9.685	1.993	16.22	2.859	23.18	2.255	15.08	0.787	30.61
Variety-seeking lifestyle propensity (VSLP)										
I think it is important to have all sorts of new experiences and I am always trying new things	1.159	13.78	2.374	26.41	3.676	35.80	2.631	19.33	0.930	22.40
Looking for adventures and taking risks is important to me	1.195	2.45	2.468	2.33	3.834	2.17	1.739	2.67	1.033	23.83
I love to try new products before anyone else	0.910	6.67	1.859	7.37	2.934	7.38	1.908	6.88	0.704	2.69
Green lifestyle propensity (GLP)										
When choosing my commute mode, there are many things that are more important than being environmentally friendly (inverse scale)	1.045	15.37	1.860	16.49	2.746	15.00	0.988	12.66	0.158	1.84
I don't give much thought to saving energy at home (inverse scale)	0.708	10.87	1.182	16.44	2.203	25.18	1.910	21.34	0.132	1.80

4.3.2 Residential Location and Vehicle Availability

Residential location and vehicle availability are modeled as endogenous variables so that we can control for self-selection effects when analyzing the impacts of these variables on ride-hailing behavior. Interestingly, as shown in Table 4-6, after controlling for the latent variable effects, there were few other sociodemographic variables having a direct impact on residential location and vehicle availability (though sociodemographic variables have an indirect effect through their impacts on the latent variables).

In terms of latent variable impacts on residential density, individuals who are tech-savvy and pursue a green lifestyle appear to prefer to reside in higher density suburban and urban areas rather than in a rural area. Access to ICT is generally more limited in rural areas, which may explain the negative effect of tech-savviness on rural living. Also, GLP is measured in our study in terms of concern about transportation and energy footprint, which may not be a priority for rural dwellers. On the other hand, the results indicate that individuals with a high variety-seeking lifestyle propensity (VSLP) tend to be more likely to live in an urban area relative to other areas, presumably because urban areas offer easy access to a diverse portfolio of activities and products. In addition to the indirect sociodemographic effects through the latent variable effects just discussed, the direct sociodemographic effects on residential location choice reveal that the youngest segment of individuals prefer more urbanized living relative to their older peers, presumably a reflection of wanting to have a variety of activity opportunities in close proximity to satisfy a heightened need for social interactions. Part-time employees tend to be located in urban areas, while self-employed individuals are more likely to reside in rural and urban areas rather than in suburban neighborhoods. These results may be associated with the nature of part-time and self-employed/independent jobs compared to full-time work arrangements. For example, many part-time jobs are associated with services in urban areas. Self-employed individuals in the service-oriented industry may also benefit from being located in areas with a high density of individuals (clients) and activities, while self-employed workers in the production industry (such as farmers) may be more likely to reside in rural areas in close proximity to their work sites. As expected, households with income above $150K dollars per year are less likely than those with lower incomes to be located in rural areas compared to suburbs and urban areas. Finally, individuals living alone show a higher propensity to locate in urban areas, consistent with the age effect discussed earlier.

Vehicle availability is positively impacted by privacy-sensitivity, which is expected since the automobile is the most private transportation mode. In contrast, tech-savviness has a negative effect on vehicle availably, plausibly because these lifestyle variables facilitate the use of, and draw toward, multi-modal travel options (Astroza et al., 2017). As anticipated, households with high incomes and with fewer workers have a higher vehicle availability, the first effect due to higher car ownership levels in households with high incomes and the second effect simply a manifestation of how we created the vehicle availability variable. Finally, households residing in the high-density urban areas of the DFW area have a lower vehicle availability, a reflection of the reduced need for vehicles in such areas because of good multi-modal transportation service as well as better access to activity opportunities within a compact geographic footprint. Importantly, this urban living effect is a "true" built environment effect after controlling for residential self-selection effects through the impacts of the latent attitudinal lifestyle variables on both residential location and vehicle availability.

4.3.3 Ride-Hailing Experience

The results of the ride-hailing experience model are presented in the third column of Table 5. The latent variable effects have the expected direction, with privacy-sensitive individuals less likely to have experience with pooled service and tech-savvy individuals most likely to have experience with private ride-hailing only. On the other hand, variety-seeking individuals are most likely to have the pooled service experience. Interestingly, GLP does not seem to play a role in ride-hailing adoption.

In addition to the indirect socio-demographic influences through the latent variable effects just discussed, there are quite a few direct socio-demographic effects on ride-hailing experience. This is unlike the case for residential location density and vehicle ownership where there are relatively fewer direct sociodemographic effects after controlling for latent variable effects. This disparity makes sense because ride-hailing is a relatively recent phenomenon and individuals are still in the process of exploring the many dimensions of this service. That is, ride-hailing preferences are still in a formative stage, with the impacts of attitudes and lifestyles not yet as deeply entrenched as for residential location density and vehicle availability (the latter choices have been available to individuals over a much longer period of time)[11]. During these initial exploratory/formative stages of preference, it is the immediate demographic lifecycle considerations that dictate and drive ride-hailing experience and frequency. Earlier studies in the social psychology literature support this notion that the effects of attitudes/lifestyle toward preference for a service/product take time to materialize and stabilize (see, for example, Hoeffler and Ariely, 1999; Amir and Levav, 2008).

Table 4-6 indicates that age has a direct negative effect on ride-hailing experience, with younger individuals more likely than their older counterparts to have used ride-hailing both in the private as well as pooled arrangements. While this is consistent with some earlier studies (Smith, 2016; Kooti et al., 2017), our study indicates that this effect is beyond the negative effect

[11] Note that the attitudinal and lifestyle latent variables and indicators used in this study do not reflect individual's direct attitudes, beliefs and perceptions about ride-hailing services. Instead, they reflect more general lifestyle dimensions.

of age on ride-hailing experience through the tech-savviness and variety-seeking effects. This direct effect may be a result of younger individuals having more exposure to new services and products through larger social networks (English and Carstensen, 2014).

The results also show that non-Hispanic Whites are less likely to have used pooled services, even after accounting for indirect race/ethnicity effects through privacy-sensitivity and VSLP, and controlling for income effects. The reason behind this race/ethnicity effect is not clear and calls for more qualitative studies investigating the willingness to share rides. Higher education appears to increase the experience with pooled ride-hailing, and employment status shows significant direct effects on private ride-hailing experience but not on the pooled option. Specifically, part-time employees are less likely to have experienced private ride-hailing services relative to full-time employees. Similar results were observed by Dias et al. (2017).

In terms of household level variables, a higher household income increases experience with both private and pooled ride-hailing, beyond the positive effect of household income through tech-savviness and VSLP (and while individuals with a household income over $200,000 have a higher privacy sensitivity, and privacy sensitivity negatively impacts pooled ride-hailing experience, this indirect negative effect gets swamped by the magnitude of the positive direct effect in Table 4-6). Considering that attitudinal and lifestyle factors are being controlled for, the direct income effect is probably an indicator of higher consumption power, though there is still a distinct preference for private ride-hailing over pooled ride-hailing within this high income group. Individuals living alone are more likely to have used private ride-hailing service relative to individuals in other household types, while those in single-worker multi-person households are the least likely to have used both private and pooled services. Even after controlling for self-selection effects, individuals living in more urbanized locations are more likely than their counterparts in less urbanized locations to have used both private and pooled ride-hailing. A similar result holds for individuals in households with more than one vehicle per worker. The result that a higher private vehicle availability leads to a higher experience with ride-hailing suggests that, in an area such as DFW where almost all households own at least one vehicle, ride-hailing serves as more of a convenience feature for those one-off trips rather than being an accessibility facilitator for routine trips (though, as we will see in the next section, increasing vehicle availability has a negative effect on ride-hailing frequency).

Table 4-6 Results of the residential location, vehicle availability, ride-hailing experience, and ride-hailing frequency model components

Variables (base category)	Residential location (base: Suburban)				Vehicle availability		Ride-hailing experience (base: none)				Ride-hailing frequency	
	Rural		Urban		(ordinal)		Private only		Pooled		(ordinal)	
	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat
Latent variables												
Privacy-sensitivity	--	--	--	--	0.405	2.25	--	--	-0.422	-1.95	--	--
Tech-savviness	-0.142	-3.51	--	--	-0.124	-2.54	0.251	2.61	0.165	1.80	--	--
VSLP	--	--	0.081	1.79	--	--	0.068	1.68	0.271	1.83	0.166	2.17
GLP	-0.376	-1.89	0.203	1.85	--	--	--	--	--	--	--	--
Exogenous effects												
Age (≥65 years)												
18 to 34	-0.789	-4.54	0.657	4.57	--	--	0.739	6.97	0.677	2.72	--	--
35 to 44	--	--	--	--	--	--	0.508	7.93	0.432	1.90	--	--
45 to 54	--	--	--	--	--	--	0.213	3.83	0.326	1.85	--	--
55 to 64	--	--	--	--	--	--	0.161	2.99	--	--	--	--
Race/ethnicity (other)												
Non-Hispanic White	--	--	--	--	--	--	--	--	-0.148	-1.87	--	--
Education (≤ undergraduate degree)												
Graduate degree	--	--	--	--	--	--	--	--	0.186	4.54	--	--
Employment (full-time)												
Part-time employee	--	--	0.369	9799	--	--	-0.135	-2.71	--	--	--	--
Self-employed	0.188	3.04	0.242	7.42	--	--	--	--	--	--	--	--
Household income (< \$100,000)												
\$100,000-\$149,999	--	--	--	--	0.519	6.16	0.326	6.67	--	--	--	--
\$150,000-\$199,999	-0.106	-2.87	--	--	0.519	6.16	0.546	11.39	0.146	1.85	--	--
\$200,000 or more	-0.106	-2.87	--	--	0.883	7.27	0.913	15.35	0.434	1.96	0.427	3.37
Household composition (multi-worker)												
Single person	-0.106	-2.52	0.189	3.74	0.532	6.22	0.386	8.50	--	--	--	--
Single worker multi-person	--	--	--	--	1.638	15.94	-0.176	-2.94	-0.243	-2.25	--	--
Endogenous effects												
Residential location (rural)												
Suburban	n/a	n/a	n/a	n/a	--	--	0.332	2.03	0.392	1.93	--	--
Urban	n/a	n/a	n/a	n/a	-0.175	-2.23	0.668	4.24	0.777	4.50	0.190	1.74
Vehicle availability (< 1 per worker)												
1 per worker	n/a	n/a	n/a	n/a	n/a	n/a	--	--	--	--	-0.239	-1.79
> 1 per worker	n/a	n/a	n/a	n/a	n/a	n/a	0.084	2.30	0.183	4.70	-0.239	-1.79
Pooled ride-hailing user (no)												
Yes	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	0.655	4.80
Constant	-0.759	-2.24	-0.876	-4.47	0.680	7.88	-1.172	-4.73	-1.702	-8.26	0.246	1.6
Thresholds												
Threshold 2	n/a	n/a	n/a	n/a	1.688	28.91	n/a	n/a	n/a	n/a	0.870	13.97
Threshold 3	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	1.338	17.13
Threshold 4	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	2.037	17.51

4.3.4 Ride-Hailing Frequency

Our model, similar to that of Alemi et al. (2018a), shows that few variables have an impact on ride-hailing frequency. Among the latent variable effects, only VSLP has a significant impact. This effect may be a result of individuals with a high VSLP experimenting and exploring different travel options and different activity pursuits (see, for example, Rieser-Schüssler and Axhausen, 2012).

Among other demographic effects, individuals in households with very high income (above $200K dollars per year) have a high ride-hailing frequency propensity, as also observed by Dias et al. (2017). Although using ride-hailing is usually cheaper than calling a taxi, frequent use can incur significant costs that may be more easily afforded by those in the high income segments. Living in an urban area (relative to living in suburbs or rural areas) also contributes to a higher propensity associated with ride-hailing trip frequency, even after controlling for self-selection effects. There are at least three possible reasons for this result. First, urban areas have more parking restrictions, increasing the benefit of being dropped-off at a destination. Second, distances are shorter, compared to more spread-out suburbs and rural areas, limiting the costs of the trips. Third, in urban areas, the supply of drivers is higher, increasing the overall reliability of the service, which is possibly an essential condition for maintaining a demand of frequent users. As also observed by Alemi et al. (2018a), higher vehicle availability rates reduce the propensity underlying the frequency of ride-hailing usage. Combined with the earlier finding of the positive effect of vehicle availability on ride-hailing experience, the results perhaps suggest that individuals in households with high vehicle availability make generally many more out-of-home trips (including those one-off trips to the airport and other recreational sites) and so are more likely to have used ride-hailing at some point as a convenience mode. However, it still holds that higher vehicle availability reduces the overall ride-hailing dependence. Another endogenous effect is that users of pooled ride-hailing have higher frequency propensities. Pooled trips offer lower fares, which may be a key element for ride-hailing services to maintain regular users.

4.3.5 Model Fit Comparison

The improved data fit from jointly modeling the four choice dimensions in the individual-level model system may be assessed by comparing the GHDM model with an Independent Heterogeneous Data Model (IHDM) that does not consider the jointness in the four dimensions (that is, the covariances engendered by the stochastic latent constructs in the GHDM model are ignored). In this IHDM model, we introduce the exogenous variables (sociodemographic variables) used to explain the latent constructs as exogenous variables in the choice dimension equations. In this way, the contribution to the observed part of the utility due to sociodemographic variables is still maintained (and is allowed to vary relative to the GHDM to absorb, to the extent possible, the GHDM covariances due to unobserved effects). The resulting IHDM may be compared to the GHDM using the composite likelihood information criterion (CLIC) introduced by Varin and Vidoni (2005). The CLIC takes the following form (after replacing the composite marginal likelihood (CML) with the maximum approximate CML (MACML)):

$$\log L^{*}_{MACML}(\hat{\ddot{\boldsymbol{\theta}}}) = \log L_{MACML}(\hat{\ddot{\boldsymbol{\theta}}}) - tr\left[\hat{\boldsymbol{J}}(\hat{\ddot{\boldsymbol{\theta}}})\hat{\boldsymbol{H}}(\hat{\ddot{\boldsymbol{\theta}}})^{-1}\right] \tag{1}$$

The model that provides a higher value of CLIC is preferred. The $\log L_{MACML}(\hat{\ddot{\boldsymbol{\theta}}})$ values for the GHDM and IHDM models were estimated to be –394,131 and –398,801, respectively, with the corresponding CLIC statistic values of –395,982 and –400,229. These CLIC statistics clearly favor the GHDM over the IHDM.

The ordinal indicator variables used in the measurement equation are included solely for the purpose of model identification and do not serve any purpose in predicting the endogenous choice bundle of interest once the model is estimated. Therefore, we can also use the familiar non-nested likelihood ratio test to compare the two models. To do so, we evaluate a predictive log-likelihood value of both the GHDM and IHDM models using the parameter values at the GHDM convergent values by excluding the indicator variables and focusing only on the four endogenous variables of interest. Then, one can compute the adjusted likelihood ratio index of each model with respect to the log-likelihood with only the constants as follows:

$$\bar{\rho}^2 = 1 - \frac{\mathscr{L}(\hat{\ddot{\boldsymbol{\theta}}}) - M}{\mathscr{L}(c)}, \tag{2}$$

where $\mathscr{L}(\hat{\ddot{\boldsymbol{\theta}}})$ and $\mathscr{L}(c)$ are the predictive log-likelihood functions at convergence and at constants, respectively, and *M* is the number of parameters (not including the constant(s) for each dimension and not including the ordinal indicators) estimated in the model. If the difference in the indices is $(\bar{\rho}_2^2 - \bar{\rho}_1^2) = \tau$, then the probability that this difference could have occurred by chance is no larger than $\Phi\left\{-[-2\tau\mathscr{L}(c)+(M_2-M_1)]^{0.5}\right\}$ in the asymptotic limit. A small value for the probability of chance occurrence indicates that the difference is statistically significant and that the model with the higher value for the adjusted likelihood ratio index is to be preferred. The $\mathscr{L}(\hat{\hat{\theta}})$ values (number of parameters) for the GHDM and IHDM models were computed to be -2,728.85 (number of parameters = 85) and –2,726.12 (number of parameters = 94), respectively. The $\mathscr{L}(c)$ value was –2,915.55. The non-nested adjusted likelihood ratio test returns a value of $\Phi(-4.64)$, which is literally zero, clearly rejecting the IHDM model in favor of the GHDM model and underscoring the importance of considering the stochastic latent constructs that engender covariation among the choice dimensions.

4.4 Trip-level Characteristics Model Results

This section analyzes the model results corresponding to the four dimensions of the individual's last ride-hailing trip: purpose, time-of-day, companion, and mode substituted. In this trip-level analysis, we included the latent psycho-social constructs as exogenous variables. However, except for the VSLP construct (and that too only for the trip purpose dimension), no other latent variable turned out to be statistically significant in explaining trip-level ride-hailing choices. This result is consistent with our notion earlier that trip-level choices regarding ride-hailing are likely more affected by unobserved factors associated with the nature of activities and trips rather than individuals' psychological and lifestyle factors.

4.4.1 Trip Purpose

The results of the model component representing trip purpose are presented in Table 4-7. In the first category of latent constructs, only the VSLP variable influences trip purpose, with individuals with a higher VSLP more inclined to participate in recreation relative to other purposes. This is reasonable simply because recreation intrinsically captures a sense of variety and exploration relative to the other more sustenance and maintenance activity purposes.

Although women are usually responsible for more personal, family, and shopping errands than men (Fan, 2017), being a woman is associated with a lower likelihood of using ride-hailing for these purposes, probably indicating that ride-hailing is not the preferred option when it comes to completing these routine commitments. By way of summarizing the effects of other socio-demographic effects, we observe that students and those with lower vehicle availability are more likely than their peers to have pursued errands in their last ride-hailing trip rather than other activity purposes, while millennials and those with lower vehicle availability are more likely to have pursued work-related travel rather than airport travel in their most recent ride-hailing trip. These results perhaps are indicative of the use of ride-hailing as an "accessibility mobility tool" to compensate for limited access to routine activities using other mobility options. On the other hand, millennials and non-Hispanic Whites are most likely to have pursued recreation (relative to all other activity purposes) in their last ride-hailing trip, presumably a reflection of the use of ride-hailing here as a "convenience mobility tool". The results also indicate that frequent ride-hailing users are more likely to have pursued work relative to other activity purposes. Finally, we also observe that living in more urbanized areas decreases the probability of having pursued other activities compared to going to the airport in the last trip. Above all, this result shows that individuals living in rural areas do not use ride-hailing to go to the airport, probably because the associated costs are still higher than parking at an airport.

4.4.2 Time of Day

The earlier ride-hailing literature indicates that the peak period of ride-hailing trips occurs during the night and does not coincide with the commuting and traffic peak periods (see Kooti et al., 2017; Komanduri et al., 2018). However, our descriptive statistics indicate otherwise; as discussed earlier in Section 2.2.2, the evening period (which includes the afternoon commute period) is when the overall intensity of ride-hailing activity is highest. But there are variations across individuals regarding when they are most likely to make a ride-hailing trip (at least based on their most recent trip). Not surprisingly, millennials (18-34 years of age) make most of their ride-hailing trips during the night period, consistent with this group more likely to socialize during the night period (see Garikapati et al., 2016). High-income individuals, on the other hand, are the least likely to ride-hail during the evening and night periods. Individuals living in single-worker multi-person households (relative to those in other households) tend to ride-hail during the morning and evening periods, while those residing in suburbs and urban areas (relative to those residing in rural areas) appear to ride-hail more during the morning and mid-day periods, presumably due to the convenience to get to work by ride-hailing in dense areas.

Table 4-7 Trip characteristics model results

Variables (base category)	Purpose (base: airport)						Time (base: mid-day)					
	Errands		Recreation		Work		Morning		Evening		Night	
	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat
Latent variables												
VSLP	--	--	0.279	4.23	--	--	--	--	--	--	--	--
Exogenous effects												
Gender (male)												
Female	-0.139	-3.43	--	--	--	--	--	--	--	--	--	--
Age (≥35 years)												
18 to 34	--	--	0.390	8.82	0.221	4.72	-0.113	-2.54	--	--	0.461	12.56
Race/ethnicity (other)												
Non-Hispanic White	--	--	0.342	10.10	--	--	--	--	--	--	--	--
Education (< undergraduate degree)												
Undergraduate degree	-0.294	-6.20	-0.180	-4.83	--	--	--	--	--	--	--	--
Graduate degree	-0.294	-6.20	-0.238	-5.88	-0.092	-2.21	--	--	--	--	--	--
Student (no)												
Yes	0.562	8.73	--	--	--	--	--	--	--	--	--	--
Employment (full-time)												
Part-time employee	--	--	--	--	--	--	--	--	--	--	--	--
Self-employed	--	--	--	--	--	--	--	--	--	--	--	--
Household income (< $50,000)												
$50,000-$99,999	--	--	--	--	--	--	--	--	--	--	--	--
$100,000-$149,999	-0.236	-5.14	--	--	--	--	--	--	--	--	--	--
$150,000-$199,999	-0.403	-6.86	--	--	--	--	--	--	--	--	--	--
$200,000 or more	--	--	--	--	-0.155	-3.06	--	--	-0.082	-2.07	-0.247	-5.25
Household composition (multi-worker)												
Single person	--	--	--	--	-0.303	-4.85	--	--	--	--	--	--
Single worker multi-person	--	--	--	--	--	--	0.217	5.54	0.217	5.54	--	--
Residential location (rural)												
Suburban	-0.361	-4.85	-0.330	-5.62	-0.373	-5.18	--	--	-0.323	-5.32	-0.215	-3.09
Urban	-0.361	-4.85	-0.163	-2.63	-0.373	-5.18	--	--	-0.399	-6.20	-0.358	-4.85
Vehicle availability (< 1 per worker)												
1 per worker	-0.248	-4.71	--	--	-0.162	-3.17	--	--	--	--	--	--
> 1 per worker	-0.248	-4.71	--	--	-0.162	-3.17	--	--	--	--	0.214	6.26
Ride-hailing frequent user (no)												
Yes	-0.283	-5.42	--	--	0.264	6.69	--	--	--	--	-0.109	-3.15
Pooled ride-hailing user (no)												
Yes	--	--	--	--	--	--	--	--	--	--	--	--
Endogenous Effects												
Trip purpose (airport)												
Errands							--	--	--	--	--	--
Recreational							--	--	1.033	29.86	1.268	34.24
Work							0.284	5.61	0.284	5.61	-0.209	-2.83
Constant	0.251	2.78	0.101	1.54	-0.202	-2.48	-0.058	-2.50	0.193	3.23	-0.250	-3.57

"--" = not statistically significantly different from zero at the 90% level of confidence and removed from the specification.

Frequent ride-hailing users appear to do so during the daytime. On the other hand, individuals who ride-hail during the nights appear to be from households with high vehicle availability and do so primarily for recreation, suggesting that these effects may be related to not wanting to drink and drive.

4.4.3 Companionship

The trip-level companionship results in Table 4-8 reveal some similarities with the individual-level pooled ride-hailing results, as one would expect. For example, middle-aged individuals are more likely than their peers to have had a companion on their most recent ride-hailing trip, while non-Hispanic Whites are more likely to have traveled alone. These are consistent with the results for pooled ride-hailing experience. Interestingly, a highly educated individual is more likely to have traveled alone during her/his last ride-hailing trip, though highly educated individuals are in general more likely to have had a pooled ridesharing experience. This perhaps is simply a reflection of highly educated individuals using a combination of private and pooled modes as they see best fit for specific trips, even if they are more open to pooled ride-hailing in general. Part-time employees and individuals from low-income households are more likely than their peers to have traveled with others. As expected, individuals who live alone, and individuals running errands or going to work are more likely to have traveled alone during their previous ride-hailing trip, while individuals pursuing recreation are more likely to have traveled with others during their previous ride-hailing trip. Finally, ride-hailing trips made during the morning peak serve mostly individuals traveling alone, which may have a negative implication on traffic congestion during this period.

4.4.4 Mode Substituted by Ride-Hailing

The results for this component of the trip-level model are presented in the last column of Table 4-8. The base category is the "private car". Women, more than men, appear to substitute active travel or transit usage by ride-hailing (at least based on the most recent ride-hailing trip). Non-Hispanic Whites, those with graduate-level education, students, part-time employees, and individuals living in medium and high income households have a higher tendency than their peers to substitute ride-hailing for taxi trips, while millennials, self-employed individuals, individuals living in non-rural locations, individuals in households with one vehicle per worker, and individuals making their trip in the evening period are the least likely to substitute ride-hailing for taxi trips. In the context of active/public transportation (APT) modes, individuals younger than 65 years of age, those with a bachelor's degree or higher, and individuals with experience with pooled ride-hailing tend to replace APT modes with ride-hailing (see also Alemi et al., 2017), while high income individuals and frequent ride-hailing users are not very likely to replace their APT travel with ride-hailing. Obviously, while one can explain these results in more ways than one, there is a clear need to investigate these effects in much more detail in future studies within the context of overall activity-travel patterns. However, the result regarding age effects does suggest that one potential detrimental effect of ride-hailing is a reduction in public health benefits, due to the substitution of active forms of transportation by ride-hailing among the substantial fraction of the population that is below the age of 65 years.

The last sub-column of the "Mode substituted by ride-hailing" corresponds to "no trips", which essentially implies that ride-hailing generated a new trip that would not have occurred otherwise. The demographic effects specific to this alternative indicate that young adults (18-44 years of age) are more likely than their older peers to have generated a new trip in their most recent ride-hailing experience, although it is more likely that these adults (relative to senior adults over the age of 65 years) switched to ride-hailing from active/public transportation. Also, part-time employees, self-employed individuals and those that live in multi-worker households appear to generate new ride-hailing trips more so than individuals in other households, perhaps a reflection of the added convenience to pursue activities due to ride-hailing. New trips are also more likely to occur among those living in non-rural areas. The generation of new trips in dense areas can, in the long term, intensify traffic congestion problems due to increased automobile usage. The new generated trips seem to be for the purposes of running errands and pursuing recreational activities, and are more likely to happen during the non-evening periods. The implied newly generated ride-hailing trips during the morning commute needs to be investigated more carefully, because the trips may add to traffic congestion as well as traffic crashes (the morning commute period is a traffic crash-prone period of the day due the combination of traffic congestion as well as the need to get to work on-time, which leads to aggressive driving during this period; see Paleti et al., 2010).

Table 4-8 Trip characteristics model results (cont.)

Variables	Trip companion (base: alone)		Mode substituted (base: own car)					
	Not alone		Taxi		Active or transit		No trip	
	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat
Exogenous effects								
Gender (base: male)								
Female	0.093	2.85	--	--	-0.255	-5.73	--	--
Age (base: ≥65 years)								
18 to 34	--	--	-0.273	-6.74	0.639	4.61	0.280	5.60
35 to 44	0.077	2.41	--	--	0.862	6.30	0.280	5.60
45 to 54	0.077	2.41	--	--	0.557	4.08	--	--
55 to 64	--	--	--	--	0.557	4.08	--	--
Race/ethnicity (base: other)								
Non-Hispanic White	-0.225	-6.17	0.092	2.79	--	--	--	--
Education (base: < undergraduate)								
Completed undergraduate degree	--	--	--	--	0.135	2.64	--	--
Completed graduate degree	-0.153	-4.46	0.089	2.97	0.135	2.64	--	--
Student (base: no)								
Yes	--	--	0.201	2.62	--	--	--	--
Employment type (base: full-time)								
Part-time employee	0.291	4.26	0.550	8.39	0.581	8.20	0.180	2.78
Self-employed	--	--	-0.133	-2.60	--	--	0.180	2.78
Household income (base: < $50,000)								
$50,000-$99,999	0.123	3.69	0.383	5.96	--	--	--	--
$100,000-$149,999	--	--	0.383	5.96	-0.331	-7.38	--	--
$150,000 or more	--	--	0.497	7.33	-0.331	-7.38	--	--
Household composition (base: multi-worker)								
Single person	-0.618	-13.35	--	--	--	--	-0.295	-3.52
Single worker multi-person	--	--	--	--	--	--	-0.316	-4.22
Residential location (base: rural)								
Suburban	--	--	-0.176	-3.28	--	--	0.600	4.19
Urban	--	--	-0.269	-4.65	--	--	0.600	4.19
Vehicle availability (base: < 1 per worker)								
1 per worker	--	--	-0.111	-3.95	--	--	--	--
> 1 per worker	--	--	--	--	--	--	0.366	6.40
Ride-hailing frequent user (base: no)								
Yes	--	--	--	--	-0.332	-6.65	--	--
Pooled ride-hailing user (base: no)								
Yes	--	--	--	--	0.182	3.72	--	--
Endogenous Effects								
Trip purpose (base: airport)								
Errands	-0.484	-7.65	--	--	--	--	1.061	15.88
Recreational	0.928	23.94	--	--	--	--	0.527	7.52
Work	-0.671	-12.36	--	--	--	--	--	--
Trip time (base: mid-day)								
Morning	-0.144	-3.41	--	--	--	--	--	--
Evening	--	--	-0.138	-4.12	-0.140	-2.85	-0.254	-4.38
Night	--	--	--	--	-0.295	-5.42	--	--
Trip companion (base: alone)								
Not alone			--	--	0.218	5.00	-0.366	-5.50
Constant	0.194	4.01	-0.292	-3.60	-1.286	-9.31	-1.974	-13.25

4.4.5 Dependence between Alternatives and Choice-Dimensions

The estimated covariance matrix corresponds to the differenced error terms for each dimension (the error term of the utility of an alternative minus that of the utility of the base alternative for that dimension). In our analysis, we could not reject the hypothesis that the error terms for each of the trip-level dimensions were independently and identically distributed (in differenced error terms, we could not reject the hypothesis that all the diagonal terms in the covariance matrix of the differenced error terms were 1.0 and all the off-diagonal elements were 0.5). However, there were two covariances statistically different from zero across dimensions, all associated with the taxi alternative in the "mode substituted by ride-hailing" dimension, corresponding to (1) the morning alternative in the time-of-day dimension and the taxi alternative (= –0.181, t-stat of –3.14), and (2) the "with others" alternative in the companionship dimension and the taxi alternative (= 0.147, t-stat of 2.49). If we assume that the error term in the base alternative in each dimension is independent of the error terms of all other alternatives in other dimensions, the implication of the first covariance is that unobserved factors that increase taxi substitution also decrease the likelihood of the ride-hailing trip occurring during the morning, while the second covariance factor suggests that unobserved factors that increase taxi substitution increase the propensity to travel with others. The first effect may be related to the overall lower share of taxi trips in the morning compared to other modes (especially drive alone), while the second effect may be a consequence of the reduced costs in using ride-hailing relative to a taxi, especially in the pooled form of ride-hailing.

4.4.6 Model Fit Comparison

The statistically significant covariance effects, even if only two in number, point to the importance of developing a joint model at the trip-level. To further examine model fit, we compare the log-likelihood of the final model (= –10,747.52), and that of the model which ignores the two covariances discussed in the previous section (= 10,751.23). The log-likelihood ratio test statistic of comparison between the two nested models is 7.42. This value is greater than the table chi-squared value with two degrees of freedom at even a 0.025 level of significance.

4.5 Policy Implications

4.5.1 From the Individual-Level Ride-hailing Experience and Frequency Model

To examine how the adoption of ride-hailing is currently occurring and how it may inform transportation planners and policy makers, we compare the magnitudes of the determinant variables in our models. This is achieved by computing average treatment effects (ATEs) of the variables on ride-hailing experience and frequency. In these ATE computations, we consider the latent psycho-social variables too as explicit determinant variables, rather than translating these latent variable effects into corresponding exogenous demographic variable effects through the structural equation model results of Table 4-4. That is, we do not combine the direct demographic effects and the indirect demographic effects (through the latent variables); rather, we compute the ATEs for the direct demographic effects and the ATEs for the latent variables. This is because, while the overall (indirect plus direct) demographic effects provide ride-hailing tendencies by demographic segment, they do not provide insights that may help in formulating policies. For example, one of the overall demographic effects is that non-Hispanic Whites are less likely to use pooled ride-hailing. However, this does not provide us additional insights on why this may be so. By including latent variables in the ATE computation, we may find, for example, that privacy sensitivity is one of the most important determinant variables in terms of the magnitude of effect on the use of pooled ride-hailing. If so, and because non-Hispanic Whites are likely to be more privacy sensitive relative to individuals of other race/ethnicity groups (according to our structural equation model results), it provides additional insights on how to position pooled ride-hailing information campaigns directed toward this segment of the population. One additional note regarding the computation of ATE effects for the latent variables. We compute these effects by examining the impact of changing each latent variable from its minimum value (the base) to its maximum value (that is, the continuous latent variable values are changed to two discrete values for the ATE computations; the minimum expected value representing the base category).

The ATE measure for the ride-hailing experience variable (which is a nominal variable in our analysis) provides the expected difference in ride-hailing experience for a random individual if s/he were in a specific category i of the determinant variable as opposed to another configuration $k \neq i$. The ATE is estimated as follows for each determinant variable:

$$\hat{\text{ATE}}_{ikj} = \frac{1}{Q}\sum_{q=1}^{Q}\left(\left[P(y_q = j \mid a_{qi} = 1) - P(y_q = j \mid a_{qk} = 1)\right]\right) \tag{3}$$

where a_{qi} is the dummy variable for the category *i* of the determinant variable for the individual *q*, y_q stands for the ride-hailing experience nominal variable, and *j* represents a specific nominal category of ride-hailing experience. Thus, $\hat{\text{ATE}}_{ikj}$ above represents the estimate of the expected value change in the nominal category *j* of ride-hailing because of a change from category *k* of the determinant variable to category *i* of the determinant variable. In computing this effect, we first assign the value of the base category for each individual in the sample (that is, we assign the value of $a_{qk} = 1$ to the determinant variable of each individual to compute $P(y_q = j \mid a_{qk} = 1)$) and then change the value of the variable to $a_{qi} = 1$ to compute $P(y_q = j \mid a_{qi} = 1)$).

The ATE measures may be computed for each nominal category *j* of the ride-hailing experience variable as well as each combination of *i* and *k* for the determinant variable. In our analysis, we compute the ATE measures for the nominal categories of “private only” and “pooled” ride-hailing experience, and for only two categories of the determinant variables. The base category for each determinant variable is used as the category to change from (as denoted by index *k* in Equation (3)) and a single non-base category of the determinant variable is selected as the category to change to (as denoted by index *i* in Equation (3)). For example, in the case of age, the base category is the “≥65 years” age group, while the changed category corresponds to the “18-34 years” age group. Similarly, for race/ethnicity, the base category is the “other” race/ethnicity (including individuals of Hispanic ethnicities and non-White races) and the changed category is the “non-Hispanic White” race/ethnicity. As already indicated, in the case of the latent psychosocial variables, the base “category” corresponds to the minimum expected (that is, deterministically predicted) value of the variable, and the changed “category” corresponds to the maximum value of the variable. Table 4-9, which provides the ATE values, shows the base category as well as the “changed category” for each determinant variable.

For the ride-hailing frequency ordinal variable, we assign cardinal values to each of the frequency ordinal levels, and then compute the ATE of determinant variables (in the same binary categorizations as discussed earlier for ride-hailing) on the expected total number of ride-hailing trips per month. The cardinal value assignments for the ordinal frequency levels in the model are as follows: (1) no ride-hailing trips: 0 trips in the past month, (2) 1-3 ride-hailing trips: 2 trips, (3) 4-5 ride-hailing trips: 4.5 trips, (4) 6-10 ride-hailing trips: 8 trips, and (5) more than 10 trips: 12 trips. With these assignments, the ATE corresponding to ride-hailing frequency for any determinant variable that is changed from category *k* to category *i* is computed as follows:

$$\hat{\text{ATE}}_{ik} = \frac{1}{Q}\sum_{q=1}^{Q}\left(\sum_{h=1}^{5} c_h \cdot \left[P(freq_q = h \mid a_{qi} = 1) - P(freq_q = h \mid a_{qk} = 1)\right]\right) \tag{4}$$

where c_h is the cardinal value assignment corresponding to the ordinal ride-hailing frequency level *h*, and $freq_q$ corresponds to the ordinal ride-hailing frequency of individual *q* in the 30 days prior to the survey.

Table 4-9 Treatment effects of different variables on ride-hailing adoption and frequency

Variable	**Categories compared (base versus changed)**	**Private only**		**Pooled**		**Frequency**	
		Est.	St. err	Est.	St. err	Est.	St. err.
Latent variables							
Privacy-sensitivity	Min vs. Max	0.000		-0.038	(0.021)	0.000	
Tech-savviness	Min vs. Max	0.160	(0.016)	0.029	(0.018)	0.000	
VSLP	Min vs. Max	0.007	(0.007)	0.028	(0.019)	0.626	(0.151)
GLP	Min vs. Max	0.000		0.000		0.000	
Sociodemographic variables							
Gender	Male vs. female	0.000		0.000		0.000	
Age	65+ vs. 18 to 34	0.223	(0.012)	0.034	(0.013)	0.000	
Race/ethnicity	Other vs. Non-Hispanic White	0.000		-0.020	(0.012)	0.000	
Education	< bachelor's vs. graduate	0.000		0.029	(0.007)	0.000	
Employment type	Full-time vs. part-time	-0.040	(0.017)	0.000		0.000	
Income	< $50,000 vs. $200,000+	0.290	(0.028)	-0.021	(0.014)	1.194	(0.134)
Household composition	Multi-worker vs. single-worker multi-person	-0.032	(0.011)	-0.012	(0.007)	0.000	
Endogenous variables							
Residential location	Rural vs. urban	0.160	(0.037)	0.067	(0.020)	0.580	(0.210)
Vehicle availability	<1 vs. >1 per worker	0.011	(0.004)	0.023	(0.006)	-0.087	(0.022)
Pooled ride-hailing user	No vs. yes	0.000		0.000		2.062	(0.234)

To calculate the ATE values in Equations (3) and (4), a realization of random draws is constructed by appropriately drawing from the sampling distribution of all the relevant parameters. Then the values of all dependent variables are calculated appropriately by following the chain of causal effects among the endogenous variables. The ATE values are computed for 1000 different draws (for each individual) so that standard errors are obtained. All results are presented in Table 4-9.

Among the latent variables, tech-savviness seems to be the strongest predictor of private ride-hailing experience with an ATE coefficient of 0.16. That is, if 100 random individuals increased their level of tech-savviness from the minimum to the maximum sample value, there would be 16 more individuals with private ride-hailing experience. In terms of pooled ride-hailing experience, privacy-sensitivity appears to be the most important deterrent, which suggests the need for concerted efforts to better understand the fundamental origins of high privacy-sensitivity, especially within the wealthiest population segment and non-Hispanic Whites (because these two groups have the highest privacy sensitivity). As importantly, qualitative research (such as focus groups) to identify how individuals may be steered toward being less privacy-sensitive in general (when in the presence of strangers in a ride-hailing trip), and especially within the wealthiest population segment and non-Hispanic Whites, would be beneficial. For instance, based on the prejudice literature within the social psychology field (see for example, Zebrowitz et al., 2008; Barlow et al., 2012), greater exposure may reduce people's aversion to strangers as long as experiences are positive. Thus, breaking the inertia barrier and encouraging people to experiment with pooled services even if only temporarily (through substantial cost incentives or convenience incentives) may naturally reduce privacy concerns and have a snow-balling effect on the use of future pooled ride-sharing. In this regard, understanding better the cost-privacy sensitivity trade-off would be a particularly valuable research pursuit to position pooled ridesharing services, especially to promote pooled ride-hailing within the low income segment and the non-Hispanic White population. Another important insight from our results is the negative correlation between green lifestyle propensity (GLP) and privacy sensitivity, which suggests that targeting individuals with a high GLP (women, non-millennials, and individuals with a graduate degree) and positioning information campaigns about the environmental benefits of pooled ride-hailing may be effective through the low privacy sensitivity prevalent in these population subgroups. While such campaigns should immediately increase pooled ride-hailing in women and in the group of individuals with a graduate degree (the second group is already pre-disposed toward pooled ride-hailing, as we will discuss later), our results suggest that information campaigns targeted toward non-millennials (and especially the oldest group of 65+ years) would be more effective if also combined with efforts to make this group of the population more tech-savvy, as discussed next.

The effects of the other latent variables in Table 7 indicate that tech-savviness and variety-seeking latent propensity (VSLP) have a positive impact on ride-hailing in general and pooled ride-hailing in particular. The positive impact on pooled ride-hailing adoption provides additional important policy insights. Tech-savviness levels in the population are generally increasing, thanks to information and communication technologies permeating into our routine daily lives. However, as evidenced in the results of Table 4, older and lower income segments seem to be falling behind and may need additional support to become "technologically-included". This calls for informational campaigns targeted at these population segments on how ride-hailing services function and how to use smartphone apps. The positive impact of VSLP on pooled ride-hailing suggests that perhaps one other way to promote pooled ride-hailing would be to promote the notion of VSLP through the use of multiple travel options, including pooled ride-hailing. Further, appealing to VSLP may be particularly effective in increasing pooled ride-hailing within the group of high income individuals and young adults (18-44 years of age), given that these groups have an intrinsically higher VSLP than their peers. Promoting VSLP in the context of pooled ride-hailing also has a substantial impact on ride-hailing frequency (see the last column of Table 7 corresponding to VSLP, which shows that if 100 random individuals were to have their VSLP levels increased from the minimum expected value of VSLP to the maximum expected value, there would be an additional 63 pooled ride-hailing trips over a period of 30 days); while ride-hailing frequency could not be split further into private and pooled modes in our study because of the very few number of individuals who reported the use of at least one pooled ride-hailing trip during the past 30 days (see Section 2.1.2), one would expect an increase in pooled ride-hailing frequency too through the promotion of VSLP.

The ATEs corresponding to the direct impacts of socio-demographic variables and the other endogenous variables, when combined with the latent variable effects just discussed, point to millennials, individuals belonging to races/ethnicities other than the non-Hispanic White with a graduate degree or higher, and those residing in urban areas as being the most likely to adopt pooled ride-hailing. In particular, the direct positive effect on pooled ride-hailing of being a millennial complements the indirect positive effect through the high tech-savviness and VSLP prevalent among millennials, while the direct positive effect of being of a race/ethnicity other than non-Hispanic White complements the low privacy sensitivity in races/ethnicities other than non-Hispanic White (as discussed earlier, privacy sensitivity appears to be the most important consideration in the use of pooled ride-hailing). Similarly, the direct positive effect of being a non-rural area resident complements the indirect positive effect through the high tech-savviness, VSLP and GLP among non-rural residents. The direct effects of income suggest that pooled ride-hailing is likely to be more adopted among individuals in low income households, which reinforces the positive indirect effect on pooled ride-hailing through the low privacy sensitivity in this low income group; however, this low income group also is less tech-savvy and has a low VSLP, both of which take away from the positive direct income effect. But promoting pooled ride-hailing in this low income group should still be effective, especially when combined with efforts to make individuals in this group more comfortable with the use of smartphone apps. More generally, the positive direct effect of low income on pooled ride-hailing is likely a reflection of the cost of ride-hailing services, which are still high. After controlling for the latent variable effects, the number of monthly ride-hailing trips would increase by an average of 1.2 trips (over a 30-day period) if a random individual were transferred from the lowest to the highest household income category, which indicates that ride-hailing use by the overall employed population can increase quite substantially if ride-hailing costs significantly drop. In that sense, the introduction of self-driving ride-hailing fleets, which promise to reduce ride-hailing trip costs, may play an important role in increasing the demand for ride-hailing services in general, and pooled ride-hailing services in particular.

Two important additional points about ride-hailing experience and frequency. First, policies that have the result of increasing the number of individuals who have experienced pooled ride-hailing immediately have the effect of increasing ride-hailing frequency too. According to the results in Table 4-9, a pooled ride-hailing user is likely to make about two more monthly ride-hailing trips than an individual who has had no experience with pooled ride-hailing. Thus, our results suggest that getting an individual to try pooled ride-hailing that one time can have a lasting impact on the frequency of pooled ride-hailing over the longer term. Second, we also computed ATEs based on the IHDM model so that we can evaluate the magnitude of any self-selection effects of residential choice and vehicle availability on ride-hailing experience and frequency. As expected, ignoring these self-selection effects (as the IHDM model does) led to a higher magnitude of effect of urban living and vehicle availability on both private and pooled ride-hailing, as well as on ride-hailing frequency. Similarly, the effect of being a pooled ride-hailing user on ride-hailing frequency was also over-estimated in the IHDM model. However, as also anticipated in Section 1.2, these overestimations from the IHDM model were marginal and statistically very insignificant. The important insight is that, at least at the current point in time, ride-hailing is a relatively new mobility option within the larger time scale at which residential choice and vehicle ownership decisions are made. Thus, at least in the very near term, studies may assume residential location choice and vehicle ownership decisions as being exogenous to ride-hailing choices, with reasonable confidence that, in doing so, the effects of residential location and vehicle ownership choices are still "true" causal effects. Of course, over time, this could change, with ride-hailing not just viewed as a component of MaaS, but as one element of a much broader lifestyle choice that includes residential choice and vehicle ownership. The analysis framework used in this study is thus very general, and can accommodate the more expansive lifestyle choice bundle context that is likely to unfold over time.

4.5.2 From the Trip-Level Model

The trip-level model in this paper is more of an exploratory nature, and thus the variable effects on the many dimensions of ride-hailing should be viewed with much more caution than for the individual-level model of the previous section. However, there are still some important insights from the results that we briefly summarize in this section.

An observation from the trip purpose results is that women are rather unlikely to use ride-hailing for routine errand trips, even though the women in a household are primarily responsible for personal, family, and shopping errands. At the same time, the "mode substituted" model results reveal that many of the new trips generated by the availability of ride-hailing (and that would not have been made otherwise) are for running errands. The implication is that, while ride-hailing provides more access to activity opportunities, it is also not the most convenient for running errands. This is perhaps because running errands typically involves chaining of multiple activities in the same sojourn from home and/or involves carrying and storing food and other perishable goods during the trip, and ride-hailing is not the most convenient because it is more of a pure trip-based consumption service as opposed to a broader transportation option that allows a cost-effective time-based consumption service (in which the same vehicle is available to pursue multiple activities and over an extended period of time). Perhaps ride-hailing providers (and more broadly, MaaS providers) need to be thinking about providing a time-based option too, which effectively would combine today's ride-hailing and car-sharing services into one service. As the mobility landscape moves more toward automated vehicles, this integration of trip-based and time-based consumption options may become even easier to implement.

Another interesting result pertaining to ride-hailing as a MaaS solution relates to commuting. Commuting encompasses a significant share of daily trips, and hence, these trips should be accommodated by a MaaS system. Despite the lower numbers of work trips captured in our sample (compared to trips to the airport and trips to recreational activities), the model results show that frequent users are likely to use ride-hailing for work trips (from the trip purpose model), and work trips by ride-hailing are typically made alone (based on the trip companion model) during the morning and evening periods (as per the time-of-day model). The net result is that many ride-hailing trips for work during the morning and evening are undertaken in private ride-hailing mode as opposed to pooled ride-hailing mode. There is substantial opportunity for ride-hailing services as well as employers to work together to increase vehicle occupancy during the commute periods, through low cost pooled ride-hailing services (such as Uber's most recently introduced "Express Pool" service) and subsidizing the use of such services. Also, appealing to the range of co-travelers one has the possibility to meet, alongside campaigns to reduce privacy-sensitivity among individuals of non-White Hispanic race/ethnicity and high income individuals, may be additional policy instruments available to promote pooled ride-hailing. Compared to traditional car-pooling arrangements that typically have scheduled times of arrival and departure, pooled ride-hailing would offer more time flexibility for workers, and would not necessitate any traveling individual driving (except, of course, the ride-hailing service driver).

The results on substitution of trips made earlier by active modes or transit and now replaced by ride-hailing also provide additional insights. First, the results reveal that people younger than 65 years of age are more likely than those 65 years of age or older to substitute active travel/transit by ride-hailing. This can further reduce the physical activity levels of individuals, and pose additional public health problems given that regular physical activity levels have now been proven to be effective as preventive medicine for a number of obesity-related diseases (Ku et al., 2018; Stamatikes et al., 2018). This is particularly of concern, since about 85.5% of the US population is under the age of 65 years (U.S. Census Bureau, 2018e) and obesity-related problems earlier on in life do lead to poor health outcomes later on in life (Cheng et al., 2016; Zheng et al., 2017). Thus, from both a traffic congestion standpoint as well as a health standpoint, policies that discourage the substitution of short-distance "walkable" trips by ride-hailing (such as a pricing scheme that more heavily prices the first mile, except if the patron is mobility-challenged) would be particularly valuable. Second, pooled ride-hailing users are more likely to have been drawn away from transit and active travel. The door-to-door travel convenience and relatively low cost differential (between pooled ride-hailing and transit) appears to lead to the substitution of transit by ride-hailing (see also Clewlow and Mishra, 2017). Of course, to increase the efficiency and sustainability of a MaaS system, the relationship between (pooled) ride-hailing and transit should be one of complementarity rather than substitution. Yet, it is reasonable to expect that a service that can be used for door-to-door trips will not be used for first- and last-mile connectivity to transit hubs, unless low cost and well-integrated MaaS systems are designed.

In addition to the substitution of transit, another negative externality of ride-hailing is that most new induced trips are generated by individuals in suburban and urban areas (rather than rural areas), serve a single passenger, and occur in the morning commute period as well as the mid-day and night periods (see the last column of Table 4-9). In other words, ride-hailing is generating more "drive alone" trips in the already-congested suburban and urban areas of the DFW area, contrary to the main transportation and environmental goals of MaaS systems. This is a serious concern in an era of dwindling real estate and financial assets to build new roads, along with increasing urban populations. There is a need for more consideration of congestion pricing schemes that discourage private ride-hailing (especially in the morning commute period), as well as a need to re-visit the criteria and fee structure for the use of managed lanes (for example, a high-occupancy vehicle may have to be defined as 3+ individuals in the vehicle as opposed to 2+).

Finally, on a more positive note, ride-hailing can provide more access to activity opportunities for individuals who do not own vehicles and/or those with limited driving capabilities. Our model results provide initial evidence for this, as we observe that students, individuals with low vehicle availability, and individuals from low-income households are generally more likely than their peers to use ride-hailing to run errands. Thus, ride-hailing can assume a welfare role, but fares would need to be revisited to fit the needs of these more financially challenged segments of our society.

4.6 Conclusions

The objective of MaaS systems is to reduce drive alone trips and promote multi-modal travel behavior. In an ideal scenario, cities would have robust transit systems as a centerpiece of a transportation system that can be accessed and egressed through other integrated travel modes such as shared bicycles or even cars. However, the reality in many metropolitan areas in the U.S. is that they are not dense enough for viable and effective public transit systems. An alternative that is being extensively studied on the supply side is the use of shared vehicles and pooled rides. The recent growth of ride-hailing services may suggest the feasibility of such an alternative, especially when considering that, in the near future, self-driving vehicles will be available, and the costs of rides should decrease. In this study, we undertook a comprehensive analysis of ride-hailing travel behavior by developing multivariate models of experience, frequency and trip characteristics as functions of lifecycle, lifestyle and built-environment variables. These analyses serve as inputs to two broader travel behavior questions: (1) Is pooled ride-hailing a feasible MaaS solution in currently car-dominated cities?, and (2) Is there evidence of the presence of positive and negative externalities of ride-hailing adoption?

Our results show that, from a behavioral perspective, a service-based transportation future where people predominantly travel using hailed pooled rides instead of their own vehicles is probably still distant. The evaluation of ATEs showed that, in isolation, each variable has only a marginal effect on the adoption of pooled ride-hailing. Thus, a complex combination of actions is required to promote the use of these services. Among these actions, we identified the need for campaigns to (a) increase technology awareness among older individuals and individuals from lower income households, and (b) reduce privacy-sensitivity among non-Hispanic Whites. However, such efforts would still need to be complemented by a decrease in service fares. In this regard, understanding better the cost-privacy sensitivity trade-off would be a particularly valuable research pursuit to position pooled ridesharing services. Additionally, even after accounting for self-selection effects, and considering that our area of analysis has a high share of suburban land (DFW), the key ingredient to ride-hailing use, and especially pooled ride-hailing use, still seems to be urban density.

The trip-level model results suggest that ride-hailing providers (and more broadly, MaaS providers) need to be thinking about providing a time-based consumption option, which effectively would combine today's ride-hailing and car-sharing services into one service. Doing so would likely make the use of ride-hailing for running errands more convenient. Ride-hailing services as well as employers also can work together to increase vehicle occupancy during the commute periods, through subsidized and reliable pooled ride-hailing services. The results also suggest a need for policies that discourage the substitution of short-distance "walkable" trips by ride-hailing (to reduce traffic congestion as well as not take away from active modes of transportation), and a need for low cost and well-integrated MaaS systems to avoid substitution of transit trips by ride-hailing.

More generally, the results in this study reveal that ride-hailing is fundamentally changing the spatial, temporal, and modal activity-travel landscape of individuals. Socio-demographics, through their direct and indirect effects (though the latent psycho-social constructs), influence this landscape. Thus, as socio-demographics change, so will the activity-travel patterns of individuals. It is important for planning agencies to collect data on ride-hailing and incorporate ride-hailing behavior (and more generally MaaS system features) within their activity-travel modeling systems. Doing so is not only important for forecasting activity-travel patterns, but also to design good MaaS systems through an understanding of how ride-hailing may be integrated with other travel modes.

CHAPTER 5. Modeling Individual Preferences for Ownership and Sharing of Autonomous Vehicle Technologies

The majority of the content of this chapter has been previously published in Lavieri, P.S., V.M. Garikapati, C.R. Bhat, R.M. Pendyala, S. Astroza, and F.F. Dias (2017). Modeling Individual Preferences for Ownership and Sharing of Autonomous Vehicle Technologies. *Transportation Research Record*, Vol. 2665, pp. 1-10

5.1 Introduction

In this chapter, we analyze travelers' interest in adopting AV technology and determine the extent to which they are inclined to acquire such vehicles for private ownership or use them in a shared mobility service configuration. We develop a model system that accounts for the endogeneity of household and individual choices such as current vehicle ownership, use of car-sharing and ride-hailing services, and residential location. To better understand consumer preferences for, attitudes toward, and adoption and potential use of AVs, we use a travel survey data set collected in 2014-2015 in the Puget Sound region of the State of Washington in the United States. Through the computation of elasticity measures, potential early adopters of AVs, both in a private ownership mode as well as a shared mode, are identified. Based on these results, we identify socio-demographic, land-use and behavioral elements that may contribute to the success of MaaS systems and discuss necessary measures to increase the potential of such schemes.

5.2 Earlier Studies

Although there is some literature that addresses user preferences, concerns, and adoption of automated vehicle technologies, much remains to be learned in this particular domain (see Becker and Axhausen, 2017, for a recent literature review). Kyriakidis et al. (2015) and Haboucha et al. (2017) reported that there is considerable heterogeneity in preferences and willingness to pay for automated vehicles, with those who drive more being more amenable to adopting and paying for automated vehicles. Bansal et al. (2016) conducted a survey of 347 people in the city of Austin, Texas, and found that more than 80 percent of the respondents are interested in owning and using fully automated vehicles. Also based on data from Austin, Zmud et al. (2016) identified that most respondents had preferences for owning rather than sharing

AVs. Preferences for SAV were also investigated by Krueger et al. (2016), who conducted a stated choice experiment in which individuals choose between SAV, PSAV, and transit. However, since ownership was not an option, their results do not allow for conclusions regarding what is going to be the true preference of current auto owners.

While the literature cited above provides some insights on consumer preferences for advanced transportation technologies and services, there is limited understanding of how human attitudes and lifestyle factors affect potential adoption and use of these technologies. Schaefers (2013) uses interviews and qualitative analysis methods to investigate the motivations behind car-sharing usage. She concludes that sense of community and identification with the lifestyle of other users are important motivating factors for car-sharing membership. More recently, de Almeida Correia and van Arem (2016) noted that despite recent signs of shifts in car ownership and travel patterns brought on by the shared economy, "owning and using an automobile is still linked to both instrumental and symbolic-affective motives". The few studies that incorporate attitudinal variables when modeling AV adoption preferences (such as Haboucha et al., 2017) do not use an integrated framework (that is, the latent variables are computed through confirmatory factor analysis and their expected value is calculated and included as an exogenous variable in the final model) and do not control for self-selection effects. As discussed in Chapter 2, lifestyle preferences, consumer attitudes, and perceptions need to be taken into account when modeling consumer adoption and use of transformative transportation technologies in an integrate manner so that both taste heterogeneity and self-selection affects are accounted for.

5.3 Methodology

This section presents the behavioral framework followed by a brief overview of the modeling methodology.

5.3.1 Behavioral Framework

In this analysis, consumer interest in the adoption and use of AVs is modeled as a function of individual lifestyle preferences, attitudinal factors, and current use of disruptive transportation services. The current choices that are assumed to affect the interest in AV adoption include the use of car-sharing and/or ride-sourcing services, vehicle ownership, and density of the residential location. It may be expected that individuals who currently own vehicles are more likely to favor private ownership of AVs over shared use.

Among underlying lifestyle factors that may affect the propensity to adopt AVs, two key aspects are considered in this analysis. These include "green lifestyle propensity" and "technology savviness". These factors have been identified in the literature as important determinants of transport choices (Schaefers, 2013; Wolf and Seebauer, 2014; Seebauwer et al., 2015). Consistent with the literature, we hypothesize that individuals who are green lifestyle oriented and technology-savvy are more likely to adopt AVs in private ownership mode or shared mobility-on-demand service mode, or both.

The use of latent lifestyle factors is critical to explaining traveler choices in different contexts. Lifestyle constructs are modeled in our framework as a function of demographic characteristics, as well as a function of characteristics unobserved to the analyst. Assuming lifestyle variables as independent variables in choice models, when in fact they are stochastic functions of socio-economic and demographic variables, will result in inconsistent model parameter estimates and erroneous inferences regarding the magnitude of the impacts of various factors on choice behaviors (Bhat and Dubey, 2014). At the same time, treating lifestyle factors as determinants of choice variables requires the specification and estimation of joint model systems (such as the one used in this study) capable of accounting for unobserved exogenous factors that jointly affect multiple endogenous outcomes. The joint model system also recognizes that individuals may be selecting a lifestyle package or bundle where a multitude of choices are made together. Figure 5-1 shows a simplified representation of the behavioral framework adopted in this study. The two lifestyle factors, green lifestyle and tech-savviness, are assumed to affect both current mobility choices as well as interest in AV adoption and use in the future.

The factor that represents the propensity for a green lifestyle corresponds to a number of variables present in the survey data set. These include the following:

- Frequency of transit usage, measured on a seven-point scale
    - Never
    - Have used transit, but not in the past month
    - 1-3 times per month
    - 1 day per week
    - 2-4 days per week
    - 5 days per week
    - 6-7 days per week
- Importance of a walkable neighborhood and being close to activities in choice of home location (five-point scale "very unimportant" to "very important")
- Importance of being close to public transit in choice of home location (same scale as above)

- Importance of being within a 30-minute commute to work in choice of home location (same scale as above)

The factor that captures tech-savviness corresponds to the following variables present in the survey data set:

- Smartphone ownership, measured on a three-point scale
    - Do not have and do not plan to buy a smartphone
    - Do not have but plan to buy a smartphone
    - Have a smartphone
- Frequency of use of smartphone apps for travel information, measured on a seven-point scale (same as scale used for "frequency of transit use" above)
- Frequency of use of in-vehicle GPS, measured on a seven-point scale (same as scale used for "frequency of transit use")

The choice variables are modeled as a bundle within a simultaneous equations modeling framework with latent constructs and socio-demographic variables serving as explanatory variables. There are five simultaneous choice models for the following endogenous outcomes:

- One multinomial choice variable representing interest in future adoption/use of AV
    - No interest
    - AV sharing only
    - AV ownership only
    - Both AV sharing and ownership
- Three binomial choice variables representing current choices including:
    - Has ever used car-sharing service (yes/no)
    - Has ever used ride-sourcing service (yes/no)
    - Household resides in high-density area (yes/no)
- One count variable representing household vehicle ownership

The endogenous outcomes are also allowed to directly impact one another following the directionality presented in Figure 5-1. A number of model specifications were tested, and the final model specification was selected based on statistical significance and fit, behavioral intuitiveness of the model structure/relationships, and desired sensitivity in the model system.

5.3.2 Modeling Approach

The modeling methodology adopted in this study is based on the GHDM approach proposed by Bhat (2015a) and discussed in Chapter 2. This model enables the consideration of multiple ordinal, count, continuous, and nominal variables jointly using a latent variable structural equation model that ties latent constructs to exogenous variables, and a measurement model that links the latent variables and possibly other explanatory variables to a set of different types of outcomes.

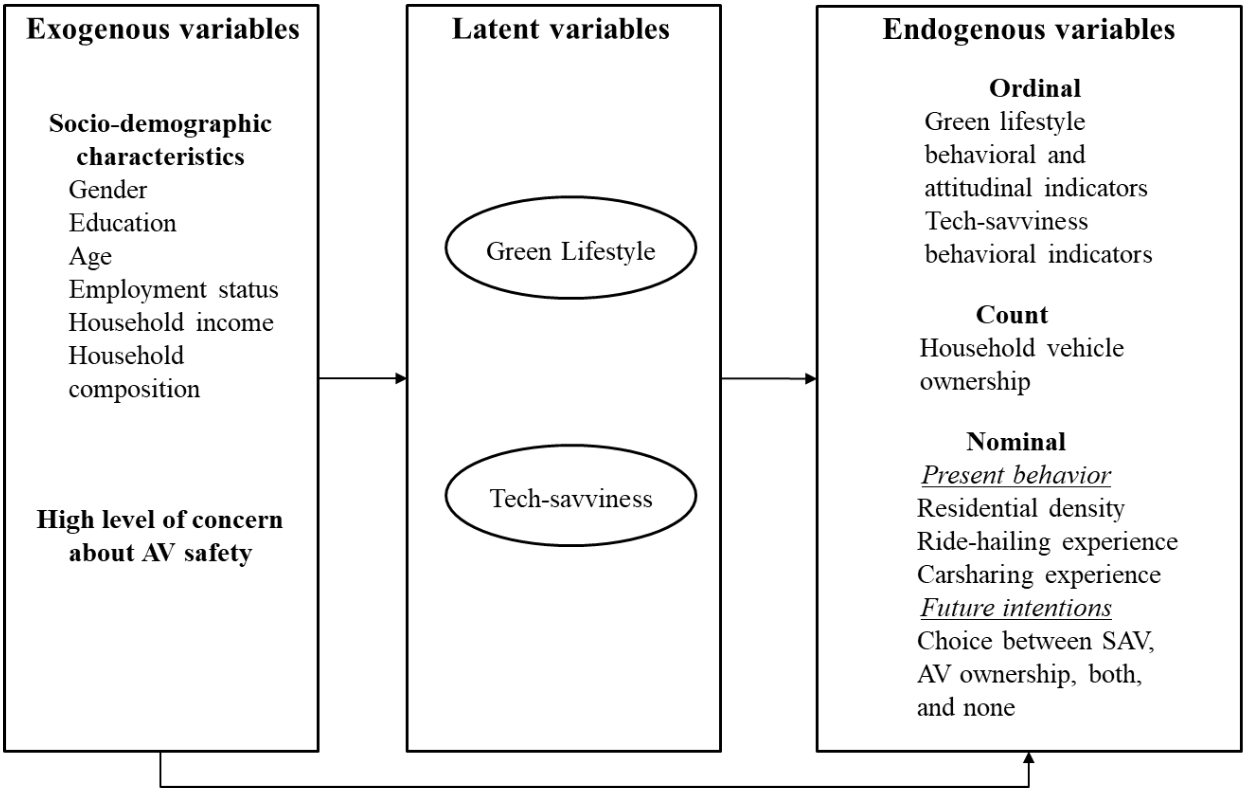


Figure 5-1 Simplified representation of the modeling framework

5.4 Data

Data for this study is derived from the Puget Sound Regional Travel Study that involved survey data collection efforts in 2014 and 2015. The survey data includes detailed information about socio-economic, demographic, activity-travel characteristics, attitudes and preferences. For this research effort, the subset of households that provided complete data in both years was used. Within this sample, individuals less than 18 years of age and individuals whose survey responses were collected through the use of a proxy were excluded. The final sample includes 1,832 individuals.

Autonomous vehicles were defined in the survey as follows: "Autonomous cars, also known as "self-driving" or "driverless" cars, are capable of responding to the environment and navigating without a driver controlling the vehicle". The survey included five questions about level of interest in AV adoption and usage. Two of the questions were used to construct a four-alternative multinomial choice variable that captures the level of interest in AV use. The two variables are:

- Level of interest in owning an autonomous car (five-point scale: "not at all interested”, “somewhat uninterested”, “neutral”, “somewhat interested”, and “very interested”)

- Level of interest in participating in a SAV system for daily travel (same scale as above).

A descriptive analysis showed a substantial percent of respondents in the “not at all interested” category, with an additional small percentage in the “somewhat uninterested” category. Further, because of the ambiguity of the “neutral” category, the ordinal expression of interest was collapsed into a binary variable. Individuals who were somewhat interested and very interested were considered as being interested in the technology, while all others were treated as being uninterested. It should be noted that including the individuals that have a neutral interest in using AV technology in the category of “not interested” would provide a conservative estimate of adoption rates if the model is used for prediction purposes. Since the survey did not offer detailed explanations about the AV technology and service characteristics to respondents, being conservative and leaning towards a lower adoption rate estimate was considered prudent in this context. The binary indicators of the levels of interest in AV ownership and AV sharing were combined into a single multinomial variable with four alternatives as follows: (1) Not interested in AV sharing or AV ownership (68.5%); (2) Interested in AV sharing only (7.6%); (3) Interested in AV ownership only (8.5%); and (4) Interested in AV sharing and AV ownership (15.4%).

In addition, the survey collected information about the general level of concern that individuals had with respect to AV technology. Five questions captured the level of concern related to AV equipment and system safety, system and vehicle security, ability to react to the environment, performance in poor weather or other unexpected conditions, and legal liability for drivers or owners. The highest level of concern expressed on any of the questions except for the last one (related to liability) was considered the level of concern with AV technology, while the level of concern on the liability question was considered separately. AV technology concern was tested as an endogenous variable, but the influence of tech-savviness was found to be insignificant; hence it was treated as an exogenous variable in the final model specification.

The current usage of car-sharing or ride-sourcing services is represented by two binary dependent variables. Individuals who used either service at least once in their lifetime were considered “users” as opposed to those who had never used either service (“non-users”). Residential density was calculated for each census block, with blocks that had a density of 3,000 households per square mile or more treated as high-density locations. For the sake of brevity, a detailed table of descriptive statistics of the sample is not furnished. Overall, the sample exhibits characteristics that render it suitable for a modeling exercise such as that undertaken in this chapter. It was found that 51% of the respondents resided in low-density neighborhoods, 12% resided in zero-vehicle households, and 39% resided in one-vehicle households. Those between 18 and 44 years of age constitute 37.5% of the sample. Among other characteristics of interest, 60% of respondents are workers, 44% are males, 93% have a driver's license, 38% have an undergraduate degree, and 29% have a graduate degree. Within the sample, 14% used a ride-sourcing service at least once, while 9.2% used a car-sharing service at least once in their lifetime. With respect to smartphone ownership, 67.7% of the respondents own a smartphone, and 28.5% of the respondents stated that they do not currently have a smartphone and have no plans to buy one.

5.5 Model Results

This section provides a brief discussion of the model estimation results, which are furnished in Tables 5-1 and 5-2.

5.5.1 Structural Equation Model Results

The results of the structural equation model component are presented in the top half of Table 5-1. Both green lifestyle propensity and tech-savviness are associated with a higher level of education attainment. This finding is consistent with the prior literature; for example, Bhat (2015b) found education to be associated with green lifestyle, while Seebauer et al. (2015) found a strong association between education level and technology adoption/use. Younger individuals show a greater propensity towards a green lifestyle, consistent with the findings of Garikapati et al. (2016) who find that millennials use alternative modes of transportation more than other generations. Gender was not found to be significant in explaining lifestyle preferences.

Lower income households are more likely to be associated with a green lifestyle. Indeed, Bhat et al. (2016) noted that a lower overall consumption level and higher alternative mode use in these households places them into the green lifestyle category relative to higher income households that tend to have a larger carbon footprint. On the other hand, lower income respondents tended to be less technology-oriented, which is consistent with expectations as there may be cost barriers involved. Workers are more prone to be tech-savvy, consistent with the notion that such individuals are likely to be exposed to technology in the workplace (Morahan-Martin and Schumacher, 2007). Respondents in households with children are less likely to be associated with a green lifestyle; this finding is consistent with Bhat (2015b) who notes that households with children tend to favor suburban residential locations with larger homes and open spaces, leading to a less green lifestyle.

The correlation between tech-savviness and green lifestyle propensity was found to be statistically insignificant. It appears that the model specification captured the key variables associated with green lifestyle propensity and tech-savviness, resulting in an insignificant error correlation for the structural equation model component. Alternatively, the specification may have been such that positive and negative correlations caused by unobserved factors may have cancelled out, leading to the result found here.

5.5.2 Measurement Equation Results

The second half of Table 5-1 provides estimation results for the non-binary and non-multinomial endogenous variables of the measurement equation component. There are seven ordinal indicators (four indicators corresponding to green lifestyle propensity and three indicators corresponding to tech-savviness) and one count variable corresponding to number of vehicles (automobiles) in the household. The constant indicates the overall proclivity of the survey respondents, but does not have a behavioral interpretation per se. Focusing on the factor loadings, it can be seen that a green lifestyle is associated with a higher frequency of transit use, and a higher level of importance for living in a walkable neighborhood, close to transit, and within a 30-minute commute of work. Tech-savvy individuals exhibit a greater frequency of the use of apps for travel information, tend to own smartphones, and are more prone to using GPS for travel information (as also observed by Seebauer et al., 2015).

Table 5-2 presents estimation results for the measurement equation component associated with the binary/multinomial variables. With respect to AV use, it appears that the respondent sample is generally *not* inclined to use AV as evidenced by the negative alternative specific constants. Males are more inclined (than females) to be interested in both AV-ownership and sharing, while education does not have a statistically significant impact (though education does play a role through the latent constructs). Younger adults aged 18-24 years old appear to be less inclined towards AV ownership than adults 25 years or older. However, both age groups show a positive propensity to both own and share AVs. Note that these age effects go beyond those permeating to AV choice through the latent lifestyle constructs. As expected, lower levels of vehicle ownership are associated with a greater proclivity towards AV-sharing.

Fewer current vehicle holdings and residing in higher density neighborhoods lead to a higher propensity for AV sharing relative to no interest at all in AV, interest in AV ownership only, or interest in both AV ownership and AV sharing. Those residing in higher density neighborhoods are likely to favor AV sharing as they do not need to travel long distances to access destinations and may experience parking constraints. Individuals who have experienced car-sharing are less likely to favor ownership in an AV era, a finding that is consistent with that reported by Clewlow (2016). Similarly, those who have used ride-sourcing services are more likely to favor AV-sharing, or AV-ownership coupled with sharing services. As expected, those who have a higher level of concern about AV technology are less likely to adopt it.

The latent variables have the expected impacts on future AV use, with a green lifestyle favoring AV sharing, and tech-savviness leading to a higher likelihood of embracing AV technology in general, and especially a combination of both AV ownership and AV sharing. The effects of these latent variables create heteroscedasticity and covariances across the utilities of the AV adoption alternatives in ways that are not likely to be as readily obvious as a covariance specification if a direct multinomial probit type model were to be estimated for the future AV use outcome. At the same time, the latent variables also impact current car-sharing and ride-sourcing experience, and current residential density living choice. This indicates that the effects of these latter variables on future AV use would be over-estimated if the stochastic latent variables were not included in the model system (and instead, car-sharing and ride-sourcing experience, and residential location density, were introduced directly as exogenous variables in the future AV choice component).

Table 5-1 Estimation Results for Structural and Non-Nominal Measurement Equations

Structural Equation Component	Green Lifestyle		Tech-savviness	
Variable	**Coefficient**	**(t-stat)**	**Coefficient**	**(t-stat)**
Education (base: less than Bachelor's degree)				
Bachelor's degree	0.363	(3.76)	0.180	(1.37)
Graduate degree	0.607	(4.41)	0.180	(1.37)
Age (base: 65+ years old)				
18 to 24 years old	0.986	(7.49)	1.196	(3.07)
25 to 44 years old	0.986	(7.49)	0.837	(6.07)
45 to 64 years old	0.482	(4.11)	--	--
Income (base: $75,000 or more per year)				
Under $24,999 per year	0.464	(4.70)	-0.769	(-1.23)
$25,000-$34,999 per year	0.464	(4.70)	-0.358	(-2.06)
$35,000-$74,999 per year	--	--	-0.358	(-2.06)
Employment status(base: non-worker)				
Worker	--	--	0.595	(1.73)
Household Composition (Base: no kids)				
Kids under 5 years old	-0.503	(-3.34)	--	--
Kids 5-17 years old	-0.743	(-5.01)	--	--
Correlation between latent variables	--			

Latent variables	**Indicators/outcomes**	**Constant**	**(t-stat)**	**Factor loading**	**(t-stat)**
	Ordinal				
Green Lifestyle	Frequency that uses transit	0.002	(0.02)	0.889	(8.60)
Green Lifestyle	Importance of having a walkable neighborhood	1.439	(11.97)	0.586	(16.57)
Green Lifestyle	Importance of being close to public transit	0.692	(4.37)	1.085	(16.19)
Green Lifestyle	Importance of being within a 30-min commute to work	1.048	(12.51)	0.360	(9.33)
Tech-savviness	Frequency of smartphone app use for travel info	-3.450	(-1.76)	3.374	(5.89)
Tech-savviness	Smartphone ownership	0.386	(0.10)	2.523	(6.72)
Tech-savviness	Frequency of GPS use for travel info	-0.701	(-3.24)	0.248	(2.32)
	Count				
Green Lifestyle	Number of vehicles in the household	0.540	(1.24)	-0.322	(-3.22)

Exogenous variables impacting the number of vehicles in the household (count outcome)	
Number of adults in the household	0.806 (2.79)
High residential density of household census bock (more than 3000hh/mi^2)	-0.653 (-1.69)

(--) coefficient was not different from zero at the 90% level of confidence and was removed from the model.

Table 5-2 Model Estimation Results for Binary/Multinomial Endogenous Variables

Variable	Coef	t-stat	Coef	t-stat	Coef	t-stat
Type of AV Use (Base: not interested in AV)	**AV sharing**		**AV Ownership**		**Both**	
Constant	-1.578	(-8.70)	-0.707	(-6.19)	-2.876	(-9.61)
Gender (base: female)						
Male	--	--	--	--	0.337	(5.60)
Education (base: less than Bachelor's degree)						
Bachelor's degree	0.091	(1.40)	0.091	(1.40)	0.091	(1.40)
Graduate degree	0.091	(1.40)	0.091	(1.40)	0.091	(1.40)
Age (base: 45+ years old)						
18 to 24 years old	--	--	-0.168	(-1.89)	0.827	(3.75)
25 to 44 years old	--	--	--	--	0.827	(3.75)
Vehicles in the household (base: 2 or more)						
No vehicle	0.409	(6.08)	--	--	--	--
One vehicle	0.121	(2.01)	--	--	--	--
Residential density of household census bock (base: less than 3000hh/mi^2)						
High density	0.223	(5.05)	--	--	--	--
Has Experienced Carsharing (base: never)						
Used	--	--	-0.167	(-3.08)	--	--
Has Experienced Ride-hailing (base: never)						
Used	0.424	(4.42)	--	--	0.424	(4.42)
Concern about AV technol. problems (base: low)						
High level of concern	-0.088	(-2.17)	-0.088	(-2.17)	-0.088	(-2.17)
Latent Variable: Green Lifestyle Propensity	0.114	(1.33)	--	--	--	--
Latent Variable: Tech-savviness	0.207	(1.61)	0.132	(1.92)	0.300	(1.61)
Carsharing Experience (Base: never used)	**Used at least once**					
Constant	-5.632	(-10.15)				
Gender (base: female)						
Male	0.187	(3.22)				
Driver's license (base: doesn't have a license)	1.887	(9.74)				
Vehicles in the household (base: 2 or more)						
No vehicle	1.811	(9.86)				
One vehicle	0.486	(6.67)				
Residential density of household census bock (base: less than 3000hh/mi^2)						
High density	0.650	(8.29)				
Latent Variable: Green Lifestyle Propensity	0.454	(4.43)				
Latent Variable: Tech-savviness	0.706	(4.73)				
Ride-hailing Experience (Base: never used)	**Used at least once**					
Constant	-3.470	(-6.75)				
Vehicles in the household (base: 2 or more)						
No vehicle or one vehicle	0.213	(2.92)				
High residential density of household census bock (base: less than 3000hh/mi^2)						
High density	0.931	(11.63)				
Latent Variable: Green Lifestyle Propensity	0.451	(4.40)				
Latent Variable: Tech-savviness	0.941	(5.32)				
Residential Density (Base: < 3000hh/mi^2)	**High density**					
Constant	-0.869	(-11.95)				
Latent Variable: Green Lifestyle Propensity	0.990	(13.03)				
Latent Variable: Tech-savviness	--	--				

(--) coefficient was not different from zero at the 90% level of confidence and was removed from the model.

Results consistent with expectations are found in the other endogenous variables models. In the model of car-sharing experience, it is found that males are more likely than females to have tried car-sharing. Those with a driver's license, those residing in households with fewer vehicles, and those in high density neighborhoods are more likely to have utilized car-sharing services. Similar indications are found in the model of ride-sourcing experience, except that gender and driver's license holding do not appear to be significant in the ride-sourcing model. As a driver's license is not needed to use ride-sourcing services, it is not surprising that this variable is insignificant in this specific model component. Green lifestyle propensity and tech-savviness are positively associated with the current use of car-sharing and ride-sourcing services.

5.6 Model assessment and computation of pseudo-elasticities

This section presents an assessment of model performance and offers pseudo-elasticity measures that may be used to determine the sensitivity of the adoption and use of AV technology to various factors. Table 5-3 presents results of the model assessment and Table 5-4 the pseudo-elasticity computations.

Table 5-3 Model Assessment

Summary Statistics	**GHDM**	**IHDM**
Composite Marginal log-likelihood value at convergence	-241,784.0	-277,212.7
Composite Likelihood Information Criterion (CLIC)	-242,606.7	-278,257.6
Log-likelihood at constants	-10,097.2	
Predictive log-likelihood at convergence	-9,466.4	-9,555.2
Number of parameters	97	112
Number of observations	1,832	1,832
Predictive adjusted likelihood ratio index	0.046	0.032
Non-nested adjusted likelihood ratio test between the GHDM and IHDM	$\Phi[-63.11] << 0.0001$	

Disaggregate Goodness -of-fit				
Overall probability of correct prediction	0.53			
Shares of Level of Interest				
	Not interested	AV sharing	AV ownership	Both
Real sample shares	68.50%	7.64%	8.46%	15.39%
Predicted shares	68.98%	7.20%	7.96%	15.86%
Absolute percentage bias	0.70%	5.79%	5.92%	3.03%
Predicted shares for the population (after applying weights)	70.30%	4.88%	7.76%	17.06%

The performance of the GHDM structure used here may be compared to one that assumes independence across the many endogenous outcomes (that is, across the current choices and future intentions shown in Figure 5-1). To arrive at a good initial specification for the second structure, an independent heterogeneous data model (IHDM) is estimated in which the determinants of the latent constructs are included directly as exogenous variables. This is an independent model because error term correlations across the choice dimensions are ignored. The GHDM and the IHDM models are not nested, but may be compared using the composite likelihood information criterion (CLIC) (Bhat, 2015b). The model that provides a higher value of CLIC is preferred. The two models can also be compared through a non-nested adjusted likelihood ratio test as described in Bhat (2015b). The results of these disaggregate data fit evaluations are provided in the first part of Table 3. The CLIC values clearly favor the GHDM over the IHDM. The same result is obtained when comparing the predictive likelihood values and adjusted likelihood ratio indices, and computing the non-nested likelihood ratio statistic.

Next, to examine the performance of the GHDM more intuitively, an "average probability of correct prediction" measure is computed for the future AV multinomial choice dimension of the model system. This is calculated to be 0.53. At the aggregate level, the actual sample shares and GHDM predicted shares are computed for the different alternatives related to future AV use and adoption. The predicted shares are computed by drawing 1,000 samples of 1,832 observations from a multivariate normal distribution and taking an average over the predictions. The absolute percent bias values in the predicted shares are quite small, suggesting that the model is able to recover overall shares quite well.

Elasticity measures were computed to identify early adopters of AV technology in general, and to identify market segments that may favor one form of AV adoption over another (i.e., sharing versus ownership or both). The elasticity results in Table 5-3 represent the percentage change in the probability of being in one of the four user categories. For example, being a worker increases the probability of an individual being interested in AV sharing by 20% (from 0.072 to 0.086). Overall, early adopters of AV technology are likely to be those with a higher level of education, individuals between 18 and 44 years of age, and workers. In particular, individuals in the youngest age group of 18-24 years show the greatest propensity for AV sharing and an aversion towards the AV ownership-only alternative. Individuals with a higher level of education are also more likely to adopt AV sharing as opposed to ownership or both, as evidenced by the higher elasticity measures within the AV sharing column. Lower income individuals appear to be largely averse to the adoption of AV technology in any form with those in the lowest income category showing the greatest level of resistance to adoption. While experience with the use of ride-sourcing services is associated with a propensity to adopt AV sharing and both sharing and ownership, experience with car-sharing services does not contribute to adoption of AV. High density neighborhood residents are also more inclined to adopt AV sharing services as opposed to any model that involves ownership.

Table 5-4 Elasticity Computations

Elasticities				
Variable	Not interested	AV sharing	AV ownership	Both
Bachelor's degree (base: less than Bachelor's degree)	-2.33%	15.68%	4.94%	1.20%
Graduate degree (base: less than Bachelor's degree)	-2.91%	21.77%	4.94%	1.20%
Age 18 to 24 (base: ≥ 65 years)	-14.86%	24.24%	-42.86%	118.18%
Age 25 to 44 (base: ≥ 65 years)	-16.08%	12.12%	-10.71%	109.09%
Age 45 to 64 (base: ≥ 65 years)	-1.22%	12.12%	--	0.91%
Annual income < $25,000 (base: > $75,000)	6.62%	-10.67%	-20.00%	-11.45%
Annual income $25-35,000 (base: > $75,000)	3.09%	1.33%	-14.12%	-6.25%
Annual income $35-75,000 (base: > $75,000)	2.94%	-12.00%	-12.94%	--
Worker (base: non-worker)	-4.23%	20.31%	18.06%	6.67%
Kids under 5 years old (base: no kids)	2.17%	-6.62%	1.41%	2.31%
Kids 5-17 years old (base: no kids)	3.04%	-7.94%	2.09%	3.30%
Experienced carsharing (base: never)	4.29%	--	-40.96%	--
Experienced ride-hailing (base: never)	-9.86%	92.31%	-17.07%	18.75%
High density household census block (base: <3,000 hh/mi^2)	-5.59%	44.86%	--	-5.96%

(--) coefficient was not different from zero at the 90% level of confidence and was removed from the model.

5.7 Conclusions

It is difficult to account for the potential impacts of AV technologies on transportation without an adequate understanding of how these vehicles might be adopted and used in the marketplace. There have undoubtedly been a few attempts to model the impacts of AVs on travel demand and transportation network performance, but these scenario tests often make exogenous assumptions about the level of penetration of AVs in the market, thus rendering the forecasts largely driven by speculative assumptions about how these vehicles will be adopted. There is very little research on consumer preferences for and potential adoption and use of AV technologies. This chapter aims to contribute to this critical gap through a systematic modeling effort aimed at unraveling relationships underlying this behavioral phenomenon.

To better understand the level of interest of consumers in AV ownership and/or AV sharing, we utilize travel survey data from the Puget Sound Region Travel Study to estimate a model that is capable of reflecting the bundle of mobility choices that people make simultaneously. Variables representing attitudes towards the built environment and technology use are used to construct two lifestyle factors, namely, green lifestyle propensity and technology-savviness. These latent lifestyle constructs are explicitly incorporated in models of current mobility choices and future intended use of AVs.

The model system presented in this chapter identifies the market segments that are likely to be early (or late) adopters and the users inclined to sharing rather than ownership of AVs. Through this understanding, public and private entities can target specific information campaigns or policy interventions to bring about more socially and environmentally desirable outcomes. It is important for public agencies to identify users inclined to adopt different AV ownership and sharing paradigms because the impacts of AV technology on the transportation system are likely to be very different depending on the AV usage paradigm that prevails in the market. For instance, AV private ownership may lead to a larger increase in empty-vehicle-miles traveled because the vehicles may drop users and seek inexpensive parking in peripheral areas or go serve other household members in different parts of the city. In addition, being able to spend time in the comfort of one's own AV while making a trip may drastically reduce AV users value of travel time. Significant reductions in value of travel time could negate network efficiency gains brought about by AV platooning and even lead to an increase in congestion. On the other hand, a greater adoption of the SAV model may help reduce empty-vehicle-miles and parking space requirements, while providing the ability to vary fares and avoid drastic reductions in value of travel time that could contribute to an increase in vehicle miles of travel (see discussion in Chapter 1).

This analysis provides important insights for planners and modelers regarding the current use of shared mobility services and future AV adoption preferences. First, the results indicate that individuals with green lifestyle preferences and who are tech-savvy are more likely to adopt car-sharing services, use ride-sourcing services, and embrace SAV in the future. Further, the importance of considering these latent lifestyle constructs is clear from the rejection of the IHDM model relative to the GHDM model. Second, notwithstanding the need for more research on psychological motivations and factors to target those who may be positively disposed toward specific new mobility technologies and services, the results from this research effort show that younger, urban residents with a high level of education are more likely to be early adopters of AV technologies, with a greater proclivity towards the use of vehicle-sharing services, *after controlling for lifestyle preferences*. Third, individuals who currently eschew vehicle ownership, and have already experienced car-sharing or ride-hailing services, are especially likely to be early adopters of SAV services. On the other hand, individuals who currently own vehicles, and have not yet experienced car-sharing services, are more inclined to adopt AV technologies in an ownership or combined ownership and sharing mode. While ignoring lifestyle preferences would exaggerate the impacts of current vehicle ownership and current mobility choices on future AV adoption, the results clearly show impacts of current mobility choices even after controlling for self-selection. Fourth, the elasticity effects in Table 5-3 indicate that perhaps the most effective way to move AV adoption toward a sharing model and elicit MaaS systems (rather than an ownership model) is to enhance neighborhood densification. The fact that this effect prevails even after any residential self-selection effect brought on by the green lifestyle propensity (that increases the likelihood of locating in dense neighborhoods and adopting AV-sharing in the future) is very significant. It motivates the consideration of neo-urbanist land-use policies in an entirely new light relative to the traditional focus of such policies as a potential way to solely reduce motorized private car travel. This is especially so because, separate from a direct neighborhood effect, densification increases AV sharing adoption propensity through a reduction in vehicle ownership. Fifth, and related to the first point, green lifestyle is an important determinant of high density living and is associated with walking and public transit use, while also directly and indirectly (through high density living) influencing adoption of AV sharing. This suggests that a goal of increased AV sharing may be advanced through campaigns that increase awareness of the benefits of green living (especially targeted towards demographic groups who are traditionally not "green").

A larger issue to examine in the context of AV adoption in general, and SAV in particular, is whether these new mobility options will reduce bicycling and walking, and the use of public transportation (PT) services. Those who are "green" and those who reside in high density residential neighborhoods today are the very individuals most likely to currently use non-motorized and PT services. These individuals are also most likely to embrace SAV. It may be conjectured then that SAVs will take modal share away from walking, bicycling, and PT. As a result, the overall purpose of developing MaaS systems should be affected and VMT or GHG reductions may not be realized (through SAV) as expected. Reduced walking and bicycling due to increased adoption of SAV services may also have adverse public health implications.

This research effort not only provides important insights into future AV adoption, but also presents a model component that can be implemented within an agent-based microsimulation model system to predict adoption of AV technologies in the future. By considering latent (and stochastic) psychological constructs, it provides "true" estimates of the effects of current residential and mobility choices on future AV-related choices. Combined with the structural equation system that "connects" the latent constructs to observed demographic variables, the future AV adoption component of the joint model system provides a platform to forecast AV impacts under alternative future scenarios.

Future research efforts should strive to address the data limitations of this study. In this research effort, the intended AV use is derived from survey questions in which respondents express their level of interest in owning/using such technology in the future. The survey does not constitute a full-fledged stated choice experiment in which respondents are provided detailed descriptions of various AV options and attributes, pricing levels, and any incentives for owning or sharing AVs. A fruitful direction for future research involves an application of the modeling framework of this study to stated choice data to gain further insights into user preferences for adoption/use of AV technologies.

CHAPTER 6. Modeling Individuals' Willingness to Share Trips with Strangers in an Autonomous Vehicle Future

6.1 Introduction

From a supply perspective, dynamic ridesharing and micro-transit are receiving significant attention from researchers (for some recent studies, see Frei et al., 2017, Levin et al., 2017, and Wang et al., 2018). These and related simulation-based studies have explored future scenarios where autonomous vehicles (AVs) are available and ride services are provided by TNCs operating shared autonomous vehicles (SAVs) fleets. The studies suggest that dynamic ridesharing through SAVs (also known as PSAV) has good potential to quite substantially reduce overall VMT relative to the case of privately owned AVs, and also that additional travel times due to pick-up and drop-off of multiple passengers could be compensated by reductions in congestion if shared rides are massively adopted by users.

Although simulation-based studies are optimistic about the potential for dynamic ridesharing systems, the performance of these services in terms of matching users, reducing pick-up waiting times, and increasing vehicle occupancy is directly dependent on public acceptance and consequent penetration rates. Unfortunately, historical data shows that sharing rides (in all different forms) has not been popular among U.S. travelers (Chan and Shaheen, 2012). Scheduling constraints have admittedly been an important barrier to the acceptance of traditional carpooling, since trips had to be identified a priori and both drivers and passengers had relatively little flexibility to make last minute changes in travel plans (Chan and Shaheen, 2012). While this reduced flexibility of carpooling has been solved by real-time scheduling and ride-hailing features, users still need to accept the potentially longer travel times of a shared ride due to pick-up/drop-off of additional passengers. In addition, another apparent obstacle to the expansion of dynamic ridesharing is the users' willingness-to-share rides with strangers. Recent studies indicate that travelers are hesitant about being in an automobile environment with unfamiliar faces due to distrust, security and privacy concerns (see, for example, Morales Sarriera et al., 2017 and Amirkiaee and Evengelopoulos, 2018).

In this context, future planning towards SAVs and MaaS systems in U.S. cities and studies examining the potential impacts of dynamic ridesharing on transportation networks could benefit from a deeper understanding of behavioral aspects associated with the acceptance of shared rides by travelers. Specifically, understanding psychosocial and financial trade-offs associated with preferences toward fare discounts, travel times, and presence of strangers in the vehicle can help identify segments of the population that are more (and less) prone to adopting dynamic ridesharing. To address this need, the current study develops the notion of willingness to share (WTS), which represents the money value (willingness to pay or WTP) attributed by an individual to traveling alone (i.e., to not share) compared to riding with strangers. Individuals' WTS is examined together with their values of travel time (VTT), enabling a comparison between people's sensitivities to delays (associated with serving multiple passengers) and their concerns about being in a car with strangers.

To investigate WTS and VTT, we develop a joint model of current ride-hailing experience and future intentions regarding the use of driver-less SAV services for commute and leisure trip purposes. Current ride-hailing experience is represented as a nominal dependent variable with three categories: (1) no experience with ride-hailing services, (2) experience only with private services (the individual traveled alone or with people s/he knew), and (3) experience with private and pooled services (the individual has, at least once, traveled with strangers for a cheaper fare). The future intention outcomes are represented as two binary outcomes corresponding to the choices between: (1) shared-ride and solo-ride in a SAV for a commute trip, and (2) shared-ride and solo-ride in a SAV for a leisure trip (both stated choice outcomes have three repeated choice occasions). The three outcomes (current ride-hailing experience and the two future SAV use choices) are jointly modeled as functions of unobserved psycho-social stochastic latent constructs, and observed transportation-related choices and sociodemographic variables. The current level of ride-hailing experience is assumed to affect the future choices of riding solo or sharing rides, which enables the evaluation of how current exposure to shared (or solo) rides may affect individuals' future intentions. The joint approach allows for the underpinning of the true effect of the current experience since we are able to control for common unobserved factors underlying all choice dimensions through the stochastic latent constructs. The modeling methodology is a special case of Bhat's (2015a) Generalized Heterogeneous Data Model, where the outcomes include one nominal outcome and two binary outcomes. However, unlike earlier implementations of the GHDM, we have a combination of one cross-sectionally observed variable (this is the nominal variable corresponding to current ride-hailing experience) and two variables with repeated choice observations (these correspond to the future intention outcomes).

Three stochastic psychological latent constructs representing *privacy-sensitivity, time-sensitivity,* and *interest in productive use of travel time (IPTT)* are modeled as functions of socio-demographic characteristics and used to create dependency among the nominal outcome and binary outcomes, and across the multiple choice-occasions. Additionally, the stochastic latent constructs are interacted with two attributes of the stated choice alternatives (time and number of additional passengers) to accommodate individual heterogeneity in VTT and WTS.

The data used is drawn from an online survey, developed and administered by the authors in the fall of 2017, of 1,607 commuters in the Dallas-Fort Worth-Arlington Metropolitan Area (DFW) of Texas, U.S. DFW is the largest metropolitan area in Texas in terms of population and the fourth largest in the U.S. It has more than 7.4 million inhabitants and is the fastest growing metropolitan area in the country (U.S. Census Bureau, 2018a). DFW is a car-dominated urban area where more than 81% of commute trips are undertaken using the drive alone mode and another 10% are pursued by a private car even if not alone. The current drive alone-dominated modal split and limited transit infrastructure in the DFW area makes it suitable as a potentially good location for the use dynamic ridesharing as a core component to facilitate the development of a MaaS system.

The remainder of this paper is organized as follows. The next section provides a detailed description of the survey, stated choice experiment, and sample used in the study. Next, in Section 3, we introduce the conceptual and analytic framework, including the procedure to compute VTT and WTS. Section 4 presents the results of the model, while Section 5 discusses policy implications. Conclusions and future research recommendations are provided in the final section.

6.2 Data

The data used for the analysis was obtained through a web-based survey. The distribution was achieved through mailing lists held by multiple entities (local transportation planning organizations, universities, private transportation sector companies, non-profit organizations, and online social media). To focus on individuals with commute travel, the survey was confined to individuals who had their primary work place outside their homes. The final sample used in the current paper includes information on 1,607 respondents.

To obtain information on the respondents' experience with ride-hailing services, the survey first provided definitions of both ride-hailing ("Ride-hailing services use websites and mobile apps to pair passengers with drivers who provide passengers with transportation in the driver's non-commercial vehicle; Examples are Uber and Lyft."), and pooled ride-hailing services ("In the carpooling option of ride-hailing, additional passengers with similar routes get picked and dropped off in the middle of the customer's ride; Customers receive discounted rates when they choose this option"). Then, before the stated choice experiments, respondents were presented with the definition of autonomous vehicles, as "Self-driving vehicles, also known as *autonomous cars* or *driverless cars*, are capable of responding to the environment and navigating without a human driver controlling the vehicle. In the following questions, whenever you read the term *self-driving vehicle*, imagine a car with no steering wheel that operates like a personal chauffeur". Respondents also were provided the option to watch a 90-second educational animation video about how AV-technology works and how the user experience might be.

Considering the uncertainties associated with the AV future, the stated choice experiment design focused on simple scenarios that would allow the simultaneous investigation of VTT and WTS without imposing too many assumptions about changes in urban mobility. Respondents were presented with situations with only binary alternatives, and both alternatives involving the use of an SAV (corresponding to traveling in an SAV alone or with strangers in a PSAV). Five trip attributes characterized each scenario, and were varied across scenarios: (1) travel time (which was associated with a specific distance for fare calculation purposes), (2) fare structure, (3) reduced cost amount for sharing, (4) additional travel time associated with sharing, and (5) the number of additional passengers. All the attributes and their respective levels are presented at the top of Figure 6-1.The levels for the travel time attributes (the first and the fourth attributes above) were defined with the objective of keeping the scenarios realistic, while also providing an instrument to engender adequate time variability in the attribute values across scenarios. For the second attribute, fare structure, a three-level scheme was used. The first level assumed that there would be no change in the non-pooled fare structure compared to today (this fare structure was based on Uber's non-pooled distance-based and time-based fare structure at the survey time; see UberEstimator, 2017). The other two levels (reflecting an autonomous vehicle future) assumed that service fees would no longer be necessary (because of the absence of human drivers) and that there would be a certain percentage reduction in the distance-based fare (relative to the current Uber fare structure). For the third attribute, corresponding to the reduced cost due to sharing, no specific source of information about current TNC procedures was readily available, but the anecdotal experience of several students at the University of Texas suggested significant variability. Hence, three levels corresponding to 20%, 40%, and 60% reduction (relative to the solo-SAV rate) were used in the stated choice experiments. The number of additional passengers was defined considering that standard autonomous cars would accommodate comfortably up to four passengers (similar to today's passenger vehicles, leading to three levels for this attribute, corresponding to one, two, and three additional passengers). In all, there were 243 (5 attributes corresponding to the five columns in Figure 6-1 and 3 levels corresponding to the three rows of Figure 6-1, for a total of $3^5 = 243$) possible combinations between the attribute levels. From these combinations, 27 different scenarios were chosen with the focus on isolating main effects and keeping orthogonality. As illustrated at the bottom of Figure 6-1, the respondent was presented with two alternatives and the information available for each alternative was the total travel time, cost, and, in the case of shared rides, the additional number of passengers. In other words, the discount rates and additional travel times due to pooling were not explicitly shown, but incorporated in the travel time and cost of the shared alternative. Each individual responded to six scenarios evenly split between commute and leisure trip purposes.

The survey also collected socio-demographic and attitudinal data from the respondents. Table 6-1 presents descriptive statistics of the socio-demographic characteristics of the sample (a discussion of the attitudinal information collected, and the corresponding descriptive statistics, is deferred until Section 6.3.1)[12]. A comparison of our sample with the employed population of DFW (as characterized by the U.S. Census Bureau, 2018b) indicates that the sample has an overrepresentation of men (58.4% in the survey compared to 54.0% from the Census data), individuals between 45 and 64 years of age (53.2% compared to 35.8%), Non-Hispanic Whites (75.0% compared to 51.0%), and individuals with bachelor's or post-graduate degrees (75.6% compared to 33.7%). We also observe that the majority of the sample corresponds to full time-employees (81.6%).

[12] Note that the sample used in this analysis is the same used in Chapter 4. To improve readability, we repeat some information and discussion presented earlier.

Experimental Design Attributes and Levels				
Solo option		***Shared option***		
Fare structure	**Travel time**	**Discount**	**Additional travel time**	**Additional passengers**
Base fare: $1 Cost per minute: $0.1 Cost per mile: $0.91 Service fee: $2.45	10 minutes	20%	4 minutes	1
Base fare: $1 Cost per minute: $0.1 Cost per mile: $0.70 Service fee: $-	15 minutes	40%	8 minutes	2
Base fare: $1 Cost per minute: $0.1 Cost per mile: $0.40 Service fee: $-	20 minutes	60%	10 minutes	3

Scenario Example	
Imagine that ride-sourcing services (similar to Uber and Lyft) use self-driving vehicles for all of their clients. Imagine also that you plan to go out on a leisure activity and you will use one of these ride-sourcing services. In the three scenarios described below, which option would you choose?	
Option 1	***Option 2***
Call a private self-driving cab service (similar to Uber/Lyft)	Call a shared self-driving cab service (similar to UberPool/LyftLine)
Travel time: 15 min Cost: $16.5 No additional passenger	Travel time: 23 min Cost: $10.0 Additional passengers: 1

Figure 6-1 Stated Choice Experiment Design Components and Scenario Example

Finally, among the socio-demographic characteristics, we are unable to compare the statistics from our survey with the Census data for the household income and household composition variables, because the Census data provides income and household composition data only for all households (while our survey is focused on households with at least one worker with a primary workplace outside home). However, the sample statistics do suggest a skew toward individuals from higher income households and multi-worker households. Overall, there are many possible reasons for the socio-demographic differences between our sample and the Census data. For example, the main topic of the survey was self-driving vehicles, which may be of more interest to highly educated men. In addition, the survey was conducted strictly through an online platform and the largest mailing list used in the distribution was of toll-road users, who are likely to be individuals with higher values of time that then correlates with the specific characteristics of our sample. In any case, while the general descriptive statistics of the dependent variables of interest cannot be generalized to the DFW population, the individual level models developed in this paper still provide important insights on the relationship between travel behavior and socio-demographic/lifestyle characteristics.

In addition to socio-demographics, we also use a set of three long and medium-term transportation-related variables as exogenous variables: residential location (characterized by urban versus non-urban living), vehicle availability (whether the number of motorized vehicles in the household was less than, equal to, or greater than the number of workers), and commute mode choice (traveling to work by driving alone, non-solo car, or non-car modes). While it can be reasoned that these transportation-related variables are influenced by common unobserved factors affecting the main outcomes, we tested this issue in our model specifications by considering these three variables also as endogenous variables. These three transportation-related variables were not significantly impacted by the latent constructs (at any reasonable statistical level) and, therefore, are treated as exogenous. There are many possible reasons for this result, from lack of variability in the actual variable (for example, only 3.5% of the sample does not drive to work) to inadequacy in the ability of latent variables to explain medium and long-term transportation-related choices (the latent variables, and therefore their indicators, used in this study are directed toward capturing trip-related attitudes in the context of an uncertain future transportation landscape, as discussed in more detail in Section 6.3.1; long and medium-term choices, on the other hand, are usually associated with overall lifestyles, such as a green-lifestyle or a luxury-orientation, as observed by Bhat, 2015b and in Chapter 5). The descriptive statistics of the three transportation-related variables are provided toward the bottom of Table 6-1, and reveal a sample with more than three-fourth of the respondents living in non-urban areas, more than 50% owning motorized vehicles equal to the number of workers in the respondent's household, and a predominance of the drive alone mode to commute to work.

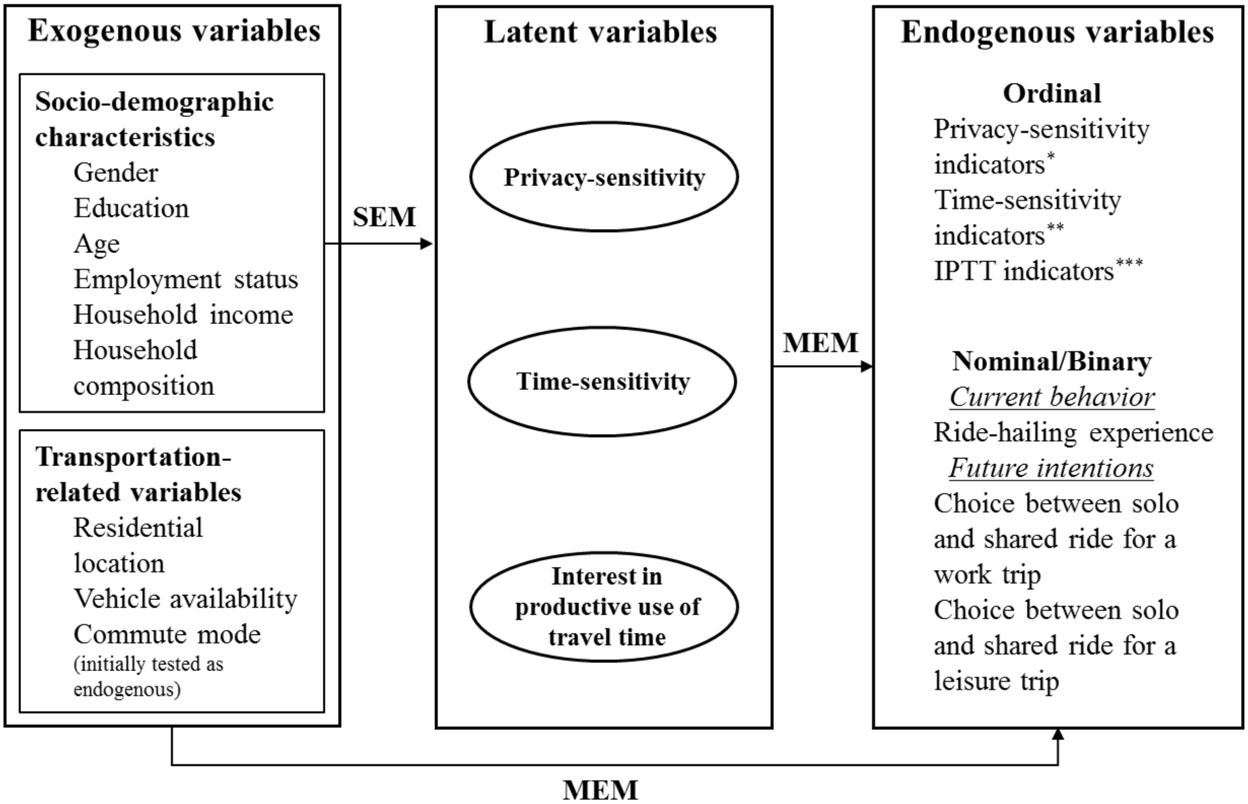
Table 6-1 Sample Socio-Demographic Characteristics and Transportation Related Exogenous Variables

Variable	Count	%
Gender		
Female	668	41.57
Male	939	58.43
Age		
18 to 34	261	16.24
35 to 44	360	22.4
45 to 54	432	26.88
55 to 64	423	26.32
65 or more	131	8.16
Race/ethnicity		
Non-Hispanic White	1205	74.98
Non-Hispanic Black	102	6.35
Hispanic	109	6.78
Asian/Pacific Islander	101	6.29
Other	90	5.60
Education		
Completed high-school	238	14.82
Completed technical school/associates degree	154	9.58
Completed undergraduate degree	724	45.05
Completed graduate degree	491	30.55
Employment type		
Full-time employee	1312	81.64
Part-time employee	138	8.59
Self-employed	157	9.77
Household income		
Under $49,999	184	11.45
$50,000-$99,999	443	27.57
$100,000-$149,999	496	30.86
$150,000-$199,999	269	16.74
$200,000 or more	215	13.38
Household composition		
Single person household	191	11.89
Single worker multi-person household	265	16.49
Multi-worker household	1151	71.62
Residential location		
Suburban, rural or small town	1232	76.67
Urban (downtown or central area)	375	23.33
Vehicle availability		
< 1 per worker	236	14.69
= 1 per worker	817	50.84
> 1 per worker	554	34.47
Commute mode		
Non-car	56	3.48
Car non-solo	146	9.09
Drive alone	1405	87.43

A note on data-related issues before moving to the description of the analytic framework. First, as mentioned earlier, the survey is not representative of the population of employed individuals in DFW and is skewed toward high-income individuals, which may result in inflated VTT and WTS. Second, it is well documented in the literature that stated choice data should be anchored to actual revealed choice values to reduce hypothetical bias and increase the external validity of WTP values (Hensher, 2010). The situation investigated in this study did not have a plausible revealed choice analogous, so WTP is not ‘calibrated’ by observed choices. Instead, to avoid drawing conclusions directly about actual VTT and WTS values, we direct our analysis toward relative comparisons between these two values for different segments of the population. Finally, while VTT may change from the current case of human-driven vehicles to the situation when individuals are no longer required to drive because of a number of reasons (see Cyganski et al. 2015, Krueger et al., 2016, and Das et al., 2017), we confine our attention in this study on VTT effects associated with being interested in using travel productively, as discussed next.

6.3 Analytic Framework

Figure 2 provides the conceptual structure for our joint model of ride-hailing experience and stated choice of SAV service for work and leisure trip purposes. Exogenous socio-demographic and transportation-related characteristics (left-side box in Figure 6-2), and three endogenous stochastic latent constructs representing psycho-social characteristics of the individual (middle box of Figure 2) are used as determinants of the three endogenous variables of interest (ride-hailing experience, and the choices between solo and shared SAV rides for work and leisure trip purposes). Together with these three endogenous outcomes (shown under the label “Nominal/Binary” in the right box of Figure 6-2), seven attitudinal indicators (representing indicators of privacy-sensitivity, time-sensitivity, and IPTT) help to characterize the three stochastic latent psycho-social constructs. The latent constructs create the dependency structure among all outcomes. A discussion of these latent constructs follows.



"*"I1: I don't mind sharing a ride with strangers if it reduces my costs.
I2: Having privacy is important to me when I make a trip.
I3: I feel uncomfortable sitting close to strangers.
"**"I4: Even if I can use my travel time productively, I still expect to reach my destination as fast as possible.
I5: With my schedule, minimizing time traveling is very important to me.
"***"I6: Self-driving vehicles are appealing because they will allow me to use my travel time more effectively.
I7: I would not mind having a longer commute if I could use my commute time productively.

Figure 6-2 Model Structure

6.3.1 Psychosocial Latent Constructs

Three psychosocial latent constructs are considered in our framework: privacy-sensitivity, time-sensitivity, and interest in productive use of travel time (IPTT). These are identified based on earlier studies in transportation and behavioral psychology, and focus on capturing underlying unobserved behavioral aspects that may influence individual's valuation of shared ride attributes. The first latent construct, privacy-sensitivity (characterized by the three attitudinal indicators identified under "*" at the bottom of Figure 6-2 and labeled as I1-I3 in Figure 6-3), represents individuals' levels of discomfort and privacy concerns when sharing a vehicle with a stranger. Previous studies have identified that the desire for personal space, aversion to social situations, distrust, and concerns about security are the most relevant behavioral barriers to ridesharing and carpooling services/programs that involve matching between strangers (for example, see

Tahmasseby et al., 2016, Morales Sarriera et al., 2017, and Amirkiaee and Evangelopoulos, 2018). Such factors have also been found to be relevant in studies on public transit use (Haustein, 2012 and Spears et al., 2013). Hence, the privacy-sensitivity latent construct is a key element in our model and is hypothesized to have negative impacts on individuals' experience with pooled ride-hailing and choice for shared rides in a SAV context. Additionally, we expect its negative effects to increase with the number of additional passengers (this is a case of the latent variable moderating the effect of an exogenous variable).

The second latent construct is time-sensitivity (see under "**" in Figure 6-2 and the indicators I4 and I5 of this latent construct in Figure 6-3). The objective of this construct is to capture people's perceptions of time scarcity and desire in reducing travel time. It is often assumed in transportation studies that an individual's goal is to minimize time traveling. However, as discussed by previous authors (see, for example, Ory and Mokhtarian, 2005), the extent to which traveling is perceived as a disutility may vary among individuals and trip purposes, depending on lifestyle and lifecycle factors and associated activity-scheduling constraints. This latent construct is introduced in the model both as a direct effect on the endogenous variables as well as a moderating effect of the influence of travel time, thereby engendering both observed and unobserved individual heterogeneity in the valuation of travel time.

The final latent construct, interest in the productive use of travel time (IPTT), identified under "***" in Figure 6-2 and labeled by indicators I6 and I7 in Figure 3, originates in the notion that the ability to use travel time productively may reduce perceived disutilities associated with traveling. This negative effect of time productivity on travel time disutility has been confirmed in the context of rail travel (Gripsrud and Hjorthol, 2012, Frei et al., 2015), and is likely to be relevant in the approaching AV future, as individuals may no longer need to drive and pay attention to traffic (Cyganski et al. 2015, Malokin et al., 2017). This latent construct too is introduced in the model both as a direct effect on the endogenous variables as well as a moderator of travel time effects on the endogenous variables.

All the latent construct indicators are measured on a five-point Likert scale and are modeled as ordinal variables. As may be observed from Figure 3, the sample shows a general tendency toward being privacy-sensitive, time-sensitive, and interested in the productive use of travel time.

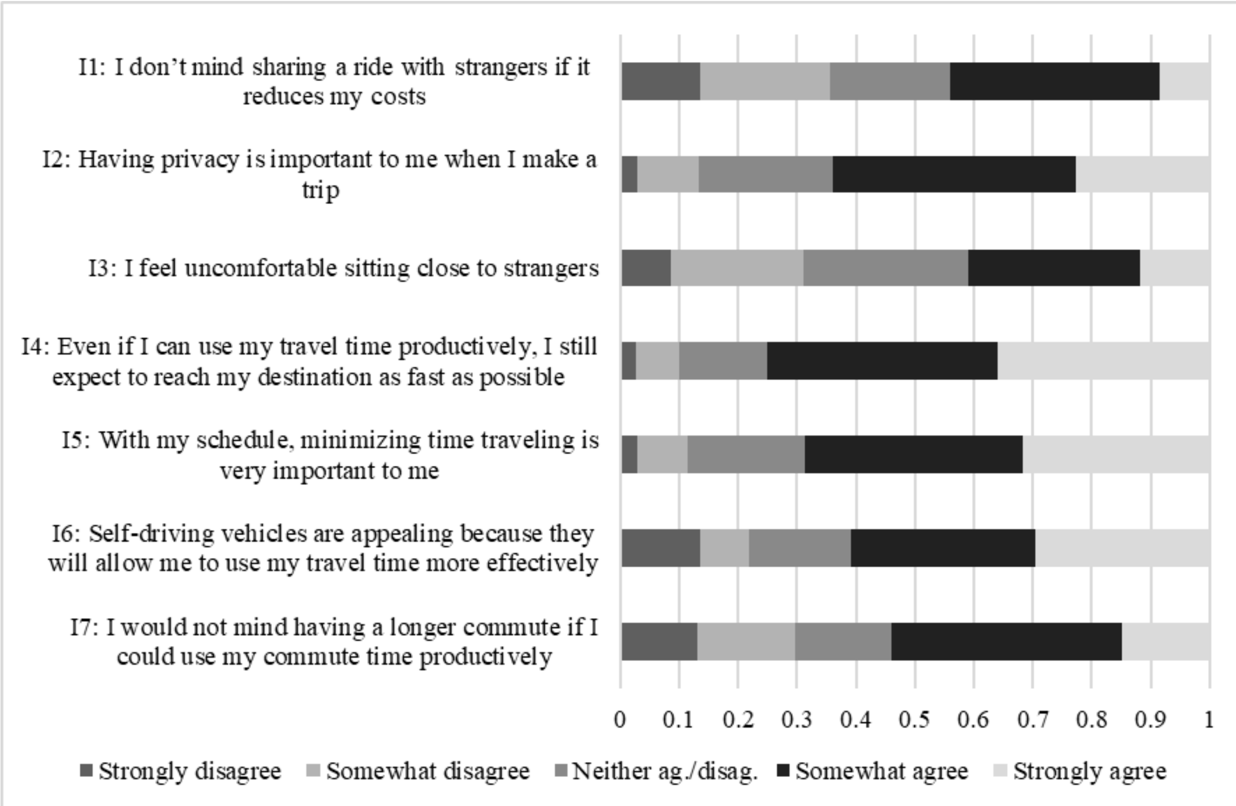


Figure 6-3 Sample Distribution of Attitudinal and Behavioral Indicators

6.3.2 Main Outcome Variables

As previously discussed, there are three main discrete choice outcomes in our model associated with individuals' ride-hailing experience (multinomial choice) and the stated choices of SAV service for work and leisure trip purposes (two binary choices). In terms of ride-hailing experience, about 56.4% of the sample (n=906) reported using ride-hailing services at least once in their lifetimes, although only about 10.0% of the sample (n=157) reported experience with the pooled version of the service. Accordingly, ride-hailing experience is represented in the three nominal categories of *no experience* (43.6%; n=701), *experience with private rides only* (46.6%; n=906-157=749), and *experience with pooled rides* (9.8%; n=157; note that this group may have had experience with private rides too). In terms of stated choices for SAV services (n=4821=1607 individuals × 3 choice occasions per individual), we observe that different trip purposes may be associated with different preferences toward sharing. In 48.3% of the choice occasions associated with work trip scenarios, respondents chose to ride alone, while this fraction is higher for leisure trip scenarios, reaching 54.0%. The outcome representing current ride-hailing experience is assumed to impact the stated SAV-service so that we can evaluate how current experiences are shaping future intentions in terms of sharing, while simultaneously controlling for the latent constructs effects on all three choice dimensions.

6.3.3 Modeling Approach

The model employed in our analysis is a special case of the Generalized Heterogeneous Data Model discussed in Chapter 2. As explained earlier, unobserved psycho-social constructs serve as latent factors that provide a structure to the dependence among the many endogenous variables, while the constructs themselves are explained by exogenous variables and may be correlated with one another in a structural relationship. In this approach, attitudinal indicators are treated as ordinal variables, while the main choice outcomes are nominal or binary. The presence of the stochastic latent variables captures not only the covariances between the attitudinal indicators, but also (a) among the indicators and the observed behaviors of interest as well as (b) between pairs of the observed endogenous variables of interest. Such an approach enables controlling for self-selection effects in the impact of current ride-hailing choice behavior on future intentions in an econometrically consistent fashion. Additionally, the stochastic latent factors serve as a parsimonious approach to incorporating observed and unobserved individual heterogeneity in variables of interest, which is done by interacting the latent factors with exogenous variables. As already indicated, in our application, we interact privacy-sensitivity with the number of additional passengers (strangers) in the shared ride alternatives, and both time-related latent variables with the travel time attribute.

There are two components to the GHDM model: (1) the latent variable structural equation model (SEM), and (2) the latent variable measurement equation model (MEM). As illustrated in Figure 2, the SEM component defines latent variables as functions of exogeneous variables. In the MEM component, the endogeneous variables are described as functions of both latent variables and exogeneous variables. The error terms of the structural equations (which define the latent variables) permeate into the measurement equations (which describe the outcome variables), creating a parsimonious dependence structure among all endogenous variables. These error terms are assumed to be drawn from multivariate normal distributions (with the dimension equivalent to the number of latent variables). The measurement equations have different characteristics depending on the type of dependent variable, following the usual ordered response formulation with standard normal error terms for the ordinal indicator variables, and the typical random utility-maximization model with a probit kernel for the nominal/binary outcomes of primary interest (see Bhat and Dubey, 2014, and Bhat, 2015a, for details of the formulation and estimation). The latent constructs are created at the individual level (as a stochastic function of individual demographics and transportation-related variables). These stochastic latent constructs influence the current ride-hailing experience endogenous variable in a cross-sectional setting (one revealed observation per individual from each of the 1607 respondents for n=1607) as well as each of the stated choice outcomes (one for commute travel and another for leisure travel) associated with the use of future SAV services in each of the three repeated choice occasions. Doing so immediately and parsimoniously captures not only unobserved factors impacting the indicator and endogenous outcomes of interest (as discussed earlier), but also accommodates covariations among the three choice occasions of the same individual. The resulting GHDM model is estimated using Bhat's (2011) MACML approach. To conserve on space, we do not provide the details of the estimation methodology, which is available in Bhat (2015a).

6.3.4 Value of Travel Time and Willingness to Share

Within the scope of discrete-choice models, WTP for travel attributes, including time (VTT), corresponds to the ratio of the estimated attribute and cost coefficients. Considering that WTP varies across the population, observed individual heterogeneity is addressed by interaction terms between attributes/cost and socio-demographic characteristics. Unobserved heterogeneity, on the other hand, is usually accommodated by specifying mixing distributions on the attribute coefficients and/or the cost coefficient, or by specifying mixing distributions on the actual WTP ratio coefficient (see Train and Weeks, 2005). A challenge associated with such approaches is that they are profligate in the number of parameters to be estimated. The current study deviates from the traditional WTP and VTT literature by adopting an alternative method to introduce individual heterogeneity in VTT and WTS. Instead of a mixing approach, we use stochastic latent variables as moderators of attributes in the choice utilities, thus capturing both observed and unobserved individual heterogeneity. In addition to a parsimonious structure, this method has the behavioral appeal of partitioning individual heterogeneity in VTT and WTS into specific psycho-social construct effects.

For each individual q, the computations of the expected values of VTT and WTS, and the corresponding variances, occur as follows:

$$E(VTT_q) = \frac{\beta_{TT_1} E(z^*_{TS_q}) + \beta_{TT_2} E(z^*_{IPTT_q}) + \beta_{TT_3}}{\beta_{COST}} \quad , \quad Var(VTT_q) = \frac{1}{\beta^2_{COST}} \left(\beta^2_{TT_1} + \beta^2_{TT_2} \right) \tag{1}$$

$$E(WTS_q) = \frac{\beta_{AP_1} E(z^*_{PS_q}) + \beta_{AP_2}}{\beta_{COST}} , \quad Var(WTS_q) = \frac{\beta^2_{AP_1}}{\beta^2_{COST}} \tag{2}$$

where β_{TT_1} is the coefficient on the interaction of the time-sensitivity latent construct $(z^*_{TS_q})$ and travel time, β_{TT_2} is the coefficient on the interaction of the interest in the productive use of travel time (IPTT) latent construct $(z^*_{IPTT_q})$ and travel time, β_{TT_3} is the coefficient on travel time, β_{AP_1} is the coefficient on the interaction of the privacy-sensitivity $(z^*_{PS_q})$ latent construct and the additional number of passengers (ADD) variable, β_{AP_2} is the coefficient on the ADD variable, and β_{COST} is the coefficient on trip cost. The expected values of the stochastic latent constructs are computed based on the SEM model results.[13]

6.4 Results

The final model specification was obtained based on a systematic process of testing alternative combinations of explanatory variables and eliminating statistically insignificant ones. Also, for continuous variables such as respondent age and respondent's household income, a number of functional forms were tested in the sub-models for each endogenous outcome variable, including a linear form, a dummy variable categorization, as well as piecewise spline forms. But the dummy variable specification turned up to provide the best data fit in all cases, and is the one adopted in the final model specification. Also, in the final model specification, some variables that were not statistically significant at a 95% confidence level were retained due to their intuitive interpretations and important empirical implications. In this regard, the methodology used involves the estimation of a large number of parameters, so the statistical insignificance of some coefficients may simply be a result of having only 1,607 respondents. Also, the effects from this analysis, even if not highly statistically significant, can inform specifications in future ride-hailing investigations with larger sample sizes.

[13] The variance formulas arise as given because the latent construct variances are normalized to one for identification in the estimation. Also, to keep the presentation simple, we do not consider the sampling variance of the estimated coefficients in the variance computation.

In the next section, we discuss the results of the SEM model component of the GHDM, as well as the latent constructs' loadings on the attitudinal indicators (which are one part of the MEM). In subsequent sections, we discuss the MEM relationships corresponding to the effects of socio-demographic and transportation-related characteristics, and the latent constructs, on the three main outcomes of interest.

6.4.1 Attitudinal Latent Constructs

The structural relationships between socio-demographic variables representing lifecycle stages and the latent constructs are presented in Table 6-2. Gender shows no significant effect on the individual's level of privacy-sensitivity and interest in the productive use of travel time (IPTT). Yet, women display higher levels of time sensitivity, which is expected considering that working women are more likely to experience time scarcity relative to men, attributable to lingering gender disparities in household-related activities, including childcare and chauffeuring activities (Fan, 2017, Motte-Baumvol et al., 2017). Younger adults display greater levels of privacy-sensitivity and IPTT. The latter effect is probably associated with higher levels of tech-savviness and ICT usage among younger adults, which facilitates the productive use of travel time (Astroza et al., 2017, Malokin et al., 2017). The first effect, on the other hand, seems less obvious and requires further investigation; however, it may also be related to higher levels of technology use, especially smartphones, by younger generations. There is growing evidence that the use of smartphones is creating a "portable-private bubble" phenomenon, which makes individuals more estranged from their surroundings and less interested in potential social interactions in public spaces (Hatuka and Toch, 2014). Along the same lines, higher smartphone usage also seems to be associated with higher social anxiety and lower social capital building (Bian and Leung, 2015, Kuss et al., 2018). We also observe that individuals between 35 and 44 years of age are more time-sensitive than their younger and older peers. This age range is associated with the beginning of the career peak cycle, and also increased responsibilities associated with raising children and looking after family elders (Nael and Hammer, 2017). Non-Hispanic White individuals tend to be more privacy-sensitive relative to other races/ethnicities, a result that aligns with the higher levels of drive-alone travel and vehicle ownership by this ethnic group (Giuliano, 2003, Klein et al., 2018). As expected, individuals who are more highly educated show greater interest in the productive use of travel time. Higher levels of education are associated with higher tech-savviness and ICT usage (Astroza et al., 2017), as well as greater opportunity to work outside the traditional work place (Singh et al., 2013), which can contribute to the ability to work and be productive while traveling. Being a part-time employee or self-employed is associated with lower time sensitivity, presumably because these employment arrangements provide greater time flexibility than full-time employment. Finally, individuals from households with very high incomes (above US$200,000 per year) show greater privacy and time-sensitivity, and are also more interested in using their travel time productively. The higher privacy-sensitivity among the wealthiest segment of individuals can be a direct result of having more access to private property and/or a need to signal exclusivity through separation and differentiation from others (Chevalier and Gutsatz, 2012, Bhat, 2015b). These individuals may also focus on privacy due to concerns associated with safety and preservation of material assets. High-income individuals also have stronger feelings of time pressure (DeVoe and Pfeffer, 2011, Chen et al., 2015), which are dictated by perceived opportunity costs, among other factors, such as increased occupation responsibilities. Such characteristics explain the positive impacts of income in the two time-related latent constructs.

All three correlations corresponding to the three pairs of latent variables are statistically significant (see Table 6-2), even if only medium-to-low in magnitude. Privacy-sensitivity is positively associated with time-sensitivity, and negatively related to IPTT. Time-sensitivity is also negatively associated with IPTT. The implication of these correlation results is that, when dealing with individuals who are intrinsically privacy and time-sensitive (due to unobserved personality characteristics), an environment that is conducive to the productive use of travel time will have little to no effect on increasing their tolerance to increased travel times and/or additional passengers.

The SEM estimation is made possible through the observations of the endogenous variables (far right block of Figure 6-3), which include the latent variable indicators and the three endogenous outcomes of interest. The loadings of the latent variables on their indicators are represented at the bottom of Table 6-2 and have the expected signs. Thresholds and constants associated with the ordinal response equations characterizing the indicators were also estimated but are omitted to conserve on space.

Table 6-2 Determinants of Latent Variables and Loadings on Indicators

Variables (base category)	**Structural Equations Model Component Results**					
	Privacy-sensitivity		Time-sensitivity		IPTT	
	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat
Gender (male)						
Female	--	--	0.183	4.27	--	--
Age (≥55 years)						
18 to 34	0.168	1.84	--	--	0.326	4.87
35 to 44	0.137	4.09	0.265	5.26	0.256	4.54
45 to 54	--	--	--	--	--	--
Race/ethnicity (other)						
Non-Hispanic White	0.131	3.76	--	--	--	--
Education (≤ undergraduate degree)						
Graduate degree	--	--	--	--	0.133	4.32
Employment (full-time)						
Part-time employee	--	--	-0.382	-4.71	--	--
Self-employed	--	--	-0.119	-1.97	--	--
Household income (< $150,000)						
$150,000-$199,999	--	--	--	--	0.092	2.84
$200,000 or more	0.350	5.16	0.298	4.26	0.092	2.84
Correlations between latent variables						
Privacy-sensitivity	1.000	n/a				
Time-sensitivity	0.241	7.59	1.000	n/a		
IPTT	-0.115	-2.67	-0.071	-2.71	1.000	n/a
Attitudinal Indicators	**Loadings of Latent Variables on Indicators (MEM component)**					
I don't mind sharing a ride with strangers if it reduces my costs (inverse scale)	0.847	13.98				
Having privacy is important to me when I make a trip	0.477	17.49				
I feel uncomfortable sitting close to strangers	0.347	3.16				
Even if I can use my travel time productively, I still expect to reach my destination as fast as possible			0.755	40.40		
With my schedule, minimizing time traveling is very important to me			1.329	57.60		
Self-driving vehicles are appealing because they will allow me to use my travel time more effectively					1.183	7.26
I would not mind having a longer commute if I could use my commute time productively					0.751	4.49

"--" = not statistically significantly different from zero at the 90% level of confidence and removed from the specification.
"n/a" = not applicable

6.4.2 Ride-Hailing Experience

The results of the ride-hailing experience model are presented in the first column of Table 6-3. The coefficients represent the effects of variables on the utilities of private only ride-hailing and shared (or pooled) ride-hailing, with the base alternative being the case of no ride-hailing experience.

The latent variable effects have the expected directionality of effects, with privacy-sensitive individuals less likely to have experience with pooled ride-hailing service and IPTT increasing the probability of both types of ride-hailing experience. This latter result suggests that interest in using travel time more productively is an important factor currently guiding ride-hailing adoption.

In addition to the indirect socio-demographic influences through the latent variable effects just discussed, there are direct socio-demographic effects on ride-hailing experience. Table 3 indicates that age has a direct negative effect on ride-hailing experience, with younger individuals more likely than their older counterparts to have used ride-hailing both in the private as well as pooled arrangements, which is consistent with some earlier studies (Smith, 2016, Kooti et al., 2017). Note that this direct negative age effect more than compensates for the average indirect positive age effects on experience with both private and pooled services through the privacy-sensitivity latent construct. Thus, for example, the average indirect age effect indicates that an individual 18-34 years of age (relative to a person 65 years of age or older) has a lower pooled ride-hailing utility valuation of the order of 0.168 (the coefficient on the "18 to 34 years" of age variable corresponding to privacy sensitivity in Table 6-2) times the average expected value of the privacy-sensitivity latent variable (0.246) multiplied by -0.131 (the magnitude of the coefficient on the privacy-sensitivity construct on pooled ride-hailing experience in Table 6-3) yielding an average indirect age effect between the "18 to 34 years" age group and the ">=65 years age group" of -0.005 (=0.168*0.246*(-0.131)). The corresponding direct age effect is 0.843, which swamps the indirect age effect, resulting in younger adults distinctly more likely to adopt the pooled form of ride-hailing compared to their older peers. In terms of the indirect age effects through the IPTT latent construct, these reinforce the negative direct age effects on experience with ride-hailing services (in both private only and pooled arrangements). Again, though, the direct age effect dominates over the indirect age effect through the IPTT latent construct (for example, the indirect age effect through the IPTT construct for the same two age groups as just discussed before is 0.326*0.184*0.151=0.009 for pooled service utility relative to no experience with ride-hailing compared to the corresponding direct effect of 0.843).

The results also show that non-Hispanic Whites are less likely to have used pooled services, even after accounting for the indirect negative effect (through the privacy-sensitivity construct) of being non-Hispanic White (relative to individuals of other race/ethnicity categories) and after controlling for income effects. The reason behind this race/ethnicity effect is not clear in the literature and calls for more qualitative studies investigating cultural influences on the willingness to share rides. However, on a related note, there is evidence that immigrants are more likely to carpool, especially if living in immigrant neighborhoods (Blumenberg and Smart, 2010). Similar to what was observed by Dias et al. (2017), part-time employees are less likely to have experienced private ride-hailing services relative to full-time employees and self-employed individuals.

In terms of household level variables, a higher household income increases experience with both private and pooled ride-hailing, beyond the positive effect of household income through IPTT (and while individuals with a household income over $200,000 have a higher privacy sensitivity, and privacy sensitivity negatively impacts pooled ride-hailing experience, this indirect negative effect gets swamped by the magnitude of the positive direct effect in Table 3; this may be observed by doing a similar computation as for the age effects discussed earlier). Considering that attitudinal and lifestyle factors are being controlled for, the direct positive income effect is probably an indicator of higher consumption power, though there is still a distinct preference for private ride-hailing over pooled ride-hailing in the higher income groups. As we will see later in Section 5.2, the magnitude of the coefficients on the household income variables on the private only and pooled ride-hailing utilities imply that an increase in household income tends to lead to a higher probability of private only ride-hailing experience, at the expense of drawing away from both the pooled ride-hailing and no ride-hailing experience categories. Individuals living alone are more likely to have used private ride-hailing service relative to individuals in other household types, while those in single-worker multi-person households are the least likely to have used both private and pooled services. Individuals living in more urbanized locations are more likely than their counterparts in less urbanized locations to have used both private and pooled ride-hailing. A similar result holds for individuals in households with more than one vehicle per worker. This latter suggests that, in an area such as DFW where almost all households own at least one vehicle, ride-hailing serves as more of a convenience feature for those one-off trips rather than being an accessibility facilitator for routine trips. Still, individuals who commute by non-car modes are more likely to have experience with both private and pooled ride-hailing.

6.4.3 Private versus Shared Rides for Work and Leisure Travel

The second and third columns of Table 6-3 present the estimated coefficients based on the stated choice between a solo ride and a shared ride for commuting scenarios and leisure trip-purpose scenarios, respectively. There is very limited literature in the context of SAVs to which we can compare our model results. This is because, although there have been multiple studies investigating individual intentions to adopt SAVs (see for example, Zmud et al., 2016, Haboucha et al., 2017, Lavieri et al., 2017), there is little research modeling the choice between riding solo in a SAV use and sharing a ride in a SAV use. The few studies on this topic have an exclusive focus on the investigation of VTT (see for example, Krueger et al., 2016). To our knowledge, there is no current study that models WTS.

As expected, privacy-sensitivity significantly reduces the likelihood of choosing to share a ride in an SAV. The other two latent variables do not show significant direct effects after accounting for their interaction with travel time attributes (as discussed later in this section). Women and young adults exhibit a lower tendency to choose shared rides in a commuting context, but gender and age do not show effects on the decision to share trips for leisure purposes. Women are usually responsible for most household chauffeuring and shopping activities, which are usually chained with into work commutes (Buddelmeyer et al., 2017; Fan, 2017; Motte-Baumvol et al., 2017). This may explain the lower tendency of women to choose the PSAV mode for the work trip. The negative inclination to use the PSAV commute mode among younger adults (relative to older adults) is intriguing, especially given that younger adults are distinctly more likely to use the pooled form of ride-hailing today (as discussed earlier). It is possible that, in today's ride-hailing setting with a human driver, millennials feel somewhat more comfortable traveling with strangers because they view the human driver as a professional "guardian" during their pooled commute trips, while these same individuals (relative to their older peers) are much more wary of sharing rides in SAVs without a "guardian" human driver. There are no statistically significant direct race/ethnicity effects in the stated choice models; yet, we observe indirect race/ethnicity effects (through privacy-sensitivity and ride-hailing experience) which indicate that Non-Hispanic Whites are less likely to opt for shared rides. Individuals with graduate degrees have lower interest in sharing rides to reach leisure activities, while self-employment, compared to part-time and full-time employment, reduces the interest in sharing commute trips.

In terms of household level variables, a higher household income decreases the propensity to choose the shared ride AV mode for both activity purposes, even after accounting for indirect effects through current ride-hailing experience and beyond the indirect effects through privacy-sensitivity. This result may be an indication of the higher consumption power and a desire for personalized SAV services among higher income individuals. Finally, in the set of demographic variables, individuals living in multi-worker households (compared to living alone or in a single-worker household) are more likely to share SAV rides for both activity purposes.

The transportation-related variables also reveal intriguing effects on the stated choices of SAV services. While living in urban areas (compared to living in the suburbs or rural areas) has a significant positive association with pooled ride-hailing experience, the opposite is observed in the SAV stated choice model. This result certainly needs further investigation in the future, though it may reflect the same perception of enhanced security (as for young individuals) with a human driver present (as opposed to not having an additional individual in the form of the human driver) when traveling with strangers in and around urban areas. Household vehicle availability seems to reduce the inclination toward sharing rides for commute purposes, while not affecting leisure trip-purposes. This effect corroborates the findings in Chapter 4 in the context of current pooled ride-hailing behavior in the DFW area. Next, the model shows that commuting with other individuals today reduces the interest in sharing SAV commute trips, but increases it for leisure trips. Indeed, sharing rides with strangers when already escorting family members or acquaintances may be perceived as a challenge. However, it is interesting to note that individuals who do not drive alone to work seem more open to sharing rides in situations that they would potentially be alone, such as trips to leisure activities.

Table 6-3 Results of the Ride-Hailing Experience and SAV Choice Model Components

Variables (base category)	Ride-hailing experience (base: none)				SAV: work purpose (base: solo)		SAV: leisure purpose (base: solo)	
	Private only		Pooled		Shared		Shared	
	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat
Latent variables								
Privacy-sensitivity	--	--	-0.131	-1.90	-1.348	-5.11	-1.251	-7.87
Time-sensitivity	--	--	--	--	--	--	--	--
IPTT	0.151	2.55	0.151	2.55	--	--	--	--
Socio-demographic variables								
Gender (male)								
Female	--	--	--	--	-0.174	-5.23	--	--
Age (≥65 years)								
18 to 34	0.978	9.19	0.843	11.61	-0.311	-1.84	--	--
35 to 44	0.699	7.10	0.564	8.83	-0.257	-3.15	--	--
45 to 54	0.321	4.09	0.336	5.46	--	--	--	--
55 to 64	0.158	2.38	--	--	--	--	--	--
Race/ethnicity (other)								
Non-Hispanic White	--	--	-0.205	-5.69	--	--	--	--
Education (≤ undergraduate degree)								
Graduate degree	--	--	--	--	--	--	-0.086	-3.67
Employment (full-time)								
Part-time employee	-0.277	-10.12	--	--	--	--	--	--
Self-employed	0.114	4.40	--	--	-0.232	-5.07	--	--
Household income (< $50,000)								
$50,000-$99,999	--	--	--	--	--	--	-0.132	-3.85
$100,000-$149,999	0.353	14.92	--	--	-0.396	-10.00	-0.692	-11.74
$150,000-$199,999	0.605	13.53	0.203	6.90	-0.396	-10.00	-0.692	-11.74
$200,000 or more	0.986	16.80	0.485	10.29	-0.396	-10.00	-0.692	-11.74
Household composition (multi-worker)								
Single person	0.362	14.50	--	--	-0.193	-4.55	--	--
Single worker multi-person	-0.171	-6.26	-0.241	-7.93	-0.435	-8.71	-0.279	-8.49
Transportation-related variables								
Residential location (rural/ suburban)								
Urban	0.363	21.64	0.413	16.35	-0.092	-2.86	-0.086	-3.43
Vehicle availability (< 1 per worker)								
= 1 per worker	--	--	--	--	-0.339	-7.58	--	--
> 1 per worker	0.059	3.79	0.144	4.06	-0.151	-3.53	--	--
Commute mode (drive alone)								
Car not-alone	-0.042	-2.00	0.053	2.04	-0.092	-2.22	0.086	2.69
Non-car	0.242	7.34	0.395	10.02	--	--	--	--
Ride-hailing experience (no)								
Private only	n/a	n/a	n/a	n/a	-0.173	-5.42	-0.420	-11.51
Pooled	n/a	n/a	n/a	n/a	-0.049	0.81	0.193	2.98
Trip attributes								
Cost [US$]	n/a	n/a	n/a	n/a	-0.294	-13.3	-0.263	-14.5
Travel time [minutes]	n/a	n/a	n/a	n/a	-0.141	-13.6	-0.102	-13.8
Additional passengers	n/a	n/a	n/a	n/a	-0.139	-8.6	-0.218	-10.0
Travel time*Time-sensitivity	n/a	n/a	n/a	n/a	-0.007	-2.0	-0.007	2.8
Travel time*IPTT	n/a	n/a	n/a	n/a	0.066	9.6	0.006	2.1
Additional passengers*Privacy-sensitivity	n/a	n/a	n/a	n/a	-0.017	-1.3	-0.073	-2.4
Constant	-0.884	-9.31	-1.214	-13.03	1.130	11.01	0.903	9.6

"--" = not statistically significantly different from zero at the 90% level of confidence and removed. "n/a" = not applicable

Finally, the endogenous variable representing ride-hailing experience also shows very interesting effects on the stated choice outcomes. Current experience with "private ride-hailing only" (relative to having no experience with ride-hailing at all or having pooled ride-hailing experience) has a negative effect on choosing to share AVs for both activity purposes. In other words, it appears that people who have used "private ride-hailing only" appreciate the convenience and flexibility of the private arrangement based on the actual experience, and are loath to sharing the travel experience with strangers (either with current pooled ride-hailing or with PSAVs in the future). Particularly intriguing here is the implication that it may be easier to "convert" individuals who have never used ride-hailing into future PSAV users than to attempt to convince current "private ride-hailing only" users to become future PSAV users. From this standpoint, part-time employees appear to be a promising demographic group to court for future PSAV travel, given, based on our ride-hailing model results of the previous section, that they are one of the most likely groups to have never experienced ride-hailing. The fraction of part-time employees is also quite significant in today's workforce, and this fraction is only projected to increase over time (Trading Economics, 2018). Perhaps understanding their needs better (such as other household responsibilities they may shoulder) can lead to the provision of pooled ride-hailing services today as well as future PSAV services that can assuage their concerns about these services meeting up to their needs. On the other hand, current pooled ride-hailing users appear to be the prime segment for promoting PSAV use, especially for trips for leisure purposes. However, it does appear from our results that PSAVs are not viewed in the same light as current pooled ride-hailing use by some population segments, such as young individuals and those residing in urban areas. If this is indeed because of the comfort/security of having a human "guardian" during the trip, then it becomes incumbent that AV design pay attention to security features, such as having an emergency "911-like" button accessible to each passenger. Also, it then suggests that AV security features be advertised particularly to young individuals, high income individuals, and urban area residents to allay their anxiety toward PSAV travel. In any case, our results call for a deeper investigation into attitudes and perceptions associated with having a human driver versus not having one in the context of pooled ride-hailing travel. Similarly, a better understanding of why non-Hispanic Whites, in particular, shy away from pooled ride-hailing travel today can be beneficial to bringing them to the "shared-ride" fold and potentially increasing the pool of individuals who may use PSAVs in the future. Further, any efforts to provide additional opportunities for, and promote the use of, pooled ride-hailing today appears will have positive pay-offs for the future use of PSAVs. That is, there may be merit to, for example, considering the provision of deep discounts for pooled ride-hailing today (or at least for a small window of time just before the large-scale advent of AVs) as a means to attract individuals to the use of pooled ride-hailing, even if these deep discounts may not be justifiable from an economic standpoint in the short-term.

In terms of trip attribute effects and interaction effects of trip attributes and latent constructs (see toward the bottom of Table 6-3), all the coefficients have the expected signs. In the specific context of the interaction effects, time-sensitive individuals place a higher premium on travel time for both the work and leisure purposes, individuals with high interest in the productive use of travel time have a lower sensitivity to travel time (particularly for the work purpose), and privacy-sensitive individuals have an increasing reluctance for PSAV travel as the number of passengers in the shared arrangement increases (this last effect is particularly so for leisure travel). However, it is also important to note that these interaction effects generally pale in comparison to the main effects. Thus, for example, the utility difference per minute between the individual in the sample with the highest expected value of the time sensitivity latent construct and the lowest expected value of the time sensitivity construct is 1.066 (this is computed based on the SEM model predictions; the range of the expected value of the time sensitivity construct is from -0.263 to 0.803), which translates to an expected travel time sensitivity difference between these two individuals of 0.007*1.006=0.0075. This difference is less than 6% of the main travel time effect of 0.141 for the work purpose and less than 8% of the main travel time effect of 0.102 for the leisure purpose. Similar computations reveal that (a) the travel time sensitivity difference between the two individuals with the minimum and maximum expected IPTT values is 22% of the main travel time effect for the work purpose, but less than 3% of the main travel time effect for the leisure purpose, and (b) the negative additional passenger utility effect on sharing between the two individuals with the minimum and maximum expected privacy sensitivity values is about 9% of the negative valuation of the main additional passenger utility effect for the work purpose and 24% of the main additional passenger utility effect for the leisure purpose. Overall, the strongest interaction effects correspond to travel time variations due to IPTT for the work purpose, and the (dis-)utility attributable to additional passengers based on the level of privacy sensitivity for the leisure purpose.

We also tested the interaction between privacy-sensitivity and PSAV travel time to examine if the presence of strangers increases the disutility of time traveling, but this effect was not statistically significant. Similarly, we also tested the interaction effect of additional passengers with travel time, but again this interaction effect was not statistically significant. That is, individuals seem to have a fixed dis-utility to having a stranger travel with them, which is independent of travel time.

6.4.4 Model fit evaluation

In this section, we present the data fit results of an independent heterogeneous data model (IHDM) model that excludes the latent psychological constructs and compare this IHDM model to the proposed GHDM model. The IHDM model essentially is a set of independent models (one for each outcome, including attitudinal indicators) and ignores the jointness in the outcomes (that is, the covariances engendered by the stochastic latent constructs are ignored). The IHDM model includes the exogenous determinants of the latent constructs directly as explanatory variables as well as considers all statistically significant demographic and transportation-related variables impacting the outcome variables in the GHDM model. The GHDM and the IHDM models are not nested, but they may be compared using the composite likelihood information criterion (CLIC)[14]. The model that provides a higher value of CLIC is preferred. Another way to examine the performance of the two models is to compute the equivalent GHDM predictive likelihood value for the three main outcomes (that is, for the current revealed preference ride-hailing experience nominal variable and the repeated stated binary choice observations of SAV use (or not) for the commute purpose and the leisure purpose). The corresponding IHDM predictive log-likelihood value may also be computed. Then, one can compute the adjusted likelihood ratio index of each model with respect to the log-likelihood with only the constants. To test the performance of the two models statistically, the non-nested adjusted likelihood ratio test may be used (see Ben-Akiva and Lerman, 1985, page 172). This test determines if the adjusted likelihood ratio (ALR) indices of two non-nested models are significantly different. In particular,

[14] The CLIC, introduced by Varin and Vidoni (2005), takes the following form (after replacing the composite marginal likelihood (CML) with the maximum approximate CML (MACML)): $\log L^{*}_{MACML}(\hat{\hat{\boldsymbol{\theta}}}) = \log L_{MACML}(\hat{\hat{\boldsymbol{\theta}}}) - tr\left[\hat{\boldsymbol{J}}(\hat{\hat{\boldsymbol{\theta}}})\hat{\boldsymbol{H}}(\hat{\hat{\boldsymbol{\theta}}})^{-1}\right]$.

the test determines the probability that the difference in the ALR indices could have occurred by chance in the asymptotic limit. A small value of the probability of chance occurrence indicates that the difference is statistically significant and that the model with the higher value of adjusted likelihood ratio index is to be preferred.

We also evaluate the data fit of the two models intuitively and informally at both the disaggregate and aggregate levels. To do so, we focus on the predictions for the 12 different combinations of ride-hailing experience (three alternatives), work purpose SAV use (two alternatives), and leisure purpose SAV use (two alternatives). We then compute multivariate predictions for these 12 (=3×2×2) combinations. At the disaggregate level, for the GHDM model, we estimate the probability of the observed multivariate outcome for each individual and compute an average (across individuals) probability of correct prediction at this three-variate level. Similar disaggregate measures are computed for the IHDM model. At the aggregate level, we design a heuristic diagnostic check of model fit by computing the predicted aggregate share of individuals in each of the 12 combination categories. The predicted shares from the GHDM and the IHDM models are compared to the actual shares, and the absolute percentage error (APE) statistic is computed.

The composite marginal likelihoods of the GHDM and IHDM models came out to be –52,4983.3 and –52,9193.4, respectively. Other measures of fit are provided in Table 6-4. The GHDM shows a better goodness-of-fit on the basis of the CLIC statistic, the predictive log-likelihood values and the predictive adjusted likelihood ratio indices. The same result is obtained from the non-nested likelihood ratio statistic; the probability that the adjusted likelihood ratio index difference between the GHDM and the IHDM models could have occurred by chance is literally zero. The average probability of correct prediction is 0.1740 for the GHDM model, and 0.1545 for the IHDM model. At the aggregate level, the shares predicted by the GHDM model are either superior to the IHDM model or about the same as the IHDM model for each of the 12 multivariate combinations. Across all the 12 combinations, the average APE is 10.69 for the GHDM model compared to 30.00 for the IHDM. The aggregate fit measures in Table 6-5 reinforce the disaggregate level results in Table 6-4. In summary, the results show that the GHDM model proposed here outperforms the IHDM model in data fit, providing support for our modeling of the revealed preference current ride-hailing experience choice and the stated choices of future SAV use as a joint package.

Table 6-4 Disaggregate Measures of Goodness-of-Fit

Summary Statistics	Model	
	GHDM	IHDM
Composite Marginal log-likelihood value at convergence	-524,196.0	-528,710.0
Composite Likelihood Information Criterion (CLIC)	-524,983.3	-529,193.4
Predictive log-likelihood at convergence	-9,847.68	-10,133.67
Constants only predictive log-likelihood at convergence	-11,220.60	
Number of parameters	120	87
Predictive adjusted likelihood ratio index	0.113	0.090
Non-nested adjusted likelihood ratio test between the GHDM and IHDM	Φ[21.75]<<0.0001	

Table 6-5 Aggregate Measures of Goodness-of-Fit

Multivariate Combination	**Sample**		**GHDM**		**IHDM**	
Ride-hailing experience, Leisure Purpose, Work Purpose	Count	Share (%)	Predicted Share (%)	APE (%)	Predicted Share (%)	APE (%)
No, Solo, Solo	675	14.00	14.69	4.93	8.60	38.55
No, Solo, Shared	343	7.11	7.22	1.50	9.92	39.46
No, Shared, Solo	294	6.10	6.59	8.06	9.86	61.69
No, Shared, Shared	791	16.41	16.01	2.40	14.69	10.49
Private, Solo, Solo	854	17.71	17.02	3.94	12.44	29.76
Private, Solo, Shared	528	10.95	11.04	0.80	12.21	11.52
Private, Shared, Solo	291	6.04	4.43	26.65	9.87	63.59
Private, Shared, Shared	574	11.91	14.43	21.21	11.91	0.01
Pooled, Solo, Solo	128	2.66	2.63	1.12	1.92	27.55
Pooled, Solo, Shared	78	1.62	1.17	27.95	2.16	33.35
Pooled, Shared, Solo	88	1.83	1.46	20.18	2.52	38.10
Pooled, Shared, Shared	177	3.67	3.32	9.59	3.89	5.89
Average APE				10.69		30.00
Average Probability of Correct Prediction			0.1740		0.1545	

6.5 Implications of Results

In this section, we examine the imputed values of travel time (VTT) and willingness to share (WTS) from our results, as well as discuss treatment effects and implications.

6.5.1 VTT and WTS analysis

The expected values of VTT and WTS values are computed for each individual as discussed in Section 3.4. These expected values may be averaged across any demographic sub-sample or across the entire sample to obtain corresponding mean values and standard deviations. Overall, the VTT sample average estimate is $26.5 for work travel and $23.2 for leisure travel, which are rather high but may be attributed to the sample being skewed toward high-income households[15]. The higher sample average VTT for work travel compared to leisure travel is consistent with findings from previous studies (for example, Axhausen et al., 2008; Börjesson and Eliasson, 2014). Interestingly, we find a lower variation in the leisure VTT relative to the work travel VTT. In terms of the WTS estimates, the results indicate that individuals are willing to pay, on average, about 50 cents (48.71 cents is the actual point value) not to have an additional passenger for commute travel, and this willingness to pay not to have an additional passenger rises to 90 cents (89.71 cents in the actual point value) on average, for leisure travel. This is, of course, consistent with the estimation results that individuals are more sensitive to additional passengers for leisure travel relative to commute travel. As already discussed, this willingness to pay to avoid traveling with strangers represents a fixed cost, and appears to be independent of travel time. That is, the notion that individuals may be more willing to share rides for short travel times in an AV, but not long travel times, is not supported by our analysis. Another perspective on these results is that individuals are willing to pay 14% [((26.5-23.2)/23.2)x100] more to reduce a minute in a commute trip compared to a leisure trip, while they are willing to pay 84% more to avoid an additional passenger in a leisure trip compared to a commute trip. The implications of these results for transportation planning and policy are that, from a shared economy perspective, it may be easier to promote PSAV use for commute trips than for leisure trips. Given that commute trips are the ones that overload the system during the peak period, there may be an opportunity to alleviate some of this peak period congestion. At the same time, there does not seem to be any difference in sensitivity to riding with others in an SAV based on travel time, which suggests that promoting PSAV use for short-distance trips will be likely as difficult as promoting PSAV use for long-distance trips, both for commute and leisure travel. Still, since value of time is somewhat higher for commute trips, efforts need to be focused on minimizing delays caused by serving multiple passengers during the peak period.

A further examination of the ratios between WTS and VTT for each trip purpose provides additional insights. In particular, for commute travel, reducing one passenger in a commute trip has the same monetary value as reducing the travel time by 1.10 minutes. For a leisure trip, the

[15] The average household income in the sample is $125,000 and the majority of the individuals live in multi-worker households. Using the estimate of 1.7 workers per household from our sample and an average work duration of about 37 hours/week in the sample, and considering that each respondent works 52 weeks per year, a worker would earn, on average, $38.2 per hour, which means that the work-trip VTT is equivalent to 69% of the hourly wage and the leisure travel VTT is about 60% of the hourly wage rate.

equivalent value is 2.33 minutes. Once again, this is a fixed time cost of an additional passenger, regardless of travel time. Overall, these values are low when compared to actual delays caused by an additional passenger in a ride. Thus, our results suggest that delays are a greater barrier to PSAV adoption than the actual presence of strangers[16]. This result reinforces the idea that privacy concerns may not be a barrier too difficult to overcome and dynamic ridesharing may have a large market penetration potential, especially for commute trips, as long as operated efficiently with minimal detour and pick-up/drop-off delays. Of course, it is possible that the perceptions associated with the experience of sharing a ride is abstract to a large group of respondents in the sample, because of the small share of the sample that has experienced pooled ride-hailing. Thus, it may be a fruitful avenue of further research to design experiments that mimic the travel experience in a more realistic manner (using pictures or even virtual reality). Nonetheless, our results provide important insights into SAV use in the future.

6.5.2 Treatment Effects and Policy Implications

To examine differences in preferences for sharing among different population segments, we compute average treatment effects (ATEs) of the socio-demographic variables on ride-hailing experience and on sharing intentions in the SAV scenarios, as well as VTT and WTS. The ATE measure for the choice outcomes provides the expected difference in ride-hailing experience or SAV-service choice for a random individual if s/he were in a specific category *i* of the determinant variable as opposed to another configuration $k \neq i$. The ATE is estimated as follows for each determinant variable:

$$\hat{\text{ATE}}_{ikj} = \frac{1}{Q}\sum_{q=1}^{Q}\left(\left[P(y_q = j \mid a_{qi} = 1) - P(y_q = j \mid a_{qk} = 1)\right]\right) \tag{3}$$

where a_{qi} is the dummy variable for the category *i* of the determinant variable for the individual *q*, y_q stands for the choice variable, and *j* represents a specific choice alternative. Thus, $\hat{\text{ATE}}_{ikj}$ above represents the estimate of the expected value change in the nominal category *j* of the choice outcome because of a change from category *k* of the determinant variable to category *i* of the determinant variable. In computing this effect, we first assign the value of the base category for each individual in the sample (that is, we assign the value of $a_{qk} = 1$ to the determinant

[16] Note that from an experimental design perspective, the range of additional time per individual varied from 1.66 to 10 minutes. Our results regarding the equivalent time value of an additional passenger is at the bottom of this range.

variable of each individual to compute $P(y_q = j \mid a_{qk} = 1))$ and then change the value of the variable to $a_{qi} = 1$ compute $P(y_q = j \mid a_{qi} = 1))$.

In our analysis, we compute the ATE measures for only two categories of the determinant variables. The base category for each determinant variable is used as the category to change from (as denoted by index *k* in Equation (3)) and a single non-base category of the determinant variable is selected as the category to change to (as denoted by index *i* in Equation (3)). For example, in the case of age, the base category is the "≥65 years" age group, while the changed category corresponds to the "18-34 years" age group. Similarly, for race/ethnicity, the base category is the "other" race/ethnicity (including individuals of Hispanic and non-White races/ethnicities) and the changed category is the "non-Hispanic White" race/ethnicity. We follow the same process of comparing a base and a non-base category of the determinant variables to evaluate percentage changes in VTT and WTS for the two trip purposes investigated. The results are presented in Table 5. Using employment type as an example, the ATE effect of -0.08 on private ride-hailing experience is interpreted as follows: if 100 random individuals moved jobs from full-time employment to part-time employment, there would be 8 fewer individuals with private ride-hailing experience.

The results in Table 6-6 indicate that high-income individuals, millennials, and individuals who live alone are the segments most likely to adopt private ride-hailing, while lower income millennials, individuals living in multi-worker households and individuals who are not non-Hispanic Whites are the most likely to have experience with pooled ride-hailing. Overall, age and income are the strongest predictors of ride-hailing experience and sharing intentions. As discussed earlier, millennials are more likely than those 65+ years of age to adopt pooled ride-hailing today, but are also more reluctant to indicate intent to use PSAVs in the future. Millennials also have a higher WTS value relative to those 65+ years of age, indicating an aversion to sharing rides in SAVs. Why these results are so is an important avenue for further research, especially because millennials just became the majority of the population in the U.S. and the success of SAVs and MaaS are critically dependent on this segment's adoption.

Although individuals living in high-income households are the most likely to use private ride-hailing services, they demonstrate high sharing aversion in all dimensions. An interesting and worrisome result is that the interest in the productive use of travel time for work travel reduces travel time disutility for this group, which then tempers the higher time-sensitivity of this group. The net result is that there is no statistically significant difference in VTT between the low and high income categories for work travel (and the difference in VTT is rather marginal even for leisure travel), as may be observed in the VTT percentage change columns for the income row in Table 5. With reduced VTT, high sharing aversion and high economic power, these individuals may have significant increase in “ride-alone VMT” when AVs become available. Encouraging high-income individuals to share rides will be challenging, but could be encouraged by upscale services offering additional comfort features for a higher price.

Table 6-6 Treatment Effect of Socio-Demographic Variables on Main Outcomes, VTT and WTS

Variable	Categories Compared (base versus changed)	Change in Probability								Percentage Change							
		Ride-hailing experience				Work purpose		Leisure purpose		Work purpose				Leisure purpose			
		Private only		Shared		Shared		Shared		VTT (%)		WTS (%)		VTT (%)		WTS (%)	
		Est.	St. err	Est.	St. err	Est.	St. err.	Est.	St. err.	Est.	St. err.	Est.	St. err.	Est.	St. err.	Est.	St. err.
Gender	**Male vs. female**	--	--	--	--	-0.032	0.006	-0.006	0.003	1.029	0.217	--	--	1.255	0.264	--	--
Age	**65+ vs. 18 to 34**	0.316	0.026	0.049	0.006	-0.021	0.008	-0.102	0.015	-16.221	3.436	2.069	1.373	-1.891	0.398	5.487	3.634
Race/ ethnicity	**Other vs. Non-Hispanic White**	0.021	0.004	-0.040	0.007	-0.028	0.006	-0.040	0.008	--	--	1.616	0.410	--	--	4.291	1.070
Education	**< bachelor's vs. graduate**	0.007	0.003	-0.002	0.001	0.028	0.007	-0.015	0.006	-6.614	1.663	--	--	-0.764	0.191	--	--
Employment	**Full-time vs. part-time**	-0.080	0.009	0.021	0.003	0.011	0.004	0.020	0.006	-2.177	0.445	--	--	-2.652	0.539	--	--
Income	**< $50,000 vs. $200,000+**	0.337	0.019	-0.023	0.006	-0.269	0.029	-0.133	0.015	--	--	4.288	0.765	1.565	0.554	11.266	1.947
Households composition	**Multi-worker vs. single-worker**	0.137	0.011	-0.032	0.003	-0.034	0.007	-0.013	0.002	--	--	--	--	--	--	--	--
Residential location	**Rural/suburban vs. urban**	0.098	0.007	0.042	0.004	-0.027	0.005	-0.017	0.004	--	--	--	--	--	--	--	--
Vehicle availability	**< 1 per worker vs. > 1 per worker**	0.008	0.005	0.021	0.006	-0.025	0.007	--	--	--	--	--	--	--	--	--	--
Commute mode	**Drive alone vs. Non-car**	0.049	0.008	0.051	0.007	--	--	--	--	--	--	--	--	--	--	--	--
Ride-hailing experience	**No vs. Pooled**	n/a	n/a	n/a	n/a	-0.008	0.008	0.039	0.009	--	--	--	--	--	--	--	--

"--" = not statistically significantly different from zero at the 90% level of confidence.

Transferring individuals from rural and suburban environments and encouraging commute by non-car modes instead of drive alone shows a positive impact on both private and pooled ride-hailing experience. In fact, together with age, both living in an urban area and commuting by a non-car mode are the strongest positive predictors of pooled ride-hailing. Yet, similar to millennials, despite the experience with pooled ride-hailing, urban residents seem less interested in sharing rides in SAVs for both work and leisure purposes. From an operational perspective, urban (dense) areas are the most suitable environment to the efficient operation of dynamic ridesharing (because the demand is concentrated and thus matching becomes easier), thus further investigation of this negative effect observed herein is necessary.

6.6 Conclusions

There is growing evidence that ridesharing will be a key element to ensure a sustainable future to urban transportation in an AV future. In this context, in the current chapter we proposed and applied a multivariate modeling framework to investigate the extent to which individuals are willing to share rides with strangers in a SAV future. A joint model of current ride-hailing experience and stated intentions regarding the use of shared rides for trips to work and to leisure activities was estimated and VTT and WTS (money value of traveling alone compared to riding with strangers) were computed for each individual in the sample. The model relied on three stochastic psychosocial latent constructs representing privacy-sensitivity, time-sensitivity and interest in productive use of travel time to create dependency among the three nominal outcomes and to moderate the effects of trip attributes (time and number of additional passengers) for each individual.

The use of psychosocial latent constructs as a key component in our model provides important insights regarding transportation planning and policy. First, we identified that privacy concerns are currently discouraging individuals (mostly non-Hispanic Whites) from experimenting pooled ride-hailing services, and such concerns also create a significant aversion to future PSAV services, which can be deterring to the idea of MaaS in currently car-dominated cities. Privacy-sensitivity may also be worsened by security concerns in a PSAV context where individuals see themselves alone with a stranger in the vehicle (since there is not a driver to serve as a "professional guardian" during the trip). Although we did not investigate security concerns directly, we did observe that current pooled ride-hailing users may be reticent to using shared rides in a SAV, which could be preliminary evidence of this issue. Hence, a comprehensive examination of privacy and safety concerns of current pooled ride-hailing users may be a necessary step to prevent this group from moving to private rides as SAVs become available. Social-network-based ridesharing schemes can be an interesting solution to privacy and security concerns in shared rides. This type of scheme has been recently proposed and simulated from a supply standpoint, but is still to be implemented (see Richardson et al., 2016, and Wang et al., 2017). In that sense, MaaS-oriented travel behavior research efforts can help investigate consumer's interest and potential demand to this new type of service. Second, the latent variable representing the interest in productive use of travel time provided evidence that this is an important factor currently guiding ride-hailing adoption. Considering the current interest by transportation researchers in understanding the impacts of automation on VTT, the evidence obtained in the current study is very important. Ride-hailing services can be an important proxy SAV services and can provide valuable data to measure potential changes in individual's VTT due to productive use of travel time (even as a tool for naturalistic experiments). We also observed that providing an environment that is conducive to productive use of travel time may increase high-income individual's tolerance to increased travel times, which may incur in increased transportation equity problems. High-income individuals are currently the main users of private ride-hailing and demonstrate high sharing aversion in all dimensions. Thus, if their VTT decreases due to productive use of travel time, they may have a disproportional increase in "ride-alone VMT". Encouraging high-income individuals to share rides will be challenging and calls for future research. Yet, this group could be encouraged to share if upscale services are offered within MaaS packages. Third, we observed that when dealing with individuals who are intrinsically privacy and time-sensitive, an environment that is conducive to the productive use of travel time will have little to no effect on increasing their tolerance to increased travel times and/or additional passengers. This indicates that despite the potential of automation in reducing VTT, there are population segments that are unlikely to become less time-sensitive, such as full-time employed women between the ages of 35 and 44 years old.

In terms of actual measures of VTT and WTS, our results point to the importance of distinguishing trip purposes. For instance, individuals seem to be less sensitive to the presence of strangers in a commute trip than in a leisure trip, but the sensitivity to time is the opposite. The implications of these results for transportation planning and policy are that, from a shared economy perspective, it may be easier to promote PSAV use for commute trips than for leisure trips. Given that commute trips are the ones that overload the system during the peak period, there may be an opportunity to alleviate some of this peak period congestion. At the same time, there does not seem to be any difference in sensitivity to riding with others in an AV based on travel time, which suggests that promoting PSAV use for short-distance trips will be likely as difficult as promoting PSAV use for long-distance trips, both for commute and leisure travel. Still, since value of time is somewhat higher for commute trips, efforts need to be focused on minimizing delays caused by serving multiple passengers during the peak period. A further examination of the ratios between WTS and VTT reinforced the idea that privacy concerns may not be a barrier too difficult to overcome and dynamic ridesharing may have a large market penetration potential, especially for commute trips, as long as operated efficiently with minimal detour and pick-up/drop-off delays. This result points to a potential bright future for PSAV based MaaS systems in car-dominated environments.

The current study is just a first step to an important travel behavior topic. A similar framework to the one proposed herein can be enhanced by the inclusion of a fourth latent variable representing individuals' sensitivities to travel monetary costs. As largely discussed in the VTT and WTP literature, accommodating variability in the cost coefficient is important to avoid erroneously attributing variation to WTP. Additionally, a new experimental design that captures individuals current VTT would allow the identification of biases in the values estimated in this study.

CHAPTER 7. Conclusions

Society is experiencing the initial stages of a technological revolution that promises to disrupt urban transportation as known today and induce behavioral and social changes. The main factors guiding the transformation of urban mobility are the growth of Information and Communication Technology (ICT)-enabled transportation services and the development of self-driving automotive technologies. The popularization of ICTs is not only allowing instantaneous and ubiquitous remote access to people and information, but is also facilitating the integration between different transportation modes and the development of on-demand transportation services, such as bicycle sharing, car sharing, and ride-hailing. Within this context, Mobility as a Service (MaaS) systems, which provide users with multiple options of personalized trip plans and packages that facilitate multimodal door-to-door travel, have a great potential to enable convenient, cost-effective, and environmentally sustainable alternatives to the use of private cars and drive alone mode. Such potential should be further enhanced by the development of autonomous vehicles, which will enable greater flexibility to ride-hailing services at a reduced cost (compared to today's ride-hailing services) as drivers will no longer be necessary.

The automation of vehicles is also expected to provide direct road capacity improvements due to crash reductions and platooning capabilities. Yet, these gains can be offset by latent demand effects. That is, car users may experience increased comfort due to both changes in vehicle design and elimination of the need to drive, which should allow for the meaningful use of the time spent traveling (socializing, working or sleeping, for example) and multitasking. Such factors may reduce the disutility commonly attributed to traveling (especially driving) and, thus, decrease an individual's valuation of travel time (VTT). The consequences may be the increase in the number of activities pursued and/or the distance between activities, resulting in the growth of vehicle miles traveled. These indirect effects may work against the objective of MaaS systems to reduce private car usage, and may compromise network efficiency gains generated by the direct technological effects of automation. As a consequence, congestion levels and energy consumption could actually increase. In that sense, proactive planning and policy guided towards promoting the use of shared vehicles and pooled rides is important to facilitate the development of MaaS systems and minimize possible negative externalities of automation. To inform such planning, a good and deep understanding of the current use of ride-hailing services, together with an examination of individual's preferences regarding AV adoption, is critical.

Motivated by the discussion above, the main objectives of this dissertation research were to develop a better understanding of the adoption of current and future mobility technologies and services, and to provide evidence (from a travel behavior perspective) to the viability of MaaS systems in environments where transportation is currently primarily based on private car usage. A research framework containing four independent but related analysis components was developed to allow a comprehensive investigation of travelers' characteristics and behaviors associated with ride-hailing use and preferences regarding AVs. While the first two analyses focused on users' current ride-hailing behavior, the other two simultaneously investigated current travel behaviors and future intentions to use automated vehicle (AV)-based services. As described in Chapter 3, the first analysis applied a two-step aggregate modeling approach to investigate the generation and distribution of daily ride-hailing trips in the city of Austin, Texas. Multivariate models were used to predict how many trips would be generated from a specific traffic analysis zone (TAZ) on both weekdays and weekend days, and to identify characteristics of zones that attract ride-hailing trips. The second analysis (Chapter 4) complemented the first by modeling the multiple choices associated with the use of ride-hailing at the individual level (instead of trip counts per TAZ) based on data from the Dallas-Fort Worth Metropolitan Area, Texas. The multiple outcomes in this second analysis component included the choice to use ride-hailing, the frequency of both solo and pooled rides, and the characteristics (purpose, time of the day, companion, and mode substituted) of the latest ride-hailing trip of survey respondents. These multiple outcomes are jointly modeled as functions of socio-demographic characteristics, latent constructs representing attitudes and lifestyles, and endogenous variables representing residential location and vehicle availability. The third analysis (Chapter 5) modeled preferences regarding the adoption of AVs in the Seattle Metropolitan Area, Washington. Based on the person's lifecycle, lifestyle (represented by latent constructs), and current transportation related behavior, the model explained whether an individual had the intention to purchase an AV or use only shared AVs (or both or none) in the future. In addition to the AV preferences, the main endogenous variables considered were residential location density, vehicle ownership, and experience with car-sharing and ride-hailing services. The final analysis, developed in Chapter 6, also used data from DFW and focused on individuals' perceptions toward pooling (or sharing) rides. The current experience with ride-hailing services was modeled together with stated choices between hiring a solo and a pooled ride for commute and leisure trips in a shared AV (or SAV) future. Again, latent constructs representing attitudes and socio-demographic characteristics were used to explain the current behavior and stated intentions.

In this concluding chapter, we summarize and compare the results from the four different analysis components based on the six research questions posed in Chapter 1, Section 1.5. It is important to mention that results obtained in the current dissertation are not generalizable; however, some of the questions can be addressed by the cumulative evidence from multiple analysis components, which also allows for a rich contrast among the three locations investigated (Austin, DFW Metropolitan Area, and Seattle Metropolitan Area). The focus here is to summarize the evidence that contributes to addressing each question rather than discussing in detail the policy implications of the results. The reader is referred to each chapter for a complete discussion of policy implications. We close this chapter with final recommendations on how to promote a MaaS-oriented future where individuals rely on shared services and shared rides.

7.1 Discussion of Research Questions

7.1.1 What Segments of the Population Already Use Ride-Hailing Services? Who is Sharing Rides? Who Are the Frequent Users?

Age and income appear as the strongest socio-demographic predictors of ride-hailing use across all four analyses. The results show that ride-hailing users are predominantly wealthy young adults and this characteristic is common to the three different cities studied. In some cases, these socio-demographic effects are manifested directly, while, in others, they are expressed indirectly through the psycho-social latent constructs, especially tech-savviness. Based on the DFW data, we also observe that being young is the main socio-demographic determinant of pooled ride-hailing experience. Having a higher level of education shows a positive effect on overall ride-hailing experience in the case of Seattle, while, for the Dallas sample, education had significant effects only on pooled ride-hailing experience. The difference between these two results may be attributed to the high share of individuals with tertiary education in the Dallas sample, but may also be a consequence of the temporal difference between the two data sets. The data collection in Seattle took place in spring 2015, while in Dallas the survey was conducted in fall 2017. It is possible that education played a key role in the early adoption of ride-hailing and as time passes it is becoming less relevant (especially in analyses that control for income effects). In that sense, since the pooled version of ride-hailing is a newer service, high education may be again a characteristic of early adopters. Non-Hispanic Whites show a lower propensity to use ride-hailing in Austin and DFW, and, in DFW, we also observe that this group has a lower tendency to partake in pooled rides. The race/ethnicity effect is expressed both directly and indirectly through the latent constructs representing VSLP and privacy-sensitivity.

Overall, the analyses conducted in this dissertation reveal that psycho-social or lifestyle characteristics play an important role in describing ride-hailing users and should be incorporated in future studies that plan to characterize such groups. Tech-savviness, for example, seems to be a necessary condition for ride-hailing use, and hence future studies investigating ride-hailing adoption should also measure indicators of individuals' familiarity and everyday use of ICT and other technologies. Similarly, privacy-sensitivity (or aversion to strangers) is identified as a major deterrent to pooled ride-hailing use. Finally, in terms of ride-hailing frequency, again we observed that age and income are the most important predictors. That is, being young and having a high income are the main contributing factors to higher frequencies of ride-hailing usage (the results for DFW are based on the ride-hailing frequency model component, and the results for Austin assume that number of trips are a proxy to frequency). In this case, the latent variable representing variety-seeking lifestyle also showed a relevant explanatory contribution, suggesting again the importance of considering psycho-social factors in the characterization of users.

7.1.2 What Land use and Transportation Aspects Contribute to the Use of Ride-Hailing?

First, in terms of land use, we observe that urban density is a key element to ride-hailing adoption and use in all four analyses (being always among the strongest predictors in the models). Even after controlling for self-selection effects, individuals living in more urbanized locations are more likely than their counterparts in less urbanized/dense locations to have used both private and pooled ride-hailing. The TAZ based analysis in Austin indicates that there is a concentration of trips in areas with higher residential density and activity intensity (proportion of retail and employment opportunities) on both weekdays and weekend days. Weekday trips are even more localized, and zones containing universities, parks, bars, and restaurants are responsible for generating and attracting the most trips. Second, in terms of transportation, in the Austin-based analysis we observe a negative influence of transit supply on ride-hailing trip rates, which suggests that ride-hailing decreases as transit service improves. Another perspective is that ride-hailing tends to get used more in areas with relatively poor transit service. We also observe the direct substitution of public transit by ride-hailing, especially pooled ride-hailing, in DFW. Although we do not have details about the public transit conditions for these trips, it is plausible that ride-hailing is compensating for a deficient supply system since it is being able to attract customers despite its higher costs.

7.1.3 Is There Evidence of Positive and Negative Externalities of Ride-Hailing Adoption?

To answer this question, we identify four types of ride-hailing externalities: (1) substitution of other modes and associated consequences, (2) impacts on vehicle ownership, (3) impacts on accessibility, and (4) induction of new trips.

7.1.3.1 Modes Substituted

Based on the DFW survey, we observe that ride-hailing is drawing users from all modes, especially from taxi and personal car. The high number of ride-hailing trips attracted by the airport TAZ in Austin also suggests that both taxi and personal car are being substituted by ride-hailing, since these are the main modes used to reach the airport in the city. As discussed earlier, in terms of public transit, evidence from Austin suggests substitution effects between transit and ride-hailing, since ride-hailing decreases as transit service improves (or ride-hailing tends to get used more in areas with relatively poor transit service). The analysis based on the DFW data also shows evidence that individuals younger than 65 years of age, those with a bachelor's degree or higher but lower income, and individuals with experience with pooled ride-hailing who are infrequent users tend to replace active/public transportation modes with ride-hailing. Further, the analysis of Seattle data suggests that those who are "green" and those who reside in high density residential neighborhoods today are the individuals most likely to embrace ride-hailing as well as the individuals most likely to currently use non-motorized and public transit services. Therefore, it may be conjectured that ride-hailing is taking modal share away from active/public transportation modes in this case as well. In order to identify which travel mode is being most affected by the popularization of ride-hailing, it would be necessary to identify the rates of substitution of each mode relative to the overall mode share. None of the analyses performed in the current dissertation allow such comparisons, but there seems to be evidence that, as ride-hailing becomes less expensive, more active/public transit trips may be substituted by this mode, potentially resulting in negative traffic and public health externalities.

7.1.3.2 Impacts on Vehicle Ownership

All the analyses conducted in this dissertation involve cross-sectional datasets, which hinders the examination of whether ride-hailing impacts vehicle ownership or whether vehicle ownership influences ride-hailing. Still, we observe a negative association between vehicle ownership and ride-hailing use across all areas studied. Even in DFW, where vehicle ownership rates are higher than the national average, we observe that frequent ride-hailing users tend to have more limited household vehicle availability than infrequent or non-users. Despite the inability to infer causality, these results suggest that ride-hailing has both the potential to increase access to car travel for those who cannot afford owning a vehicle (or prefer not to own a vehicle) as well as to reduce vehicle ownership rates. Still, it is important to emphasize that the negative association between ride-hailing and vehicle ownership does not necessarily imply the reduction of car travel (VMT). Indeed, as discussed in the previous section, ride-hailing can increase car usage if substituting active/public transit travel modes.

7.1.3.3 Impacts on Accessibility

Ride-hailing can provide more access to activity opportunities for individuals who do not own vehicles and/or those with limited driving capabilities. Based on the DFW data, we observe that students and those with lower vehicle availability are more likely than their peers to have pursued errands in their last ride-hailing trip rather than other activity purposes, while millennials and those with lower vehicle availability are more likely to have pursued work-related travel rather than airport travel in their most recent ride-hailing trip. These results perhaps are indicative of the use of ride-hailing as an "accessibility mobility tool" to compensate for limited access to routine activities using other mobility options. Indeed, the Austin result that suggests that ride-hailing tends to get used more in areas with relatively poor transit service corroborates this finding. On the other hand, millennials and non-Hispanic Whites are most likely to have pursued recreation (relative to all other activity purposes) in their last ride-hailing trip, presumably a reflection of the use of ride-hailing here as a "convenience mobility tool". Overall, we observe that ride-hailing is not the preferred option when it comes to completing routine commitments. While ride-hailing provides more access to activity opportunities to certain segments, it is also not the most convenient for conducting activities that involve trip chaining, for example, running errands. Since running errands typically involves chaining of multiple activities in the same sojourn from home and/or involves carrying and storing food and other perishable goods during the trip, ride-hailing is not the most convenient because it is more of a pure trip-based consumption service as opposed to a broader transportation option that allows a cost-effective time-based consumption service (in which the same vehicle is available to pursue multiple activities and over an extended period of time). Perhaps greater gains in accessibility would be achieved if ride-hailing was provided in a time-based option as well, which effectively would combine today's ride-hailing and car-sharing services into one service. As the mobility landscape moves more toward automated vehicles, this integration of trip-based and time-based consumption options may become even easier to implement.

7.1.3.4 Generation of New Trips

The only dataset that contains information that allows the investigation of the generation of new trips due to ride-hailing availability is the one based on the DFW survey. As discussed in Chapter 4, the demographic effects indicate that young adults (18-44 years of age) are more likely than their older peers to have generated a new trip in their most recent ride-hailing experience. Also, part-time employees, self-employed individuals and those that live in multi-worker households appear to generate new ride-hailing trips more so than individuals in other households, perhaps a reflection of the added convenience to pursue activities due to ride-hailing. New trips are also more likely to occur among those living in non-rural areas. The generation of new trips in dense areas can, in the long term, intensify traffic congestion problems due to increased automobile usage. The new generated trips seem to be for the purposes of running errands and pursuing recreational activities, and are more likely to happen during the non-evening periods. Complementing the discussion in the previous section, we observe that the generation of new trips, in some cases, reflects an increase in the ability to access activity opportunities for individuals who do not own vehicles and/or those with limited driving capabilities. For example, this is the case with students, individuals with low vehicle availability, and individuals from low-income households who generate new trips associated with running errands. Thus, ride-hailing can assume a welfare role, but fares would need to be revisited to fit the needs of these more financially challenged segments of our society. On the negative side, we observe that most ride-hailing induced trips are generated by individuals in suburban and urban areas, serve a single passenger, and occur in the morning commute period as well as the mid-day and night periods. In other words, ride-hailing is generating more "drive alone" trips in the already-congested suburban and urban areas of the DFW.

7.1.4 What Segments of the Population Have the Intention to Adopt AVs? Who Wants to Share Vehicles? Who Wants to Own? And Who Wants Both?

The responses to these questions are based on the Seattle MSA data and explained in detail in Chapter 5. Note that, in this analysis, we investigate individuals' willingness to share vehicles and not rides, as examined in Chapter 6. Overall, early adopters of AV technology are likely to be those with a higher level of education, individuals between 18 and 44 years of age, and workers. In particular, individuals in the youngest age group of 18-24 years show the greatest propensity for AV sharing and an aversion towards the AV ownership-only alternative. Individuals with a higher level of education are also more likely to adopt AV sharing as opposed to ownership or both. Lower income individuals appear to be largely averse to the adoption of AV technology in any form with those in the lowest income category showing the greatest level of resistance to adoption. Individuals who currently eschew vehicle ownership, and have already experienced car-sharing or ride-hailing services, are especially likely to be early adopters of SAV services. On the other hand, individuals who currently own vehicles, and have not yet experienced mobility on demand services, are more inclined to adopt AV technologies in an ownership or combined ownership and sharing mode. Even after controlling for self-selection effects, high-density neighborhood residents are also more inclined to adopt AV sharing services as opposed to any model that involves ownership. The latent variables representing lifestyles also so important explanatory power and indicate that green lifestyle is associated with favoring AV sharing, and tech-savviness leads to a higher likelihood of embracing AV technology in general, and especially a combination of both AV ownership and SAV services.

7.1.5 How Much Individuals Would Be Willing to Pay to Not Share Rides in a SAV Scenario? How Does the Willingness-to-Pay to Not Share Relate with the Value of Travel Time?

To answer these questions, we rely on the analysis in Chapter 6, which is based on the DFW sample. The results are not generalizable, but still provide guidance for future studies and planning efforts. Overall, our results point to the importance of distinguishing trip purposes. For instance, individuals seem to be less sensitive to the presence of strangers in a commute trip than in a leisure trip, but the sensitivity to time is the opposite. The VTT sample average estimate is

$26.5 for work travel and $23.2 for leisure travel, which are rather high but may be attributed to the sample being skewed toward high-income households. Interestingly, we find a lower variation in the leisure VTT relative to the work travel VTT. In terms of the WTS estimates, the results indicate that individuals are willing to pay, on average, about 50 cents (48.71 cents is the actual point value) not to have an additional passenger for commute travel, and this willingness to pay not to have an additional passenger rises to 90 cents (89.71 cents in the actual point value) on average, for leisure travel. This willingness to pay to avoid traveling with strangers represents a fixed cost, and appears to be independent of travel time. That is, the notion that individuals may be more willing to share rides for short travel times in a SAV, but not long travel times, is not supported by our analysis.

Another perspective on these results is that individuals are willing to pay 14% more to reduce a minute in a commute trip compared to a leisure trip, while they are willing to pay 84% more to avoid an additional passenger in a leisure trip compared to a commute trip. The implications of these results for transportation planning and policy are that, from a shared economy perspective, it may be easier to promote PSAV use for commute trips than for leisure trips. A further examination of the ratios between WTS and VTT for each trip purpose provides additional insights. In particular, for commute travel, reducing one passenger in a commute trip has the same monetary value as reducing the travel time by 1.10 minutes. For a leisure trip, the equivalent value is 2.33 minutes. Overall, these values are low when compared to actual delays caused by an additional passenger in a ride. Thus, our results suggest that delays are a greater barrier to PSAV adoption than the actual presence of strangers. This result reinforces the idea that privacy concerns may not be a barrier too difficult to overcome and dynamic ridesharing may have a large market penetration potential, especially for commute trips, as long as operated efficiently with minimal detour and pick-up/drop-off delays.

7.1.6 What Are the Impacts of Current Ride-Hailing Experience on the Intentions to Adopt AVs, SAVs and PSAVs?

The analysis in Chapter 5 provides evidence on the impacts of ride-hailing experience on the general intention to adopt AVs and SAVs. The analysis in Chapter 6 complements the previous analysis by investigating the impacts of ride-hailing experiences on choices between SAVs and PSAVs. In both cases, we are able to identify the "true effects" of ride-hailing experience since we control for self-selection though the use of a joint modeling framework and stochastic psycho-social latent constructs. Even after accounting for tech-savviness and green lifestyle propensity, the ride-hailing experience seems to contribute to the overall interest in AV adoption and a stronger preference for exclusive use of SAVs or for a combination of personally owned and SAVs.

In terms of impacts of ride-hailing experience on the intention to use PSAVs compared to SAVs, that is, pooling rides instead of riding alone, are less promising. Current experience with "private ride-hailing only" (relative to having no experience with ride-hailing at all or having pooled ride-hailing experience) has a negative effect on choosing to share rides in SAVs for both activity purposes. In other words, it appears that people who have used "private ride-hailing only" appreciate the convenience and flexibility of the private arrangement based on the actual experience, and are loath to sharing the travel experience with strangers (either with current pooled ride-hailing or with PSAVs in the future). Particularly intriguing here is the implication that it may be easier to "convert" individuals who have never used ride-hailing into future PSAV users than to attempt to convince current "private ride-hailing only" users to become future PSAV users. On the other hand, current pooled ride-hailing users appear to be the prime segment interested in PSAV use, especially for trips for leisure purposes. However, it does appear from our results that PSAVs are not viewed in the same light as current pooled ride-hailing use by some population segments, such as young individuals and those residing in urban areas. If this is because of the comfort/security of having a human "guardian" during the trip, then it becomes incumbent that AV design pay attention to security features. Note that the negative impact of current ride-hailing experiences on the intention to use PSAVs may be particular to DFW. Unfortunately, there was no data available to investigate such effect in the context of Seattle. Thus, it is unquestionable the need for further research on the impact of current ride-hailing experiences on the intention to adopt PSAVs and especially on attitudes and perceptions associated with having a human driver versus not having one in the context of pooled travel.

7.2 Recommendations for a Shared Future

The results from the analyses undertaken in this dissertation show that, from a behavioral perspective, a service-based transportation future where people predominantly travel using shared vehicles and pooled rides instead of their own vehicles is on its way but still distant. A complex combination of actions is required to promote the use of shared services both today and in an AV future. Among these actions, we identified the need for campaigns to (a) increase technology awareness among older individuals and individuals from lower income households, and (b) reduce privacy-sensitivity among non-Hispanic Whites and millennials. However, such efforts would still need to be complemented by a decrease in service fares. In this regard, understanding better the cost-privacy sensitivity trade-off would be a particularly valuable research pursuit to position pooled ride-hailing and PSAV services.

Even after accounting for self-selection effects, the four analyses in this dissertation point to urban density as the most effective ingredient to promoting the use of shared vehicles and shared rides today and in the future. The fact that this effect prevails even after any residential self-selection is very significant. It motivates the consideration of neo-urbanist land-use policies in an entirely new light relative to the traditional focus of such policies as a potential way to reduce motorized private car travel. This is especially so because, separate from a direct neighborhood effect, densification shows the potential to increase ride-hailing adoption and AV sharing adoption propensity through a reduction in vehicle ownership. Along those lines, our results also suggest a need for policies that discourage the substitution of short-distance "walkable" trips by ride-hailing and SAVs (to reduce traffic congestion as well as not take away from active modes of transportation), and a need for low cost and well-integrated MaaS systems to avoid substitution of transit trips by ride-hailing and SAVs. Pooled services that offer relatively lower costs have an even stronger potential to draw users from active/public transit modes. Of course, to increase the efficiency and sustainability of a MaaS system, the relationship between (pooled) ride-hailing and transit should be one of complementarity rather than substitution. Yet, it is reasonable to expect that a service that can be used for door-to-door trips will not be used for first- and last-mile connectivity to transit hubs, unless low cost and well-integrated MaaS systems are designed.

Finally, for a shared future to be successful, MaaS systems must accommodate the majority of commute trips, since commuting corresponds to a substantial share of daily trips and entails the majority of the peak-period demand. Despite the lower numbers of work trips captured in our sample (compared to trips to the airport and trips to recreational activities), the model results show that frequent users are likely to use ride-hailing for work trips (from the trip purpose model), and work trips by ride-hailing are typically made alone (based on the trip companion model) during the morning and evening periods (as per the time-of-day model). The net result is that many ride-hailing trips for work during the morning and evening are undertaken in private ride-hailing mode as opposed to pooled ride-hailing mode. There is substantial opportunity for ride-hailing services as well as employers to work together to increase vehicle occupancy during the commute periods, through low cost pooled ride-hailing services (such as Uber's most recently introduced "Express Pool" service) and subsidizing the use of such services. In that sense, when analyzing WTS in a SAV future, we observed that individuals seem to be less sensitive to the presence of strangers in a commute trip than in a leisure trip, but the sensitivity to time is the opposite. Thus, there is clear evidence that it may be easier to promote PSAV use for commute trips than for leisure trips as long as the services provide are efficient and minimize delays associated with serving multiple passengers.

REFERENCES

Alemi, F., Circella, G., Handy, S., Mokhtarian, P., 2017. What influences travelers to use Uber? Exploring the factors affecting the adoption of on-demand ride services. Presented at the 96th Annual Meeting of the Transportation Research Board, Washington, D.C., January, Paper No. 17-05630.

Alemi, F., Circella, G., Mokhtarian, P., Handy, S., 2018a. On-demand ride services in California: Investigating the factors affecting the frequency of use of Uber/Lyft. Presented at the 97th Annual Meeting of the Transportation Research Board, Washington, D.C., January, Paper No. 18-05563.

Alemi, F., Circella, G., Sperling, D., 2018b. Adoption of Uber and Lyft, factors limiting and/or encouraging their use and impacts on other travel modes among Millennials and Gen Xers in California. Presented at the 97th Annual Meeting of the Transportation Research Board, Washington, D.C., January, Paper No. 18-06713.

Amir, O., Levav, J., 2008. Choice construction versus preference construction: The instability of preferences learned in context. Journal of Marketing Research 45(2), 145-158.

Amirkiaee, S.Y. and Evangelopoulos, N., 2018. Why do people rideshare? An experimental study. *Transportation Research Part F*, 55, 9-24.

Astroza, S., Garikapati, V.M., Bhat, C.R., Pendyala, R.M., Lavieri, P.S., Dias, F.F., 2017. Analysis of the impact of technology use on multimodality and activity travel characteristics. Transportation Research Record: Journal of the Transportation Research Board 2666, 19-28.

Astroza, S., Bhat, P.C., Bhat, C.R., Pendyala, R.M., Garikapati, V.M., 2018. Understanding activity engagement across weekdays and weekend days: A multivariate multiple discrete-continuous modeling approach. Journal of Choice Modelling (forthcoming).

Axhausen, K.W., Hess, S., König, A., Abay, G., Bates, J.J. and Bierlaire, M., 2008. Income and distance elasticities of values of travel time savings: New Swiss results. Transport Policy, 15(3), 173-185.

Bansal, P., Kockelman, K.M. and Singh, A., 2016. Assessing public opinions of and interest in new vehicle technologies: An Austin perspective. Transportation Research Part C: Emerging Technologies, 67, pp.1-14.

Barlow, F.K., Paolini, S., Pedersen, A., Hornsey, M.J., Radke, H.R., Harwood, J., Rubin, M., Sibley, C.G., 2012. The contact caveat: Negative contact predicts increased prejudice more than positive contact predicts reduced prejudice. Personality and Social Psychology Bulletin 38(12), 1629-1643.

Becker, F. and Axhausen, K.W., 2017. Literature review on surveys investigating the acceptance of automated vehicles. Transportation, 44(6), pp.1293-1306.

Ben-Akiva, M.E., and Lerman, S.R., 1985. *Discrete Choice Analysis: Theory and application to travel demand* (Vol. 9). The MIT Press, Cambridge, MA.

Bhat, C.R., 2011. The maximum approximate composite marginal likelihood (MACML) estimation of multinomial probit-based unordered response choice models. Transportation Research Part B: Methodological, 45(7), pp.923-939.

Bhat, C.R., 2015a. A new generalized heterogeneous data model (GHDM) to jointly model mixed types of dependent variables. Transportation Research Part B: Methodological, 79, pp.50-77.

Bhat, C.R., 2015b. A Comprehensive Dwelling Unit Choice Model Accommodating Psychological Constructs within a Search Strategy for Consideration Set Formation. Transportation Research Part B, Vol. 79, pp. 161-188.

Bhat, C.R. and Dubey, S.K., 2014. A new estimation approach to integrate latent psychological constructs in choice modeling. Transportation Research Part B: Methodological, 67, pp.68-85.

Bhat, C.R., Dubey, S.K. and Nagel, K., 2015. Introducing non-normality of latent psychological constructs in choice modeling with an application to bicyclist route choice. Transportation Research Part B: Methodological, 78, pp.341-363.

Bhat, C., Paleti, R., Pendyala, R., Lorenzini, K., Konduri, K., 2013. Accommodating immigration status and self-selection effects in a joint model of household auto ownership and residential location choice. *Transportation Research Record: Journal of the Transportation Research Board* 2382, 142-150.

Bhat, C.R., Paleti, R. and Singh, P., 2014. A spatial multivariate count model for firm location decisions. Journal of Regional Science, 54(3), pp.462-502.

Bhat, C.R., Astroza, S., Bhat, A.C. and Nagel, K., 2016. Incorporating a multiple discrete-continuous outcome in the generalized heterogeneous data model: Application to residential self-selection effects analysis in an activity time-use behavior model. Transportation Research Part B: Methodological, 91, pp.52-76.

Bian, M. and Leung, L., 2015. Linking loneliness, shyness, smartphone addiction symptoms, and patterns of smartphone use to social capital. Social Science Computer Review, 33(1), pp.61-79.

Börjesson, M. and Eliasson, J., 2014. Experiences from the Swedish value of time study. *Transportation Research Part A*, 59, 144-158.

Blumenberg, E. and Smart, M., 2010. Getting by with a little help from my friends… and family: immigrants and carpooling. Transportation, 37(3), pp.429-446

Buddelmeyer, H., Hamermesh, D.S. and Wooden, M., 2017. The stress cost of children on moms and dads. European Economic Review.

Castro, M., Paleti, R. and Bhat, C.R., 2012. A latent variable representation of count data models to accommodate spatial and temporal dependence: Application to predicting crash frequency at intersections. Transportation research part B: methodological, 46(1), pp.253-272.

Chan, N.D. and Shaheen, S.A., 2012. Ridesharing in North America: Past, present, and future. *Transport Reviews*, 32(1), 93-112.

Chen, C., Xu, H. and Zheng, J., 2015. The relationship between economic value of time and feelings of time pressure. *Social Behavior and Personality: An international journal*, 43(8), 1395-1407.

Cheng, H.L., Medlow, S., Steinbeck, K., 2016. The health consequences of obesity in young adulthood. Current Obesity Reports 5(1), 30-37.

Chevalier, M., Gutsatz, M., 2012. Luxury Retail Management: How the World's Top Brands Provide Quality Product and Service Support. John Wiley & Sons, Singapore.

Clewlow, R.R., 2016. Carsharing and sustainable travel behavior: Results from the San Francisco Bay Area. Transport Policy, 51, pp.158-164.

Clewlow, R., Mishra, G.S. and Center, P.E.E., 2017, January. Shared mobility: Current adoption, use, and potential impacts on travel behavior. In Annual Meetings of the Transportation Research Board, Washington DC.

Cyganski, R., Fraedrich, E. and Lenz, B., 2015, January. Travel-time valuation for automated driving: A use-case-driven study. In Annual Meeting of the Transportation Research Board (Vol. 15, No. 4259, pp. 11-15).

Das, S., Sekar, A., Chen, R., Kim, H.C., Wallington, T.J. and Williams, E., 2017. Impacts of Autonomous Vehicles on Consumers Time-Use Patterns. Challenges, 8(2), p.32.

DeVoe, S.E. and Pfeffer, J., 2011. Time is tight: How higher economic value of time increases feelings of time pressure. Journal of Applied Psychology, 96(4), p.665.

Delhomme, P. and Gheorghiu, A., 2016. Comparing French carpoolers and non-carpoolers: Which factors contribute the most to carpooling?. Transportation Research Part D: Transport and Environment, 42, pp.1-15.

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

de Almeida Correia, G.H. and van Arem, B., 2016. Solving the User Optimum Privately Owned Automated Vehicles Assignment Problem (UO-POAVAP): A model to explore the impacts of self-driving vehicles on urban mobility. Transportation Research Part B: Methodological, 87, pp.64-88.

English, T., Carstensen, L.L., 2014. Selective narrowing of social networks across adulthood is associated with improved emotional experience in daily life. International Journal of Behavioral Development 38(2), 195-202.

Enoch, M.P., 2015. How a rapid modal convergence into a universal automated taxi service could be the future for local passenger transport. Technology Analysis & Strategic Management, 27(8), pp.910-924.

Fan, Y., 2017. Household structure and gender differences in travel time: spouse/partner presence, parenthood, and breadwinner status. Transportation 44(2), 271-291.

Fernandes, P. and Nunes, U., 2012. Platooning with IVC-enabled autonomous vehicles: Strategies to mitigate communication delays, improve safety and traffic flow. IEEE Transactions on Intelligent Transportation Systems, 13(1), pp.91-106.

Ford Motor Company, 2018. Looking further: Ford will have a fully autonomous vehicle in operation by 2021. [online] Available at: http://corporate.ford.com/innovation/autonomous-2021.html [Accessed 02/07/2018].

Franzen, A., Vogl, D., 2013. Two decades of measuring environmental attitudes: A comparative analysis of 33 countries. Global Environmental Change 23(5), 1001-1008.

Frei, C., Mahmassani, H.S. and Frei, A., 2015. Making time count: Traveler activity engagement on urban transit. Transportation Research Part A, 76, 58-70.

Frei, C., Hyland, M., and Mahmassani, H. S., 2017. Flexing service schedules: Assessing the potential for demand-adaptive hybrid transit via a stated preference approach. Transportation Research Part C: Emerging Technologies, 76, 71-89. doi:10.1016/j.trc.2016.12.017

Garikapati, V.M., Pendyala, R.M., Morris, E.A., Mokhtarian, P.L. and McDonald, N., 2016. Activity patterns, time use, and travel of millennials: a generation in transition?. Transport Reviews, 36(5), pp.558-584.

Gerte, R., Konduri, K.C., Eluru, N., 2018. Is there a limit to adoption of dynamic ridesharing systems? Evidence from analysis of Uber demand data from New York. Presented at the 97th Annual Meeting of the Transportation Research Board, Washington, D.C., January, Paper No. 18-03265.

Gibbs, N., 2017. Automakers face big, costly challenges to make self-driving cars a reality. Automotive News Europe Monthly Magazine [online]. Available at: http://europe.autonews.com/article/20170604/ANE/170529784/automakers-face-big-costly-challenges-to-make-self-driving-cars-a [Accessed 02/07/2018]

Gifford, R., Nilsson, A., 2014. Personal and social factors that influence pro-environmental concern and behaviour: A review. International Journal of Psychology 49(3), 141-157.

Gilg, A., Barr, S., Ford, N., 2005. Green consumption or sustainable lifestyles? Identifying the sustainable consumer. Futures 37(6), 481-504.

Giuliano, G., 2003. Travel, location and race/ethnicity. Transportation Research Part A: Policy and Practice, 37(4), pp.351-372.

Gourieroux, C., Monfort, A. and Trognon, A., 1984. Pseudo maximum likelihood methods: Theory. Econometrica: Journal of the Econometric Society, pp.681-700.

Gripsrud, M. and Hjorthol, R., 2012. Working on the train: from 'dead time'to productive and vital time. Transportation, 39(5), pp.941-956.

Gutiérrez, J.L.G., Jiménez, B.M., Hernández, E.G., Pcn, C., 2005. Personality and subjective well-being: Big five correlates and demographic variables. *Personality and Individual Differences* 38(7), 1561-1569.

Haboucha, C.J., Ishaq, R. and Shiftan, Y., 2017. User preferences regarding autonomous vehicles. Transportation Research Part C: Emerging Technologies, 78, pp.37-49.

Hampshire, R.C., Simek, C., Fabusuyi, T., Di, X. and Chen, X., 2017. Measuring the Impact of an Unanticipated Suspension of Ride-sourcing in Austin, Texas. https://ssrn.com/abstract=2977969 or http://dx.doi.org/10.2139/ssrn.2977969. [Accessed 07/26/2017]

Hatuka, T. and Toch, E., 2016. The emergence of portable private-personal territory: Smartphones, social conduct and public spaces. *Urban Studies*, 53(10), 2192-2208.

Haustein, S., 2012. Mobility behavior of the elderly: an attitude-based segmentation approach for a heterogeneous target group. Transportation 39(6), 1079-1103.

Helsper, E.J., Eynon, R., 2010. Digital natives: Where is the evidence? British Educational Research Journal 36(3), 503-520.

Hensher, D.A., 2010. Hypothetical bias, choice experiments and willingness to pay. Transportation Research Part B: Methodological, 44(6), pp.735-752.

Hensher, D.A. and Greene, W.H., 2011. Valuation of travel time savings in WTP and preference space in the presence of taste and scale heterogeneity. Journal of Transport Economics and Policy (JTEP), 45(3), pp.505-525.

Hensher, D.A., Greene, W.H. and Rose, J.M., 2006. Deriving willingness-to-pay estimates of travel-time savings from individual-based parameters. Environment and Planning A, 38(12), pp.2365-2376.

Hensher, D.A., Rose, J.M. and Greene, W.H., 2005. Applied choice analysis: a primer. Cambridge University Press.

Hess, S., Bierlaire, M. and Polak, J.W., 2005. Estimation of value of travel-time savings using mixed logit models. Transportation Research Part A: Policy and Practice, 39(2-3), pp.221-236.

Hoeffler, S., Ariely, D., 1999. Constructing stable preferences: A look into dimensions of experience and their impact on preference stability. Journal of Consumer Psychology 8(2), 113-139.

Howard, D. and Dai, D., 2014. Public perceptions of self-driving cars: The case of Berkeley, California. In Transportation Research Board 93rd Annual Meeting (Vol. 14, No. 4502).

Jittrapirom, P., Caiati, V., Feneri, A.M., Ebrahimigharehbaghi, S., González, M.J.A. and Narayan, J., 2017. Mobility as a Service: a critical review of definitions, assessments of schemes, and key challenges. Urban Planning, 2(2), p.13.

Klein, N.J., Guerra, E., Smart, M.J., 2018. The Philadelphia story: Age, race, gender and changing travel trends. Journal of Transport Geography 69, 19-25.

Komanduri, A., Wafa, Z., Proussaloglou, K., 2018. Assessing the impact of app-based ride share systems in an urban context: Findings from Austin. Presented at the 97th Annual Meeting of the Transportation Research Board, Washington, D.C., January, Paper No. 18-01356.

Kooti, F., Grbovic, M., Aiello, L.M., Djuric, N., Radosavljevic, V., Lerman, K., 2017. Analyzing Uber's ride-sharing economy. Proceedings of the 26th International Conference on World Wide Web Companion, pp. 574-582. International World Wide Web Conferences Steering Committee.

Krueger, R., Rashidi, T.H. and Rose, J.M., 2016. Preferences for shared autonomous vehicles. Transportation research part C: emerging technologies, 69, pp.343-355.

Kuss, D.J., Kanjo, E., Crook-Rumsey, M., Kibowski, F., Wang, G.Y. and Sumich, A., 2018. Problematic mobile phone use and addiction across generations: The roles of psychopathological symptoms and smartphone use. Journal of Technology in Behavioral Science, pp.1-9.

Kyriakidis, M., Happee, R. and de Winter, J.C., 2015. Public opinion on automated driving: Results of an international questionnaire among 5000 respondents. Transportation research part F: traffic psychology and behaviour, 32, pp.127-140.

Lavieri, P.S., Dias, F.F., Juri, N.R., Kuhr, J., Bhat, C.R., 2018. A model of ridesourcing demand generation and distribution. *Transportation Research Record: Journal of the Transportation Research Board.*

Lavieri, P.S., Garikapati, V.M., Bhat, C.R., Pendyala, R.M., Astroza, S. and Dias, F.F., 2017. Modeling individual preferences for ownership and sharing of autonomous vehicle technologies. *Transportation Research Record: Journal of the Transportation Research Board*, 2665, 1-10.

Leistner, D.L., Steiner, R.L., 2017. Uber for seniors? Exploring transportation options for the future. Transportation Research Record: Journal of the Transportation Research Board 2660, 22-29.

Levin, M.W., Kockelman, K.M., Boyles, S.D. and Li, T., 2017. A general framework for modeling shared autonomous vehicles with dynamic network-loading and dynamic ride-sharing application. Computers, Environment and Urban Systems, 64, pp.373-383.

Li, Z., Hong, Y., Zhang, Z., 2017. An empirical analysis of on-demand ride sharing and traffic congestion. Presented at the 2016 International Conference on Information Systems (ICIS 2016), Dublin, September.

Liu, N., Yu, R., 2017. Identifying design feature factors critical to acceptance and usage behavior of smartphones. Computers in Human Behavior 70, 131-142.

Liu, X., Vedlitz, A., Shi, L., 2014. Examining the determinants of public environmental concern: Evidence from national public surveys. Environmental Science & Policy 39, 77-94.

Lyons, G. and Urry, J., 2005. Travel time use in the information age. Transportation Research Part A: Policy and Practice, 39(2-3), pp.257-276.

Maddala, G.S., 1983. *Limited-Dependent and Qualitative Variables in Econometrics* (No. 3). Cambridge University Press, Cambridge, MA.

Malokin, A., Circella, G. and Mokhtarian, P.L., 2015. How do activities conducted while commuting influence mode choice? Testing public transportation advantage and autonomous vehicle scenarios (No. 15-1179).

Malokin, A., Circella, G. and Mokhtarian, P.L., 2017. Do multitasking millennials value travel time differently? a revealed preference study of northern california commuters(No. 17-00891).

Martinez, L.M., Correia, G.H. and Viegas, J.M., 2015. An agent-based simulation model to assess the impacts of introducing a shared-taxi system: an application to Lisbon (Portugal). Journal of Advanced Transportation, 49(3), pp.475-495.

Milojev, P., Sibley, C.G., 2017. Normative personality trait development in adulthood: A 6-year cohort-sequential growth model. *Journal of Personality and Social Psychology* 112(3), 510.

Mokhtarian, P.L. and Chen, C., 2004. TTB or not TTB, that is the question: a review and analysis of the empirical literature on travel time (and money) budgets. Transportation Research Part A: Policy and Practice, 38(9-10), pp.643-675.

Morahan-Martin, J. and Schumacher, P., 2007. Attitudinal and experiential predictors of technological expertise. Computers in Human Behavior, 23(5), pp.2230-2239.

Morales Sarriera, J., Escovar Álvarez, G., Blynn, K., Alesbury, A., Scully, T. and Zhao, J., 2017. To Share or Not To Share: Investigating the Social Aspects of Dynamic Ridesharing. Transportation Research Record: Journal of the Transportation Research Board, (2605), pp.109-117.

Motte-Baumvol, B., Bonin, O. and Belton-Chevallier, L., 2017. Who escort children: Mum or dad? Exploring gender differences in escorting mobility among Parisian dual-earner couples. Transportation, 44(1), 139-157.

Neal, M.B. and Hammer, L.B., 2017. Working couples caring for children and aging parents: Effects on work and well-being. Psychology Press.

Nikitas, A., Kougias, I., Alyavina, E. and Njoya Tchouamou, E., 2017. How Can Autonomous and Connected Vehicles, Electromobility, BRT, Hyperloop, Shared Use Mobility and Mobility-As-A-Service Shape Transport Futures for the Context of Smart Cities?. Urban Science, 1(4), p.36.

NHTS, 2017. 2017 National Household Travel Survey, U.S. Department of Transportation, Federal Highway Administration. https://nhts.ornl.gov

NVIDIA, 2018. NVIDIA, AUDI Partner to Put World's Most Advanced AI Car on Road by 2020 [online]. Available at: https://nvidianews.nvidia.com/news/nvidia-audi-partner-to-put-world-s-most-advanced-ai-car-on-road-by-2020 [Accessed 02/07/2018]

Ory, D.T. and Mokhtarian, P.L., 2005. When is getting there half the fun? Modeling the liking for travel. Transportation Research Part A: Policy and Practice, 39(2-3), pp.97-123.

Paleti, R., Eluru, N., Bhat, C.R., 2010. Examining the influence of aggressive driving behavior on driver injury severity in traffic crashes. *Accident Analysis & Prevention* 42(6), 1839-1854.

Ralph, K.M., 2017. Multimodal millennials? The four traveler types of young people in the United States in 2009. Journal of Planning Education and Research, 37(2), pp.150-163.

Rayle, L., Dai, D., Chan, N., Cervero, R. and Shaheen, S., 2016. Just a better taxi? A survey-based comparison of taxis, transit, and ridesourcing services in San Francisco. Transport Policy, 45, pp.168-178.

Richardson, M., Petrescu, A. and Finch., M., 2016. Event-based ridesharing. http://appft1.uspto.gov/netacgi/nph-Parser?Sect1=PTO1&Sect2=HITOFF&d=PG01&p=1&u=/netahtml/PTO/srchnum.html&r=1&f=G&l=50&s1=20160026936.PGNR. [Accessed: 06/30/2018]

RideAustin. Data File and Code Book. Available at: https://data.world/ride-austin/ride-austin-june-6-april-13. Accessed July 26, 2017.

Rieser-Schüssler, N., Axhausen, K., 2012. Investigating the influence of environmentalism and variety seeking on mode choice. Transportation Research Record: Journal of the Transportation Research Board 2322, 31-41.

Santos, D.O. and Xavier, E.C., 2015. Taxi and ride sharing: A dynamic dial-a-ride problem with money as an incentive. Expert Systems with Applications, 42(19), pp.6728-6737.

Schaefers, T., 2013. Exploring Carsharing Usage Motives: A Hierarchical Means-End Chain Analysis. Transportation Research Part A, Vol. 47, pp. 69-77.

Schwartz, S.H., 1992. Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. *Advances in Experimental Social Psychology* 25, 1-65.

Schwartz, S.H., Rubel, T., 2005. Sex differences in value priorities: Cross-cultural and multimethod studies. *Journal of Personality and Social Psychology*, 89(6), 1010-1028.

Schwartz, S.H., Melech, G., Lehmann, A., Burgess, S., Harris, M., Owens, V., 2001. Extending the cross-cultural validity of the theory of basic human values with a different method of measurement. *Journal of Cross-Cultural Psychology* 32(5), 519-542.

Seebauer, S., Stolz, R. and Berger, M., 2015. Technophilia as a driver for using advanced traveler information systems. Transportation Research Part C: Emerging Technologies, 60, pp.498-510.

Shaheen, S.A., Chan, N.D. and Gaynor, T., 2016. Casual carpooling in the San Francisco Bay Area: Understanding user characteristics, behaviors, and motivations. Transport Policy, 51, pp.165-173.

Siddiqi, Z. and Buliung, R., 2013. Dynamic ridesharing and information and communications technology: past, present and future prospects. Transportation Planning and Technology, 36(6), pp.479-498.

Singh, P., Paleti, R., Jenkins, S. and Bhat, C.R., 2013. On modeling telecommuting behavior: option, choice, and frequency. Transportation, 40(2), pp.373-396.

Smart, M.J., 2015. A nationwide look at the immigrant neighborhood effect on travel mode choice. Transportation, 42(1), pp.189-209.

Smith, A. 2016. Shared, collaborative and on demand: the new digital economy. Pew research Center [online]. Available at: http://www.pewinternet.org/2016/05/19/the-new-digital-economy/ [Accessed 02/07/2018].

Sivakumar, A. and Bhat, C., 2002. Fractional split-distribution model for statewide commodity-flow analysis. Transportation Research Record: Journal of the Transportation Research Board, (1790), pp.80-88.

Spears, S., Houston, D. and Boarnet, M.G., 2013. Illuminating the unseen in transit use: A framework for examining the effect of attitudes and perceptions on travel behavior. Transportation Research Part A: Policy and Practice, 58, pp.40-53.

Stamatakis, E., de Rezende, L.F., Rey-López, J.P., 2018. Sedentary behaviour and cardiovascular disease. In *Sedentary Behaviour Epidemiology*, pp. 215-243, Springer.

Statista, 2018a. Number of Uber users in the United States as of April 2017, by device (in millions) [online]. Available at: https://www.statista.com/statistics/715236/us-uber-users-by-device/ [Accessed 02/07/2018]

Statista, 2018b. North American carsharing member growth from 2000 to 2016 (in 1,000s) [online]. Available at: https://www.statista.com/statistics/263847/carsharing-growth-members-in-north-america/ [Accessed 02/07/2018]

Statista, 2018c How often do you use bicycles you borrowed or rented, independently from whom? [online]. Available at: https://www.statista.com/statistics/699861/percentage-of-people-in-households-who-borrow-or-rent-bikes/ [Accessed 02/07/2018].

Texas Government, 2017. Schedule Data for All Fixed Route MetroBus Lines and MetroRail [online]. Available at: https://data.texas.gov/Transportation/GTFS/r4v4-vz24. [Accessed 07/26/ 2017].

Tahmasseby, S., Kattan, L. and Barbour, B., 2016. Propensity to participate in a peer-to-peer social-network-based carpooling system. Journal of Advanced Transportation, 50(2), pp.240-254.

Tientrakool, P., Ho, Y.C. and Maxemchuk, N.F., 2011, September. Highway capacity benefits from using vehicle-to-vehicle communication and sensors for collision avoidance. In Vehicular Technology Conference (VTC Fall), 2011 IEEE (pp. 1-5). IEEE.

Trading Economics, 2018. United States Part Time Employment https://tradingeconomics.com/united-states/part-time-employment [Accessed: 08/01/2018].

Train, K. and Weeks, M., 2005. Discrete choice models in preference space and willing-topay 592 space. Applications of simulation methods in environmental and, 593, pp.1-16.

Tudela, A., Habib, K.M.N., Carrasco, J.A., Osman, A.O., 2011. Incorporating the explicit role of psychological factors on mode choice: A hybrid mode choice model by using data from an innovative psychometric survey. Presented at 2nd International Choice Modelling Conference, Leeds.

Twenge, J.M., 2013. Does online social media lead to social connection or social disconnection? *Journal of College and Character* 14(1), 11-20.

Uber Blog. Uber Hits 5 Billion Rides Milestone. Available at www.uber.com/en-SG/blog/uber-hits-5-billion-rides-milestone/ [Accessed: 07/26/2017]

UberEstimator, 2017. Uber Estimator: Uber in Dallas-Fort Worth. Available at: https://uberestimator.com/cities/dallas [Accessed: 08/13/2017].

U.S. Census Bureau, 2018a. New Census Bureau Population Estimates Show Dallas-Fort Worth-Arlington Has Largest Growth in the United States. Release Number: CB18-50. Available at: https://www.census.gov/newsroom/press-releases/2018/popest-metro-county.html [Accessed: 03/23/2018].

U.S. Census Bureau, 2018b. Means of Transportation to Work by Selected Characteristics, Table S0802. 2012-2016 American Community Survey 5-Year Estimates. Available at: https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_16_5YR_S0802&prodType=table [Accessed: 03/23/2018].

U.S. Census Bureau, 2018c. Employment Status, Table S2301. 2012-2016 American Community Survey 5-Year Estimates. Available at: https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_16_5YR_S2301&prodType=table [Accessed: 03/23/2018].

USDOT-FTA, 2018. Shared Mobility Definitions [online]. Available at: https://www.transit.dot.gov/regulations-and-guidance/shared-mobility-definitions [Accessed 02/07/2018].
Van Acker, V., Mokhtarian, P.L., Witlox, F., 2014. Car availability explained by the structural relationships between lifestyles, residential location, and underlying residential and travel attitudes. *Transport Policy* 35, 88-99.
van Wee, B., 2014. The unsustainability of car use. *Handbook of Sustainable Travel*, pp. 69-83, Springer, Netherlands.
Varin, C., Vidoni, P., 2005. A note on composite likelihood inference and model selection. Biometrika, 92(3), 519-528.
Wang, Y., Winter, S. and Ronald, N., 2017. How much is trust: The cost and benefit of ridesharing with friends. Computers, Environment and Urban Systems, 65, pp.103-112.
Wang, Y., Zheng, B. and Lim, E.P., 2018. Understanding the effects of taxi ride-sharing—A case study of Singapore. Computers, Environment and Urban Systems, 69, pp.124-132.
Ward, J.S. and Barker, A., 2013. Undefined by data: a survey of big data definitions. arXiv preprint arXiv:1309.5821.
Waymo, 2018. Be an early rider [online]. Available at: https://waymo.com/apply/ [Accessed 02/07/2018].
Winkle, T., 2016. Safety Benefits of Automated Vehicles: Extended Findings from Accident Research for Development, Validation and Testing. In Autonomous Driving (pp. 335-364). Springer, Berlin, Heidelberg.
Wolf, A. and Seebauer, S., 2014. Technology adoption of electric bicycles: A survey among early adopters. Transportation Research Part A: Policy and Practice, 69, pp.196-211.
Ye, R., Titheridge, H., 2017. Satisfaction with the commute: The role of travel mode choice, built environment and attitudes. *Transportation Research Part D* 52, 535-547.
Zebrowitz, L.A., White, B., Wieneke, K., 2008. Mere exposure and racial prejudice: Exposure to other-race faces increases liking for strangers of that race. *Social Cognition* 26(3), 259-275.
Zhang, W., Guhathakurta, S., Fang, J. and Zhang, G., 2015. Exploring the impact of shared autonomous vehicles on urban parking demand: An agent-based simulation approach. Sustainable Cities and Society, 19, pp.34-45.
Zhang, W. and Guhathakurta, S., 2017. Parking spaces in the age of shared autonomous vehicles: How much parking will we need and where?. Transportation Research Record: Journal of the Transportation Research Board, (2651), pp.80-91.
Zheng, Y., Manson, J.E., Yuan, C., Liang, M.H., Grodstein, F., Stampfer, M.J., Willett, W.C., Hu, F.B., 2017. Associations of weight gain from early to middle adulthood with major health outcomes later in life. *JAMA* 318(3), 255-269.
Zheng, H., Chen, X., Chen, X.M., 2018. How does on-demand ridesplitting influence vehicle use and ownership? A case study in Hangzhou, China. Presented at the 97th Annual Meeting of the Transportation Research Board, Washington, D.C., January, Paper No. 18-04327.
Zmud, J., Lee-Gosselin, M., Munizaga, M. and Carrasco, J.A. eds., 2013. Transport survey methods: Best practice for decision making. Emerald Group Publishing Limited.
Zmud, J., Sener, I.N. and Wagner, J., 2016. Self-Driving Vehicles: Determinants of Adoption and Conditions of Usage. Transportation Research Record: Journal of the Transportation Research Board, (2565), pp.57-64.