Final Report

Investigation of the Role of Attitudinal Factors on the Adoption of Emerging Automated Vehicle and Safety Technologies

Prepared for **Teaching Old Models New Tricks (TOMNET) Transportation Center**

By,

Fred Mannering
Professor, Executive Director
Email: flm@usf.edu | ORCID: 0000-0002-2803-4582
Department of Civil and Environmental Engineering
University of South Florida, 4202 E. Fowler Avenue, Tampa, FL 33620

Michael Maness
Assistant Professor
Email: manessm@usf.edu | ORCID: 0000-0001-5780-8666
Department of Civil and Environmental Engineering
University of South Florida, 4202 E. Fowler Avenue, Tampa, FL 33620

With

Abdul Pinjari, Yu Zhang, Suryaprasanna Kumar Balusu, Natalia Barbour, Nikhil Menon, Parvathy Vinod Sheela and Divyakant Tahlyan

August 2018

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
N/A	N/A	N/A

4. Title and Subtitle	5. Report Date
Investigation of the Role of Attitudinal Factors on the Adoption of Emerging Automated Vehicle and Safety Technologies	August 2018
	6. Performing Organization Code N/A
7. Author(s) Fred Mannering, https://orcid.org/0000-0002-2803-4582 Michael Maness, https://orcid.org/0000-0001-5780-8666 Abdul Pinjari, https://orcid.org/0000-0002-3056-4259 Yu Zhang, Suryaprasanna Balusu, https://orcid.org/0000-0003-0726-4036 Natalia Barbour, https://orcid.org/0000-0002-0787-3993 Nikhil Menon, https://orcid.org/0000-0002-7035-204X Parvathy Vinod Sheela, https://orcid.org/0000-0001-8099-1235 Divyakant Tahlyan, https://orcid.org/0000-0002-1129-6172	**8. Performing Organization Report No.** N/A
9. Performing Organization Name and Address University of South Florida, 4202 E. Fowler Avenue, Tampa, FL 33620	**10. Work Unit No. (TRAIS)** N/A
	11. Contract or Grant No. 69A3551747116
12. Sponsoring Agency Name and Address U.S. Department of Transportation, University Transportation Centers Program, 1200 New Jersey Ave, SE, Washington, DC 20590	**13. Type of Report and Period Covered** Research Report (2017 – 2018)
	14. Sponsoring Agency Code USDOT OST-R

15. Supplementary Notes
N/A

16. Abstract
This research project addresses a series of issues relating to autonomous vehicle adoption, the potential impact of temporal instability (with an application to vehicle safety), and the role of social learning processes as they relate to future travel behavior. The report of research results begins with a study of the effect shared autonomous vehicles and their potential impacts on household vehicle ownership (Chapter 2). Focus is directed toward the potential impact that fleets of shared autonomous vehicles might have on household vehicle ownership. Chapter 3 provides a market-segmentation approach consumers' perceptions towards automated vehicles and their intended adoption. Chapter 4 provides a study of consumer use likelihoods and concerns with shared automated vehicles. Chapter 5 presents an extensive discussion of issues related to the potential temporal instability in statistical models with an application to the analysis of highway accident data. Chapter 6 provides a discrete choice modeling framework to differentiate opinion neutral and unsure responses in surveys with Likert scale attitudinal questions. Lastly, the project report concludes with Chapter 7 that looks at issues relating to the transferability of integrated choice and latent variable models.

17. Key Words	18. Distribution Statement
Autonomous Vehicles, Vehicle Ownership, Perceptions, Attitudes, Hybrid Choice Modeling, Travel Demand Forecasting	No restrictions.

19. Security Classif.(of this report)	20. Security Classif.(of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	139	N/A

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The report is funded, partially or entirely, by a grant from the U.S. Department of Transportation's University Transportation Centers Program. However, the U.S. Government assumes no liability for the contents or use thereof.

TABLE OF CONTENTS

Chapter 1

1.1 Introduction

There is considerable interest among transportation planners in forecasting the ownership, market penetration, and travel behavior impacts of automated vehicles. Similarly, transportation engineers are interested in predicting the safety and driving behavior impacts of vehicle safety technologies and how these might change over time. From a transportation systems planning standpoint, it is important to forecast potential future scenarios with automated vehicles, including the market penetration of these technologies, the different forms of ownership and usage (such as owning automated vehicles for personal use versus using automated vehicles as a shared mobility service), and the impact of these technologies on travel demand.

This research project addresses a series of issues relating to autonomous vehicle adoption, the potential impact of temporal instability (with an application to vehicle safety), and the role of social learning processes as they relate to future travel behavior. The report of research results begins with a study of the effect shared autonomous vehicles and their potential impacts on household vehicle ownership (Chapter 2). Focus is directed toward the potential impact that fleets of shared autonomous vehicles might have on household vehicle ownership. To obtain initial insights into this issue, a sample of university personnel and members of the American Automobile Association is used to determine how likely they would be to consider relinquishing one of their household's personal vehicles if shared autonomous vehicles were available (thus reducing their household vehicle ownership level by one). For single-vehicle households, this would be relinquishing their only vehicle, and for multi-vehicle households (households owning two or more vehicles) this would be relinquishing just one of their vehicles. Possible responses to the question about relinquishing a household vehicle if shared autonomous vehicles are present are: extremely unlikely, unlikely, unsure, likely, and extremely likely. To determine the factors that influence this response, random parameters ordered probit models are estimated to account for the likelihood that considerable unobserved heterogeneity is likely to be present in the data. The findings show a wide range of socio-economic factors affects people's likelihood of vehicle relinquishment in the presence of shared autonomous vehicles. Key among these are gender effects, generational elements, commuting patterns, and respondents' vehicle crash history and experiences. While people's opinions of shared autonomous vehicles are evolving with the continual introduction of new autonomous-vehicle technologies and shifting travel behavior, the results of this study provide important initial insights into the likely effects of shared autonomous vehicles on household vehicle ownership.

Chapter 3 provides a market-segmentation approach consumers' perceptions towards automated vehicles and their intended adoption. Using extensive survey data, cluster analysis is applied to better understand consumers' perceptions toward potential benefits and concerns relating to automated vehicles with regard to factors influencing their autonomous-vehicle adoption likelihoods. Four market segments are identified which and are classified as benefits-dominated, concerns dominated, uncertain, well-informed. A random parameters multinomial logit model is then estimated to identify factors influencing the probability of respondents belonging to one of these four specific market segments. Among other influences (such as socio-economic, and current travel characteristics), it is found that millennials had a higher probability of belonging to the well-informed market segment, Gen-X-ers have a lower probability of belonging to the uncertain market segment, and baby boomers had a higher probability of belonging to the concerns-dominated market (relative to the golden generation). To gain a further understanding, a study of individuals' expressed likelihood of autonomous vehicle adoption using separate random parameters ordered probit estimations for each of the four market segments is undertaken. The substantial and statistically significant differences across each automated-vehicle consumer market segment underscores the potentially large impact that different consumer demographics may have on new technology adoption, and the need for targeted marketing to achieve better market-penetration outcomes with regard to autonomous vehicles.

Chapter 4 provides a study of consumer use likelihoods and concerns with shared automated vehicles. Because shared automated vehicles could be a disruptive transportation modal alternative, understanding the factors that may affect the likelihood of using and possible concerns is extremely important. To do so, this chapter uses a survey of American Automobile Association members to ask whether or not survey respondents were willing to use shared automated vehicles if they became available, and members are also asked their main concerns associated with this technology (safety, privacy, reliability, travel time or travel cost). Two random-parameter logit models were estimated to gain insights into the likely usage/concerns processes. Some of the key variables playing statistically significant roles in the willingness to use of shared automated vehicles process were ethnicity, household size, daily travel times, and vehicle crash history. With regard to shared automated vehicle concerns, we were also able to identify the characteristics of respondents who were more or less likely to be concerned with safety, reliability, privacy, and travel time/travel cost. While the opinions and perceptions towards shared automated vehicles are likely to fluctuate in the coming years as more and more information relating to the potential of such sharing becomes available, the findings in this chapter provide an important initial assessment before this technology becomes widely available to the public. The more that is known about shared automated vehicles and their early adopters, the better and seamless the potential modal transition can be. Learning what groups of people are more or less willing to use this technology will help to improve the overall mobility of all. Combining the significant variables provides a rough profile description of early users of shared automated vehicles and their environment. This helps to prioritize possible investments (urban vs. rural) and allows the policy and auto makers to identify the critical needs of the users. This initial assessment provides the characteristics of early adopters and their travel behavior. The model estimation results clearly show that different socio-demographic groups value different aspects and have different concerns relating to shared automated vehicles.

Chapter 5 presents an extensive discussion of issues related to the potential temporal instability in statistical models with an application to the analysis of highway accident data. With regard to vehicle safety, virtually every statistical analysis of highway safety data is predicated on the assumption that the estimated model parameters are temporally stable (the same is true of travel demand models). That is, in the case of vehicle safety, the assumption that the effect of the determinants of accident likelihoods and resulting accident-injury severities do not change over time. The material in this chapter draws from research previously conducted in fields such as psychology, neuroscience, economics, and cognitive science to build a case for why we would not necessarily expect the effects of explanatory variables to be stable over time. The review of this literature suggests that temporal instability is likely to exist for a number of fundamental behavioral reasons, and this temporal instability is supported by the findings of several recent accident-data analyses. The chapter goes on to discuss the implications of this temporal instability on contemporary accident-data modeling methods (unobserved heterogeneity, data driven, traditional, and causal inference methods) and concludes with a discussion of how temporal instability might be addressed and how its likely presence can be used to better interpret accident data-analysis findings.

Chapter 6 provides a discrete choice modeling framework to differentiate opinion neutral and unsure responses in surveys with Likert scale attitudinal questions. Socio-psychological factors, namely perceptions, attitudes, beliefs, and norms, are often measured using Likert scale questions. These Likert scales responses are modeled on a continuum from one extreme (e.g. unlikely) to another extreme (e.g. likely), and thus the middle/neutral response acts as a transition point between the two polar options. Psychometric research has found that the neutral group of respondents is not homogeneous and does not act as a transition group between extremes for all respondents. The middle option can also be chosen either for expressing a lack of knowledge or opinion. Capturing this heterogeneity can have important considerations for policy analysis and forecasting as well as the design of information awareness campaigns. In this work, a framework is developed to distinguish opinion neutrality from lack of knowledge/opinion using integrated choice and latent variable models. A case study on intended autonomous vehicle use is used to explore the framework's properties, since familiarity with AVs is not high among the general public due to the novelty of the concept. Using 1245 responses from AAA–South members, the framework was able to clearly distinguish between neutrality-familiarity types and associate each with group with sociodemographic patterns. The chapter concludes by describing how the framework is flexible enough to model neutrality and familiarity in Likert scales without a neutral option and Likert scales that include a "no opinion" option. Additionally, the framework can be modified to use discrete latent classes rather than continuous latent variables.

Lastly, the project report concludes with Chapter 7 that looks at issues relating to the transferability of integrated choice and latent variable models. It has been postulated that travel forecasting models that are more behaviorally realistic and those that better capture heterogeneity in travel behavior are more transferable than traditional models. Incorporating attitudes and perception variables into discrete choice models is most widely performed using integrated choice and latent variable (ICLV) models. ICLV models offer greater insights into the decision making process by including additional information through measurement equations for the latent variables. Existing work has examined whether ICLV models are more behaviorally sound and offer better predictions, but the spatial transferability of ICLV models has not been fully explored. This paper focuses on assessing the impact of incorporating attitudinal and perception variables on the spatial transferability of travel forecasting models. Specifically, this paper compares the spatial transferability, in an empirical setting, for two model structures: multinomial logit (MNL) and ICLV models. From this study, it is anticipated that incorporating attitudinal/perception variables through ICLV model structure will improve transferability across regions. In a case study of intended usages of autonomous vehicles, this study had mixed results in finding improvements from using transferred ICLV models versus locally estimated MNL models.

Chapter 2: Shared Autonomous Vehicles and their Potential Impacts on Household Vehicle Ownership: An Exploratory Empirical Assessment

2.1 Introduction

Emerging automotive and transportation technologies, such as autonomous vehicles, have created revolutionary possibilities with regard to future travel. Several prominent automotive and technology companies have presented their versions of autonomous vehicles, and are predicting that autonomous vehicle technology, with the capability of being fully self-driving, will be available to the general public in the near future (Fagnant and Kockelman, 2015a; Menon et al., 2016). With fully self-driven vehicles, users may not need to be engaged in the driving process and could, therefore, be involved a host of other activities such as working, talking to friends, sleeping or reading (Le Vine et al., 2015).

As the technological development is progressing rapidly, governmental agencies are grappling with how to plan transportation systems for such technologies. Considering the high initial cost of owning these technologies, there is a significant discussion on the possible emergence of shared autonomous vehicle fleets as an alternative to owning individual autonomous vehicles. Testing of shared autonomous vehicles has gathered momentum with Uber, nuTonomy, and Lyft evaluating these technologies on city streets (Bliss, 2016; Boston, 2017). Additionally, the entry of innovative transit companies such as Navya and EasyMile into college campuses and cities for testing and research purposes is further evidence of the growing interest in shared autonomous vehicles (Hawkins, 2017; Motion Digest, 2017). Shared autonomous vehicles have the potential to be an inexpensive on-demand mobility service that could play a key role in the future transportation systems. For instance, shared autonomous vehicles could provide convenient last-mile (transporting people from transit drop-offs to final destinations) solutions to support multimodal transportation systems (Krueger et al., 2016). In fact, recent literature modeling different scenarios with shared autonomous vehicle fleets show significant cost benefits in comparison to individually owned and operated vehicles (Fagnant and Kockelman, 2015a).

Past studies on understanding household vehicle ownership trends have provided interesting insights on what triggers the acquisition as well as the relinquishment of vehicles. There has been a downward trend in vehicle purchases over the last few years among younger generations (Millard-Ball and Schipper, 2011) and, over the years, the influence of life events on household vehicle relinquishments has been well documented (Dargay and Hanly, 2007; Oakil et al., 2014; Clark et al., 2015). Even without automation, there is increasing evidence that the emergence of vehicle-sharing services is leading to a reduction in household vehicle ownership (Martin et al., 2010; Elliott and Shaheen, 2011). For instance, individuals who currently own vehicles out of necessity, rather than preference, are likely to switch to vehicle-sharing (Ohta et al., 2013), if provided at a cost comparable to owning a personal vehicle. There is an increasing possibility of higher levels of vehicle relinquishment at the household level when technologies take the task of driving away from the driver.

Recent news on the emergence of popular vehicle-sharing services such as Uber and Lyft (Kosoff, 2016), have supported the need to understand possible shifts in household vehicle ownership trends with the introduction of shared autonomous vehicles. While a relatively large number of previous studies have focused on understanding people’s preferences for autonomous vehicles and their intended adoption (Schoettle and Sivak, 2014; Menon et al., 2016), only a few studies have explicitly dealt with the adoption of shared autonomous vehicles. Examples include

Haboucha et al., (2015), who conducted a stated preference questionnaire to 800 individuals living in Israel and North America to develop a joint ownership and choice model that included shifting to a fleet of shared autonomous vehicles among other options (retain vehicle, buy and ride in an autonomous vehicle). And Bansal et al., (2016), who analyzed individuals' frequency of use of shared autonomous vehicles under different pricing scenarios and identified characteristics of potential shared autonomous vehicle users. Furthermore, studies generally do not explicitly address households' tendency to relinquish vehicles in the presence of shared autonomous vehicles. Yet, people's willingness to relinquish household vehicles in the presence of shared autonomous vehicles is a key to the success of shared autonomous vehicle systems. Therefore, the objective of this study is to understand the factors influencing households' intentions to relinquish their own vehicles in the presence of shared autonomous vehicles.

To this end, we conduct a survey of two different target groups of interest: faculty, students, and staff from a large university (University of South Florida); and the members of the American Automobile Association (AAA) Foundation of the southeastern United States. We develop a survey instrument asking them how likely they would be to consider relinquishing one of their household's personal vehicles if shared autonomous vehicles were available (thus reducing their household vehicle ownership level by one). University members were chosen because universities are often a fertile ground for testing and early adoption of new technologies. Additionally, university respondents are often some of the earliest adopters (and sometimes vocal critics) of emerging technologies thereby making them an interesting demographic to consider for the purpose of this chapter. Also, with approximately every one in four households in the United States being AAA members (American Automobile Association, 2017), the results from this study would be representative of a broad cross-section of American society. However, it is important to note that the intent of this chapter is exploratory and, as such, we do not seek a nationally representative sample. In a time when opinions and attitudes toward autonomous vehicles are changing rapidly as technologies advance and consumers process available information, even a fully representative national sample would provide findings that would not be temporally stable. Thus our focus on a select sub-sample of potential consumers is intended to provide some initial insights and a demonstration of a methodological approach that can be used to guide future studies on the subject.

Possible responses to the question of interest relating consumer intentions to relinquish their own vehicles in the presence of shared autonomous vehicles are: extremely unlikely, unlikely, unsure, likely, and extremely likely. For single-vehicle households, this would be relinquishing their only vehicle, and for multi-vehicle households (households owning two or more vehicles) this would be relinquishing one of their vehicles. Therefore, two different random parameters ordered probit models are estimated to analyze the factors that influence the households' likelihood of relinquishing one of their vehicles; one model for single-vehicle households and the other model for multi-vehicle households. While people's opinions of shared autonomous vehicles will likely evolve (as well as fluctuate) with the increasing penetration of new autonomous vehicle technologies and the realization of their benefits (or negative impacts), the model results provide important initial insights into the likely effects of shared autonomous vehicles on household vehicle ownership in the short term.

The remainder of this chapter starts, in section 2.2, with an assessment of recent trends in vehicle acquisition and relinquishment and goes on, in Section 2.3, to a discussion of ideas relating to shared autonomous vehicles and their potential impacts on vehicle ownership. Section 2.4 describes the data used for the analysis. Section 2.5 presents the random parameters ordered probit modeling methodology used to study possible household vehicle relinquishment. Section 2.6

discusses the statistical results, and Section 2.7 deliberates their implications for vehicle ownership (vehicle relinquishment, to be precise) in a shared-autonomous-vehicle environment. Section 2.8 concludes the chapter.

2.2. Vehicle Ownership Trends

Since the turn of the millennium, vehicle ownership levels have seen a steady decline among the young (Millard-Ball and Schipper, 2011; Kuhnimhof et al., 2013; Metz, 2013). Recent studies have shown that this growing trend among millennials (those who are born in the 1980s and 1990s) would make them own fewer vehicles, drive less and be less likely to obtain driving licenses (Polzin et al., 2014). The reasons for this decline in vehicle purchases have been attributed to many factors including changing preferences in urban living, increased transit use, increased environmental awareness, and shifting economic circumstances (McDonald, 2015; van Wee, 2015). While several studies have pointed to the role of new technologies in reducing travel (Martin et al., 2010) and therefore a decline in vehicle ownership levels (van Wee, 2015), others take the more skeptical view that new technologies can often create new travel demand and more travel, not less (Mokhtarian, 2002, 2009; Blumenberg et al., 2012).

Past research has shown that the acquisition and relinquishment of motorized vehicles is a complex intertemporal decision-making process (Mannering and Winston, 1985) and can often be the result of a life-changing event that typically leads to changes in travel behavior and vehicle utilization (Dargay and Hanly, 2007; Beige and Axhausen, 2012; Chatterjee et al., 2013; Clark et al., 2015). As an example, Oakil et al. (2014) examined households in the Netherlands and found an association between vehicle relinquishments and childbirth in households. Another study by Zhang et al. (2014) conducted in Japan shows how vehicle ownership changes are influenced by residential moves than by changes in education or employment. Other studies show the complex influence of household-level changes (job relocation of family members, presence of children, household member(s) leaving the household, and so on) and travel attributes on the decision of buying and selling vehicles (Rashidi et al., 2011).

2.3. Vehicle Ownership in the Presence of Shared Autonomous Vehicles

Vehicle-sharing is considered a flexible mobility option that offers the flexibility of a private vehicle without the responsibilities associated with private vehicle ownership (Shaheen and Cohen, 2013). The potential benefits envisioned with vehicle-sharing include the facilitation of multi-modal travel behavior (Nobis, 2006) and eventually the reduction in vehicle ownership levels (Martin et al., 2010; Firnkorn and Muller, 2012).

Vehicle-sharing with autonomous vehicles has the potential to revolutionize travel with respect to conventional vehicle- and ride-sharing paradigms. Because shared autonomous vehicles will be able to drive up to potential passengers, walking times to access shared vehicles could potentially be almost reduced to zero. Conventional vehicle-sharing has suffered from availability concerns for one-way vehicle-sharing users because there may not always be a vehicle available for use at the destination once travelers finish their activity. Thus, conventional vehicle sharing requires substantial labor costs to rebalance the potential mismatch of supply and demand. A shared autonomous vehicle-based vehicle-sharing model has the potential to avoid such issues (Fagnant and Kockelman, 2014; Firnkorn and Muller, 2015).

Ridesharing with a shared autonomous vehicle fleet could alleviate many of the adverse environmental impacts of current on-demand mobility services. For example, a recent simulation-based study of a shared autonomous vehicle fleet in Austin, Texas (Fagnant and Kockelman, 2015b) showed that the excess vehicle kilometers traveled due to empty vehicle relocation could be reduced by almost 50% with shared autonomous vehicle ridesharing relative to current, conventional ridesharing services. In addition, implementing ridesharing services with the use of shared autonomous vehicles would eliminate the transaction costs involved with having a driver operate the vehicle from origin to destination (Krueger et al., 2016).

While there is ample literature on the potential users of autonomous vehicles, there is substantially less information on potential user groups when it comes to shared autonomous vehicles. Past research points towards shared autonomous vehicles becoming an attractive mobility option for subgroups of the population such as the elderly or individuals who are currently unwilling and/or unable to drive (Rosenbloom, 2001; Alsnih and Hensher, 2003; Fagnant and Kockelman, 2015a; Shaheen et al., 2016). For example, research by Sikder and Pinjari (2012) found that while elderly may become immobile due to physical and cognitive limitations, their desire to continue to be mobile remains. Thus, shared autonomous vehicles could act as an elderly mobility alternative with the possibility of providing convenient and flexible mobility at a lower cost without the burden of driving. It should be pointed out, however, that it has been shown that population subgroups, such as elderly cohorts, are highly heterogeneous and vary considerably with respect to their motives for travel and the use of different modes (Haustein, 2012). In addition to the elderly, shared autonomous vehicles could be thought of as an age-appropriate mobility alternative for travelers who do not have access to private transportation, regardless of their age (Anderson et al., 2014; Krueger et al., 2016).

There is very little academic literature on the impact of shared autonomous vehicles on future household vehicle ownership trends in terms of both acquisitions and relinquishments. Although, recent discussions on potential vehicle ownership impacts have been fueled by the investment of vehicle-sharing and ride-sharing companies like Uber and Lyft in the autonomous vehicle market. With regard to the impacts of the emerging shared-autonomous-vehicle business models on future vehicle ownership, Lyft predicts that vehicle ownership will all but end by 2025 (Kosoff, 2016). And, Jaynes (2016) provides a comprehensive discussion on this topic by explaining the various scenarios that may arise regarding vehicle ownership in a driverless era. For example, Jaynes argues that it is very likely that the ownership model will never change for luxury vehicle buyers. However, it seems likely that luxury vehicle brands may start offering different ownership programs to cater to a driverless world, besides the traditional model of full ownership, with a more flexible fractional ownership model where the people pay a price depending on their usage. Other possible models of ownership that would arise in a driverless world with shared autonomous vehicles could include an own-plus-share model where people could still be tied to the traditional vehicle ownership but be able to opt into a sharing program where their vehicles would autonomously drive and chauffeur people around during its idle time (Jaynes, 2016).

From a market-impact perspective, a number of studies have found that shared autonomous vehicles have the potential to displace conventional vehicles (Wang et al., 2006; Spieser et al., 2014; Fagnant and Kockelman, 2014), but the magnitude of this displacement has been estimated to vary widely and is not well understood. Still, individuals' willingness to relinquish their conventional household vehicles in the presence of available shared autonomous vehicles is critical to measuring the impact and success of shared autonomous vehicles.

Given the above discussions, it is clear that future household vehicle ownership decisions in the presence of shared autonomous vehicles are going to be complex, and involve individual perceptions with regard to technology, potential benefits, likely costs, and so on. The objective of the current chapter is to develop some insights into these decisions by studying the willingness of people to relinquish a currently held household vehicle when shared autonomous vehicles become available.

2.4. Data

To understand the factors that may influence people's willingness to relinquish a household-owned vehicle in the presence of shared autonomous vehicles (thus reducing their household vehicle ownership level by one), a web-based survey was conducted to target population groups. The first targeted group is the students, faculty, and staff of the University of South Florida (USF) system (all three campuses – Tampa, St. Petersburg, and Sarasota-Manatee), and the second targeted group is members of American Automobile Association (AAA) South. The customized surveys consisting of 94 (USF) and 75 questions (AAA) were disseminated for data collection during April and June 2015, respectively. Some university related questions, such as working status at university, international student, on-campus residence, university campus were removed in AAA survey questionnaire. Meanwhile, additional questions on number of children in the household, and when the most severe crash occurred, were added into the AAA questionnaire at the request of AAA personnel (for the analysis of travel-related matters of interest to their association).

Part A of the survey collected general information including respondent demographics, current travel characteristics, crash history, and vehicle inventories. Part B elicited information on consumers' perceptions of autonomous vehicles. Questions included respondent familiarity with autonomous vehicles, likelihood of certain benefits and concerns with autonomous vehicles, willingness to pay and use autonomous vehicles, understanding of on-board safety/automation features. The last part of the multi-population surveys gathered information on the anticipated travel-related impacts of autonomous vehicles including individuals' willingness to use shared autonomous modes for their trips. Part C also collected information relating to people's willingness to relinquish one of their household vehicles given the availability of shared autonomous vehicles.

The willingness to relinquish a vehicle in the presence of shared autonomous vehicles presents respondents with a difficult hypothetical choice. First, individuals do not currently have a good grasp of autonomous vehicle technology and its operational characteristics in a shared environment. Second, because household vehicle decisions involve a complex intertemporal decision-making process that includes number of vehicles, type of vehicles, individual vehicle utilizations, intertemporal discounting, etc. (Mannering and Winston, 1985), the willingness to relinquish will have a temporal dynamic that will be impossible to completely capture in a hypothetical survey. And third, there is ample evidence from fields such as psychology, neuroscience, economics, and cognitive science that suggests that the introduction of a new choice option (such as shared autonomous vehicles) will result in an extended period where individual preferences will be highly unstable as they gather information, develop attitudes, potentially polarize in their preferences, etc. (Mannering, 2018). Because it is virtually impossible to account for the above factors in hypothetical choices of shared autonomous vehicle preferences, our forthcoming analysis will be limited in this regard. However, even with these limitations, our analysis will provide some potentially important initial insights into individual preferences for shared autonomous vehicles.

Using data collected from both the target groups, a total of 1214 observations were available to study people's willingness to relinquish their household vehicles in the presence of a shared autonomous vehicles (for the 417 single-vehicle households this would be relinquishing their only vehicle, for the 797 multi-vehicle households, households owning two or more vehicles, this would be relinquishing just one of their vehicles). At the time the survey was conducted, and even today, the exact specifications and attributes of shared autonomous vehicle systems are not yet fully known or understood. Therefore, a stated preference survey about hypothetical scenarios would be saddled with a hypothetical bias as has been found in previous literature (Chang et al., 2009; Carlsson, 2010). In light of this, the approach we adopt (one without a stated preferences and the additional details of a shared autonomous vehicle system) still provides important initial insights into respondent intentions for relinquishing one household vehicle and partaking in a shared-vehicle environment.

In our data, 27.5% of respondents indicated their likelihood of relinquishing a household vehicle in the presence of shared autonomous vehicles as extremely unlikely, 26.7% as unlikely, 19.4% as unsure, 18.6% as likely and 7.3% as extremely likely. Table 2.1 provides summary statistics for some key elements of the sample. This table shows that roughly one-fifth of those surveyed were millennials (20.7%) and that 37.1% of the respondents possessed a graduate degree. Nearly one-fourth of the respondents belonged to households with an annual income below $50,000 (24.1%) and traveled a one-way commute distance of fewer than 10 miles (25.8%). However, a majority of the respondent households owned multiple vehicles (65.7%) and had been involved in a crash prior to taking the survey (74%).

2.5. Methodology

Several statistical/econometric modeling approaches are available to capture the influence of multiple factors that may affect vehicle ownership decisions in the presence of shared autonomous vehicles. In the current study, we will estimate a random-parameter ordered probit model where the dependent variable (peoples' willingness to relinquish a household vehicle, thus

Table 2.1. Descriptive statistics of the variables of interest in understanding respondent's willingness to relinquish a household vehicle with the introduction of shared autonomous vehicles for single-vehicle households (multi-vehicle household values in parentheses).

	Descriptive Statistics	
Variable Description	**Mean**	**Standard Deviation**
Male Respondent Indicator (1 if respondent is male, 0 otherwise)	0.420 (0.605)	0.494 (0.489)
Millennial Indicator (1 if respondent is classified as a millennial, 0 otherwise)	0.393 (0.109)	0.489 (0.312)
White Respondent Indicator (1 if respondent is classified as white, 0 otherwise)	0.822 (0.866)	0.383 (0.341)
Post Graduate Indicator (1 if respondent's highest educational qualification is a post graduate degree, 0 otherwise)	0.372 (0.371)	0.484 (0.483)
Multi-Person Household Indicator ((1 if respondent is a member of a household with more than 3 persons, 0 otherwise)	0.086 (0.252)	0.281 (0.435)
Single Licensed Driver Household Indicator (1 if respondent is a member of a household with only one licensed driver, 0 otherwise)	0.465 (0.080)	0.499 (0.266)
Vehicle Ownership Indicator (1 if respondents is a member of a household that owns three or more vehicles, 0 otherwise)	— (0.407)	— (0.491)

Moderate Commute Distance Indicator (1 if respondent travels a one-way distance less than 10 miles for their commute, 0 otherwise)	0.348 (0.211)	0.477 (0.408)
High Daily Travel Time Indicator (1 if respondent travels more than 90 minutes every day for all their trips, 0 otherwise)	0.158 (0.156)	0.365 (0.363)
Low Parking Time Indicator (1 if respondent spends 5 minutes or less in order to park their vehicle, 0 otherwise)	0.465 (0.650)	0.499 (0.477)
Crash Indicator (1 if respondent has been involved in a traffic crash in the past, 0 otherwise)	0.688 (0.766)	0.464 (0.423)
Complete Vehicle Damage Indicator (1 if respondent was in a crash that resulted in their vehicles suffering complete damage, totaled, 0 otherwise)	0.216 (0.231)	0.412 (0.422)
No Injury Severity Indicator (1 if the respondent was involved in one or more crashes, but no respondent-involved crashes resulted in injury, 0 otherwise)	0.676 (0.640)	0.468 (0.480)

reducing their household vehicle ownership level by one, in the presence of shared autonomous vehicles) is modeled as ordinal data (where respondents indicate their willingness to relinquish as; extremely unlikely, unlikely, unsure, likely, extremely likely).

With such ordered data (extremely unlikely, unlikely, unsure, likely, extremely likely to relinquish), an ordered probability modeling approach is appropriate (Greene, 1997; Washington et al., 2011). An ordered probability model is derived by defining an unobserved variable, z, which is used as a basis for modeling the ordinal ranking of data. This unobserved variable is specified as a linear function,

$$z_n = \boldsymbol{\beta}\mathbf{X}_n + \varepsilon_n \tag{2.1}$$

where $\mathbf{X}$ is a vector of explanatory variables determining the discrete ordering for observation n, $\boldsymbol{\beta}$ is a vector of estimable parameters, and ε is a disturbance term. Using this equation, observed ordinal data, y_n, are defined as (with 1 = extremely unlikely, 2 = unlikely, 3 = unsure, 4 = likely, and 5 = extremely likely),

$$\begin{aligned} y_n &= 1 && \text{if } z_n \leq \mu_0 \\ &= 2 && \text{if } \mu_0 < z_n \leq \mu_1 \\ &= 3 && \text{if } \mu_1 < z_n \leq \mu_2 \\ &= 4 && \text{if } \mu_2 < z_n \leq \mu_3 \\ &= 5 && \text{if } z_n \geq \mu_3, \end{aligned} \tag{2.2}$$

where μ's are estimable parameters (referred to as thresholds) that define y_n and are estimated jointly with the model parameters $\boldsymbol{\beta}$. The estimation problem then becomes one of determining the probability of the five specific ordered responses for each observation n. This is done by making an assumption on the distribution of ε_n in Equation 2.1. If ε_n is assumed to normally distributed across observations an ordered probit model results (alternatively, if ε_n is assumed to logistic distributed an ordered logit model results). Note that without loss of generality μ_0 can be set equal to zero requiring estimation of three thresholds, μ_1, μ_2, and μ_3.

Assuming the disturbance terms are normally distributed (Washington et al., 2011), the ordered category selection probabilities can be written as (removing subscripting *n* for notational convenience),

$$\begin{aligned} P(y=1) &= \Phi(-\boldsymbol{\beta}\mathbf{X}) \\ P(y=2) &= \Phi(\mu_1-\boldsymbol{\beta}\mathbf{X}) - \Phi(-\boldsymbol{\beta}\mathbf{X}) \\ P(y=3) &= \Phi(\mu_2-\boldsymbol{\beta}\mathbf{X}) - \Phi(\mu_1-\boldsymbol{\beta}\mathbf{X}) \\ P(y=4) &= \Phi(\mu_3-\boldsymbol{\beta}\mathbf{X}) - \Phi(\mu_2-\boldsymbol{\beta}\mathbf{X}) \\ P(y=5) &= 1 - \Phi(\mu_{I-1}-\boldsymbol{\beta}\mathbf{X}), \end{aligned} \tag{2.3}$$

where $\Phi(.)$ is the cumulative normal distribution.

For model interpretation, a positive value of $\boldsymbol{\beta}$ implies that an increase in $\mathbf{X}$ will increase the probability of getting the highest response (extremely likely) and will decrease the probability of getting the lowest response (extremely unlikely), but to interpret the intermediate categories (to estimate the direction of the effects of the interior categories of unlikely, unsure and likely) and the probability effect of any variable in the vector $\mathbf{X}$ on each outcome category, average marginal effects are computed at the sample mean as Equation 2.4 below (Washington et al., 2011).

$$\frac{P(y=n)}{\partial \mathbf{X}} = \left[\phi(\mu_{n-1}-\boldsymbol{\beta}\mathbf{X}) - \phi(\mu_n-\boldsymbol{\beta}\mathbf{X})\right]\boldsymbol{\beta}, \tag{2.4}$$

where $P(y = n)$ is the probability of outcome n, μ represents the thresholds, and $\phi(.)$ is the probability mass function of the standard normal distribution. The computed marginal effects quantify the effect that a one-unit change of an explanatory variable will have on outcome category *n*'s selection probability.

Finally, there is likely unobserved heterogeneity present in the data that would result in the effect of explanatory variables to vary across individual observations or groups of observations. To account for this possibility, in the transportation literature, researchers have used random parameters models, latent class (finite mixture) models, Markov switching models, or combinations of these approaches. Using a model structure that can potentially account for unobserved heterogeneity is important because constraining parameters to be fixed across observations when they actually vary across observations can lead to inconsistent, inefficient and biased parameter estimates (Mannering et al., 2016). In this chapter, the possibility of parameters varying across observations is considered by estimating a random parameters formulation with,

$$\boldsymbol{\beta}_i = \boldsymbol{\beta} + \varphi_i, \tag{2.5}$$

where $\boldsymbol{\beta}_i$ is a vector of observation parameters and φ_i is a randomly distributed term (for example, normally distributed term with mean zero and variance σ^2). Estimation of this random parameters formulation is done by simulated maximum likelihood estimation, and we will use a 500 Halton-draw sequencing approach for the simulation as is commonly done in the literature (Bhat, 2003; Anastasopoulos and Mannering, 2009).

2.6. Model Estimation Results

Peoples' willingness to relinquish one of their household's vehicles in the presence of shared autonomous vehicles is likely to be much different in a single-vehicle household than it is in a multi-vehicle household (households owning two or more vehicles). This is because, among other possible reasons, single-vehicle households may have stronger resistance of relinquishing their only vehicle so as to be exposed to more uncertainty with regard to the effectiveness of shared autonomous vehicle as a transportation mode relative to conventional vehicle ownership, especially during hurricane, earthquake or other natural disasters. To test if separate statistical models should be estimated for single- and multi-vehicle households, a likelihood ratio test is conducted with the test statistic $X^2 = -2[LL(\boldsymbol{\beta}_{total}) - LL(\boldsymbol{\beta}_{single}) - LL(\boldsymbol{\beta}_{multi})]$ where the $LL(\boldsymbol{\beta}_{total})$ is the log-likelihood at convergence of the model using all respondents (both single- and multi-vehicle households), $LL(\boldsymbol{\beta}_{single})$ is the log-likelihood at convergence using only respondents from single-vehicle households, and $LL(\boldsymbol{\beta}_{multi})$ is the log-likelihood at convergence using only respondents from multi-vehicle households. This test statistic is χ^2 distributed with degrees of freedom equal to the difference in the number of parameters of both of the models. The value of X^2 is 42.44, and with 21 degrees of freedom, we are more than 99% confident that the null hypothesis that the single- and multi-vehicle household respondents are the same can be rejected. Thus separate models are estimated for single- and multi-vehicle households.

A likelihood ratio test was also conducted to determine if there were significant differences between the University of South Florida and American Automobile Association respondents. In both single- and multi-vehicle household models we could not reject the null hypothesis that the two survey groups were the same at reasonable confidence levels. Thus we do not estimate separate models for these two survey groups.

Random parameters ordered probit model results of peoples' willingness to relinquish one of their household vehicles in the presence of shared autonomous vehicles are as presented in Table 2.2 (for respondents from single-vehicle households) and Table 2.3 (for respondents from multi-vehicle households). In Table 4, the average marginal effects of the individual variables are presented in order to assess the influence of specific parameters on the probabilities of the five possible outcomes (extremely unlikely, unlikely, unsure, likely, and extremely likely). Parameters

Table 2.2. Single-vehicle household (households owning only one vehicle) random parameter ordered probit model estimation of respondents' willingness to relinquish a household vehicle with the introduction of shared autonomous vehicles (extremely unlikely, unlikely, unsure, likely, extremely likely), all random parameters are normally distributed.

Variable Description	**Estimated Parameter**	***t* statistic**
Constant	1.435	6.50
Male Respondent Indicator (1 if respondent is male, 0 otherwise)	-0.211	-1.61
Standard deviation of parameter	(*1.627*)	(*12.38*)
Millennial Indicator (1 if respondent is classified as a millennial, 0 otherwise)	0.679	4.54
Post Graduate Indicator (1 if respondent's highest educational qualification is a post graduate degree, 0 otherwise)	0.119	0.92
Standard deviation of parameter	(*0.821*)	(*7.43*)
Multi-Person Household Indicator (1 if respondent is a member of a household with more than 3 persons, 0 otherwise)	0.935	4.21
Single Licensed Driver Household Indicator (1 if respondent is a member of a household with only one licensed driver, 0 otherwise)	-0.258	-1.83
Standard deviation of parameter	(*1.456*)	(*12.06*)

Moderate Commute Distance Indicator (1 if respondent travels a one-way distance less than 10 miles for their commute, 0 otherwise) *Standard deviation of parameter*	0.231 (*1.221*)	1.70 (*9.98*)
High Daily Travel Time Indicator (1 if respondent travels more than 90 minutes every day for all their trips, 0 otherwise) *Standard deviation of parameter*	-0.662 (*2.150*)	-3.44 (*9.64*)
Low Parking Time Indicator (1 if respondent spends 5 minutes or less in order to park their vehicle, 0 otherwise)	-0.592	-4.36
Crash Indicator (1 if respondent has been involved in a traffic crash in the past, 0 otherwise) *Standard deviation of parameter*	0.101 (*1.239*)	0.70 (*12.60*)
Complete Vehicle Damage Indicator (1 if respondent was in a crash that resulted in their vehicles suffering complete damage, totaled, 0 otherwise) *Standard deviation of parameter*	-0.424 (*1.121*)	-2.52 (*7.32*)
Threshold, μ_1	2.168	13.55
Threshold, μ_2	3.406	16.93
Threshold, μ_3	5.308	17.36
Number of observations	417	
Log-likelihood at convergence	-581.017	
Restricted log-likelihood	-607.209	

Table 2.3. Multi-vehicle household (households owning two or more vehicles) random parameter ordered probit model estimation of respondent's willingness to relinquish a household vehicle with the introduction of shared autonomous vehicles (extremely unlikely, unlikely, unsure, likely, extremely likely), all random parameters are normally distributed.

Variable Description	**Estimated Parameter**	***t* statistic**
Constant	1.000	6.45
Male Respondent Indicator (1 if respondent is male, 0 otherwise) *Standard deviation of parameter*	0.119 (*0.622*)	1.49 (*11.41*)
Millennial Indicator (1 if respondent is classified as a millennial, 0 otherwise)	0.593	4.33
White Respondent Indicator (1 if respondent is classified as white, 0 otherwise)	-0.346	-3.03
Post Graduate Indicator (1 if respondent's highest educational qualification is a post graduate degree, 0 otherwise)	0.305	3.76
Single Licensed Driver Household Indicator (1 if respondent is a member of a household with only one licensed driver, 0 otherwise)	-0.706	-4.47
Vehicle Ownership Indicator (1 if respondent is a member of a household that owns more than three vehicles, 0 otherwise)	-0.289	-3.54
Moderate Commute Distance Indicator (1 if respondent travels a one-way distance less than 10 miles for their commute, 0 otherwise) *Standard deviation of parameter*	0.362 (*0.386*)	3.70 (*4.50*)
High Daily Travel Time Indicator (1 if respondent travels more than 90 minutes every day for all their trips, 0 otherwise) *Standard deviation of parameter*	0.174 (*0.926*)	1.54 (*8.26*)
Low Parking Time Indicator (1 if respondent spends 5 minutes or less in order to park their vehicle, 0 otherwise)	-0.184	-2.18
Crash Indicator (1 if respondent has been involved in a traffic crash in the past, 0 otherwise) *Standard deviation of parameter*	0.272 (*0.538*)	2.33 (*11.26*)
Complete Vehicle Damage Indicator (1 if respondent was in a crash that resulted in their vehicles suffering complete damage, totaled, 0 otherwise) *Standard deviation of parameter*	-0.165 (*0.646*	-1.52 (*7.45*)
No Injury Severity Indicator (1 if the respondent was involved in one or more crashes, but no respondent-involved crashes resulted in injury, 0 otherwise)	-0.210	-2.14
Threshold, μ_1	0.816	15.14
Threshold, μ_2	1.548	22.55
Threshold, μ_3	2.737	28.03
Number of observations	797	
Log-likelihood at convergence	-1195.938	
Restricted log-likelihood	-1238.243	

producing statistically significant standard deviations for their assumed distribution are treated as parameters that vary across the population (with each observation having its own parameter), and the remaining parameters are treated as fixed parameters because the standard deviations are not significantly different from zero (one parameter for all observations).

Table 2.2 shows that for respondents from single-vehicle households, seven parameters (indicators for male respondent, post graduate, single licensed driver household, moderate commute distance, high daily travel time, crash, complete vehicle damage) were found to vary significantly across the population. Table 2.3 shows that for respondents from multi-vehicle households, five parameters (indicators for male respondent, moderate commute distance, high daily travel time, crash, complete vehicle damage) were found to vary significantly across the population. Again, a likelihood ratio test was used to statistically compare the random-parameters, and fixed parameters ordered probit models for both single- and multi-vehicle household respondents. The likelihood ratio test statistic is calculated as $X^2 = -2[LL(\boldsymbol{\beta}_{random}) - LL(\boldsymbol{\beta}_{fixed})]$ where the $LL(\boldsymbol{\beta}_{random})$ is the log-likelihood at convergence of the random-parameter ordered probit model and the $LL(\boldsymbol{\beta}_{fixed})$ is the log-likelihood at convergence of the fixed-parameter ordered probit model. The test statistic X^2 is χ^2 distributed with degrees of freedom equal to the difference in the number of parameters of both fixed and random parameters models. For respondents from single-vehicle households, the value of X^2 is 17.97, and with 7 degrees of freedom, we are more than 98% confident that the null hypothesis that the random- and fixed-parameters ordered probit models are equal can be rejected (thus justifying the use of the random parameters formulation). For respondents from multi-vehicle households, the value of X^2 is 11.97, and with 5 degrees of freedom, we are more than 97% confident that the null hypothesis that the random- and fixed-parameters ordered probit models are equal can be rejected (thus justifying the use of the random parameters formulation).

2.7. Discussion of Estimation Findings

As shown in Tables 2.2 and 2.3, gender is a statistically significant factor in relinquishing vehicle ownership in the presence of shared autonomous vehicles in both single- and multi-vehicle households. From the marginal effects in Table 2.4, being male, on average, increases the probability of being unlikely or extremely unlikely to relinquish a household vehicle in a single-vehicle household, but decreases these probabilities in multi-vehicle households, relative to their female counterparts in the presence of shared autonomous vehicles (however, in both single- and multi-vehicle households the model estimations produced a statistically significant random parameter suggesting considerable heterogeneity across the population). Although the probability influences of the male indicator variables are small on average in both models, part of the reason for this statistically significant male/female difference could be due to men being more risk averse with respect to new vehicle technologies in single-vehicle households and less risk averse in multi-vehicle households relative to females. In fact, there is a large body of literature showing gender differences in risk-taking in transportation-related decisions (Abay and Mannering, 2016).

Comparing across generations, millennials (respondents who are less than 35 years of age) are more likely or extremely likely to relinquish a household vehicle with the introduction of shared autonomous vehicles in both single- and multi-vehicle households, relative to other age groups (as shown in the marginal effects in Table 2.4). Millennials are a significant demographic in determining the course of future technology adoption as they are the largest living generation (Fry, 2016) and are set to dominate the future discussions and discourse on adoption of new

Table 2.4. Average marginal effects of the random parameter ordered probit model estimation of respondent's willingness to relinquish a household vehicle with the introduction of shared autonomous vehicles for single-vehicle households (multi-vehicle household values in parentheses).

	Average Marginal Effects				
Variable Description	**Extremely Unlikely**	**Unlikely**	**Unsure**	**Likely**	**Extremely Likely**
Male Respondent Indicator (1 if respondent is male, 0 otherwise)	0.037 (-0.034)	0.020 (-0.014)	-0.048 (0.011)	-0.009 (0.029)	-0.00026 (0.008)
Millennial Indicator (1 if respondent is classified as a millennial, 0 otherwise)	-0.108 (-0.132)	-0.088 (-0.093)	0.016 (0.016)	0.036 (0.148)	0.00015 (0.061)
White Respondent Indicator (1 if respondent is classified as white, 0 otherwise)	— (0.086)	— (0.050)	— (-0.019)	— (-0.087)	— (-0.030)
Post Graduate Indicator (1 if respondent's highest educational qualification is a post graduate degree, 0 otherwise)	-0.020 (-0.082)	-0.013 (-0.039)	0.028 (0.023)	0.005 (0.076)	0.00016 (0.022)
Multi-Person Household Indicator (1 if respondent is a member of a household with more than 3 persons, 0 otherwise)	-0.096 (—)	-0.226 (—)	0.232 (—)	0.090 (—)	0.00078 (—)
Single Licensed Driver Household Indicator (1 if respondent is a member of a household with only one licensed driver, 0 otherwise)	0.045 (0.239)	0.030 (0.028)	-0.059 (-0.096)	-0.011 (-0.143)	-0.00033 (-0.028)
Vehicle Ownership Indicator (1 if respondent is a member of a household that owns three or more vehicles, 0 otherwise)	— (0.082)	— (0.033)	— (-0.027)	— (-0.070)	— (-0.019)
Moderate Commute Distance Indicator (1 if respondent travels a one-way distance less than 10 miles for their commute, 0 otherwise)	-0.038 (-0.092)	-0.031 (-0.051)	0.054 (0.022)	0.011 (0.091)	0.00035 (0.030)
High Daily Travel Time Indicator (1 if respondent travels more than 90 minutes every day for all their trips, 0 otherwise)	0.149 (-0.046)	-0.004 (-0.023)	-0.126 (0.013)	-0.019 (0.044)	-0.00044 (0.013)
Low Parking Time Indicator (1 if respondent spends 5 minutes or less in order to park their vehicle, 0 otherwise)	0.106 (0.050)	0.053 (0.023)	-0.132 (-0.015)	-0.026 (-0.046)	-0.00086 (-0.013)
Crash Indicator (1 if respondent has been involved in a traffic crash in the past, 0 otherwise)	-0.018 (-0.080)	-0.009 (-0.028)	0.023 (0.028)	0.004 (0.064)	0.00012 (0.016)
Complete Vehicle Damage Indicator (1 if respondent was in a crash that resulted in their vehicles suffering complete damage, totaled, 0 otherwise)	0.085 (0.048)	0.018 (0.018)	-0.089 (-0.016)	-0.014 (-0.040)	-0.00037 (-0.010)
No Injury Severity Indicator (1 if respondent suffered no injuries in their most severe crash, 0 otherwise)	— (0.060)	— (0.024)	— (-0.019)	— (-0.051)	— (-0.014)

technologies. These results are also in line with recent literature that looked at generational-level differences in the adoption of new technology (Smith, 2011; Smith, 2013; Anderson, 2015), and millennials' willingness to use multiple modes of transportation to reach a destination and the differences in their overall travel behavior and preferences towards more equitable modes of transportation (APTA, 2013; Circella et al., 2016). The results also make intuitive sense considering millennial attitudes towards vehicle ownership and a sharing economy (Circella et al., 2016).

Marginal effects in Table 2.4 show that white respondents (1 if respondents are classified as white for ethnicity, 0 otherwise) tend to be more unlikely or extremely unlikely to relinquish a household vehicle in multi-vehicle households relative to other ethnicities (this indicator variable was statistically insignificant in single-vehicle households). Past literature has touched upon the higher levels of accessibility to automobiles enjoyed by whites (Berube et al., 2006), and their general reluctance to engage in shared transportation modes such as carpools (McKenzie, 2015). This seems to be particularly true in multi-vehicle households.

In contrast, respondents with a graduate degree (1 if respondents whose highest qualification was a graduate degree, 0 otherwise), in both single- and multi-vehicle households, have higher probabilities to be likely or extremely likely to relinquish a household vehicle to utilize shared autonomous vehicles when they become available in the market relative to other educational levels (see Table 2.4). However, in single-vehicle households, the effect of the variable was found to vary significantly across respondents (producing a statistically significant random variable), suggesting considerable heterogeneity across observations, whereas this variable produced a fixed parameter in the case of multi-vehicle households. In both single- and multi-vehicle households it is likely that a higher level of education exposes respondents to greater discourse and discussion on the benefits of autonomous vehicles and shared economies.

In single-vehicle households with three or more household members, respondents, on average, were found to be less unlikely or extremely unlikely (Table 2.4) to relinquish a household vehicle (this variable was statistically insignificant in the multi-vehicle household model) relative to one- and two-person households. This would seem to support the hope that shared autonomous vehicles can substantially improve mobility among larger households that are currently restricted by owning only a single vehicle.

Estimation results in both single- and multi-vehicle models show that households with a single licensed driver (1 if respondents belong to households with only one licensed driver, 0 otherwise) on average are more unlikely or extremely unlikely to give up a household vehicle with the availability of shared autonomous vehicle alternatives (Table 2.4). Interestingly, this variable produced a statistically significant random parameter in the single-vehicle case (suggesting considerable heterogeneity across the sample) and a fixed parameter in the multi-vehicle case. In both cases, it is likely that such households may have transportation patterns that make them less willing to rely on sharing.

For the case of multi-vehicle households, households owning three or more vehicles were found to be more unlikely or extremely unlikely to relinquish one of their vehicles (see marginal effects in Table 2.4) relative to their two-vehicle multi-vehicle household counterparts. It appears as though respondents in households with a large number of vehicles seem to be more entrenched in the private-vehicle ownership culture and thus less likely to relinquish in favor of shared autonomous vehicles. Another possible reason is that high-vehicle-ownership respondents may own one or more vehicles largely for enjoyment and collection purposes, which would make their relinquishment less likely. It is noteworthy that other household attributes such as household income were considered in the modeling process, but found to be statistically insignificant.

A number of model results show the impacts of current travel characteristics on vehicle ownership decisions. For example, in both single- and multi-vehicle households, if a respondent commutes a one-way distance of fewer than 10 miles, on average, they tend to be less unlikely or extremely unlikely to give up a household vehicle (Table 2.4). The effect of this variable varies across the population in both vehicle-ownership-level models (Tables 2.2 and 2.3), again implying heterogeneous effects suggesting, for example, that not all less than 10-mile commutes are the same.

In addition to commute distance, total daily travel time was found to significantly influence vehicle-relinquishment decisions (Table 2.4), with respondents from single-vehicle households who traveled more than 90 minutes on all travel in a day being more extremely unlikely to relinquish a household vehicle, and respondents from multi-vehicle households who traveled more than 90 minutes on all travel in a day being less unlikely and extremely unlikely to relinquish a household vehicle (Table 2.4). Although the effect of this variable was found to vary significantly across the respondent population in both models (as reflected by the presence of a statistically significant random parameter), the findings suggest the substantive differences in the way single- and multi-vehicle households view travel times and vehicle ownership needs.

With regard to the possible effects of parking on shared autonomous vehicle adoption, for both single- and multi-vehicle household respondents, those respondents who spent 5 minutes or less on an average to park their vehicles during their commute trips were more unlikely or extremely unlikely (Table 2.3) to relinquish a household vehicle relative to people that spend longer periods parking. This shows, as expected, that parking scarcity is likely to be a major driver in shared autonomous vehicle adoption.

Three variables relating to crash history were found to be statistically significant in the model; an indicator depicting respondents' involvement in a crash, an indicator for respondents that experienced complete vehicle damage in a crash, and an indicator for respondents that did not sustain an injury in their most severe crash. In both single- and multi-vehicle households, respondents who have been involved in a crash are, on average, more likely or extremely likely to relinquish a household vehicle with the introduction of shared autonomous vehicles (Table 2.4), although the effects of this variable are heterogeneous across the population as indicated by the significant random parameter.

Among those who were involved in one or more traffic crashes, in both single- and multi-vehicle households, respondents who suffered complete vehicle damages in one of their crashes are, on average, more unlikely or extremely unlikely to relinquish a household vehicle than those who experienced moderately severe crashes, although again the effect of this variable varies across observations. It is likely that these respondents, who have experienced extensive-damage crashes, are more skeptical of emerging vehicle technologies, such as autonomous vehicles, because of safety-related concerns. At the other extreme of crash severity, respondents in multi-vehicle households, who were in one or more crashes but did not sustain injuries in any crash, were also found to be more unlikely or extremely unlikely to relinquish a household vehicle. Since these people have had crash experiences with favorable injury outcomes, they may discount the potential safety benefits of shared autonomous vehicles and thus may be more reluctant to relinquish one of their vehicles than those who experienced moderately severe crashes.

Finally, it is noteworthy that variables such as household income and others were not found to be statistically significant in the models. It appears that the variables we have included (while obviously correlated with many variables not found to be significant) are statistically the best in terms of modeling people's vehicle relinquishment likelihoods in the presence of shared autonomous vehicles.

2.8. Summary and Conclusions

This chapter presents an initial assessment of people's likelihood of relinquishing a household vehicle (reducing their household vehicle ownership level by one) in the presence of shared autonomous vehicles. To this end, we conduct a survey of two different target groups of interest: faculty, students, and staff from a large university (University of South Florida); and the members of the AAA Foundation of the southeastern United States – asking how likely they would be to consider relinquishing one of their household's personal vehicles if shared autonomous vehicles were available (thus reducing their household vehicle ownership level by one). Possible responses to the question are: extremely unlikely, unlikely, unsure, likely, and extremely likely. For single-vehicle households, this would be relinquishing their only vehicle, and for multi-vehicle households (households owning two or more vehicles) this would be relinquishing one of their vehicles. Therefore, two different random parameters ordered probit models are estimated to analyze the factors that influence the households' likelihood of relinquishing one of their vehicles; one model for single-vehicle households and the other model for multi-vehicle households.

Our estimation results show that for single-vehicle households, seven parameters (indicators for male respondent, post graduate, single licensed driver household, moderate commute distance, high daily travel time, crash, complete vehicle damage) were found to vary significantly across the population and for multi-vehicle households, five parameters (indicators for male respondent, moderate commute distance, high daily travel time, crash, complete vehicle damage) were found to vary significantly across the population. Different influential factors relating to gender, respondent characteristics, household characteristics, current travel characteristics and crash history are statistically significant and affect the likelihood of vehicle-relinquishment with the introduction of shared autonomous vehicles. The findings from this study provide key insights regarding vehicle-relinquishment in an era of shared autonomous vehicles including the following:

1. Gender has a significant but variable impact on people's likelihood of relinquishing a household vehicle when shared autonomous vehicles become available on the market. Males on average had lower probabilities of being likely or extremely likely to relinquish a household vehicle in single-vehicle household, but higher probabilities in these categories in multi-vehicle households, relative to their female counterparts.

2. Socio-economic characteristics are significant indicators towards people's likelihood of relinquishing a household vehicle for shared autonomous vehicles. For instance, millennials and graduate degree holders are more likely to relinquish a household vehicle when shared autonomous vehicles come into the market, possibly indicating their preferences towards a more sustainable lifestyle in comparison to their older counterparts.

3. Respondent commute distances and average daily travel times have a complex effect on the likelihood of relinquishing vehicles, one that varies considerably between single- and multi-vehicle households.

4. While previous crash history usually makes respondents more likely to relinquish their vehicles to use emerging technologies like shared autonomous vehicles, a previous

experience of suffering complete vehicle damage or no-injury makes people more unlikely to relinquish their vehicles in order to use shared autonomous vehicles (than those who experienced moderately severe damages).

5. Throughout our model estimations, there are substantial and statistically significant differences between single- and multi-vehicle household respondent opinions. This underscores the potentially large impact that the traditional human-driven-vehicle culture may have on new technology adoptions.

The insights obtained from this study can be used to target demographic groups most likely to adopt shared autonomous vehicles. The study can also help better understand the sentiments of the public relating to their willingness to use such emerging technologies. However, it is important to keep in mind that people’s perception of shared autonomous vehicles is not likely to be temporally stable. As autonomous vehicle technologies unfold, personal experiences, publicity, and information gathering will undoubtedly change people’s perceptions of shared autonomous vehicles. Thus it is important to view the findings in this chapter with some caution in light of this. Future studies could examine the sentiments of the general public towards autonomous vehicles and utilizing shared autonomous vehicles when they become available in the market. Yet, the marginal effects and the initial findings from this chapter will serve as a baseline for comparison of changes in people’s intentions as additional studies are conducted in the future.

Chapter 3: A statistical analysis of consumers' perceptions towards automated vehicles and their intended adoption

3.1. Introduction

Many major automotive manufacturers, technology firms, and ridesharing companies are actively pursuing and developing automated vehicle technologies. In the U.S., testing of autonomous vehicles is currently underway in a number of cities and nearly half of the states have enacted legislation related to autonomous vehicles. However, there is still considerable uncertainty with regard to public acceptance and adoption of autonomous vehicles. As past research has shown, not all new technological innovations are immediately welcomed by the consumers (Moore, 2002; Heffner et al., 2007; Edison and Geissler, 2003), and there are potentially important psychological and behavioral tendencies that will affect peoples' attitudes and opinions toward autonomous vehicles and their eventual adoption (Sheela and Mannering, 2018).

Past research has provided some important insights into potential autonomous-vehicle adoption. For example, Bansal and Kockelman (2017) found that while over 40% of Americans indicated that autonomous vehicle would be important in future transportation systems, fewer than 20% indicating they felt comfortable allowing autonomous vehicles drive them. There has also been an abundance of literature that has looked at factors that make individuals more or less likely to adopt autonomous vehicles. For example, Haboucha et al. (2017) provided a review of recent findings from the extant literature with the intent of uncovering common findings among research efforts that have studied autonomous vehicle adoption. Their findings suggest a consensus among studies that men will be more likely to use self-driving technologies than women. However, they also uncovered numerous contradictory findings among studies specifically with regard to age, with some studies finding that younger individuals are more likely to adopt autonomous vehicles while others finding that elderly individuals would be more like to accept autonomous vehicles.

The intent of this chapter is to add to the growing body of literature that studies likelihood that consumers will adopt an autonomous vehicle, but to do so in a novel way. Using a survey that samples a large university community and members of a national automobile association, we begin our analysis by categorizing the consumer population into market segments using cluster analysis methods. Once these market segments are identified, we estimate an econometric model to understand the factors that make consumers more or less likely to belong to one of the identified market segments. Finally, for each market segment, we estimate an autonomous-vehicle likelihood-adoption model (with likelihoods of adoption ranging from *extremely unlikely* to *extremely likely*). With this multi-stage analysis we hope to provide new insights into consumer autonomous-vehicle-adoption likelihoods.

The chapter begins with a description of the survey design and data, followed by a description of the cluster-analysis approach and a discussion of the resulting identified market segments. We then present the logit-model development and estimation results that relate to understanding the factors influencing the probability of a respondent belonging to a specific market segment. After this, the results of a random parameters probit model is presented (addressing the likelihood of autonomous vehicle adoption for each market segment). The chapter concludes with a summary of key findings and suggestions for future work.

3.2. Survey Design and Data

To determine the factors that may influence the probability of a consumer belonging to an autonomous vehicle consumer market segment, a web-based survey was conducted to target two interest groups: faculty, students, and staff from a large university (University of South Florida); and the members of the American Automobile Association of the southeastern United States. These surveys were disseminated for data collection during April and June 2015, respectively. Both surveys collected a wide range of data relating to socioeconomics, and various travel-related characteristics (such as commuting behavior, vehicle crash experience, and vehicle inventory).

University members were chosen because university campuses are typically a fertile ground for the testing and early adoption of new technologies. University respondents are often some of the earliest adopters (and sometimes the most vocal critics) of these technologies, thereby making them an interesting demographic to consider for the purpose of our study. Also, with approximately 25% of the households in the United States being American Automobile Association members, the inclusion of this group gives a broad cross-section of American society.[1]

Using data collected from both surveyed groups, a total of 2,477 observations were available. Table 3.1 provides summary statistics for some key elements of our data sample. This table shows that 28.8% of those surveyed were millennials (which we define here as those less than 35 years old), and roughly one-third (33.1%) of the respondents possessed a graduate degree.

[1] Although the data from these two sources does not represent a national sample, our intent is exploratory and nature and it is important to note that even if we were to base our analysis on a national sample there would be issues because consumer preferences for new-technologies tend to be highly volatile. In fact, Menon et al. (2016) and Sheela and Mannering (2018) show preferences for autonomous-vehicle adoption have a high degree of temporal instability which would mitigate the benefits of a more representative survey. Thus, our focus on this select sub-sample of potential consumers is merely intended to provide some initial insights and a demonstration of a methodological approach that is simple and easily accessible for empirical researchers and yet can be used to guide future studies on the subject.

Table 3.1. Descriptive Statistics of the Variables of Interest in Understanding the Probability of a Consumer Belonging to a Particular Autonomous Vehicle Market Segment

Variable Description	Mean	Standard Deviation
Male Respondent Indicator (1 if respondent is male, 0 otherwise)	0.536	0.499
University Respondent Indicator (1 if respondent is classified as a university respondent, 0 otherwise)	0.269	0.442
Millennial Indicator (1 if respondent is less than 35 years old, 0 otherwise)	0.288	0.453
Baby Boomer Indicator (1 if respondent is 50 to 64 years old, 0 otherwise)	0.300	0.458
Generation X Indicator (1 if respondent is 35 to 49 years old, 0 otherwise)	0.115	0.320
Graduate Indicator (1 if respondent's highest educational qualification is a graduate degree, 0 otherwise)	0.331	0.471
Non-Commuter Indicator (1 if respondent does not undertake a commute trip, 0 otherwise)	0.225	0.418
Drive Alone Commuter Indicator (1 if respondent typically drives alone to their commute, 0 otherwise)	0.675	0.468
Very High Licensed Driver Household Indicator (1 if respondent is a member of a household that has 3 or more licensed drivers, 0 otherwise)	0.084	0.278
High Income Household Indicator (1 if respondent is a member of a household with an annual income $150,000 or more, 0 otherwise)	0.145	0.352
High Commute Distance Indicator (1 if respondent travels a one-way distance of 20 miles or more for their commute, 0 otherwise)	0.161	0.368
High Commute Time Indicator (1 if respondent spent a total of 60 minutes or more on an average for their one-way commute, 0 otherwise)	0.046	0.209
Medium Overall Daily Travel Time Indicator (1 if respondent travels 45 minutes or less on an average for their total daily travel, 0 otherwise)	0.462	0.499
High Parking Time Indicator (1 if respondent spent 10 or more minutes on an average towards finding a parking spot during their commute, 0 otherwise)	0.215	0.411
Three-Plus Vehicle Ownership Indicator (1 if respondent is a member of a household that owns more than three vehicles, 0 otherwise)	0.084	0.278
Recent New Vehicle Purchase Category Indicator (1 if respondent most recently purchased or leased a new vehicle, 0 otherwise)	0.542	0.498
Major Injury Severity Indicator (1 if the respondent was involved in one or more crashes, and respondent-involved crashes resulted in major injury, 0 otherwise)	0.176	0.381

Fifteen percent of the respondents belonged to higher income households (with an annual income of $150,000 or more), with 16% traveling a one-way commute distance of 20 or more miles. Survey results also indicated that the majority of the respondent households recently purchased/leased a new vehicle (54%). And lastly, around one-fifth of the respondents (18%) had been involved in crashes resulting in a major injury.

In addition to socio-demographics and other behavioral information, respondents' opinions were also sought on potential benefits, and concerns with autonomous vehicles (see Table 3.2). For their views on the benefits of autonomous vehicles, respondents were asked for their opinions on a five-point scale ranging from *Extremely Unlikely* to *Extremely Likely*. Respondents indicated their opinions on five potential benefits of autonomous vehicles: fewer traffic crashes and increased roadway safety; less stressful driving experience; less traffic congestion; more productive use of travel time; and increased fuel efficiency. Respondents' concerns with autonomous vehicles were similarly elicited on a five-point scale ranging from *Not at all Concerned* to *Extremely Concerned*. Respondents indicated their opinions on six potential concerns with autonomous vehicles: safety of the vehicle occupants and other road users such as pedestrians; system/equipment failure or autonomous vehicle hacking; performance in (or response to) unexpected traffic conditions, poor weather conditions; difficulty in determining who is liable in the event of a crash; privacy risks from data tracking on my travel locations and speed; and loss in human driving skill over time.

3.3. Identification of Market Segments

Cluster analysis is a multivariate technique widely used to identify data structures based on the information found in the data (Anderberg, 1973). Its primary objective is to restructure the

Table 3.2. Descriptive Statistics of Consumer's Opinions on the Proposed Benefits, and Concerns with Autonomous Vehicles

Description of Autonomous Vehicles Potential Benefits	Extremely Unlikely	Unlikely	Unsure	Likely	Extremely Likely
Potential Benefit - Fewer traffic crashes and increased roadway safety	6.8%	11.7%	22.2%	36.5%	22.8%
Potential Benefit - Less stressful driving experience	8.4%	15.3%	18.9%	34.0%	23.4%
Potential Benefit - Less traffic congestion	12.3%	28.1%	25.6%	21.9%	12.2%
Potential Benefit - More productive (than driving) use of travel time	6.6%	13.5%	22.1%	31.2%	26.5%
Potential Benefit - Increased fuel efficiency	5.7%	13.4%	27.4%	38.0%	15.6%

Description of Autonomous Vehicles Potential Concerns	Not at all Concerned	Slightly Concerned	Somewhat Concerned	Moderately Concerned	Extremely Concerned

Potential Concern - Safety of the vehicle occupants and other road users such as pedestrians, bicyclists.	5.0%	14.0%	14.8%	37.8%	28.3%
Potential Concern - System/equipment failure or automated vehicle system hacking	2.6%	11.3%	15.2%	34.4%	36.5%
Potential Concern - Performance in (or response to) unexpected traffic situations, poor weather conditions	3.1%	10.0%	15.5%	35.6%	35.8%
Potential Concern - Difficulty in determining who is liable in the event of a crash	8.6%	16.0%	21.3%	28.0%	26.1%
Potential Concern - Privacy risks from data tracking on my travel locations and speed	9.9%	16.7%	15.4%	27.4%	30.7%
Potential Concern - Loss in human driving skill over time	11.2%	17.2%	13.2%	30.8%	27.6%

data into nearly homogeneous groups (Guo et al., 2016). This technique has been applied in numerous transportation studies in the past. For example, Chang et al. (1992) used cluster analysis and discriminant analysis to determine the impact of commuter driving behavior on the rapid growth in suburban populations. Ng et al. (1998) employed cluster analysis to uncover groups of private and commercial drivers based on how much importance they placed on trip factors that influenced their commute trips. Hildebrand (2003) used cluster analysis to identify six lifestyle groups based on socio-demographics in order to model elderly travel behavior and activity engagement. Pinjari et al. (2008) used a two-step cluster analysis approach to divide more than 1000 zones in the San Francisco Bay area into bicycle-friendly, and less bicycle-friendly neighborhoods to estimate a joint model of residential neighborhood and bicycle ownership. And finally, Guo et al. (2016) employed cluster analysis to understand the correlation between truck freight carriers' operational and behavioral characteristics, and the factors that foster/impede their willingness to collaborate with rail-freight carriers.

In the following analysis, a two-step cluster analysis is used, which is preferred in our case over hierarchical or portioning cluster analyses due to its ability to simultaneously handle both categorical and continuous data. The two-step cluster analysis also has the capacity to be flexible in defining the required number of clusters (Chui et al., 2001). The two-step cluster analysis we employ identifies consumer market segments based on factors that (we hypothesize) influence the adoption likelihood of autonomous vehicles. We hypothesize that consumers' intended adoption of autonomous vehicles can be captured largely by their perceptions regarding the potential benefits/concerns with autonomous vehicles (how likely/unlikely are these benefits/concerns with autonomous vehicles?).[2] Respondents' opinions on these factors (as discussed in Table 3.2) were gathered during the survey. These factors comprise five potential benefits, and six potential concerns with autonomous vehicles as shown in Table 3.2.

3.4. Description of Market Segments

Based on the results from the two-step cluster analysis procedure applied, four different autonomous vehicle consumer market segments (clusters) are identified. To understand the intended adoption potential of these market segments, the average scores obtained for intended adoption under each cluster were correlated along with the scores obtained for the perception variables under the same cluster. The findings of the two-step cluster analysis are as shown in Table 3.3.

The benefits-dominated market segment (n=513, 19.3% of the sample) included consumers who foresee benefits with the introduction of autonomous vehicles. Respondents who identify with this market segment believe that the proposed benefits of autonomous vehicles, such as fewer traffic crashes and increased roadway safety, less stressful driving experience, more productive use of travel time, increased fuel efficiency, and less traffic congestion, are more likely to occur with the introduction of autonomous vehicles (as can be seen by their higher market-segment means). Further, as evidenced by the low mean values of this segment's opinions on the likelihood of potential concerns of autonomous vehicles, respondents in this market segment are not concerned about possible system/equipment failure, performance in unexpected traffic and weather conditions, privacy risks from data tracking, difficulty in determining liability in the event

[2] This is supported by the empirical work of Menon et al. (2016) and Sheela and Mannering (2018) who demonstrate the importance of initial perceptions with regard to autonomous-vehicle adoption.

Table 3.3. Segment Means for Each Respondent Group Based on Perception of Benefits and Concerns of Autonomous Vehicles (Bold Numbers Indicate that the Majority of Respondents in this Segment Consider this Factor Likely or Extremely Likely)*

Description of Autonomous Vehicles Perception Variables	Benefits-Dominated Cluster (N=513)	Concerns-Dominated Cluster (N=602)	Uncertain Cluster (N=732)	Well-Informed Cluster (N=811)
Potential Benefit - Fewer traffic crashes and increased roadway safety	4.65	2.47	3.08	4.14
Potential Benefit - Less stressful driving experience	4.62	2.27	2.89	4.21
Potential Benefit - Less traffic congestion	4.18	1.89	2.46	3.35
Potential Benefit - More productive (than driving) use of travel time	4.57	2.57	2.97	4.24
Potential Benefit - Increased fuel efficiency	4.21	2.69	3.07	3.85
Potential Concern - Safety of the vehicle occupants and other road users such as pedestrians, bicyclists.	2.35	4.43	3.43	4.26
Potential Concern - System/equipment failure or automated vehicle system hacking	2.77	4.73	3.48	4.40
Potential Concern - Performance in (or response to) unexpected traffic situations, poor weather conditions	2.82	4.64	3.49	4.44
Potential Concern - Difficulty in determining who is liable in the event of a crash	2.46	4.52	3.15	3.63
Potential Concern - Privacy risks from data tracking on my travel locations and speed	2.68	4.59	3.07	3.67
Potential Concern - Loss in human driving skill over time	2.46	4.49	3.34	3.45
Likelihood of adopting autonomous vehicles when they become available in the market	4.24	1.74	2.54	3.39

* Scale of responses; 1= extremely unlikely, 2 = unlikely, 3 = uncertain, 4 = likely, 5 = extremely likely.

of a crash, loss in human driving skill over time, and safety of the vehicle occupants and other road users. Given their high belief in the potential benefits (and low-level of concerns) of autonomous vehicles, it is likely that these respondents are typical early adopters of new technologies. This is supported by a high score obtained under likelihood of intended adoption of autonomous vehicles (4.24 out of 5), indicating that these respondents could very likely be the initial adopters when autonomous vehicles are available in the market.

The concerns-dominated market segment (n=602, 22.6% of the sample) consisted of consumers who have a focus on potential concerns with autonomous vehicles and do not believe in many of the potential benefits of autonomous vehicles. Respondents under this market segment are concerned about possible system/equipment failure, performance in unexpected traffic and weather conditions, privacy risks from data tracking, difficulty in determining liability in the event of a crash, loss in human driving skill over time, and safety of the vehicle occupants and other road users. A low score obtained under likelihood of intended adoption of autonomous vehicles (1.74 out of 5) means that these respondents are least likely to adopt when autonomous vehicles are available in the market.

The third market segment (n=732, 27.5% of the sample), the uncertain segment, included consumers who are either indifferent or unsure about both the potential benefits as well as the potential concerns with autonomous vehicles. It is likely that this segment has limited exposure and/or interest in the ongoing discourse on emerging vehicle technologies. Alternatively, it is also possible that this segment is usually unsure about emerging technologies (as evidenced by the clustering of their mean scores towards “unsure”). Their lower adoption scores (2.54 out of 5) are unsurprising given they are either indifferent or unsure about the potential benefits or concerns of new technologies.

The final market segment (n=811, 30.5% of the sample), the well-informed segment, include consumers who seem to equally aware of the potential benefits and concerns associated with autonomous vehicles. While consumers in this market segment feel that the proposed benefits such as more productive use of travel time, less stressful driving experience, and fewer traffic crashes are likely to occur, they are also concerned about issues such as performance of the automated vehicle in unexpected traffic situations, possible system/equipment failure, and other safety-related concerns with the introduction of autonomous vehicles. Their somewhat high adoption scores (3.39 out of 5) indicate a wait-and-see attitude before immersing themselves into adopting autonomous vehicles.

In summary, the cluster analysis suggested four different market segments based on their opinions about the potential benefits and concerns of automated vehicle technologies; benefits dominated segment, well-informed segment, uncertain segment, and concerns dominated segment. Interestingly, these market segments range from one (positive) extreme to another (negative) extreme in their mean-stated likelihood of adoption of autonomous vehicles.

3.5. Descriptive Statistics of Market Segments

Table 4 provides the descriptive statistics for each autonomous vehicle consumer market segment. While a significant proportion of the benefits-dominated consumer market segment (60%) is comprised of males, it was interesting to note the higher percentage of women under the well-informed market segment (51%) and the lower percentages under the uncertain market segment (45%). For the purposes of discussion in this chapter, we define millennials as respondents less than 35 years of age), Gen-X-ers as those between 35 and 49 years old, baby boomers as those between 50 and 64 years old, and the great generation as those 65 and older. Table 3.4 shows millennials constituted more than one-third of the well-informed market segment (38%) as well as the benefits-dominated market segment (33%). It was also seen that the highest percentage of concerns-dominated market segment comprised of baby boomers (37%) while the highest percentage of the uncertain market segment comprised of the great generation (40%). Almost equal shares of respondents from each consumer market segment belonged to low-income and high-income households respectively.

Interestingly, the uncertain and the concerns-dominated market segments had the highest percentage on non-commuters (both around 27%). In contrast, a large portion of the well-informed (72%) as well as the benefits-dominated (72%) market segments were comprised of respondents who drove alone on their commute possibly indicating their higher amounts of exposure and preference towards new technologies. A larger share of the respondents from uncertain (65%) as well the concerns-dominated (63%) market segments spent 5 minutes or less parking their car during their commute trip. Lastly, it was also seen that more than one-fourth of the concerns-dominated market segment (28%) consisted of respondents whose households owned one or more vehicles.

The two-step cluster analysis employed in this study reveals interesting insights on autonomous vehicle consumer market segments. Aside from the conventional benefits- and concerns-dominated market segments, the uncertain and the well-informed market segments create value in enhancing our understanding of the consumer demographics. This information would likely provide auto-manufacturers, and transportation professionals with market segments that

Table 3.4. Cluster-Based Descriptive Statistics of the Variables of Interest in Understanding Consumers' Likelihood of Adopting Autonomous Vehicles

	Benefits-Dominated(N=468)		Concerns-Dominated (N=567)		Uncertain (N=681)		Well-Informed (N=761)	
Variable Description	**Mean**	**Mean**	**SD**	**SD**	**Mean**	**SD**	**Mean**	**SD**
Male Respondent Indicator (1 if respondent is male, 0 otherwise)	0.596	0.519	0.500	0.491	0.551	0.498	0.499	0.500
University Respondent Indicator (1 if respondent is classified as a university respondent, 0 otherwise)	0.303	0.217	0.413	0.460	0.198	0.399	0.343	0.475
Millennial Indicator (1 if respondent is less than 35 years old, 0 otherwise)	0.331	0.224	0.417	0.471	0.213	0.410	0.376	0.485
Baby Boomer Indicator (1 if respondent is between 50 and 64 years old, 0 otherwise)	0.280	0.367	0.482	0.449	0.292	0.456	0.268	0.443
Great Generation Indicator (1 if respondent is 65 years old or older, 0 otherwise)	0.248	0.289	0.454	0.432	0.395	0.489	0.246	0.431
White Respondent Indicator (1 if respondent is classified as white, 0 otherwise)	0.823	0.852	0.356	0.382	0.860	0.347	0.820	0.384
Hispanic/Black Respondent Indicator (1 if respondent is classified as Hispanic/black, 0 otherwise	0.105	0.078	0.268	0.306	0.100	0.300	0.105	0.307
Low Income Household Indicator (1 if respondent belongs to a household that earns an annual income less than $50,000, 0 otherwise)	0.267	0.272	0.445	0.443	0.257	0.437	0.294	0.456
High Income Household Indicator (1 if respondent belongs to a household that earns an annual income of more than $100,000, 0 otherwise)	0.370	0.340	0.474	0.483	0.372	0.484	0.360	0.480
Two Person Household Indicator (1 if respondent belongs to a two person household, 0 otherwise)	0.453	0.489	0.500	0.498	0.526	0.500	0.472	0.500

Non-Commuter Indicator (1 if respondent does not commute to work, 0 otherwise)	0.175	0.266	0.442	0.381	0.273	0.446	0.181	0.386
Drive Alone Commuter Indicator (1 if respondent usually drives alone for his commute trip, 0 otherwise)	0.729	0.621	0.486	0.445	0.639	0.481	0.715	0.452
Short Commute Distance Indicator (1 if respondent travels a one-way distance less than 5 miles for their commute, 0 otherwise)	0.182	0.134	0.341	0.386	0.170	0.376	0.205	0.404
Long Commute Time Indicator (1 if respondent travels 45 minutes or more one-way for their commute, 0 otherwise)	0.135	0.106	0.308	0.342	0.103	0.304	0.142	0.349
Total Daily Travel Time Indicator (1 if respondent travels less than 30 minutes every day for all their trips, 0 otherwise)	0.263	0.259	0.439	0.441	0.261	0.440	0.242	0.428
Total Daily Travel Time Indicator (1 if respondent travels more than 90 minutes every day for all their trips, 0 otherwise)	0.150	0.109	0.312	0.357	0.103	0.304	0.154	0.361
Low Parking Time Indicator (1 if respondent spends 5 minutes or less in order to park their vehicle, 0 otherwise)	0.577	0.626	0.484	0.496	0.648	0.478	0.548	0.498
Zero Vehicle Ownership Indicator (1 if respondent is a member of a household that owns zero vehicles, 0 otherwise)	0.092	0.041	0.197	0.289	0.056	0.230	0.085	0.280
One-Plus Vehicle Ownership Indicator (1 if respondent is a member of a household that owns more than one vehicle, 0 otherwise)	0.244	0.277	0.448	0.430	0.244	0.430	0.214	0.411
Three-Plus Vehicle Ownership Indicator (1 if respondent is a member of a household that owns three or more vehicles, 0 otherwise)	0.073	0.113	0.317	0.260	0.081	0.273	0.074	0.261
Recent New-Vehicle Purchase Indicator (1 if respondent recently purchased or leased a new vehicle, 0 otherwise)	0.536	0.522	0.500	0.499	0.589	0.492	0.519	0.500
Recent Used-Vehicle Purchase Indicator (1 if respondent recently purchased a used vehicle, 0 otherwise)	0.368	0.430	0.496	0.483	0.352	0.478	0.392	0.488
Crash Indicator (1 if respondent has been involved in a traffic crash in the past, 0 otherwise)	0.741	0.709	0.455	0.438	0.706	0.456	0.735	0.442

provide different opportunities and challenges during the marketing of autonomous-vehicle technologies. Insights from this cluster analysis could also be used to devise different approaches to prepare various consumer segments for a world with autonomous vehicles.

3.6. Individual Attributes and Market-Segment Probabilities

The two-step cluster analysis employed in the previous section provided insights into the different autonomous vehicle market segments. To further understand what factors make respondents more or less likely to belong to a particular market segment, a discrete outcome modeling approach is applied. For each of the four market segments, we define a function that determines an individual's probability of being in a specific market segment. To arrive at an estimable statistical model for this, we define the function that determines the probability of individual i belonging to market segment n as (Washington et al., 2011),

$$MS_{in} = \boldsymbol{\beta}_n \mathbf{X}_{in} + \varepsilon_{in} \tag{3.1}$$

where MS_{in} is a function that determines the probability of individual i belonging to market segment n, $\boldsymbol{\beta}_n$ is a vector of estimable parameters for corresponding to market segment n, $\mathbf{X}_{in}$ is a vector of explanatory variables that affect the probability for individual i for market segment n, and ε_{in} is a disturbance term. If the disturbance terms are assumed to be generalized extreme-valued distributed, a standard multinomial logit model results as (McFadden, 1981):

$$P_i(n) = \frac{EXP\left[\boldsymbol{\beta}_i \mathbf{X}_{in}\right]}{\sum_{\forall N} EXP\left(\boldsymbol{\beta}_I \mathbf{X}_{In}\right)} \tag{3.2}$$

where $P_i(n)$ is the probability of individual i belonging to market segment n.

In model estimation, we also consider the possibility of unobserved heterogeneity across individuals (the possibility that individuals' likelihood of being in a market segment will be affected by explanatory variables differently due to unobserved reasons). To account for the possibility of having one or more of the parameter estimates in the vector $\boldsymbol{\beta}_n$ vary across individuals, we assume a distribution of these parameters and rewrite Equation 3.2 as (Washington et al., 2011)

$$P_i\ (n) = \int_{\mathbf{X}} P_i\ (n)\ f\left(\boldsymbol{\beta}_n \mid \boldsymbol{\varphi}_n\right) d\boldsymbol{\beta}_n \tag{3.3}$$

where $f(\boldsymbol{\beta}_n|\boldsymbol{\varphi}_n)$ is the density function of $\boldsymbol{\beta}_n$, $\boldsymbol{\varphi}_n$ is a vector of parameters describing the density function (mean and variance), and all other terms are as previously defined. This gives the random parameters logit model, the estimation of which is undertaken by simulated maximum likelihood approaches. For the simulation process in model estimation, previous studies have shown that Halton draws (Halton, 1960) can provide a more efficient distribution of simulation draws than purely random draws (Bhat, 2003). In our model estimations, we use 1,000 Halton draws to determine if random parameters are statistically significant (Bhat, 2003; Anastasopoulos and Mannering, 2009; Behnood and Mannering, 2016, have all shown this number to be more than adequate). A variety of parameter distributions including the normal, log-normal, uniform and exponential are tested to determine if any of the model parameters result in statistically significant standard deviations (indicating that unobserved heterogeneity is present and that parameters vary across the various group samples).[3]

To determine the effect that individual explanatory variables have on response probabilities, marginal effects are computed for each explanatory variable. The marginal effect of an explanatory variable gives the effect that a one-unit increase in an explanatory variable has on the response probabilities. Since each respondent would have their own marginal effect, we report the average marginal effect over all respondents.

3.6.1 Model estimation results

To determine the characteristics that increase or decrease respondents' probability of belonging to one of the identified market segments, a random parameters logit model is estimated. However, in all model estimations, we were unable to find statistically significant random parameters, thus all final-estimation model are fixed parameters.

There was initial concern that there might be fundamental differences between the university (University of South Florida respondents) and non-university (American Automobile Association respondents) portions of our sample. To test for this a likelihood ratio test is conducted with the test statistic $X^2 = -2[LL(\boldsymbol{\beta}_{total}) - LL(\boldsymbol{\beta}_{university}) - LL(\boldsymbol{\beta}_{non\text{-}university})]$ where the $LL(\boldsymbol{\beta}_{total})$ is the log-likelihood at convergence of the model using all respondents (both university and non-university respondents), $LL(\boldsymbol{\beta}_{university})$ is the log-likelihood at convergence using only university respondents, and $LL(\boldsymbol{\beta}_{non\text{-}university})$ is the log-likelihood at convergence using only non-university respondents. In this case, we could not reject the null hypothesis that the two survey groups were the same at reasonable confidence levels. This is an interesting result because it suggests that the sociodemographic factors associated with a person's probability of belonging to one of the four market segments (benefits-dominated, concerns-dominated, uncertain, and well-informed) do not differ substantially between the university and non-university samples, indicating the presence of such distinct (yet latent) market segments in various demographic groups of the population. Future empirical studies should explore if this result holds with other population segments as well.

[3] We also considered the possibility of heterogeneity in means and variances (Behnood and Mannering, 2017a, 2017b; Seraneeprakarn, 2017). But likelihood ratio tests showed that these formulations did not significantly improve the model estimation results.

Table 3.5. Logit-Model Estimation Results for the Probability of Belonging to a Specific Market Segment (t-statistic in Parenthesis)

Variable Description	Estimated Parameter (t statistic)	Marginal Effects by segment			
		Benefits Dominated	Concerns Dominated	Uncertain	Well Informed
Factors for the benefits-dominated market segment					
Male Respondent Indicator (1 if respondent is male, 0 otherwise)	0.361 (3.33)	0.0543	-0.0150	-0.0183	-0.0210
University Respondent Indicator (1 if respondent is classified as a university respondent, 0 otherwise)	0.405 (3.04)	0.0610	-0.0168	-0.0205	-0.0236
High Commute Distance Indicator (1 if respondent travels a one-way distance of 20 miles or more for their commute, 0 otherwise)	0.445 (3.28)	0.0670	-0.0185	-0.0225	-0.0259
Medium Overall Daily Travel Time Indicator (1 if respondent travels 45 minutes or less on an average for their total daily travel, 0 otherwise)	0.250 (2.31)	0.0377	-0.0104	-0.0127	-0.0146
High Parking Time Indicator (1 if respondent spent 10 or more minutes on an average towards finding a parking spot during their commute, 0 otherwise)	0.223 (1.75)	0.0337	-0.0093	-0.0113	-0.0130
Factors for the concerns-dominated market segment					
Constant	1.147 (7.23)				
Baby Boomer Indicator (1 if respondent is 50 to 64 years old, 0 otherwise)	0.359 (3.31)	-0.0149	0.0617	-0.0227	-0.0241
High Income Household Indicator (1 if respondent is a member of a household with an annual income $150,000 or more, 0 otherwise)	-0.263 (-1.77)	0.0109	-0.0453	0.0166	0.0177
Graduate Degree Indicator (1 if respondent's highest educational qualification is a graduate degree, 0 otherwise)	-0.239 (-2.21)	0.099	-0.0411	0.0151	0.0160
Zero Vehicle Ownership Indicator (1 if respondent is a member of a household that owns more than three vehicles, 0 otherwise)	0.353 (2.14)	-0.0147	0.0607	-0.0223	-0.0237
Recent New Vehicle Purchase Category Indicator (1 if respondent most recently purchased or leased a new vehicle, 0 otherwise)	-0.211 (-2.11)	-0.0147	0.0607	-0.0223	-0.0237
Drive Alone Commuter Indicator (1 if respondent typically drives alone to their commute, 0 otherwise)	-0.414 (-3.77)	0.0172	-0.0712	0.0262	0.0278

Major Injury Severity Indicator (1 if the respondent was involved in one or more crashes, and respondent-involved crashes resulted in major injury, 0 otherwise)	-0.295 (-2.21)	0.0123	-0.0508	0.0187	0.0199
Factors for the uncertain market segment					
Constant	0.968 (7.41)				
Generation X Indicator (1 if respondent is 35 to 49 years old, 0 otherwise)	-0.336 (-2.22)	0.0170	0.0213	-0.0660	0.0276
Very High Licensed Driver Household Indicator (1 if respondent is a member of a household that has 3 or more licensed drivers, 0 otherwise)	-0.274 (-2.32)	0.0139	0.0173	-0.0537	0.0225
Non-Commuter Indicator (1 if respondent does not undertake a commute trip, 0 otherwise)	0.257 (2.23)	-0.0130	-0.0163	0.0505	-0.0212
Factors for the well-informed market segment					
Constant	0.844 (6.77)				
Millennial Indicator (1 if respondent is less than 35 years old, 0 otherwise)	0.603 (5.76)	-0.0352	-0.0405	-0.0486	0.1253
High Commute Time Indicator (1 if respondent spent a total of 60 minutes or more on an average for their one-way commute, 0 otherwise)	0.339 (1.69)	-0.0198	-0.0228	-0.0278	0.0704
Number of observations	2477				
Log-likelihood at zero	-3433.85				
Log-likelihood at convergence	-3319.39				

3.6.2 Discussion of estimation findings

Being male, on average, was found to increase the probability of belonging to the benefits-dominated market segment relative to their female counterparts (as shown in Table 3.5).Additionally, respondents from the university portion of the sample had a higher probability of belonging to the benefits-dominated market segment relative to their American Automobile Association counterparts.

Comparing across generations, millennials (respondents who were less than 35 years of age) had a higher probability of belonging to the well-informed market segment, relative to the senior counterparts (the great generation, respondents over the age of 65 years). Model results from Table 3.5 also show that Gen-X-ers (respondents between 35 and 49 years old) had a lower probability of being in the uncertain market segment, relative to the great generation. In contrast, baby boomers (respondents between 50 and 64 years old) had a higher probability of belonging to the concerns-dominated market segment than their senior counterparts (the great generation).

Respondents with a graduate degree, those belonging to high-income households (with an annual income of $150,000 or more), and those who drove alone for their commute trips had a lower probability of belonging to the concerns-dominated market segment. It is possible that a higher education, and higher annual household income expose respondents to greater discussions and discourse on the benefits of autonomous vehicles. Respondents belonging to households with a high number of licensed drivers (respondents who are members of a household that has 3 or more licensed drivers) had a lower probability of belonging to the uncertain market segment.

Current vehicle ownership, and vehicle purchase behavior also influence the likelihood of belonging to a particular market segment. While respondents in households that own more than three vehicles had a higher probability of belonging to a concerns-dominated market segment, those who most recently purchased or leased a new vehicle had a lower probability of belonging to the same market segment. At the outset, these results look counter-intuitive. On closer look, however, it is likely that respondents belonging to households with more than three vehicles are entrenched in a driving culture. Therefore, they may be less likely to be enthused about a technology that takes the pleasure of driving away from them, and may be likely to be skeptical of its benefits. Additionally, most new vehicles are equipped with some advanced safety and automation features that may be playing a role in reducing consumers' skepticism about the potential issues with emerging technologies.

Several model results show the impact of current travel characteristics on respondents' probability of belonging to a particular market segment (see Table 3.5). For instance, respondents traveling long commute distances (an average one-way distance of 20 miles or more), or those that travel 45 minutes or less for all travel in a day (on average) had a higher probability of belonging to the benefits-dominated market segment. Meanwhile, respondents who spent a substantial amount of time on their commute (an average of 60 minutes or more for their one-way commute) had a higher probability of belonging to the well-informed market segment. In contrast, respondents who drove alone for their commute trips had a lower probability of belonging a concerns-dominated market segment. Lastly, non-commuters (respondents who did not undertake a commute trip) had a higher probability of being uncertain about the benefits and concerns regarding autonomous vehicles, which might ultimately impact their adoption of such technologies when they become available in the market.

Finally, injuries suffered in the respondent-involved crashes had a significant impact on the likelihood of belonging to a certain market segment. Respondents who were involved in crashes that resulted in major injuries had a lower probability of belonging to the concerns-dominated market segment. It is likely that these respondents, due to their past experiences, enjoy higher levels of awareness and exposure on safety and automation features that are aimed at reducing fatalities and enhancing the perception of safety in driving.

3.7. Autonomous Vehicle Adoption Likelihoods by Identified Market Segments

Given the results of the previous sections, we now look specifically at the autonomous-vehicle adoption likelihoods for each of the four market segments in order to gain an understanding of the factors affecting adoption likelihoods. For each of the four identified market segments, we seek to study survey respondents' likelihood of adopting autonomous vehicles when they become available in the market, with their ordered responses being *extremely unlikely*, *unlikely*, *unsure*, *likely*, *extremely likely*. An ordered probability modeling approach is appropriate in this case to account for the ordering of the data (Greene, 1997; Washington et al., 2011). Ordered probability models are typically specified by defining an unobserved variable, z, for each respondent i as the linear function,

$$z_i = \boldsymbol{\beta}\mathbf{X}_i + \varepsilon_i \tag{3.4}$$

where $\mathbf{X}_i$ is a vector of explanatory variables determining the discrete responses for respondent i, $\boldsymbol{\beta}$ is a vector of estimable parameters, and ε_i is a disturbance term. Using this equation, observed ordinal responses, y_i, are defined as (with 1 = *extremely unlikely*, 2 = *unlikely*, 3 = *unsure*, 4 = *likely*, and 5 = *extremely likely*),

$$\begin{aligned} y_i &= 1 \text{ if } z_i \leq \mu_0 \\ &= 2 \text{ if } \mu_0 < z_i \leq \mu_1 \\ &= 3 \text{ if } \mu_1 < z_i \leq \mu_2 \\ &= 4 \text{ if } \mu_2 < z_i \leq \mu_3 \\ &= 5 \text{ if } z_i \geq \mu_3, \end{aligned} \tag{3.5}$$

where μ's are estimable parameters (thresholds) that define y_i and are estimated jointly with the model parameters $\boldsymbol{\beta}$. With this, as shown in Washington et al. (2011) and other sources, if ε_i is assumed to be normally distributed across respondents an ordered probit model results with ordered categorical selection probabilities (removing subscripting i for notational convenience and noting that without loss of generality, μ_0 can be set equal to zero thus requiring the estimation of only three thresholds, μ_1, μ_2, and μ_3 to define all 5 selection probabilities),

$$\begin{aligned} P(y = 1) &= \Phi(-\boldsymbol{\beta}\mathbf{X}) \\ P(y = 2) &= \Phi(\mu_1-\boldsymbol{\beta}\mathbf{X}) - \Phi(-\boldsymbol{\beta}\mathbf{X}) \\ P(y = 3) &= \Phi(\mu_2-\boldsymbol{\beta}\mathbf{X}) - \Phi(\mu_1-\boldsymbol{\beta}\mathbf{X}) \\ P(y = 4) &= \Phi(\mu_3-\boldsymbol{\beta}\mathbf{X}) - \Phi(\mu_2-\boldsymbol{\beta}\mathbf{X}) \\ P(y = 5) &= 1 - \Phi(\mu_3-\boldsymbol{\beta}\mathbf{X}), \end{aligned} \tag{3.6}$$

where $\Phi(.)$ is the cumulative normal distribution.

For model interpretation, we again compute marginal effects (the effect that a one-unit change in X has on the probability of ordered outcome n) for each of the five ordered outcomes. As with the previous logit analysis these marginal effects are computed for each respondent and then averaged over all respondents to arrive at an average marginal effect for the population.

Finally, we account for the possibility of unobserved heterogeneity in the data by allowing for parameters to vary across respondents. We use a standard random parameters approach with (please see Mannering et al., 2016 for a full description of alternate heterogeneity modeling approaches),

$$\beta_{ki} = \beta_k + \varphi_{ki} \,, \tag{3.7}$$

where β_{ki} is the parameter estimate for explanatory variable k (one of the elements in the parameter vector $\boldsymbol{\beta}$) for respondent i, β_k is the mean parameter estimate for explanatory variable k, and φ_i is a randomly distributed term (for example, normally distributed term with mean zero and variance σ^2). As with the previously described random parameters logit model, estimation of the random parameters ordered probit is undertaken by simulated maximum likelihood approaches and again 1,000 Halton draws are used.

3.7.1 Adoption Likelihood Model Estimation Results

Respondents' likelihood of adopting autonomous vehicles is very likely to be different across the four consumer market segments. This is because, among other factors, members of a market segment perceive the potential benefits and concerns with autonomous vehicles differently than members of other market segments. To test if separate statistical models should be estimated for the various consumer market segments, a likelihood ratio test is conducted with the test statistic $X^2 = -2[LL(\boldsymbol{\beta}_{total}) - LL(\boldsymbol{\beta}_{benefits}) - LL(\boldsymbol{\beta}_{uncertain}) - LL(\boldsymbol{\beta}_{informed}) - LL(\boldsymbol{\beta}_{concerns})]$ where $LL(\boldsymbol{\beta}_{total})$ is the log-likelihood at convergence of the model using all respondents (from all four consumer markets segments), and $LL(\boldsymbol{\beta}_{benefits})$, $LL(\boldsymbol{\beta}_{uncertain})$, $LL(\boldsymbol{\beta}_{informed})$, $LL(\boldsymbol{\beta}_{concerns})$ are the log-likelihoods at convergence using only respondents in benefits-dominated, uncertain, well-informed, and concerns-dominated clusters, respectively. This test statistic is χ^2 distributed with degrees of freedom equal to the difference in the number of parameters of the total model and the sum of the parameters in the market-segment models. The value of X^2 is 703.77, and with 41 degrees of freedom, we are more than 99% confident that the null hypothesis that the four cluster respondents are the same, can be rejected. Thus, separate models are estimated for all the four clusters (the benefits-dominated, uncertain, well-informed, and concerns-dominated market segment).

A likelihood ratio test was also conducted to determine if there were significant differences between the University of South Florida (university respondents) and the American Automobile Association respondents. In each individual market segment model, we could not reject the null hypothesis that the two survey groups were the same at anywhere near the 90% confidence level. Thus, we do not estimate separate models for these two survey groups.

Parameters producing statistically significant standard deviations for their assumed distribution are treated as parameters that vary across the population (with each observation having its own parameter), and the remaining parameters are treated as fixed parameters because the standard deviations are not significantly different from zero (one parameter for all observations). Again, a log-likelihood ratio test was conducted to statistically compare the random parameters and the fixed parameters model for all the consumer market segments. With all random parameters, the normal distribution was used because other distribution did not result in statistically superior models as indicated by likelihood ratio tests.

Model estimation results are shown in Table 3.6 while the marginal effects across each consumer market segment are shown in Tables 3.7, 3.8, 3.9, and 3.10.

3.7.1.1 Intended Adoption for the Benefits-Dominated Market Segment

Gender was found to be insignificant in the adoption decisions concerning autonomous vehicles in the benefits-dominated market segment. Being a millennial (under the age of 35 years), or a baby boomer (between 50 and 64 years old) in this market segment increased the probability of being *extremely likely* to adopt autonomous vehicles (as indicated by the marginal effects in Table 3.7). However, the effect of these variables varied across the population indicating considerable heterogeneity among millennials and baby boomers in a benefits-dominated market segment. Respondents belonging to the great generation (65 years old or above) in the benefits-dominated market segment were more *unlikely* or *extremely unlikely* to adopt autonomous vehicles

Table 3.6. Random Parameters Ordered Probit Model Estimation of Consumers' Likelihood of Adopting Autonomous Vehicles across Different Market Segments

Variable Description	Benefits-Dominated		Concerns-Dominated		Uncertain		Well-Informed	
	Estimated Parameter	***t* statistic**	**Estimated Parameter**	***t* statistic**	**Estimated Parameter**	***t* statistic**	**Estimated Parameter**	***t* statistic**
Constant	6.515	10.54	0.426	2.58	0.510	5.11	1.971	11.35
Male Respondent Indicator (1 if respondent is male, 0 otherwise) *Standard deviation of parameter*	--	--	-0.253	-2.48	--	--	0.261 (*0.284*)	3.16 (*4.98*)
University Respondent Indicator (1 if respondent is classified as a university respondent, 0 otherwise)	--	--	0.482	3.42	--	--	-0.560	-4.62
Millennial Indicator (1 if respondent is less than 35 years old, 0 otherwise) *Standard deviation of parameter*	0.391 (*1.478*)	1.46 (*8.98*)	--	--	--	--	--	--
Baby Boomer Indicator (1 if respondent is between 50 and 64 years old, 0 otherwise) *Standard deviation of parameter*	0.730 (*2.042*)	2.82 (*9.90*)	--	--	--	--	-0.311	-2.60
Great Generation Indicator (1 if respondent is 65 years old or older, 0 otherwise)	-0.802	-2.69	0.289	2.42	--	--	-0.621	-4.87
White Respondent Indicator (1 if respondent is classified as white, 0 otherwise) *Standard deviation of parameter*	--	--	-0.273	-2.03	--	--	-0.088 (*0.141*)	-0.81 (*3.22*)
Hispanic/Black Respondent Indicator (1 if respondent is classified as Hispanic/black, 0 otherwise) *Standard deviation of parameter*	1.155 (*3.871*)	3.36 (*9.38*)	--	--	--	--	--	--
Low Income Household Indicator (1 if respondent belongs to a household that earns an annual income less than $50,000, 0 otherwise)	--	--	-0.221	-1.82	-0.316	-3.23	--	--
High Income Household Indicator (1 if respondent belongs to a household that earns an annual income of more than $100,000, 0 otherwise) *Standard deviation of parameter*	0.570	3.13	--	--	--	--	0.131 (*0.648*)	1.51 (*9.21*)

Two Person Household Indicator (1 if respondent belongs to a two person household, 0 otherwise) *Standard deviation of parameter*	--	--	--	--	--	--	-0.025 (*0.630*)	-0.30 (*10.29*)
Non-Commuter Indicator (1 if respondent does not commute to work, 0 otherwise) *Standard deviation of parameter*	0.560 (*1.872*)	2.37 (*8.42*)	--	--	--	--	--	--
Drive Alone Commuter Indicator (1 if respondent usually drives alone for his commute trip, 0 otherwise)	--	--	--	--	0.308	3.49	--	--
Short Commute Distance Indicator (1 if respondent travels a one-way distance less than 5 miles for their commute, 0 otherwise) *Standard deviation of parameter*	-0.635 (*2.153*)	-3.05 (*9.39*)	--	--	-0.187	-1.65	--	--
Long Commute Time Indicator (1 if respondent travels 45 minutes or more one-way for their commute, 0 otherwise) *Standard deviation of parameter*	0.747 (*2.282*)	2.88 (*8.02*)	--	--	--	--	0.253 (*0.479*)	2.08 (*4.34*)
Total Daily Travel Time Indicator (1 if respondent travels less than 30 minutes every day for all their trips, 0 otherwise)	--	--	--	--	0.210	2.20	--	--
Total Daily Travel Time Indicator (1 if respondent travels more than 90 minutes every day for all their trips, 0 otherwise)	--	--	-0.368	-2.21	--	--	--	--
Low Parking Time Indicator (1 if respondent spends 5 minutes or less in order to park their vehicle, 0 otherwise) *Standard deviation of parameter*	0.106 (*2.281*)	0.65 (*12.73*)	-0.332	-3.08	--	--	--	--
Zero Vehicle Ownership Indicator (1 if respondent is a member of a household that owns zero vehicles, 0 otherwise) *Standard deviation of parameter*	--	--	--	--	--	--	0.553 (*0.863*)	3.46 (*5.91*)
One-Plus Vehicle Ownership Indicator (1 if respondent is a member of a household that owns more than one vehicle, 0 otherwise)	-0.870	-4.88	--	--	--	--	--	--

Three-Plus Vehicle Ownership Indicator (1 if respondent is a member of a household that owns three or more vehicles, 0 otherwise)	--	--	--	--	-0.162	-1.66	--	--
Recent New Vehicle Purchase Indicator (1 if respondent recently purchased or leased a new vehicle, 0 otherwise)	0.963	5.59	--	--	--	--	0.264	3.08
Recent Used Vehicle Purchase Indicator (1 if respondent recently purchased a used vehicle, 0 otherwise)	--	--	-0.183	-1.80	--	--	--	--
Crash Indicator (1 if respondent has been involved in a traffic crash in the past, 0 otherwise) *Standard deviation of parameter*	0.427 (*0.947*)	2.42 (*8.92*)	--	--	--	--	--	--
Threshold, μ_1	1.864	4.57	0.732	13.86	0.671	16.65	0.922	11.57
Threshold, μ_2	3.921	8.74	1.491	18.01	1.451	28.03	1.668	18.59
Threshold, μ_3	7.818	13.16	2.159	15.63	2.600	27.35	3.277	28.78
Number of observations	468		567		681		761	
Log-likelihood at constants	-517.453		-657.672		-1006.730		-1089.447	
Log-likelihood at convergence	-488.059		-631.104		-990.081		-1061.253	

Table 3.7. Marginal Effects for Significant Parameters in the Benefits-Dominated Market Segment

Variable Description	Marginal Effects				
	Extremely Unlikely	**Unlikely**	**Unsure**	**Likely**	**Extremely Likely**
Millennial Indicator (1 if respondent is less than 35 years old, 0 otherwise)	0.0	0.0	-0.0003	-0.145	0.146
Baby Boomer Indicator (1 if respondent is between 50 and 64 years old, 0 otherwise)	0.0	0.0	-0.0005	-0.276	0.276
Great Generation Indicator (1 if respondent is 65 years old or older, 0 otherwise)	0.0	0.0000004	0.0019	0.258	-0.260
Hispanic/Black Respondent Indicator (1 if respondent is classified as Hispanic/black, 0 otherwise)	0.0	0.0	-0.0004	-0.436	0.436
High Income Household Indicator (1 if respondent belongs to a household that earns an annual income of more than $100,000, 0 otherwise)	0.0	0.0	-0.0005	-0.211	0.212
Non-Commuter Indicator (1 if respondent does not commute to work, 0 otherwise)	0.0	0.0	-0.0003	-0.214	0.215
Short Commute Distance Indicator (1 if respondent travels one-way distance less than 5 miles for their commute, 0 otherwise)	0.0	0.0000003	0.0014	0.206	-0.207
Long Commute Time Indicator (1 if respondent travels 45 minutes or more one-way for their commute, 0 otherwise)	0.0	0.00	-0.0004	-0.288	0.288
Low Parking Time Indicator (1 if respondent spends 5 minutes or less to park their vehicle, 0 otherwise)	0.0	0.00	-0.0001	0.038	-0.039
One-Plus Vehicle Ownership Indicator (1 if respondent is a member of a household that owns more than one vehicle, 0 otherwise)	0.0	0.00	0.0007	0.322	-0.323
Recent New Vehicle Purchase Indicator (1 if respondent recently purchased or leased a new vehicle, 0 otherwise)	0.0	-0.0000003	-0.0015	-0.323	0.336
Crash Indicator (1 if respondent has been involved in a traffic crash in the past, 0 otherwise)	0.0	0.00	-0.0006	-0.147	0.148

Table 3.8. Marginal Effects for Significant Parameters in the Concerns-Dominated Market Segment

Variable Description	Marginal Effects				
	Extremely Unlikely	**Unlikely**	**Unsure**	**Likely**	**Extremely Likely**
Male Respondent Indicator (1 if respondent is male, 0 otherwise)	0.100	-0.030	-0.042	-0.020	-0.007
University Respondent Indicator (1 if respondent is classified as a university respondent, 0 otherwise)	-0.191	0.044	0.083	0.045	0.019
Great Generation Indicator (1 if respondent is 65 years old or older, 0 otherwise)	-0.115	0.031	0.049	0.024	0.010
White Respondent Indicator (1 if respondent is classified as white, 0 otherwise)	0.108	-0.028	-0.047	-0.024	-0.01
Low Income Household Indicator (1 if respondent belongs to a household that earns an annual income less than $50,000, 0 otherwise)	0.086	-0.029	-0.036	-0.016	-0.006
Total Daily Travel Time Indicator (1 if respondent travels 90 minutes or more, every day, for all their trips, 0 otherwise)	0.141	-0.052	-0.057	-0.023	-0.008
Low Parking Time Indicator (1 if respondent spends 5 minutes or less to park their vehicle, 0 otherwise)	0.131	-0.037	-0.056	-0.027	-0.011
Recent Used Vehicle Purchase Indicator (1 if respondent recently purchased a used vehicle, 0 otherwise)	0.072	-0.023	-0.30	-0.014	-0.005

Table 3.9. Marginal Effects for Significant Parameters in the Uncertain Market Segment

Variable Description	Marginal Effects				
	Extremely Unlikely	Unlikely	Unsure	Likely	Extremely Likely
Low Income Household Indicator (1 if respondent belongs to a household that earns an annual income less than $50,000, 0 otherwise)	0.103	0.023	-0.035	-0.072	-0.018
Drive Alone Commuter Indicator (1 if respondent usually drives alone for his commute trip, 0 otherwise)	-0.098	-0.024	0.032	0.071	0.019
Short Commute Distance Indicator (1 if respondent commutes less than 5 miles, one-way, 0 otherwise)	0.061	0.014	-0.021	-0.043	-0.011
Total Daily Travel Time Indicator (1 if respondent travels less than 30 minutes every day for all their trips, 0 otherwise)	-0.067	-0.016	0.022	0.048	0.013
Three-Plus Vehicle Ownership Indicator (1 if respondent is a member of a household that owns three or more vehicles, 0 otherwise)	0.052	0.013	-0.017	-0.037	-0.01

Table 3.10. Marginal Effects for Significant Parameters in the Well-Informed Market Segment

Variable Description	Marginal Effects: Extremely Unlikely	Unlikely	Unsure	Likely	Extremely Likely
Male Respondent Indicator (1 if respondent is male, 0 otherwise)	-0.019	-0.049	-0.035	0.064	0.038
University Respondent Indicator (1 if respondent is classified as a university respondent, 0 otherwise)	0.048	0.108	0.064	-0.147	-0.073
Baby Boomer Indicator (1 if respondent is between 50 and 64 years old, 0 otherwise)	0.026	0.061	0.037	-0.082	-0.041
Great Generation Indicator (1 if respondent is 65 years old or older, 0 otherwise)	0.060	0.122	0.062	-0.171	-0.073
White Respondent Indicator (1 if respondent is classified as white, 0 otherwise)	0.006	0.016	0.012	-0.021	-0.013
High Income Household Indicator (1 if respondent belongs to a household that earns an annual income of more than $100,000, 0 otherwise)	-0.009	-0.024	-0.018	0.031	0.020
Two Person Household Indicator (1 if respondent belongs to a two-person household, 0 otherwise)	0.002	0.005	0.003	-0.006	-0.004
Long Commute Time Indicator (1 if respondent travels 45 minutes or more one-way for their commute, 0 otherwise)	-0.015	-0.044	-0.037	0.055	0.042
Zero Vehicle Ownership Indicator (1 if respondent is a member of a household that owns zero vehicles, 0 otherwise)	-0.026	-0.086	-0.087	0.091	0.109
Recent New Vehicle Purchase Indicator (1 if respondent recently purchased or leased a new vehicle, 0 otherwise)	-0.019	-0.049	-0.035	0.065	0.038

when they became available in the market, relative to Gen-X-ers (respondents between 35 and 49 years old). Also, from the marginal effect estimates shown in Table 3.7, Hispanic/black respondents in this market segment were more extremely likely to adopt autonomous vehicles when they become available in the market, relative to everyone else, although the finding of a statistically significant random parameter suggests considerable heterogeneity in the effect of this variable across the population.

Household income was found to be an important indicator in adoption decisions regarding autonomous vehicles. Respondents belonging to high-income households (with an annual household income of more than $100,000) in the benefits-dominated market segment were more *extremely likely* to adopt autonomous vehicles when they become available in the market (as shown in Table 3.7).

Several model results show the influence of current travel characteristics on consumers' adoption likelihood decisions with autonomous vehicles in the benefits-dominated market segment. For instance, it is interesting to note that non-commuter respondents were found to be more *extremely likely* to adopt autonomous vehicles when they become available in the market. Likewise, respondents who spent a long time on their commutes (commuting 45 minutes or more one-way, on average) were more *extremely likely* to adopt autonomous vehicles. However, this behavior was not echoed by short-distance commuters. Respondents who traveled, on average, a one-way distance of 5 miles or less for their commute in the benefits-dominated market segment were less *extremely likely* to adopt autonomous vehicles, relative to those who traveled longer commute distances. It is plausible that commuters who spent longer times on the road see the benefits of adopting these technologies in comparison to their short-distance counterparts. There was considerable heterogeneity among observations as the variables depicting current travel characteristics in the benefits-dominated market segment were random parameters, indicating that not all commuters' adoption behaviors were similar.

Parking seems to have a complex effect on respondents' likelihood of adopting autonomous vehicles in the benefits-dominated market segment. Respondents who spent 5 minutes or less to park their vehicle were less *extremely likely* to adopt autonomous vehicles when they become available in the market, however, the magnitude of the effect of this variable is quite small as indicated by the marginal effects shown in Table 3.7.

Vehicle ownership had an interesting influence on intended adoption of autonomous vehicles in this market segment. If a respondent belongs to a household that owns two or more vehicles, they were found to be less *extremely likely* to adopt autonomous vehicles when they become available in the market. At the outset, these results look counter-intuitive. On closer look, however, it is likely that respondents in households that own a large number of vehicles are likely entrenched in a driving culture. Therefore, they are likely to be less enthused about adopting a technology that takes the pleasure of driving away from the driver.

Respondents in the benefits-dominated market segment, who recently purchased or leased a new vehicle, were more *extremely likely* to adopt an autonomous vehicle when they become available in the market. Most new cars are fitted with some advanced safety/automation features that make drivers safer, and respondents seem willing to invest further in such technologies to potentially reduce crashes. Lastly, previous crash experience made respondents more *extremely likely* to adopt autonomous vehicles when they became available in the market, perhaps indicating an increased emphasis on safety in their driving.

3.7.2.2 Intended Adoption for the Concerns-Dominated Market Segment

Model estimation results for parameters that were found to be statistically significant in the concerns-dominated market segment showed that being male, on average, increased the probability of being *extremely unlikely* to adopt autonomous vehicles. Part of the reason for this statistically significant male/female difference could be due to men being more risk averse concerning new vehicle technologies in this market segment. This finding can be supported by recent literature that shows empirical evidence of gender differences in risk-taking in transportation-related decisions (Abay and Mannering, 2016).

University respondents in the concerns-dominated market segment were less *extremely unlikely* to adopt an autonomous vehicle than their American Automobile Association counterparts. Please note that while our earlier statistical tests show that the overall differences between these two groups is not significant, the significance of this indicator variable suggests there are at least some residual differences between the two groups, likely reflecting some common unobserved characteristics.

Comparing across generations, the great generation in the concerns-dominated market segment, on average, were less *extremely unlikely* to adopt an autonomous vehicle relative to their younger counterparts. These results are somewhat surprising considering that these respondents, despite foreseeing the potential issues with autonomous vehicles seem more positive about their adoption. It is possible that older generations are currently unable to use ubiquitous modes of transportation due to their advanced age and therefore potentially see autonomous vehicles as a solution to their travel needs.

In contrast, white respondents in the concerns-dominated market segment were found to be more *extremely unlikely* to adopt an autonomous vehicle, relative to other respondents. Household income was a significant parameter in adoption in the concerns-dominated market segment. Respondents belonging to low-income households in the concerns-dominated market segment were more *extremely unlikely* to adopt an autonomous vehicle (refer marginal effects in Table 3.8).

Current travel characteristics were also found to be influential aspects in adoption decisions regarding autonomous vehicles. For instance, respondents who spent 90 minutes or more, on average, traveling daily for all their trips were more *extremely unlikely* to adopt an autonomous vehicle. Spending more time on the road in an already concerns-dominated environment perhaps increases their skepticism of the effectiveness of technologies that improve safety, and related aspects.

Lastly, respondents who spent 5 minutes or less parking their vehicle during their commute trip, or those who recently purchased a used vehicle were found to be more *extremely unlikely* to adopt autonomous vehicles.

3.7.2.3 Intended Adoption of the Uncertain Market Segment

In the uncertain market segment with the possibility of limited awareness of the potential benefits and concerns of autonomous vehicles, respondent’s gender, age, and ethnicity were found to have no significant influence on the intended adoption of autonomous vehicles. However, respondents from low-income households (with an annual household income less than \$50,000), in the uncertain market segment, were more *unlikely* or *extremely unlikely* to adopt autonomous vehicles (as shown by the marginal effects in Table 3.9). This may be due to the high costs involved in adopting emerging vehicle technologies.

The influence of current travel characteristics on autonomous vehicle adoption likelihoods is evident in the uncertain market segment as well. For instance, respondents who drove alone to work in this market segment were less *unlikely* or *extremely unlikely* to adopt an autonomous vehicle when they become available in the market. It is interesting that despite the uncertainty regarding potential benefits and concerns with autonomous vehicles, these respondents perhaps see driving to work as a highly onerous task and could very well consider investing in emerging vehicle technologies to alleviate this. Similarly, respondents who traveled less than 30 minutes every day for all their trips were less *unlikely* or *extremely unlikely* to adopt an autonomous vehicle. In contrast, respondents who commute less than 5 miles were more *unlikely* or *extremely unlikely* to adopt autonomous vehicles (refer to Table 3.9). These above-mentioned parameters show the rather complex relationship between current travel characteristics and the intended adoption likelihoods of autonomous vehicles.

Lastly, respondents belonging to households with high vehicle ownership (with three or more vehicles) in the uncertain market segment were more *unlikely* or *extremely unlikely* to adopt an autonomous vehicle compared to respondents from other households.

3.7.2.4 Intended Adoption of the Well-Informed Market Segment

Gender was found to play a significant role in the adoption of autonomous vehicles in a well-informed market segment. The marginal effects from Table 3.10 show that being male, on average, increased the probability of being more *likely* or *extremely likely* to adopt an autonomous vehicle. However, in the well-informed market segment, the effect of the variable was found to vary significantly across respondents (producing a statistically significant random parameter), suggesting considerable heterogeneity across observations. Part of the reason for this statistically significant male/female difference could be due to men being less risk-averse in well-informed market segment, relative to women (Abay and Mannering, 2016). Interestingly, university members in the well-informed market segment were less *likely* or *extremely likely* to adopt an autonomous vehicle, relative to their American Automobile Association counterparts.

Comparing across generations, well-informed baby boomers and the great generation were less *likely* or *extremely likely* to adopt an autonomous vehicle, relative to their younger counterparts (the millennials, and the Gen-X-ers). Despite being equally aware of the potential benefits and concerns with autonomous vehicles, the older generations seem to want to use a wait-and-see approach before they adopt emerging vehicle technologies such as autonomous vehicles in comparison to their younger cohorts. This is consistent with previous transportation literature that points towards generational-level differences in transportation decisions and overall travel behavior (Circella et al., 2016).

White respondents in the well-informed market segment were less *likely* or *extremely likely* to adopt autonomous vehicles when they become available in the market. The effect of this ethnicity variable varied across the population in the well-informed market segment (Table 10), again implying heterogeneous effects suggesting, for example, that not all white respondents in a well-informed market segment behave in the same way.

Model estimation results showed the significance of several household-level indicators towards consumers' intended adoption of autonomous vehicles in the well-informed market segment. For instance, respondents belonging to high-income households (with annual income of $100,000 or more) in the well-informed market segment were more *likely* or *extremely likely* to adopt autonomous vehicles. It is likely that high-income households are also exposed to greater amounts of discussion and discourse on emerging vehicle technologies and therefore see the merit in early adoption of such technologies, or simply find them more affordable. However, the effect of the variable was found to vary significantly across respondents (producing a statistically significant random variable), suggesting considerable heterogeneity across respondents.

Respondents in a two-person household were less *likely* or *extremely likely* to adopt an autonomous vehicle. This variable also produced a statistically significant random parameter suggesting that not all two-person households were same in their adoption decisions. Like that observed in the benefits-dominated market segment, respondents who commuted 45 minutes or more one-way, on average, were more *likely* or *extremely likely* to adopt autonomous vehicles. This parameter also produced a statistically significant random variable suggesting that not all long-commute respondents think in this way.

Lastly, zero-vehicle households, or those that recently purchased or leased a new vehicle in a well-informed market segment, were more *likely* or *extremely likely* to adopt an autonomous vehicle. The parameter on zero-vehicle households produced a statistically significant random parameter suggesting that not all-zero vehicle households behave the same way regarding their adoption decisions. This is intuitive because not owning a car could be due to different reasons, such as economic constraints (a consequence) or green life styles (a choice), thus their intention towards adopting autonomous vehicles would be different.

3.8. Differences in Adoption Behavior Across Identified Market Segments

Table 11 summarizes intended-adoption findings by market segment with variables making respondents more likely to adopt autonomous vehicles (+), less likely to adopt autonomous

Table 3.11. Effects of Variables on Consumers' Likelihood of Adopting Autonomous Vehicles across Different Market Segments

Variable Description	**Benefits-Dominated**	**Concerns-Dominated**	**Uncertain**	**Well-Informed**
	Effect	**Effect**	**Effect**	**Effect**
Male Respondent Indicator (1 if respondent is male, 0 otherwise)	ns	-	ns	+*
University Respondent Indicator (1 if respondent is classified as a university respondent, 0 otherwise)	ns	+	ns	-
Millennial Indicator (1 if respondent is less than 35 years old, 0 otherwise)	+*	ns	ns	ns
Baby Boomer Indicator (1 if respondent is between 50 and 64 years old, 0 otherwise)	+*	ns	ns	-
Great Generation Indicator (1 if respondent is 65 years old or older, 0 otherwise)	-	+	ns	-
White Respondent Indicator (1 if respondent is classified as white, 0 otherwise)	ns	-	ns	-*
Hispanic/Black Respondent Indicator (1 if respondent is classified as Hispanic/black, 0 otherwise)	+*	ns	ns	ns
Low Income Household Indicator (1 if respondent belongs to a household that earns an annual income less than $50,000, 0 otherwise)	ns	-	-	ns
High Income Household Indicator (1 if respondent belongs to a household that earns an annual income of more than $100,000, 0 otherwise)	+	ns	ns	+*
Two Person Household Indicator (1 if respondent belongs to a two person household, 0 otherwise)	ns	ns	ns	-*
Non-Commuter Indicator (1 if respondent does not commute to work, 0 otherwise)	+*	ns	ns	ns
Drive Alone Commuter Indicator (1 if respondent usually drives alone for his commute trip, 0 otherwise)	ns	ns	+	ns
Short Commute Distance Indicator (1 if respondent travels a one-way distance less than 5 miles for their commute, 0 otherwise)	-*	ns	-	ns
Long Commute Time Indicator (1 if respondent travels 45 minutes or more one-way for their commute, 0 otherwise)	+*	ns	ns	+*
Total Daily Travel Time Indicator (1 if respondent travels less than 30 minutes every day for all their trips, 0 otherwise)	ns	ns	+	ns

Total Daily Travel Time Indicator (1 if respondent travels more than 90 minutes every day for all their trips, 0 otherwise)	ns	-	ns	ns
Low Parking Time Indicator (1 if respondent spends 5 minutes or less in order to park their vehicle, 0 otherwise)	+*	-	ns	ns
Zero Vehicle Ownership Indicator (1 if respondent is a member of a household that owns zero vehicles, 0 otherwise)	ns	ns	ns	+*
One-Plus Vehicle Ownership Indicator (1 if respondent is a member of a household that owns more than one vehicle, 0 otherwise)	-	ns	ns	ns
Three-Plus Vehicle Ownership Indicator (1 if respondent is a member of a household that owns three or more vehicles, 0 otherwise)	ns	ns	-	ns
Recent New Vehicle Purchase Indicator (1 if respondent recently purchased or leased a new vehicle, 0 otherwise)	+	ns	ns	+
Recent Used Vehicle Purchase Indicator (1 if respondent recently purchased a used vehicle, 0 otherwise)	ns	-	ns	ns
Crash Indicator (1 if respondent has been involved in a traffic crash in the past, 0 otherwise)	+*	ns	ns	ns

"+" more likely to adopt autonomous vehicles; "-" less likely to adopt autonomous vehicles; "ns" not a statistically significant effect; * random parameter

vehicles (-) or having no statistically significant effect on the adoption process (ns), and random parameters are represented by (*). Table 3.11 shows that specific variables tend to have widely different influences across market segments, reflecting the importance of our market-segment classification.

The benefits-dominated market segment (characterized by the highest score for likelihood of intended adoption, 4.24, as shown in Table 3.3) has twelve significant variables with various generations [millennials (+*), baby boomers (+*), great generation (-)], ethnicity [being Hispanic/blacks (+*)], high household income (+), several current travel characteristics [non-commuters (+*), short commute distances (-*), long commute times (+*)], parking (+*), vehicle ownership and purchasing behavior [multi-vehicle ownership (-), recent new vehicle purchase (+)], and crash history (+*) influencing the adoption of autonomous vehicles. It is interesting to note that being male, and belonging to the university population had no influence in the adoption decisions in a benefits-dominated market segment (as shown in Table 3.11).

The concerns-dominated market segment (characterized by the lowest score for likelihood of intended adoption, 1.74 out of 5, as shown in Table 3.3) has eight significant variables with gender [males (-)], belonging to the university population (+), belonging to the great generation (+), ethnicity [being white (-)], low household income (-), total daily travel time (-), parking (-), and vehicle purchasing behavior [recent used vehicle purchase (-)] influencing the adoption of autonomous vehicles (as shown in Table 3.11). Interestingly, current vehicle ownership has no impact in autonomous vehicle adoption decisions in a concerns-dominated market segment.

The uncertain market segment (characterized by their somewhat lower score for likelihood of intended adoption, 2.54 out of 5, as shown in Table 3.3) has only five significant variables that influence the adoption decisions with low household income (-), some current travel characteristics [drive alone commuter (+), short commute distances (-), total daily travel time (+)], and multi-vehicle ownership (-) influencing the adoption of autonomous vehicles (as shown in Table 3.11). It is interesting to note that all respondent demographics except income are insignificant in the adoption decisions in an uncertain market segment. The inherent uncertainty in these respondents has to do with factors other than respondent demographics.

Lastly, the well-informed market segment (characterized by their somewhat higher score for likelihood of intended adoption, 3.39 out of 5, as shown in Table 3.3) has ten significant variables that influence the adoption decisions with gender [males (+*)], belonging to the university population [-], older respondents [baby boomers (-), great generation (-)], high household income (+*), longer commute times (+*), and vehicle ownership and purchasing behavior [zero vehicle households (+*), recent new vehicle purchase (+)] influencing the adoption of autonomous vehicles. Most current travel characteristics (except longer commute times) and parking are insignificant in the adoption decisions in a well-informed market segment. The well-rounded opinions regarding autonomous vehicles seems to make the experiences during current travel negligible during the adoption decisions (as shown in Table 3.11).

3.9. Summary and Conclusions

This chapter sought to provide insights into the relationship between consumers' perceptions (of the benefits and concerns) towards autonomous vehicles on their intended adoption. To achieve this, we conduct a survey which asks respondents their opinions on the potential benefits and concerns with autonomous vehicles, and on their intended adoption of autonomous vehicles. Using these data, a two-step cluster analysis of consumers' perceptions of benefits and concerns revealed four autonomous vehicle consumer market segments (the benefits-dominated, concerns-dominated, uncertain, and well-informed market segments). Interestingly, these market segments range from one, positive extreme (the benefits-dominated segment) to another, negative extreme (the concerns-dominated segment) in their mean stated likelihood of adoption of autonomous vehicles.

A multinomial logit model was then estimated to determine the probability of respondents belonging to a specific autonomous vehicle consumer market segment. Our estimation results show that many different factors such as gender, respondent characteristics, household characteristics, current travel characteristics, and crash history influence the probability of respondents belonging to a specific autonomous vehicle consumer market segment. The analysis also suggests that the sociodemographic factors associated with a person's probability of belonging to one of the four market segments (benefits-dominated, concerns-dominated, uncertain, and well-informed) do not differ substantially between the university and non-university samples, indicating the presence of such distinct (yet latent) market segments in various demographic groups of the population.

Given the presence of distinct market segments in both the university and non-university samples and the differences among these segments in the mean stated likelihood of adoption of autonomous vehicles, we estimated separate random parameter ordered probit models of the intended adoption likelihoods for each of the four previously determined consumer market segments. Likelihood ratio tests of these (market-segmented) models against an unsegmented model on the entire sample rejected the null hypothesis that the indented adoption behavior is same across these market segments. Further, it is interesting that several factors that are significant across some market segments (such as gender, current travel characteristics, or vehicle ownership) are insignificant across others, reflecting the considerable differences among market segments.

The insights obtained from this study can be used to form initial strategies for specific consumer market segments to increase the likelihood of autonomous vehicle acceptance and adoption. The study can also help better understand the sentiments of the general public relating to their willingness to use such emerging technologies. However, it is important to keep in mind that people's perceptions of emerging vehicle technologies like autonomous vehicles are not likely to be temporally stable. People's perceptions of these technologies will undoubtedly change as autonomous vehicle technology evolves, and individuals gather additional information and experience with these vehicle technologies. Thus it is important to view the findings in this chapter with some caution in light of this.

Chapter 4: Shared Automated Vehicles: A Statistical Analysis of Consumer Use Likelihoods and Concerns

4.1. Introduction

The term sharing economy has been widely used in recent years by scientists, economists and people in both public and private sectors. It refers to new emerging business models that allow people to share underutilized resources in more effective ways. The concept of a shared economy could have considerable impact on transportation systems. Shared mobility, the shared use of a motor vehicle, bicycle, or other low-speed transportation mode, is one facet of the sharing economy (Shaheen et al., 2016) and has the potential to disrupt the current transportation system (Meyer and Shaheen, 2017). In fact, in recent years there has been a growing focus on shared mobility as a key element of a sustainable transportation paradigm. The generalization of a shared mobility typology could include bike-sharing, car-sharing, ride-sharing, and the sharing-related potential of private and public transportation network companies (Kodransky and Lewenstein, 2014). However, shared mobility also includes ride-sourcing companies such as Lyft and Uber as well as courier network services and flexible good delivery (Shaheen et al., 2016).

Low vehicle-occupancies, high crash rates combined with high levels of greenhouse gas emissions create an opportunity for shared automated vehicles to enter the market and improve some of these issues (Transportation Research Board, 2013; Greenblatt and Shaheen, 2015; Bills and Walker, 2017). Also, in cases when physical disability is a barrier to mobility and accessibility, shared automated vehicles could be a valuable transportation option. In particular, groups of users who are visually or physically impaired could find this new transportation mode most helpful (assuming that the price of the service is not a major constraint).

In terms of car-sharing, the likely forthcoming introduction of fully driverless vehicles has the potential to substantially alter the mindset with regard to sharing in the context of privately-operating vehicles (Fagnant and Kockelman, 2015). The presence of such automated vehicles opens up an opportunity for creation of a new transportation mode that would combine features of short-term on-demand utilization with self-driving capabilities and, in some applications, effectively a driverless taxi (Fagnant et al., 2015). Recent research has concluded that one-way car-sharing could reduce vehicle ownership, vehicle-miles travelled, and greenhouse gas emissions as well as contributes to modal shifts done by the users. A study in five North American cities found that 2 to 5% of members sold a vehicle due to one-way car-sharing, and another 7% to 10% did not acquire a vehicle, depending on the city (Stocker and Shaheen, 2017). Research that has focused on station-based car-sharing were able to identify multiple positive impacts such as less vehicle travel and lower emissions (Martin and Shaheen, 2011) while reducing the need for parking (Shaheen et. al, 2010). Similar benefits will be likely seen after the introduction of shared automated vehicles (Meyer and Shaheen, 2017).

However, there is likely to be considerable uncertainty with regard to the user demographic and usage trajectory of shared automated vehicles due to factors such as continuously changing general knowledge (as individuals gather experience and information relating to this new mode). To better understand the likely usage trajectory of shared automated vehicles, and the various consumer preferences and concerns that may influence this trajectory, this study uses survey data gathered from members of the

American Automobile Association Foundation from 12 U.S. states. The questionnaires given to these members (distributed in the spring of 2015) incorporated a number of detailed questions relating to shared automated vehicle preferences and concerns, as well as socio-demographic data. Note that, because our study is exploratory, and the likely presence of temporal instability in the preferences associated with the introduction of shared automated vehicles (as individuals gather information and form attitudes and opinions), a nationally representative sample was not an initial priority (see Mannering 2018 for a discussion of temporal instability and its likely effects in statistical modeling, which would be a concern for a national survey as well). Also, with regard to most emerging technologies, and especially the likely disruptive and controversial nature of shared automated vehicles, individual preferences and opinions will be highly volatile early in consumers' information-gathering process (Parvathy and Mannering, 2018). Thus it is important to stress that studies based on any data collected before a technology has reached significant maturity, in terms of market penetration, must be viewed as exploratory in nature since consumer preferences and opinions will be unstable. However, it is also worth pointing out that the American Automobile Association survey we use in this study has considerable spatial and socio-economic diversity, covering 12 States that are representative of the U.S. population and their membership is fairly inclusive with one in four families being members in the U.S.

The collected data were used to estimate statistical models of individual preferences for using shared automated vehicles, and possible concerns associated with shared automated vehicles. The chapter begins with a literature review that focuses on factors likely to play a role in shared automated vehicles use. This is followed by a description of the survey and research design and the methodological approach used to analyze the data. Model estimation results are then provided and discussed. Finally, the chapter concludes with a summary and discussion of key findings.

4.2. Usage likelihoods and Concerns

Shared automated vehicles potentially provide a new transportation concept that could merge features of traditional public and private transportation modes (Haboucha et al., 2017). Emerging automated vehicles combined with on-demand mobility may provide important alternatives to conventional transportation. While the development phase for automated vehicles is still in the early stages, several analysts have predicted potentially transformative changes to personal transportation (Greenblatt and Shaheen, 2015; Stocker and Shaheen, 2017).

Several recent research efforts have provided initial insights into the potential benefits and drawbacks of shared automated vehicles as a transportation alternative. For example, Dia and Javanshour (2017) showed that incorporating shared automated vehicles as a modal option can significantly reduce the total number of vehicles required to meet the transport needs of a community. They also argued that shared automated vehicles can potentially decrease parking requirements, which would free up this space for other purposes. However, their study also identified some possible negative impacts such as an increase in total kilometers of travel due to vehicle repositioning (which has potential implications for greenhouse gas emissions), but this can be mitigated if a large proportion of self-driving vehicles used propulsion systems that are more environmentally friendly than the internal combustion engine. Also, growing interest in on-demand mobility coupled with automated vehicles may amplify adoption of both, and further lower energy use and greenhouse gas emissions through the use of small, efficient shared automated vehicles (Greenblatt and Shaheen, 2015). Greenblatt and Saxena (2015) found that that, if a low vehicle occupancy trip was accommodated by appropriately sized vehicles, fleet average energy consumption could drop by almost a factor of two. This would cause the greenhouse gas emission to decrease especially if the vehicles are electric. Shared automated vehicles' life-cycle greenhouse gas emissions per distance driven could fall by roughly 90 % relative to today's average passenger vehicle.

In other work, to get a sense of the potential impacts of shared automated vehicles, an Atlanta-based traffic simulation study found that parking areas could be reduced by approximately 4.5% once shared automated vehicles start to serve 5% of the trips. The reduction is achieved primarily through improving vehicle utilization and reductions in private car use (Zhang and Guhathakurta, 2017). With regard to the impact that shared automated vehicles can potentially have on the presence of privately own vehicles, it was found that each car-sharing vehicle can substitute 9 to 13 vehicles from the road (Martin and Shaheen, 2011).Martin and Shaheen (2011) found that 25% of participants sold a vehicle and 25% of postponed a vehicle purchase due to car-sharing (in a sample containing approximately 9500 participants), thus it is likely that shared automated vehicles will also impact vehicle purchases.

Menon et al. (2018) looked at variables that play significant roles in relinquishing a traditional human-driven vehicle in the presence of a shared automated vehicle option. They found that gender, age, education, commute distance and daily travel time, as well as previous vehicle-crash history and vehicle inventory, all play significant roles in the likelihood of giving up a traditional vehicle. Males were found to be less likely to relinquish a household vehicle in a single-vehicle household; however they were more likely to do so in a multivehicle households, relative to their female counterparts. Millennials and people with graduate degrees were found to be more likely to give up a vehicle when shared automated vehicles became available. Also, previous involvement in a traffic crash was found to make respondents more likely to relinquish their vehicles to potentially use shared automated vehicles.

Bansal and Kockelman (2017) also provided significant insights with regard to opinions about automated vehicles and shared automated vehicles. In their work they found that a little over half of survey respondents (54.4%) agreed that automated vehicles would be useful transportation advancement, but 58.4% indicated that they would have some concerns relating to their use. Only 19.5% felt comfortable with allowing an automated vehicle to drive them independently, but 41.4% agreed that automated vehicles will be a fixture in future transportation systems.

In other work, Haboucha et al. (2017) compiled recent findings from literature and news sources and concluded that men are more likely to use self-driving technologies relative to women. They also uncovered several contradictory findings among studies. For example, they found some research papers concluded that younger individuals are more likely to use shared automated vehicles while others found that elderly individuals would be more interested in using shared automated vehicles.

Another perspective on estimating the use of shared automated vehicles was undertaken by Lavieri et al. (2017). They found that the consumer interest in automated-vehicle use is a function of lifestyle and current transportation choices. The authors identified two basic propensities; the “green lifestyle” and “technology savviness” from their study. The authors explain that, because overall lifestyle types had significant impacts on travel behavior, so that individuals who tend to lead green lifestyles and are tech-savvy would be more likely to be early users of shared automated vehicles. Furthermore, people who were younger, had higher education, lived in urban settings and already exhibited transportation sharing behaviors were more likely to use shared automated vehicles.

However, it is important to note that the use of shared automated vehicle services will take time to mature and that current shared mobility modal definitions and business models will continue to change over time. For instance, car-sharing and ride-sourcing may start to look like very similar services (Stocker and Shaheen, 2017).

The intent of this chapter is to add to the growing body of literature on shared automated vehicles willingness-to-use and concerns. Using statistical models that address potential unobserved heterogeneity in shared automated-vehicle data, we seek to uncover the complex relationships among socio-demographic characteristics and shared automated vehicle choice and provide insights that can be used to assist in the development of various marketing strategies to support shared automated vehicle use adoption. The survey approach, research design, and empirical findings we employ to do this are presented in the following sections.

4.3. Survey and Research Design

A web-based survey was used to collect the data on the potential use of shared automated vehicles and concerns associated with them. A survey consisting of 75 questions was distributed to the American Automobile Association South in the United States. The twelve states that belong to American Automobile Association South are: Florida, Georgia, Illinois, Indiana, Iowa, Michigan, Minnesota, Nebraska, North Dakota, Tennessee and Wisconsin. These states cover multiple geographic locations and different climates, thus it allowed reaching diverse populations. However, it is important to note that, because the survey was distributed in the United States, transferability of findings to other countries may be problematic.

In addition to questions on the preferences and concerns with regard to shared automated, the survey covered a variety of socio-demographic and household characteristics, travel behavior, and travel history. The survey was composed of three parts: a part that collected general information including socio-demographics, travel characteristics, crash history, and vehicle inventories; a part that gathered information on people’s perceptions of automated vehicles; and a final part that collected data on the anticipated impacts of automated vehicles and shared automated vehicles including willingness to use and concerns. For this final portion of the survey, we classified everyone who indicated an interest in any mode of shared automated vehicles (privately owned, rented, publicly owned) as belonging to the willingness-to-use group as opposed to people who indicated no interest in using. Survey respondents were also asked to rank their concerns relating to shared automated vehicles. Because ranking concerns about a technology/modal option that is not yet available is a difficult to imagine, only the most important concern for each person was used in the analysis.

The data used in the analysis includes 782 respondents. Of these 782 respondents, 467 respondents indicated no interest in using shared automated vehicles whereas 315 indicated an interest in using at least one of the alternate modes of shared automated vehicles (car sharing with or without ownership, ride sharing with or without ownership, taxi-service, or as public transit). In terms of most important concerns regarding automated vehicles, 468 respondents indicated that they were most concerned about safety, 112 respondents were most concerned about reliability, followed by 109 people concerned mainly about privacy, 57 were concerned about possible increases in travel time, and 36 where concerned about the cost of the service.

It is important to note, that in an environment where shared automated vehicles do not yet exist, understanding people's willingness to use shared automated vehicles, and understanding the various concerns that they may have with shared automated vehicles, presents obvious challenges. For example, respondents' current level of knowledge about shared automated vehicles, which is likely to vary significantly across demographic sectors of the population, will affect usage likelihoods and concerns. In addition, the willingness to use shared automated vehicles is dynamically changing as the general population is gaining familiarity with this technology. There is a large body of research that shows that preferences for new technologies are likely to be temporally unstable during the early stages of use as experiences form attitudes and preferences (Mannering, 2018). With these issues in mind, our analysis will still provide potentially important initial insights into the willingness to use shared automated vehicles and concerns associated with them.

4.4. Methodology

With regard to possible shared automated vehicle use, we seek to statistically analyze two of the responses gathered on the survey. First, we consider respondents' binary response as to whether or not they would be interested in any one of the following shared automated vehicle modes: automated vehicle car-sharing with car ownership (you own an automated vehicle and are willing to make it available to others); automated vehicle car-sharing without car ownership (you obtain an automated vehicle from individual owners or companies that offer car sharing service via car-sharing platforms such as a web page, smartphone app); automated vehicle ride-sharing with car ownership (you own an automated vehicle and are willing to share the ride with co-passengers such as colleagues, friends, or someone you might find through ride-sharing web pages or apps); automated vehicle ride-sharing without car ownership (you share the ride with an automated vehicle owner such as colleagues, friends, or someone you might find through ride-sharing web pages or apps); automated vehicle taxi service; or automated vehicle public transit. Second, we consider one of the following four concerns that individual respondents view as most concerning with regard to the use of shared automated vehicles: safety concerns, privacy concerns, reliability concerns (uncertainty as to whether they will be able to arrive at their destination on time) and other concerns (including the cost of service and potentially higher travel times which, based on a statistical analysis are combined into a single choice).

Both of the above responses are discrete; the yes/no response about willingness to use one of the shared automated vehicle modes, and the four-alternative option relating to respondents' greatest concern regarding shared automated vehicles (safety, privacy, reliability and other). To arrive at an estimable statistical model for these two responses, we define a function that either determines the probability of shared automated vehicle use (a yes or no response) or the greatest concern (safety, privacy, reliability and other) as (Washington et al., 2011),

$$F_{in} = \boldsymbol{\beta}_i \mathbf{X}_{in} + \varepsilon_{in} \tag{4.1}$$

where F_{in} is a function that determines the probability of respondent n selecting response i (yes/no or one of the four concern responses), $\boldsymbol{\beta}_i$ is a vector of estimable parameters for corresponding to discrete response i, $\mathbf{X}_{in}$ is a vector of explanatory variables that affect the probability of discrete response i for respondent n, and ε_{in} is a disturbance term. If the disturbance terms are assumed to be generalized extreme-valued distributed, a standard multinomial logit model results as (McFadden, 1981):

$$P_n(i) = \frac{EXP\left[\boldsymbol{\beta}_i \mathbf{X}_{in}\right]}{\sum_{\forall I} EXP\left(\boldsymbol{\beta}_I \mathbf{X}_{In}\right)} \tag{4.2}$$

where $P_n(i)$ is the probability of respondent n giving response i and I is the set of responses (either yes and no for the usage likelihood, or safety, privacy, reliability and other for the greatest concern likelihood).

In model estimation, we also would like to account for the possibility of unobserved heterogeneity across respondents. That is, the possibility that individual respondents will be affected by explanatory variables differently due to unobserved reasons (this is particularly likely with the introduction of new technologies such as shared automated vehicles). To account for the possibility of having one or more parameter estimates in the vector $\boldsymbol{\beta}_i$ vary across respondents, we assume a distribution of these parameters and rewrite Equation 4.2 as (Washington et al., 2011)

$$P_n(i) = \int_{\mathbf{X}} P_n(i)\; f\left(\boldsymbol{\beta}_i \,|\, \boldsymbol{\varphi}_i\right) d\boldsymbol{\beta}_i \tag{4.3}$$

where $f(\boldsymbol{\beta}_i|\boldsymbol{\varphi}_i)$ is the density function of $\boldsymbol{\beta}_i$, $\boldsymbol{\varphi}_i$ is a vector of parameters describing the density function (mean and variance), and all other terms are as previously defined. This gives the random parameters logit model (also called the mixed logit model). There are other methods of capturing unobserved heterogeneity, such as a latent class model, with the preference of one approach over another often being data-specific (see Mannering et al., 2016, for a complete discussion of this). Our estimation of a latent class formulation did not seem to track the unobserved heterogeneity in our data as well so we present only the random parameters estimation results in this chapter.

Estimation of the random parameters logit model is undertaken by simulated maximum likelihood approaches because the required integration of the logit formula over the distribution of parameters is not closed form. Previous studies have shown that Halton draws can provide a more efficient distribution of simulation draws than purely random draws (McFadden and Ruud, 1994; Bhat, 2003). We use 1,000 Halton draws in our simulated likelihood functions, a number that has been shown to be more than sufficient to provide accurate parameter estimates (Bhat, 2003; Milton et al., 2008; Anastasopoulos and Mannering, 2009; Behnood and Mannering, 2016). In this study, the normal distribution has been used in estimation of the explanatory variables because it provided the best statistical fit for our two response models (other distributions such as the log-normal, uniform, and exponential were not found to produce statistically better results than the normal distribution).

To determine the effect that individual explanatory variables will have on response probabilities, marginal effects are computed for each explanatory variable. The marginal effect of an explanatory variable gives the effect that a one-unit increase in an explanatory variable has on the response probabilities. For indicator variables (that assume values of zero or one), marginal effects will give the effect of the explanatory variable going from zero to one (Washington et al., 2011). Each respondent will have their own marginal effect and we will report the average marginal effect over all respondents.

4.5. Model Results

Table 4.1 presents summary statistics of variables found to be statistically significant. Tables 4.2 and 4.3 provide the random parameters logit models results, including parameter estimates, t-statistics and marginal effects, for the shared automated vehicle-use model and the shared automated vehicle-concern model, respectively. Please note that in these tables, explanatory variables found to be statistically significant were grouped into four categories; socio-demographic factors, household characteristics, travel behavior factors, and crash involvement. It should also be noted that likelihood ratio tests were performed to determine if random parameters logit models presented herein were statistically different than their fixed parameters counterparts. In both modeling cases we can reject the null hypothesis that fixed and random parameters models are the same with over 95% confidence. Thus only random parameters model results are presented.

4.5.1 Model estimation results: willingness to use shared automated vehicles

For socio-demographic factors, we found that respondents who identified themselves as Caucasian were less likely to use shared automated vehicles, which may be the result of cultural norms set up among this group. The higher education indicator

Table 4.1. Summary Statistics for Variables Included in Final Model Estimations

Variable Description	Mean	Standard Deviation
Socio-demographic factors		
Male indicator (1 if respondent is a male, 0 otherwise)	0.59	0.49
Older age indicator (1 if respondent is at least 60 years old, 0 otherwise)	0.42	0.49
Caucasian indicator (1 if respondent identifies as Caucasian, 0 otherwise)	0.89	0.46
Black/African American ethnicity indicator (1 if respondent is Black/African American, 0 otherwise)	0.038	0.19
High education indicator (1 if respondent has at least bachelor's degree, 0 otherwise)	0.69	0.46
Graduate level education indicator (1 if respondent has completed graduate school, 0 otherwise)	0.37	0.48
Household characteristics		
Small household indicator (1 if there are at most 2 people living in a household, 0 otherwise)	0.65	0.48
Three people household size indicator (1 if a household size is 3 people, 0 otherwise)	0.17	0.37
One-vehicle household indicator (1 if a household owns 1 vehicle, 0 otherwise)	0.21	0.41
More than four-vehicle household indicator 1 if a household owns (or leases) more than 4 vehicles, 0 otherwise)	0.13	0.33

One licensed driver household indicator (1 if a household has exactly 1 licensed driver, 0 otherwise)	0.21	0.41
Most recent vehicle leased indicator (1 if the most recent vehicle was new leased, 0 otherwise)	0.10	0.31
Travel behavior		
Drive-alone commute indicator (1 if respondent commutes to work by driving alone, 0 otherwise)	0.80	0.40
Short one-way distance to the grocery store indicator 1 (1 if the distance to grocery store is 1 mile or less, 0 otherwise)	0.09	0.29
Short one-way distance to the grocery store indicator 2 (1 distance to grocery store is less than 5 miles, 0 otherwise)	0.76	0.43
Longer one-way commute time indicator 1 (1 if respondent has an average commute time of 45 minutes or more, 0 otherwise)	0.71	0.46
Lack of commute indicator (1 if respondent doesn't commute, 0 otherwise)	0.12	0.33
Low usual parking search-time indicator (1 if respondent spends less than 5 minutes on parking, 0 otherwise)	0.69	0.46
Crash involvement		
Crash indicator (1 if respondent has been involved in a vehicle crash, 0 otherwise)	0.76	0.43

Table 4.2. Random Parameters Logit Model for Willingness (Yes or No) to Use Shared Automated Vehicles. (All Random Parameters are Normally Distributed)*

Variable Description	Estimated Parameter	t-Statistic	Marginal Effect
Constant	1.10	0.96	
Socio-demographic factors			
Caucasian indicator (1 if respondent identifies as Caucasian, 0 otherwise)	-1.88	-1.98	-0.103
High education indicator (1 if respondent has at least bachelor's degree, 0 otherwise) (*Standard deviation of parameter distribution*)	0.98 (*5.37*)	1.39 (*1.90*)	0.060
Household characteristics			
Small household indicator (1 if there are at most 2 people living in a household, 0 otherwise)	-1.89	-2.18	-0.070
One-vehicle household indicator (1 if a household owns 1 vehicle, 0 otherwise)	1.52	1.77	0.018
One-driver household indicator (1 if a household has 1 licensed driver, 0 otherwise) (*Standard deviation of parameter distribution*)	-1.69 (*7.01*)	-1.02 (*1.51*)	0.009
Travel behavior			
Drive-alone commute indicator (1 if respondent commutes to work by driving alone, 0 otherwise) (*Standard deviation of parameter distribution*)	-0.76 (*3.57*)	-0.98 (*1.96*)	-0.010

Short one-way distance to the grocery store indicator (1 distance to grocery store is less than 5 miles, 0 otherwise)	1.27	1.65	0.062
Longer one-way commute time indicator (1 if respondent has an average commute time of 45 minutes or more, 0 otherwise)	-1.40	-2.08	-0.063
Low usual parking search-time indicator (1 if respondent spends less than 5 minutes on parking, 0 otherwise)	-1.18	-2.03	-0.048
Crash involvement			
Crash indicator (1 if respondent has been involved in a vehicle crash, 0 otherwise)	0.86	2.10	0.084
Number of observations	782		
Log likelihood at zero	-542.04		
Log likelihood at convergence	-494.13		

* All explanatory variables are in the "Yes" response function with the "No" response function explicitly set to zero.

Table 4.3. Random Parameters Logit Model for Main Concerns Relating to the Use of Shared Automated Vehicle Systems (All Random Parameters are Normally Distributed)

Variable Description*	Estimated Parameter	t-Statistic	Marginal Effects			
			Safety[S]	**Reliability[R]**	**Privacy[P]**	**Other[O]**
Constant [S]	2.21	8.15				
Socio-demographic factors						
Male indicator (1 if respondent is a male, 0 otherwise) [S]	-0.77	-2.85	-0.062	0.022	0.022	0.018
Graduate level education indicator (1 if respondent has completed graduate school, 0 otherwise) [S]	-0.83	-3.11	-0.043	0.017	0.015	0.012
Black/African American ethnicity indicator (1 if respondent is Black/African American, 0 otherwise) [R]	-1.62	-1.53	0.009	-0.021	0.005	0.0006
High education indicator (1 if respondent has at least bachelor's degree, 0 otherwise) [R]	0.44	2.94	-0.016	0.037	-0.011	-0.009
Older age indicator (1 if respondent is at least 60 years old, 0 otherwise) [O]	-0.55	-2.59	0.005	0.003	0.003	-0.011
Household characteristics						
Three people household size indicator (1 if a household size is 3 people, 0 otherwise) [P]	0.57	2.21	-0.006	-0.004	0.013	-0.003
More than four-vehicle household indicator 1 if a household owns (or leases) more than 4 vehicles, 0 otherwise) [P]	0.63	2.01	-0.005	-0.003	0.011	-0.003
Most recent vehicle leased indicator (1 if the most recent vehicle was new leased, 0 otherwise) [O]	1.14	3.56	-0.009	-0.005	-0.005	0.019
Travel behavior						
Lack of commute indicator (1 if respondent doesn't commute, 0 otherwise) [S]	0.86	2.25	0.013	-0.005	-0.004	-0.004
Short one-way distance to the grocery store indicator (1 if the distance to grocery store is 1 mile or less, 0 otherwise) [R]	-0.75	-1.60	0.002	-0.005	0.002	0.001

Crash involvement						
Crash involvement indicator (1 if respondent was involved in a vehicle crash, 0 otherwise) [S] (*Standard deviation of parameter distribution*)	0.69 (*1.71*)	3.06 (*2.24*)	-0.002	0.002	0.0009	0.0004
Number of observations	782					
Log likelihood at zero	-1084.08					
Log likelihood at convergence	-838.75					

* Parameter defined for: [S] Safety; [R] Reliability; [P] Privacy; [O] Other (travel time and cost).

(respondents who had at least bachelor's degree) produced a normally distributed parameter with a mean of 0.98 and a standard deviation of 5.37. This result suggests (a normal distribution with this mean and standard deviation) that 57.2% of the respondents with higher education were more likely to be willing to use shared automated vehicles and 42.8% were less likely, reflecting considerable heterogeneity across this education group.

For household characteristics, Table 4.2 shows that smaller households (with one or two people) had lower probabilities of using shared automated vehicles (relative to households of three or more). This may reflect the fact that smaller households have their transportation needs met and they may not see immediate benefits from using shared automated vehicles. On the other hand, larger households (with three or more people) could likely benefit from having additional transportation option that would allow the accommodation of additional transportation needs for all the household members. In contrast, respondents from households with just one vehicle were more willing to use shared automated vehicles. It is possible that this reflects the fact that such respondents may not be as fully indoctrinated into the current private-vehicle-ownership culture as are households with multiple vehicles. Moreover, households with only one licensed driver produced a normally distributed parameter with a mean of -1.69 and standard deviation equal to 7.01. With this distribution, roughly 59.5% of people from households with one licensed driver will be less likely to use shared automated vehicles whereas 40.5% will be more likely. Based on this finding, the large portion of households with one licensed driver (almost 60%) do not behave in uniform way (as a fixed-parameter finding would have suggested) and that there are other unobserved factors that seem affect their decision when it comes to the adoption of shared automated vehicles.

For travel behavior factors, the indicator variable for commuters who normally drove alone produced a normally distributed parameter with a mean of -0.76 and a standard deviation of 3.57. This results in 58.4% commuters who drive alone being less likely to use shared automated vehicles and 41.6% more likely (again reflecting considerable heterogeneity within this group). Respondents whose one-way distance to the grocery store was less than 5 miles were more likely to be willing to use shared automated vehicles (a 0.062 higher probability as reflected by the average marginal effect presented in Table 4.2). This grocery-store proximity indicator may be capturing development density, reflecting higher shared automated vehicle usage in dense urban areas.

Respondents with longer commutes (45 minutes or more) were found to be less likely to be willing to use shared automated vehicles. This likely reflects uncertainties about the possible reliability of shared automated vehicles (in terms of on-time performance), which may be more critical in long commutes. In contrast, those respondents whose average parking search time during their most regular trip was less than 5 minutes had lower probabilities of using shared automated vehicles (by -0.048 as indicated by average marginal effect). Ease of parking is potentially a strong indicator of satisfaction with the current private-vehicle-ownership paradigm, making the use of shared automated vehicles less likely.

Finally, respondents who had been involved in a vehicle crash were found to be more likely to use shared automated vehicles. This may reflect the possibility that crash-involved respondents may have the expectation that shared automated vehicles will improve safety. It is plausible that people who were involved in a crash are more aware of the fact that human error plays a significant role in road incidents and as the literature says it contributes to 90% of crashes (Litman, 2018). Even during times when automated vehicles are being actively tested on public roads and the first crashes have happened, it is likely that many believe that automated vehicles and their sensor-systems are superior to human drivers.

4.5.2. Model estimation results: concerns associated with shared automated vehicles

The socio-demographic estimation results shown in Table 4.3 indicate that males were found to be less concerned about the safety of shared automated vehicles. This likely reflects deep-seeded and well-established cultural differences between genders with regard to risk estimation and vehicle safety (see Abay and Mannering, 2016). The results in Table 4.3 also show respondents with graduate degrees were found to be less concerned about the safety aspects of shared automated vehicles, which may reflect increased confidence in vehicle-safety technologies among more educated respondents. In addition, respondents who identified themselves as Black/African American were less concerned with the reliability of shared automated vehicles. One might speculate that this finding may reflect possible urban environments and shorter trips, which may make travel-time variance less of a concern in general. It is also plausible that the value of travel time may differ between races. According to the report published by Semega et al. (2017) the 2016 real median income of non- Hispanic White households was $65,041, whereas for the Hispanic-origin and Black/ African American households it was equal to $47,675 and $39,490 respectively. Lower household income, which is associated with the lower value of travel time could be a factor in the above finding as well.

Respondents who had at least bachelor's degree were more concerned with reliability, which may reflect the residential choices of these respondents, and the general increase in travel time uncertainty relating to the choices. This may be because individuals with higher education tend to hold jobs requiring more responsibility and decide to live in particular areas and prioritize reliable transportation options. Also, those respondents who were at least 60 years old were found to be less concerned with travel time and cost compared to their younger counterparts.

With regard to household characteristics, respondents belonging to three-person households as well as those whose households own more than 4 vehicles were more concerned about the privacy of shared automated vehicle compared to smaller and larger households and those owning less vehicles. Those whose most recent vehicle purchase was a new leased car, were found to be more likely to be concerned about the travel time and cost of shared automated vehicles. This is likely capturing fundamental differences households relating to vehicle-fleet size and leasing decisions (see, for example, Mannering et al., 2002 for a discussion of these points).

For travel-behavior factors, respondents who do not commute at all were found to be more concerned about safety of shared automated vehicles compared to respondents with longer commute times. And, respondents whose distance to a grocery store that was less than 1 mile were less likely to be concerned with reliability (on-time arrivals) of shared automated vehicles relative to respondents with longer distances to the grocery store (like reflecting the effects of development density).

Finally, the crash involvement indicator produced a normally distributed parameter with a mean of 0.69 and a standard deviation of 1.71 defined in safety alternative. This results in roughly 65.7% of respondents who were involved in a vehicle crash being more concerned with the safety of shared automated vehicles and 34.3% being less concerned. This finding shows considerable heterogeneity with regard to past accident involvement, which is likely due to the variance in their crash experiences (some respondents will be involved in more and less severe crashes) and other factors.

4.6. Summary and Conclusions

This research focuses on exploring the determinants of shared automated vehicle usage likelihoods and concerns. Our model estimation results show that a wide range of respondent characteristics significantly affect these. With regard to shared automated vehicle usage likelihoods (Table 2), we find that respondents who are in households with just one vehicle, are in close proximity to grocery stores, and have previously been involved in a vehicle crash are more willing to use shared automated vehicles. In contrast, respondents who identify themselves as Caucasian, live in households with 2 or fewer people, have commutes 45 minutes and longer, and require minimal time finding parking are all less likely to use shared automated vehicles. High education levels, small households (with at most 2 people) and driving alone produced results that varied across respondents making some more likely to use shared automated vehicles and others less likely).

With regard to concerns associated with shared automated vehicles, male respondents and those who have graduate education levels were less likely to be concerned about safety. In contrast, those respondents who do not commute were more likely to be concerned with safety. Respondents from three-people households and those whose households own more than four vehicles tended to be mostly concerned about the privacy aspect of shared automated vehicles. With regard to reliability (on-time performance), respondents with at least a bachelor's degree were more likely to be concerned with shared automated vehicles being dependable. In contrast, respondents who identified themselves as Black/African American, and those whose proximity to grocery store was less than 1 mile were less likely to be concerned with the reliability. Variables that were significant on "other" function (travel time and cost concerns) were older age, and whose recent vehicle purchase was a lease.

This chapter's model-estimation findings provide some initial insights into how different groups of people are likely to behave with regard to the use of shared automated vehicles. Our results show that different aspects of shared automated vehicles are important to different groups of respondents. On average, younger and more educated people from households that do not have high auto ownership levels and currently do not exhibit out-of-ordinary travel needs were found to be more willing to use shared automated vehicles.

Our findings could be used to understand and encourage shared mobility behavior, especially during the times when on-demand mobility is growing in popularity. One of the policy implications could involve incentivizing and/or subsidizing the use of shared automated vehicles as oppose to offering and maintaining free parking spaces. Knowing how different groups of people tend to make transportation decisions could provide important insights in planning for both private and public sectors. The identified factors suggest that the people who are more willing to adopt this technology lead urban lifestyles. In urban areas, one of the potential implications could be the need for expanded pick-up and drop-off areas instead of parking spaces. The retrofitting and updating the infrastructure to cater it to the future transportation needs will likely start in urban areas. Providing the necessary space that accommodates comfortable pick-up and drop-off is a significant part of safe operations of shared automated vehicle systems. And our research findings suggest that the dense urban areas contain the most receptive consumers. Our findings also suggest that elderly people, who could greatly benefit from shared automated vehicles, are less willing to use them. Knowing the elderly's unwillingness to use this technology can help policy makers address this issue in advance by allowing them to proactively develop marketing strategies to increase acceptance among the elderly. Such marketing strategies can be extended to other identified groups of people who could benefit the most from shared automated vehicle usage.

One possible concern of our study is the potential limitations of our data (our sampling is limited to the American Automobile Association sample). However, because people's opinions and perceptions are likely to be changing continuously with the introduction of an emerging technology such as shared automated vehicles (particularly in the early use phase when opinions are being formed), the advantage of a larger and more spatial diverse survey is likely to be quite limited. Still, future research could focus on expanding the data set with additional geographic diversity and variables to track the evolution of perceptions with regard to shared automated vehicles both temporally and spatially.

Chapter 5: Temporal Instability and the Analysis of Highway Accident Data

5.1. Introduction

Worldwide, the analysis of highway accident data has formed the basis for the development and implementation of a wide variety of safety policies. Without doubt, many of these policies have made highway travel substantially safer but, worldwide, with more than 1.2 million fatalities annually in highway-related accidents and an estimated 50 million more people injured, highway safety remains a tragic human health issue (World Health Organization, 2015).

While the highway safety field continues to address a wide variety of topics in an effort to reduce the carnage on world highways, in the past few decades researchers and safety analysts have struggled to explaining two longer-term phenomena; the general downward trend in fatalities per distance driven over time in most industrialized countries, and the fact that fatalities per distance driven tend to decline in economic downturns and increase in economic upturns. The general downward trend has often been attributed to improved vehicle-safety technologies, improved highway design, improved impaired driver enforcement, and driver/public education programs, etc. And, the effect of an adverse economy on fatalities per mile driven has been attributed to factors such as changes is discretionary driving patterns, changes in values of time, changes in the distances risky versus safe drivers drive, and so on.

However, there is potentially an additional element at play in these trends. That is, the fundamentals of human behavior may be changing gradually over time and there may be changes in the short-term, to these fundamentals, in response macroeconomic conditions. In fact, as will be shown, there is a vast body of literature from psychology, neuroscience, economics, cognitive science and other fields that suggests that this temporal element, typically overlooked in accident data analysis, could play a key role in explaining accident trends. This has potentially profound implications for traditional statistical analyses that use data to estimate parameters for various explanatory variables to determine the effect that these variables have on the likelihood and resulting injury severities of accidents. Traditional approaches estimate the effect of changes in explanatory variables using statistically estimated parameters (which are typically assumed to be fixed over time) to determine their likely impacts on accident likelihoods and injury severities. But what if the statistically estimated parameter values are changing in some fundamental way over time? The intent of this chapter is to explore this possibility and discuss its potential implications for safety research.

5.2. Overview Accident-Data Analysis and the Temporal Element

Over the years, researchers have applied a vast array of statistical methods to analyze accident-related data in an effort save lives, and to reduce injury severities and property damage resulting from motor-vehicle accidents. Studies that have focused on the statistical analysis of accident data have traditionally addressed one or more of three general objectives: 1) data analysis with the sole intent of quantifying the effect of statistically significant determinants (explanatory variables) on the likelihood and severity of accidents; 2) data analysis with the intent of using the resulting parameter estimates of the statistical model to forecast future accident likelihoods and severities; and 3) data analysis of before and after data to evaluate the effectiveness of a specific safety countermeasure or a change in a specific factor that may influence likelihood and severity of accidents. For all three of these objectives, researchers have most often implicitly made the assumption that the effects of the statistically identified determinants are constant over time (temporally stable). For forecasting (objective 2 above) and before-and-after analysis (objective 3 above) the passage of time makes it obvious why temporal considerations are important. However, the temporal issue may even arise in what we might consider the "cross-sectional" data likely to be used when simply seeking to quantify the effect of statistically significant determinants (objective 1 above). That is, because vehicle accidents are relatively rare events, they tend to be aggregated over time (weeks, months or years) to arrive at a sufficient number of observations for statistical analysis.[4] Thus, the passage of time between accident observations implies the notion of temporal stability has potentially profound implications on virtually all statistical analyses of accident data.

But how reasonable is the assumption of temporal stability of statistically estimated model parameters? To begin to address this question, it is important to first recognize the two general approaches to the statistical analysis of accident data that dominate the literature, because these approaches may have different implications in terms of temporal stability.[5] The first approach focuses on the likelihood of an accident in general, or the likelihood of an accident of a specified injury severity. Again, because accidents rare events, these models typically address the frequency of accidents on a roadway entity over some time period (for a review of these models, see Lord and Mannering, 2010) and, because the focus is on the likelihood of an accident, the detailed accident information available after an accident has occurred is not used in model estimation. The second approach, in contrast, focuses on the injury severity of specific accidents and can potentially make full use of the highly detailed data (such as detailed information on vehicle and occupant characteristics) available after an accident has occurred (for a review of these models, see Savolainen et al., 2011).

However, there is an abundance of relatively recent research that suggests the influence of factors affecting both the likelihood and resulting severity of highway accidents may not be stable over time. With regard to the aggregation of data over a specified time period (months or years) to gather a sufficient number of accident observations to conduct a statistical analysis, there is a growing body of empirical evidence that suggests at least some temporal instability. For example, Malyshkina et al. (2009) and Malyshkina and Mannering (2009) estimated Markov switching models (with estimated accident models alternating between two states over time) which provides some statistical support for temporal instability since ignoring the transition between states would cause a bias in parameter estimation. Malyshkina and Mannering (2010) and Xiong et al. (2014) found the similar statistical support for Markov switching in injury-severity models using the detailed data available conditioned on an accident having occurred. Other studies have looked at temporal instability of such injury-severity models over longer time periods and found that model parameter estimates were not temporally stable.[6] For example, using detailed accident-injury severity data annually from 2004 to 2012, Behnood and Mannering (2015) found that the effect

---

[4] There is also potentially a spatial consideration here as well because accidents also tend to be aggregated over space. There would thus be an implicit assumption that the effects of explanatory variables are spatially stable. See Mannering and Bhat (2014) for a discussion of spatial considerations and a review of the literature on this topic.

[5] There are also some studies that look accident data in aggregate form, such as the number of fatalities per year in a state/province or country. These studies usually apply some form of time-series modeling approach that is typically predicated on the assumption of temporal stability. These models are not addressed explicitly in this paper, but a later footnote is provided to provide additional insight into the effect that possible temporal instability would have on these models.

[6] Other fields of transportation, such as travel-activity modeling and tradition travel-demand modeling have also demonstrated temporal instability (Mannering et al., 1995; Rossi and Bhat, 2014).

that roadway characteristics, vehicle characteristics, and driver characters have on resulting driver-injury severities varied significantly from one year to the next. Subsequent work from these authors (Behnood and Mannering, 2016), showed similar temporal instability with regard to pedestrian injuries resulting from vehicle accidents in Chicago. [7,8]

Clearly, there is a growing body of empirical evidence that suggests potential temporal instability in models of accident likelihood and resulting injury severity. The consequences of ignoring possible temporal effects, and thus not accounting for potential temporal shifts in estimated parameters, could adversely affect the inferences drawn model estimations as well as their ability to be used to forecast and evaluate the effects of safety countermeasures.

What might be the underlying causes of temporal in accident data? As one would expect, factors that form the basis for temporal instability, such as changes in individual behaviors and their evolution over time, have been studied extensively in fields such as psychology, economics, neuroscience, cognitive science, and others, and countless papers have been published on topics that would support temporally instability in accident-data modeling. The case for suspecting temporal instability begins with very basic human desires for immediate versus delayed satisfaction (suggesting a temporal discounting), which is supported by observational data and neuroscience. However, the case for temporal instability evolves beyond that and into the effect that the evolution of individuals' cognitive biases, attitudes, and risk-taking, have on their driving behavior. The current chapter does not seek to provide a comprehensive assessment of the extensive literature that relates to changing human behavior over time, but instead, to present a few well-established findings and discuss their possible implications with regard to the temporal stability assumptions that almost all safety-research studies implicitly make, and the safety forecasts based on these assumptions.

The chapter begins by making a case for temporal instability based on individual driver decision-making and its potential evolution over time. The chapter then moves on to discuss the potential role that cognitive biases, the relationship between macroeconomics and individuals' driving-risk assessment, and the role that the relationship between attitudes and behavior may play in temporal instability. This is followed by a discussion of the primary safety-data analysis methods and how temporal instability may affect their ability to draw inferences and to forecast the likelihood and severity of accidents. The chapter concludes with a discussion and a summary and conclusions section.

---

[7] Temporal instability has also been observed in the demand for safety features on vehicles. For example, Mannering and Winston (1995) looked at the adoption of driver-side airbags in new vehicles in the early 1990s. They found that consumers' willingness to pay for a driver-side airbag in a new vehicle increased over time from $331 in 1990 to $512 dollars in 1993. They also found that media exposure (average number of hours spent watching television per day) and social networks (number of friends owning cars with driver-side airbags) were significant factors affecting willingness to pay.

[8] A notable exception to other findings of temporal instability is the work Malyshkina and Mannering (2008). They found temporal stability in accident-injury parameter estimates for accidents occurring on rural interstates in Indiana between the years 2004 and 2006 data, even when speed limits were increased in 2005 (they did, however, find instability in other roadway classifications such as rural multilane highway). While their approach was less sophisticated in that it did not explicitly account for unobserved heterogeneity as the more recent approaches did (more on this below), their findings may suggest that temporal instability may also vary by highway functional class.

5.3. The General Case for Temporal Instability

Virtually all fields of behavioral research, including psychology, economics, neuroscience, and cognitive science, have theories and empirical evidence that point to temporal instability in decision making.[9] The most basic case for the existence of a temporal element in models of decision making can be made from the widely observed phenomenon that intertemporal decisions involve a trade-offs between time and satisfaction, with preference given to immediate rewards and some level of temporal discounting given to future rewards (Green and Myerson, 2004). In the driving task, individuals can be viewed constantly making temporal trade-offs between immediate satisfaction and delayed satisfaction. For example, speeding and other aggressive driving behaviors provide immediate satisfaction (making trips shorter) which is constantly being weighed against the probability of a citation or accident which would adversely affect the longer term satisfaction of lower insurance rates and uninterrupted driver-license privileges.[10]

In fact, this notion of immediate versus delayed awards has its basis in neuroscience. For example, McClure et al. (2004) used magnetic resonance imaging to demonstrate that different portions of the brain are activated when individuals face immediate rewards versus facing longer-term options. Lieberman et al. (2002) argue that there are automatic and controlled responses ongoing in decision making, and that these processes can be identified by the portions of the brain being used. Brain regions supporting automatic responses include the limbic system whereas brain regions supporting controlled responses are the lateral prefrontal cortex and associated structures.

Driver decision-making can be viewed as an ongoing trade-off between the more automatic responses of the brain versus the controlled processes. The automatic processes are made instinctively with little thought or subsequent insight. For example, the sudden unexpected introduction of an object in front of a vehicle's path will trigger an immediate automatic response from the driver and, as one might imagine, based on genetics, experiences and other factors, this automatic response is likely to vary widely across the population of drivers. Drivers will also be confronted with decisions that involve controlled responses, such as observing unfolding traffic conditions well down the road and having sufficient time to come up with an appropriate decision regarding braking, lane positioning, and so on.

Most research on decision making has focused on the development and evolution of controlled processes which tend to be more logically modeled. For example, cognitive science approaches the study of controlled process by identifying traits, which are baseline personality elements that are assumed to be mostly stable over time, and state effects, which are situational factors that influence individual preferences and decisions that vary over time (Peters and Buchel, 2011). Research suggests that both trait and state effects influence decision making (Bickel at al., 2007), but preferences (which may vary based on decision contexts) can induce state-dependent shifts as individuals adjust to changing environments and changing goals. These state-dependent shifts can vary significantly across individuals even when faced with the same stimuli since individuals will have different baseline traits and state-dependent shifts.

---

[9] Economists have long recognized the possibility of temporal instability in individuals' consumption decisions. Most often this has been framed as an evolution of individuals' tastes over time (see for example, Peston, 1967; Gorman 1967; Fisher and Shell, 1969; Pollak and Wales, 1969; Pollak, 1970; Phlips, 1972; Lluch, 1974). More recently, researchers such as Camerer et al. (2005) have attempted to reconcile elements of neuroscience with more traditional economic approaches to decision making.

[10] This will be discussed in the context of risk-taking behavior later in the paper.

With regard to the above discussion and temporal stability of accident-model estimates, research suggests that individuals' controlled processes will be constantly evolving based on experiences (state-dependent shifts) and even their automatic processes will evolve over time based on accumulated experiences. At the individual level, this would imply considerable instability in model-estimated parameters over time. To be sure, existing statistical analyses of accident data attempt to account for this evolution by including a measurable variables such as age, which would likely have a reasonably strong correlation with how controlled and automatic processes evolve. But age is still likely to be a relatively crude estimate for the variability that is likely to occur across the population as drivers mature. To some extent, accident-data modeling approaches have evolved to capture this evolution by considering it as unobserved heterogeneity, which could allow for the possibility of the age effect to vary from one observation to the next, thus potentially accounting, on some level, for the variation in controlled and automatic from one individual to the next (see Mannering et al., 2016). While providing an approximate replication of existing heterogeneity in the population, such an approach does not provide insights into how individual driver behavior will evolve in the over time.

5.4. The Role of Cognitive Biases

The extant psychological literature associates the underlying personality of drivers as a factor prompting aggressive and risky-driving behavior (Constantinou et al., 2011; Sumer, 2003; Ulleberg and Rundmo, 2003). These studies suggest that personality traits and associated risk-taking behavior may be correlated with risky-driving behavior (such as driving above the speed limit or driving under the influence of alcohol). More fundamentally, drivers' perceptions of risk lie at the core of risk-related decision making, and there is an abundance of research that shows this risk perception often differs significantly from actual risk, and that this difference varies from driver to driver and potentially over time for the same driver.[11]

With regard to potential changes in an individual's perceived risk over time, there are a number of cognitive biases that are likely to come into play (Tversky and Kahneman, 1974; Fischhoff et al., 1993). Perhaps key among these is anchoring bias, which is a bias in risk perception resulting from individuals' tendency to heavily weight the first piece of information they gather concerning risk and make corrections to this only incrementally based on new information they gather over time. With regard to driving, individuals are likely to develop an initial assessment of risk based on knowledge gleaned from initial driver training, published knowledge of risk (for example, knowing that over 30,000 people die in U.S. car accidents annually), and personal experiences (and/or those of acquaintances), with regard to accident risks. Once the initial assessment is formed, the subsequent adjustment as new information is gathered will cluster about this initial assessment and thus anchoring bias provides at least some support for risk perception adjusting only slowly over time. On the other hand, the past few decades have seen unprecedented advances in vehicle technologies. Whenever a new technology is introduced, such as side impact airbags, individuals must gather new risk-related information relating to these technologies. Thus, one could argue that, in eras of rapidly developing vehicle safety technologies,

[11] There is a large body of literature that suggests individuals have overconfidence in their own abilities and thus tend to underestimate their own accident risk relative to the population as a whole (Svenson, 1981). However, work by Benoit and Dubra (2011) suggests that much of this apparent overconfidence and underestimation of personal risk is due to a natural information-gathering process and not necessarily a biased estimation of one's own abilities.

individuals' anchoring bias, with regard to overall safety, is continually shifting based on their developing perceptions of new safety technologies.[12]

Another potentially relevant source of bias in risk perception is availability bias, which results from events being disproportionally reported by the media or even disproportionately experienced and recalled by individual drivers (Kahneman et al., 1982). For example, the media may over-report roll-over accidents due to their often catastrophic injuries, and an individual witnessing a specific type of accident (such as a train-car collision) may substantially overestimate their likelihood. The temporal issues generated from this sort of bias can be problematic if there are changes in media coverage, social networks and/or global experiences of drivers over time.[13] Other potential cognitive biases relevant to risk perception, such as the tendency of individuals to overestimate lower probability events and underestimate higher probability events (often referred to as variance bias), may have somewhat ambiguous implications for temporal stability, but the fact that such a bias may vary across the driving population is an important matter to address in accident-data analysis.

What is clear is that the various cognitive biases will play a role in the potential temporal instability estimated statistical models. This is because individual drivers have different cognitive biases that affect how they gather and process information and this is likely to vary over time.

5.5. Macroeconomics and Risk-Taking Behavior

There is an abundance of literature that has documented that risk-taking behavior is strongly influenced by macroeconomic conditions and specifically investment-return experiences, (even those that occurred decades ago). For example, Malmendier and Nagel (2011) found that, while more recent investment-return experiences were better predictors of financial risk-taking, investment-return experiences from decades earlier were also influential in determining current risk-taking behavior. In other work, Guiso et al. (2013) found that, using data from Italy, risk aversion increased substantially after the 2008 global financial crisis.

The relationship between macroeconomics and risk-taking behavior has profound implications for the temporal stability accident data. For example, the empirical work of Abay and Mannering (2016) has shown that there is a significant positive correlation between financial risk-taking and risky driving.[14] Thus one cannot not only expect current macroeconomic conditions to influence risky driving, but also the macroeconomic experiences of drivers occurring decades before. With regard to current macroeconomic conditions affecting risk-taking, there is an abundance of research that shows that adverse macroeconomic conditions result in a decline in the number of motor-vehicle fatalities per distance driven (Ruhm, 2000; Peterman, 2013). While researchers have attempted to link such macroeconomic effects to possible changes in

---

[12] Winston et al. (2006) provide empirical evidence that shows how individuals may respond to new safety technologies and adjust their behavior based on perceived risks.

[13] The effect of changes in media coverage over time are supported by the previously discussed findings of Mannering and Winston (1995) where they found that consumers' willingness to pay for a driver-side airbag was significantly influenced to media exposure (average number of hours spent watching television per day) and social networks (number of friends owning cars with driver-side airbags).

[14] Abay and Mannering's empirical work shows that the level and statistical significance of this correlation also varies by gender and the intensity of involvement in financial and driving risk taking.

discretionary driving patterns, values of time, distances risky versus safe drivers drive, and so on,[15] the reduction in accident fatalities during economic downturns could be explained, at least partially, by fundamental changes values of risk.[16]

5.6. Temporal Instability and the Relationship between Attitudes and Behavior

Individual drivers are continuously gathering information from their own driving experiences, observing other drivers, as well as gathering information from various social interactions, social networks,[17] media stories, vehicle and governmental safety advertisements, and other sources. All of this information is used to from attitudes toward highway safety and to revise them as new information becomes available. One would expect that these continuously evolving attitudes toward highway safety would also affect driving behavior, including the selection of driving speeds, spacing when following vehicles, and other factors that would in turn affect the frequency and severity of accidents. In fact, there is a very large body of psychological literature that seeks to establish attitudes as predictors of behavior (Glasman and Albarracin, 2006). How strongly attitudes influence behavior has been found to depend on how attitudes were formed (with attitudes forming from experience having the greatest influence), how easily attitudes can be retrieved from memory (Fazio, 1989), and how stable the attitudes are over time (those attitudes being constantly influenced by new information have been found to be less influential).

With this said however, in general, the psychological literature has found the correlation between attitudes and behavior to be quite tenuous at times. As a result, considerable variability has been reported in terms of the ability of attitudes to actually predict behavior (Ajzen, 2001). The possible divergence between attitudes and behavior can be explained to some extent by the well-known theory of cognitive dissonance (Festinger, 1962), which has as its basis individuals knowing things that are not psychologically consistent (an attitude that suggests one thing and a behavior that suggest another). In the case of attitudes and behaviors, cognitive dissonance theory suggests people will adjust their attitudes and behaviors to reduce the dissonance between the two over time, but at any point in time each individual is likely to have some attitude/behavior dissonance, and the extent of this dissonance is likely to vary considerably from one individual to the next as they will be in different places in forming attitudes and resolving potential attitude/behavior dissonance.

There is also a substantial body of literature that looks at how beliefs and attitudes may tend polarize over time, thus also suggesting temporal instability (Benoit and Dubra 2017). In one of the early studies on this polarization tendency, Lord et al. (1979) presented the same mixed evidence on the effectives of the death penalty in deterring crime (crime observed to go down in most but not all locations that adopted the death penalty) to two types of individuals; those disposed to believe in the deterrence effect and those disposed to doubting it. With the same information both groups of individuals became more confident in their initial death-penalty/deterrence

[15]Maheshri and Winston (2016) provide some evidence that the reduction in fatality rates observed during economic downturns may be due in part to riskier drivers driving less during recessions relative to their safer-driver counterparts.

[16]The effects of possible macroeconomics on risk-taking is also supported by Behnood and Mannering (2016) where they found significant temporal instability in model parameters before, during, and after the 2007-2009 global great recession.

[17]Recent work in the transportation literature that has sought to account for social influences on travel behavior (Maness et al., 2015; Maness and Cirillo, 2016). This work provides support for possible relationships between attitudes and behavior in the traditional psychological sense.

attitudes. Those believing the death penalty deters crime cited the general upward trend in crime as the reason that all locations did not go down in absolute terms. Those doubting the deterrence of the death penalty cited the mixed results among locations as evidence that the results were inconclusive thus supporting their doubts. This polarization of populations has obvious implications with regard to safety. For example, individuals processing aggregate information on the dangers of texting and driving, may polarize into groups nearly correctly estimating the dangers and those under-estimating the dangers based on their initial beliefs.

The formation and evolution of safety attitudes, and their influence on driving behavior, clearly points toward temporal instability in accident data. Moreover, from an econometrics perspective, cognitive dissonance theory alone supports the simultaneous determination of attitudes and behavior, suggesting a complex endogenous relationship between the two.

5.7. Implications of Temporal Instability on current Accident-Data Analysis Methods

There are a number of analysis approaches that can potentially be applied to make inferences and predictions from accident data. These approaches can be broadly classified as; heterogeneity models, data-driven approaches, and traditional safety models. The choice of one approach over another has typically involved a trade-off between prediction accuracy and the causality/inference capability. Figure 5.1 illustrates the general trade-off between the two elements and the four analysis approaches identified.[18]

While the choice of analysis approach has been often dictated by issues of prediction and casual inferences in the past, it is interesting to discuss how potential temporal instability might influence the accuracy of these approaches with regard to predictions and causal inferences.[19] This discussion is provided in the following sections.

5.7.1 Unobserved Heterogeneity Models and Temporal Instability

In recent years, a wide variety of statistical methods have been applied to the analysis of accident data to capture the effects of unobserved heterogeneity (Mannering et al., 2016). Such models attempt to account for the fact that existing accident data sources cover only a small fraction of the wide variety of factors that affect the likelihood and resulting injury severities of accidents. With regard to accident likelihood models, unobserved heterogeneity has been addressed with random parameters count-data models, random parameters tobit model, random parameters generalized count models, latent-class models, Markov-switching count models, and bivariate/multivariate models with random parameters (see Table 4.2 in Mannering et al., 2016, for a recent listing and categorization of this literature). And, with regard to injury-severity models, unobserved heterogeneity has been addressed with random parameters (mixed) logit models, random parameters ordered probability models, latent-class models, latent-class models with random parameters within classes, Markov-switching models, Markov-switching models with random parameters, bivariate/multivariate models with random parameters (see Table 4.3 in Mannering et

---

[18]There are also safety models, at a more aggregate level, that apply time-series statistical approaches. While these models are less frequently used to guide specific safety policies, they would be particularly susceptible to temporal instability issues because of the extended time period considered. The complexity of the statistical correction that would be required to account for time-varying parameters in such models, in the presence of unobserved heterogeneity and other factors, would be a formidable obstacle.

[19]There is also a potential class of models that use control variables to account for possible selectivity/identification issues. This class of models is discussed in Mannering and Winston (2018). In the presence of temporal instability, the control-variable approach would be highly susceptible to both estimation error and forecasting error.

al., 2016, for a recent listing and categorization of this literature) and, more recently, random parameters models with heterogeneity in means and/or variances (Behnood and Mannering, 2017a; Seraneeprakarn et al., 2017; Behnood and Mannering, 2017b).

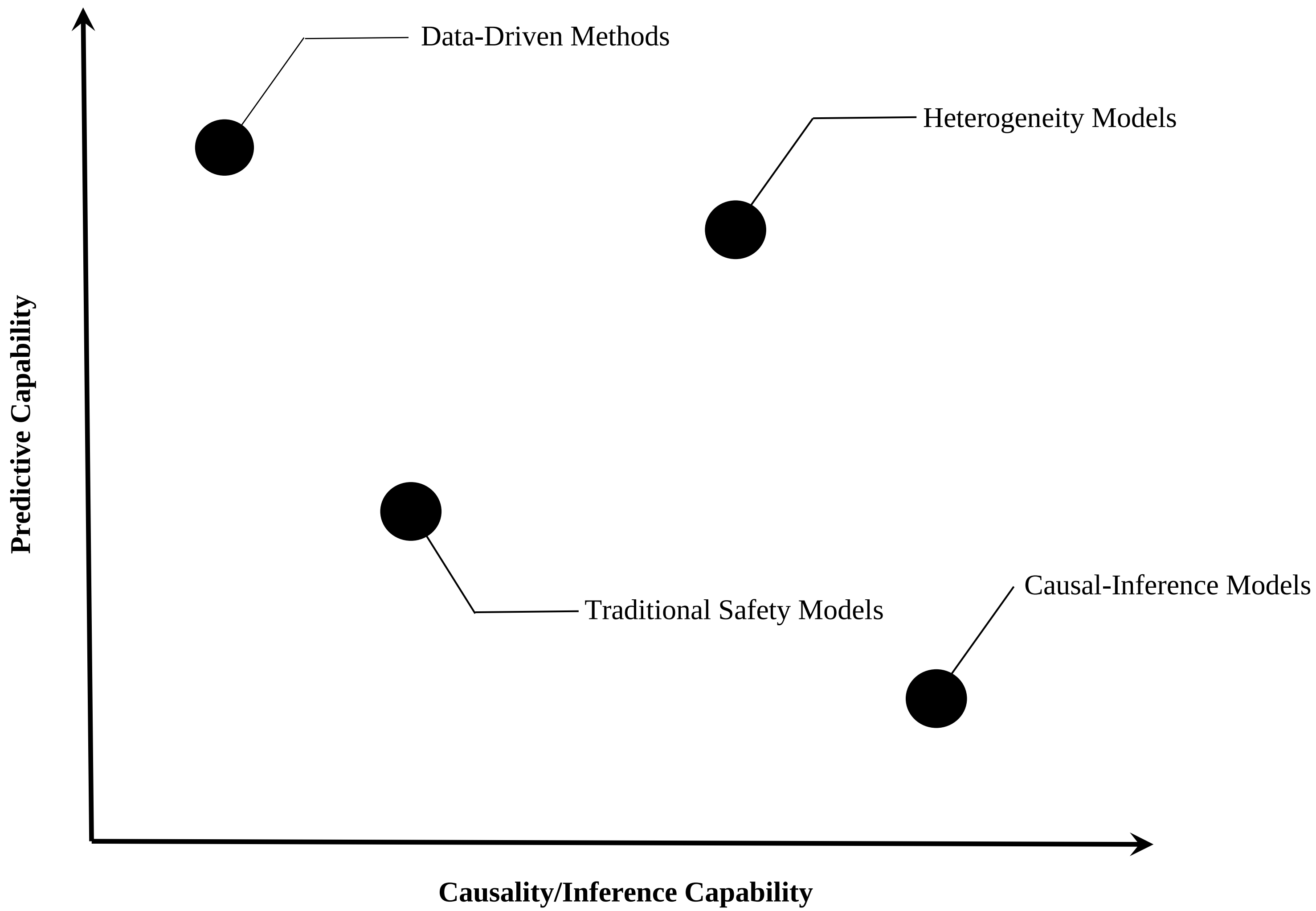

Figure 5.1. Current modeling trade-offs between predictive capability and causality/inference capability (assuming temporal stability).

By far, the most popular approach to deal with unobserved heterogeneity in the safety field and other transportation fields, such as travel behavior and choice modeling, has been the random parameters approach and its variants. Random parameters models allow for the possibility for each observation to have its own parameter value that determines the effect that individual explanatory variables have on the likelihood and/or severity of accidents. This is done, for example, by writing an expression for observation *i*'s estimable parameter on explanatory variable *k* as,

$$\beta_{ik} = b_k + \nu_{ik} \,, \tag{5.1}$$

where β_{ik} is the parameter on the *k*th explanatory variable for observation *i*, b_k is the mean parameter estimate across all observations for the *k*th explanatory variable, and ν_{ik} is a randomly distributed term that accounts for possible unobserved heterogeneity across observations. During the model estimation process, the analyst can statistically test a wide variety of distributions of ν_{ik} (for example, normal, log-normal, etc.) to determine the best model fit. For explanatory variable *k*, if the assumed distribution of ν_{ik} produces a statistically significant variance of parameter values across the population, then each observation will have its own parameter estimate. If the variance of ν_{ik} is statistically insignificant from zero, explanatory variable *k* has the same parameter across all observations, as in a traditional statistical model.[20]

Nearly all studies that have uncovered statistically significant unobserved heterogeneity have attributed it to potential unobserved variations in factors such as human behavior, human physiology, vehicle characters, roadway characters, and weather. Often overlooked is the possibility that such heterogeneity models may also be capturing temporal instability. In accident likelihood models, such as a count-data model that develops a statistical model of the number of accidents per year on specific highway segment, finding one or more random parameters may be due in some part to the different time-distribution of accidents on specific highway segments over the observation year. That is, if the effect of explanatory variables are globally changing over the observation year, a highway segment having more of its accidents early in the year may have different parameter values than a highway segment that has more of its accidents later in the year. Heterogeneity models will capture this effect by estimating different parameters across highway segments for one or more explanatory variables, but the analyst will not be able to distinguish the unobserved heterogeneity induced by temporal variations from that of other sources.

This problem is even more apparent in an injury-severity model, where each observation is an accident that occurs at a specific point in time. Finding statistically significant unobserved heterogeneity (such as significant parameter variances for one or more explanatory variables in a random parameters model) could be entirely due to temporal shifts or, more likely, a combination of temporal shifts and other traditional sources of unobserved heterogeneity.[21]

While heterogeneity models may be able to account for the potential temporal instability of parameters by capturing it as unobserved heterogeneity, forecasting in the presence of temporal instability would be problematic because such models would not be able to track the implied changes in unobserved heterogeneity over time. In some sense this is analogous to the spatial transferability issues associated with heterogeneity models. That is, such models can potentially

[20]Most random parameter models will have some explanatory variables (*k*'s) producing statistically insignificant variances (fixed parameters) and some others producing statistically significant variances (random parameters).

[21]If the model is tested for unobserved heterogeneity with a random parameters formulation, for example, and all parameters are found to be fixed (the variance of the assumed parameter distribution is statistically insignificant from zero), one could develop a strong argument for temporal stability.

define a unique explanatory-variable parameter for each observation, but determining the correct parameter for new observations (those not included in the estimation data set) is problematic.

5.7.2 Data-Driven Models and Temporal Instability

There are a number of analysis methods that have been applied to the analysis of accident data with the intent of uncovering correlations and developing accurate predictive models. These include various methods such as neural networks, decision trees, support vector machines and others and have been previously applied to the analysis of accident data (Abdelwahab and Abdel-Aty, 2001; Chong, 2005; Chang, 2005; Delen et al., 2006; Riviere et al., 2006; Xie et al., 2007; Li et al., 2008; Yu et al., 2013). Sophisticated forms of these data-driven methods have been shown to predict accident data with comparatively high accuracy (earning high predictability marks in Figure 5.1). However, uncovering causality and making substantive inferences has been an historical weakness of these approaches, often earning them a “black-box” designation because of the difficulty of unraveling how specific elements that might influence predictions with these approaches (giving it low marks in Figure 5.1). Still, these data-driven methods are likely to become increasingly popular with the emergence of high dimensional big-data in transportation safety (National Academies, 2013).

Similar to heterogeneity models, data-driven models can potentially track existing data very well, but they do not provide really provide much guidance for how the results may shift over time.

5.7.3 Traditional Safety Models and Temporal Instability

Traditional safety models attempt to model the likelihood of an accident (usually the number of accidents occurring on a highway segment or in a highway intersection over some specified time period) as well as the resulting occupant-injury severity typically gathered from police reported data in the discrete-injury categories (such as no injury, possible injury, evident injury, disabling injury or fatality). Traditional statistical approaches to this problem have included count-data models (to capture the frequency of accidents over some time period) and various discrete-outcome models to study the resulting injury severities of vehicle occupants (see Lord and Mannering, 2010, Savolainen et al., 2011, and Mannering and Bhat, 2014 for a review of studies that have addressed the likelihood and injury-severity outcomes of an accident). Count data models such as the negative binomial and injury severity models such as the ordered probit and multinomial logit have become mainstay “traditional” models. These models have also gained wide-spread use in safety practice in documents such as the Highway Safety Manual (American Association of State Highway and Transportation Officials, 2010).

As Figure 5.1 indicates, these traditional statistical approaches to accident-data analysis provide the ability to make some substantive inferences on the effect that specific variables will have on the likelihood and severity of accidents, and these inferences have been used to guide safety countermeasures in practice by local, state and federal agencies (American Association of State Highway and Transportation Officials, 2010). Unfortunately, traditional safety-modeling approaches are typically plagued by estimation-related limitations such as omitted variables bias (a statistical limitation resulting from the absence of key explanatory variables in the estimation database) and not explicitly accounting for unobserved heterogeneity (Mannering and Bhat, 2014; Mannering et al., 2016), which limits the accuracy of their inferences and predictions. However, it is important to keep in mind that these limitations are often ones of necessity because traditional models are developed for practical use and must make compromises in terms of data availability and model complexity to arrive at a model that can be used by highway agencies to mitigate the effects of vehicle accidents.

Traditional safety models are often used in an empirical Bayes approach to evaluate the before and after effect of safety interventions. This approach is typically used to determine the effect of a safety intervention (for example a new traffic signal) on the likelihood of accidents by applying a count-data model, for example, to study the frequency of accidents in the before- and after-intervention periods. The method is ostensibly designed to account for a perceived regression-to-the-mean effect, where high and low accident frequencies in specific time periods will tend to converge to a longer-term mean. This effect has considerable appeal to safety practitioners who often deal with limited data and use modeling approaches that do not account for unobserved heterogeneity. In the absence of these, safety practitioners view accident frequencies as largely random in the generation of low and high accident counts. Technically, the regression-to-the-mean concept is not valid because accidents are not random but instead are caused be specific circumstances that may not be fully known to the safety analyst. As the statistical model becomes better and better specified, the regression-to-the-mean concept becomes less and less useful as empirical construct. In essence, the regression-to-the-mean concept is simply a crude way of tracking unobserved heterogeneity and potential omitted variables bias (Mannering et al., 2016).[22,23]

In the presence of temporal stability, traditional safety models (which again usually suffer from unobserved heterogeneity issues and omitted variables bias, since many important variables are excluded because they are not available for estimation and/or forecasting) will be problematic in terms of prediction and uncovering causality. This is because estimated parameters in such models will be estimated with bias, and this bias will change over time not necessarily due to temporal shifts, but to shifts in the values of omitted variables and unobserved heterogeneity. In the presence of temporal instability, traditional safety models, and their potential application in Bayesian before and analyses, are likely to be highly inaccurate.

5.8. Discussion

The previous sections of this chapter show that there are a multitude reasons from fields such as psychology, neuroscience, economics, and cognitive science that suggest that individual driving behavior develops in complex ways over time.[24] While current safety research does not explicitly address this, it is often implicitly addressed by using measurable explanatory variables such as driver age, and potentially other socioeconomic variables such as household income and type of car owned, to serve as a proxy for complex behavioral formations, cognitive biases, risk estimations, and attitude formations. However, these underlying fundamentals likely change in complex ways over time, making the parameter estimates of proxy variables temporally unstable.

---

[22]The Bayesian approach combines prior accident information with current accident information. In empirical Bayes, prior accident information is gathered from a group of observations (highway segment, intersections, etc.) similar to those under evaluation and a statistical model of accident frequency, for example, is estimated and used to provide an improved estimate of long-term accident frequency means and variances (this approach also mitigates regression-to-the-mean issues that could affect findings).

[23]In addition to empirical Bayes safety practitioners have also applied full Bayes approaches (which considers the distribution of accident frequencies instead of just the mean and variance of empirical Bayes). See Persaud et al. (2009) for a discussion and comparison of the two approaches.

[24]As previously mentioned, the presentation of reasons is representative only, and by no means comprehensive.

To address the issue of temporal instability, one would seek to develop a modeling approach that could potentially account for how model-estimated parameters might change over time. Perhaps the simplest approach is to specify an estimated parameter to be a function of factors that are expected to affect how the estimated parameters may change over time,

$$\boldsymbol{\beta}_i = \boldsymbol{\alpha} + \boldsymbol{\lambda}\mathbf{T}_i \tag{5.2}$$

where $\boldsymbol{\beta}_i$ is a vector of estimable parameters for observation i, $\mathbf{T}_i$ is a vector of explanatory variables that determine the change in $\boldsymbol{\beta}_i$ over time, and $\boldsymbol{\alpha}$ and $\boldsymbol{\lambda}$ are vectors of estimable parameters. As discussed in the above sections, the vector $\mathbf{T}_i$ could include macroeconomic factors interacted with observation-specific elements, observation-specific attitudes and other factors that could relate to how model parameters evolve over time.

Another approach would be to view this temporal tracking in the context of a model that accounts for unobserved heterogeneity, such as a random parameters model. In this context, the temporal element could be introduced as unobserved heterogeneity in the means and variances of random parameters by allowing $\boldsymbol{\beta}_i$ be a vector of estimable parameters that varies across observations as (Seraneeprakarn et al., 2017):

$$\boldsymbol{\beta}_i = \boldsymbol{\alpha} + \boldsymbol{\Theta}_i\mathbf{T}_i + \boldsymbol{\sigma}_i EXP\left(\boldsymbol{\omega}_i\mathbf{W}_i\right)\, \boldsymbol{\xi}_i \tag{5.3}$$

where β is the mean parameter estimate across all observations i, $\mathbf{T}_i$ is a vector of explanatory variables that captures time-dependent heterogeneity in the mean, $\boldsymbol{\Theta}_i$ is a corresponding vector of estimable parameters, $\mathbf{W}_i$ is a vector of explanatory variables that captures time-dependent heterogeneity in the standard deviation $\boldsymbol{\sigma}_i$ with corresponding parameter vector $\boldsymbol{\omega}_i$, and $\boldsymbol{\xi}_i$ is a disturbance term.

However, accounting for the potential temporal variation in this manner is not as trivial as it seems. Even if the correct elements of the vector $\mathbf{T}_i$ could be obtained, the model estimation is still likely to be wrought with serious econometric issues relating to factors such measurement error and endogeneity that could potentially lead to inconsistent model estimations. To address potential estimation issues which may arise from the inclusion of variables likely to capture the temporal elements of accident data such as this, Bhat and Dubey (2014) developed an integrated latent variables approach that allows latent constructs that can potentially track temporal instability. These latent constructs (variables) can be developed from attitudinal and perception variables (that address potential temporal shifts) in a latent measurement equation model. This approach can potentially address issues of measurement error and complex error-term correlations to ensure consistent estimates of model parameters.

Markov-switching models also have the potential to account for temporal variations in accident data. Work by Malyshkina et al. (2009), Malyshkina and Mannering (2009), Malyshkina and Mannering (2010), and Xiong et al. (2014) have used such models to track heterogeneity assuming heterogeneity follows a stationary multiple-state Markov chain process. This can be extended to systematically account for shifts between multiple states by defining the state-transition probabilities to be functions of variables suspected to influence temporal instability (such as macroeconomic conditions). Such an approach could potentially provide an excellent method of tracking temporal shifts, but the complexity of the model estimation process could be cumbersome (Xiong et al., 2014).

However, it should be clear from the above examples that accounting for potential temporal shifts is potentially an exceedingly difficult statistical problem and one that will increase the complexity of modeling processes considerably. Nevertheless, it is important for the safety field to think more carefully about the temporal elements in accident data and work to improve model specifications and interpretations in this regard.

5.9. Summary and Conclusions

This chapter provides numerous examples from a variety of fields that indicate that there is strong behavioral evidence to suggest that temporal instability is likely an important issue in contemporary analyses of accident data. However, it should be clear from the summary of conventional analysis measures (section 5.7) and the subsequent discussion in section 5.8 that accounting for potential temporal shifts in a systematic and econometrically defensible way is an exceedingly difficult problem. Even with the potential temporal limitations of our current methodological approaches (in terms of analyzing accident data) the various points raised in this chapter can be useful on a number of levels:

1. The temporal elements associated with individual behavior and the aggregate trends that result from these (such as the long-term decline fatalities per mile and the phenomenon of aggregate economics affecting accident rates) are important factors to consider when developing modeling approaches and interpreting model findings. Ignoring these fundamental temporal elements can lead to erroneous conclusions and ineffective or even dangerous safety policies.[25]
2. Different data analysis methods (unobserved heterogeneity, data driven, traditional, and causal inference models) can be affected by potential temporal instability in different ways. This has to be given careful consideration in the interpretation of results.
3. While explicitly accounting for temporal elements in current modeling approaches presents a formidable technical challenge, the field must move to address this challenge even if only in an incremental way.

[25] For example, suppose that speed limits are raised during an economic downturn, and a study finds no significant change in accident rates before and after the speed-limit increase. However, as discussed in the paper, during an economic downturn accident rates would be expected to decline overall. The fact that the increased speed limits did not result in a significant decline suggests they may be more dangerous, but this would be overlooked without consideration of temporal elements.

Chapter 6: When Neutral Responses on a Likert Scale Do Not Mean Opinion Neutrality: Accounting for Unsure Responses in a Hybrid Choice Modeling Framework

6.1. Introduction

Integrated choice and latent variable (ICLV) models enhance discrete choice models by explicitly considering socio-psychological factors. These enhancements can lead to improvements in analysts' ability to explain behavior and may lead to improvements in forecasting (Vij and Walker, 2016). Although a variety of indicators can be used in ICLV models to measure the attitudes and perceptions, Likert scale questions are widely used in the transportation field to psychometrically measure attitudes and perceptions. Likert scales are bi-polar symmetric scales. Respondent are given a set of points to choose on a scale ranging from an extreme negative end to an extreme positive end. Commonly, Likert scales are given with an odd number of points. In this form, the middle option acts as a transition point between the two polar regions. Despite the use of a variety of labels, the middle point is often modeled as indicating opinion neutrality and the corresponding respondents represents a "neutral group."

But psychometric research has found that the group of respondents who choose the middle option in Likert scales is not homogeneous. These respondents are not all truly opinion neutral and thus do not act as a transition group between the two extremes (Sturgis et al., 2014, Kalton et al., 1980, Cacioppo et al., 1997 and Baka et al., 2012). The respondents who choose the neutral/middle option can often fall into two groups: (1) those individuals who possess true opinion neutrality on the issue and select the neutral option (2) those individuals without adequate knowledge or familiarity who choose the neutral option as a way of saying that they do not know or have no opinion (Sturgis et al., 2014). The latter group is not considered in existing implementations of Likert scale indicators in ICLV models. This is due to how existing models treat all neutral responses as opinion neutrality since the response is on a continuum.

In terms of policy implications, respondents with lack of knowledge may respond differently from the people with no-opinion once they gather information. Thus, segregating the two groups beforehand helps in formulating appropriate policies by avoiding biases in the measurement of policy outcomes. Moreover, identifying the no-opinion groups provides guidance on where to concentrate education efforts during information awareness campaigns.

Determining who has no-opinion rather than opinion neutrality is primarily done through survey design. A demographic link has not been well established. Kalton et al. (1980) studied the effect of sociodemographic characteristics on the neutral responses but failed to find a clear relation between them. More recently, Krosnick et al. (2002) found that respondents with low levels of education were more likely to choose the middle option. Some past studies have attributed this to social desirability bias as respondents may feel it is more appropriate to express a neutral opinion rather than explicitly admit ignorance or to not respond to the question. Nowlis et al. (2002) suggest:

> *"If researchers use odd-point scales, they might want to develop methods to distinguish between the kinds of responses that underlie the selection of a neutral position because of their different implications."* (p. 332)

This study proposes the development of an ICLV modeling framework to identify such different groups of neutral position respondents by considering the individual's level of neutrality and familiarity with a topic area. The framework uses a confirmatory latent variable model approach where the group are formed through combining expected middle option choices and varying topic familiarity levels. A case study is conducted about consumers' perception and intended adoption of autonomous vehicles (AVs) to identify an individual's preferred way of using AVs once they become available. People are categorized into different multiple groups assuming that they will have varying levels of neutrality and familiarity towards AVs. The four groups are opinion-neutral, no-opinion, familiar opinionated, and unfamiliar opinionated. Three latent variables are used to measure respondent's propensity to be in these groups (while controlling for one group). The model was able to identify the relationships between a respondent's demographic characteristics and their tendency to be in such neutrality-familiarity groups.

The remainder of the paper is organized as follows. First, the dataset is described with emphasis on the attitudinal Likert scale questions used and the distribution of neutral responses. Then the ICLV model framework's methodology is described specifically for the case study presented. Next, model results are presented and analyzed. The final section discusses the ways to enhance the model structure, modification due to data differences, and additional reasons for middle responses.

6.2. Data and Survey Design

A web-based survey was administered to collect data from a sample of American Automobile Association-South (AAA-South) members in the United States, and a total 2,338 respondents were obtained (Menon et al., 2016 and Menon et al., 2018). After considering various factors such as the consent to participate in the survey, incomplete responses, and premature completion (if the respondent didn't answer the choice question under consideration), a reduced sample size with 1235 respondents were available for the estimation process. Table 6.1 summarizes the design of the survey. The survey consisted of four sections: respondent and household characteristics, vehicle characteristics, autonomous vehicle (AV) perceptions, and anticipated impacts of AVs.

Table 6.1 Summary of Survey Methods

Characteristic	Description
Time Frame	June 2015
Target Population	US Household across 11 states in the South and Midwest
Sampling Frame	Households with internet access in 11 states across the US South and Midwest
Sample Design	Non-probability sample via convenience sample of AAA South members
Use of Interviewer	Self-administered
Mode of Administration	Self-administered via the computer and internet for remaining respondents
Computer Assistance	Web-based survey
Reporting Unit	One person age 18 or older per household reports for the entire household
Time Dimension	Cross-sectional survey
Frequency	One one-month phase of collecting responses
Levels of Observation	Household, person, vehicle

For the latent variable model with latent variables (a) the benefits and (b) the concerns and (c) AV technology familiarity indicator questions used are listed below (with corresponding Likert scale given in parentheses):

a) How likely do you think the following benefits will occur when using Autonomous Vehicles (AVs)? (extremely unlikely, unlikely, don't know/ can't say, likely, and extremely likely)
    1. Fewer traffic crashes
    2. Less traffic congestion
    3. Less stressful driving experience
    4. More productive use of travel time
    5. Lower car insurance rates
    6. Increased fuel efficiency
    7. Lower vehicle emissions
b) How concerned are you about the following issues when using Autonomous Vehicles (AVs)? (not at all concerned, not very concerned, don't know/ can't say, somewhat concerned, extremely concerned)
    1. Safety of vehicle occupants and other road users
    2. System equipment failure
    3. Performance in unexpected and extreme conditions
    4. Giving up control of steering wheel
    5. Loss in human driving skill over time
    6. Privacy risks from data tracking
    7. Difficulty in determining crash liability
    8. Motion sickness
c) How familiar were you about Autonomous Vehicles (AVs) before you participated in this survey? (not at all familiar, slightly familiar, moderately familiar, extremely familiar)

As can be seen in the concerns chart (Figure 6.1), the descriptive statistics show that the proportion of neutral/middle responses is similar between all the concerns categories. The benefits chart (Figure 6.1) shows a similar relationship although not as strong. The top three categories show similar neutral/middle response proportions, whereas the remaining categories show a smaller proportion of neutral responses (but similar proportion between these categories). The stability of the neutrality proportions motivates the use of this dataset to explore differentiating opinion neutrality and the lack of knowledge/opinions among respondents.

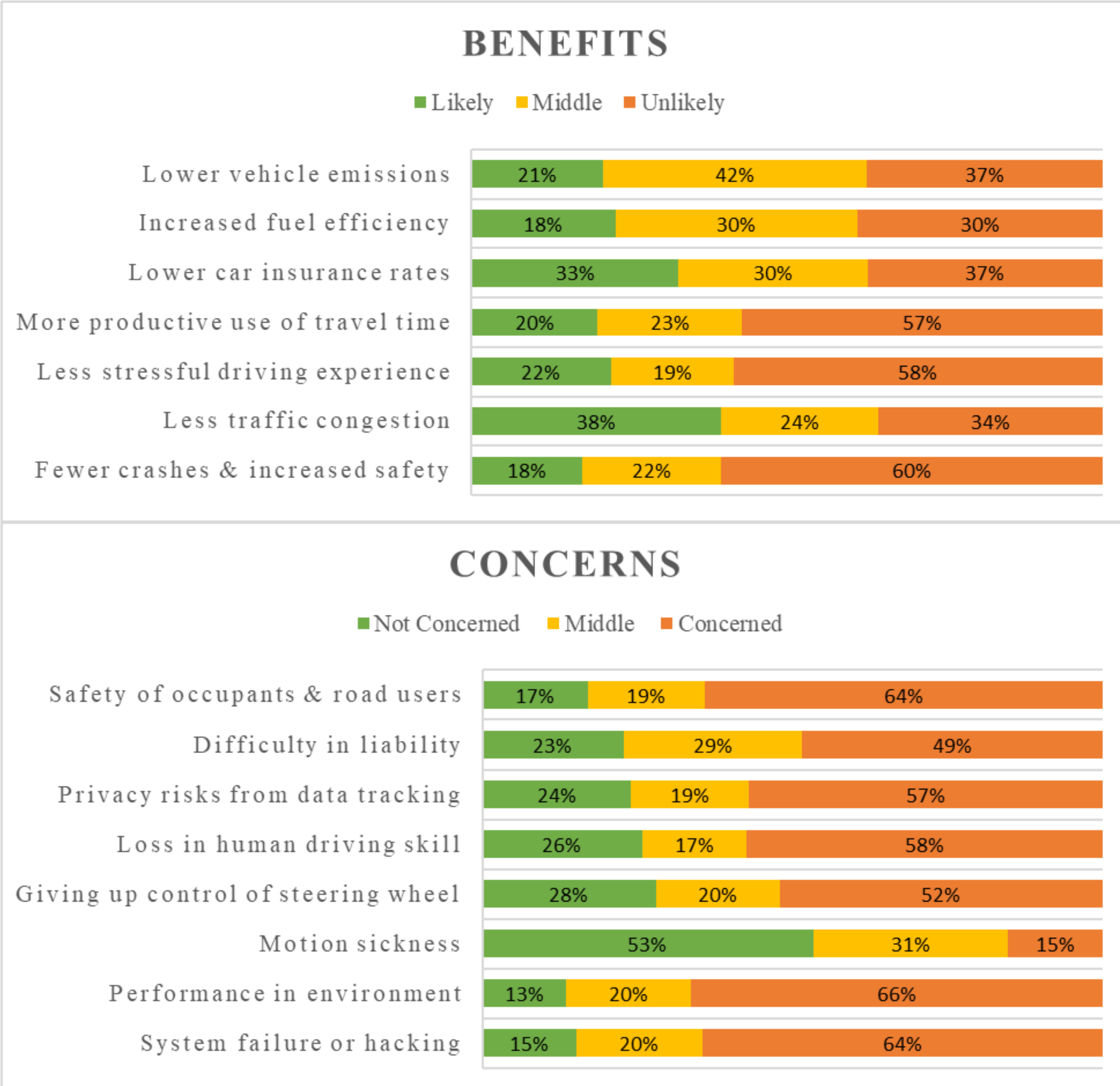

Figure 6.1. Response Distribution for Benefit Likelihood and Concern Level Questions

For the choice model an extensive list of foreseeable ways of using a AVs were provided to the respondents. The question used for AV usage intent choice model is: “What would be your most preferred way to use AVs that can fully drive by themselves without your involvement (when they become available).” The options given in the survey were:

- Own (purchase or lease) AVs and use them only for personal use or use by family members
- Own (purchase or lease) an AV and earn extra income on the side by making it available to other drivers when not needed
- Own (purchase or lease) an AV and earn extra income on the side by providing rides for fellow passengers when you use it
- Rent an AV as the need arises

- Use AVs in the form of transportation (taxi, or public transit) provided by a service provider
- Neither interested in investing in an AV nor using AVs as a transportation service

For this study, these categories were simplified into three alternatives as follows:

- Own (purchase or lease) AVs
- Rent an AV as the need arises and use AVs as part of a service
- Neither interested in investing in an AV nor using AVs as a transportation service

The descriptive statistics for the model relevant characteristics are as follows:

- Gender:
  - Woman: 38.47%
- Age range:
  - Under 35 years: 3.45%
  - Between 35 and 65: 60.64%
  - 65 years or older: 35.40%
- Education
  - College degree: 77.43%
- Household size
  - Mean: 2.32
- Income
  - Less than \$25,000: 3.29%
  - \$25,000 - \$100,000: 54.54%
  - More than \$100,000: 42.17%
- Travel history
  - Drive alone: 66.67%
  - Automobile for commute: 72.9%
  - Involved in Auto Crash: 76.2%
  - Most recent purchase/lease was new vehicle: 62.9%

6.3. Methodology

In this study, respondents are assumed to have varying levels of neutrality and familiarity. Four distinct categories are formed: opinion-neutral, no opinion, familiar opinionated, and unfamiliar opinionated. The *opinion-neutral* respondents are assumed to pick middle Likert scale responses often because they have formed an opinion that is between two poles (e.g. in this study, they assess AV technology as neither a benefit nor concern). The *no opinion* respondents are assumed to pick middle Likert scale response often because they are unfamiliar with AV technology and choose the middle option to express being unsure. The *familiar opinionated* and *unfamiliar opinionated* respondents tend to not choose the middle response thus expressing some polarization of opinions on benefits and concerns, yet they are expressing this opinion from a position of knowledge versus lack of knowledge respectively.

The framework uses a latent variable model to evaluate the propensity of respondents to be in three of these four groups (*opinion-neutral* (*t*), *no opinion* (*o*), and *unfamiliar opinionated* (*p*)). The measurement equations for these group depend on sixteen indicators each: 15 indicator functions for choosing the third/middle option in the seven benefit and eight concern Likert scale questions and 1 ordered response to the self-reported AV familiarity question. The fourth group, *familiar opinionated*, is not represented but assumed to have the remaining respondents. The neutrality-familiarity type propensities use the following measurement equations:

$i_{t,n} = D_t t_n^* + \eta_{t,n}$ (opinion-neutral type propensity indicators)
$i_{o,n} = D_o o_n^* + \eta_{o,n}$ (no-opinion type propensity indicators)
$i_{p,n} = D_p p_n^* + \eta_{p,n}$ (unfamiliar opinionated type propensity indicators)

Where:
$i_{t,n}, i_{o,n}, i_{p,n}$ = (1x16) vectors of indicators for the *opinion-neutral*, *no opinion*, and *unfamiliar opinionated* latent variables respectively
t_n^*, o_n^*, p_n^* = *opinion-neutral*, *no opinion*, and *unfamiliar opinionated* latent variables
D_t, D_o, D_p = (1x16) vectors of parameters relating latent variables t_n^*, o_n^*, p_n^* to their respective indicators
$\eta_{t,n}, \eta_{o,n}, \eta_{p,n}$ = (1x16) vectors of independent error terms, ~Normal(0,1)

For all three latent variables, the 15 middle option indicators are modeled as independent binary probit models. The familiarity indicators are modeled as ordered probit models. Although the three latent variables use the same set of indicators, the expected directionality of their *D* parameters should vary. This is enforced through estimating the parameter in an exponential form and setting its directionality. The latent variable types have the following measurement equation parameter schemes:

- *Opinion-neutral*: The middle option indicators measurement equations have positive D_t with the familiarity indicator also having a corresponding positive D_t. This represents how *opinion-neutral* respondents are assumed to be familiar with the technology but have propensity to have non-polar opinions.
- *No-opinion*: The middle option indicators have positive D_t but the familiarity indicator has a negative D_t relationship.
- *Unfamiliar Opinionated*: The middle option indicators have negative D_t with the familiarity indicator also having a negative D_t relationship. This represents how *unfamiliar opinionated* respondents are uninformed about AV technology yet have formed a non-neutral opinion about AV benefits and concerns.

The three latent variables have three corresponding measurement equations as follows:
$t_n^* = A_t w_n + v_{t,n}$ (opinion-neutral type propensity)
$o_n^* = A_o w_n + v_{o,n}$ (no-opinion type propensity)
$p_n^* = A_p w_n + v_{p,n}$ (unfamiliar opinionated type propensity)

Where:
w_n = (7xN) matrix of respondent characteristics
A_t, A_o, A_p = (1x7) vectors of parameters relating respondent characteristics to respective latent variables
$v_{t,n}, v_{o,n}, v_{p,n}$ = independent normally distributed error terms with mean zero and unknown variances $\sigma_t, \sigma_o, \sigma_p$ respectively

For the choice task involved, intended AV usage, a fourth latent variable is estimated corresponding to respondents' *positive assessment* of AV technology. This *positive assessment* latent variable is measured using 15 ordered responses from the survey's seven benefit and eight concern Likert scale questions. The assessment latent variable has the following measurement and structural equations:

$i_{a,n} = D_a a_n^* + \eta_{a,n}$ (positive assessment indicators)

$a_n^* = A_a z_n + v_{a,n}$ (positive assessment)

Where:

$i_{a,n}$	= (1x15) vector of indicators for the *positive assessment* latent variable
a_n^*	= *positive assessment* latent variable
D_a	= (1x15) vector of parameters that relate a_n^* to $i_{a,n}$
$\eta_{a,n}$	= (1x15) vector of independent error terms, ~Normal(0,1)
z_n	= (5xN) matrix of respondent characteristics
A_a	= vector of parameters relating respondent characteristics to the *positive assessment* latent variable
$v_{a,n}$	= independent error terms, ~Normal(0,σ_a)

For the positive assessment measurement equations, the directionality of D_a is expected to be positive for the benefit questions and negative for the concern questions. This directionality was not forced through an exponential form but was observed through the data. The 15 indicators are modeled through ordered probit formulations.

The intended AV usage choice is represented through a utility maximizing choice with three alternatives (own AV, rent AV or use AV as service, and do not use AV). The three alternatives have the following utility functions:

$U_{n,own} = B_{own} x_n + b_{own,a}\, a_n^* + b_{own,t}\, t_n^* + b_{own,o}\, o_n^* + b_{own,p}\, p_n^* + \varepsilon_{own,n}$

$U_{n,rent} = B_{rent} x_n + b_{rent,a} a_n^* + b_{rent,t}\, t_n^* + b_{rent,o}\, o_n^* + b_{rent,p}\, p_n^* + \varepsilon_{rent,n}$

$U_{n,no\ AV} = 0$

The *do not use AV* option is assumed to be the base alternative with normalized utility. The choice is assumed to depend on respondents' characteristics and their assessment of AV technology. Additionally, the neutrality-familiarity type propensity latent variables are added in a linear form. The propensities can be thought as introducing heterogeneity in the ASCs in accordance with the individual's likely neutrality-familiarity type.

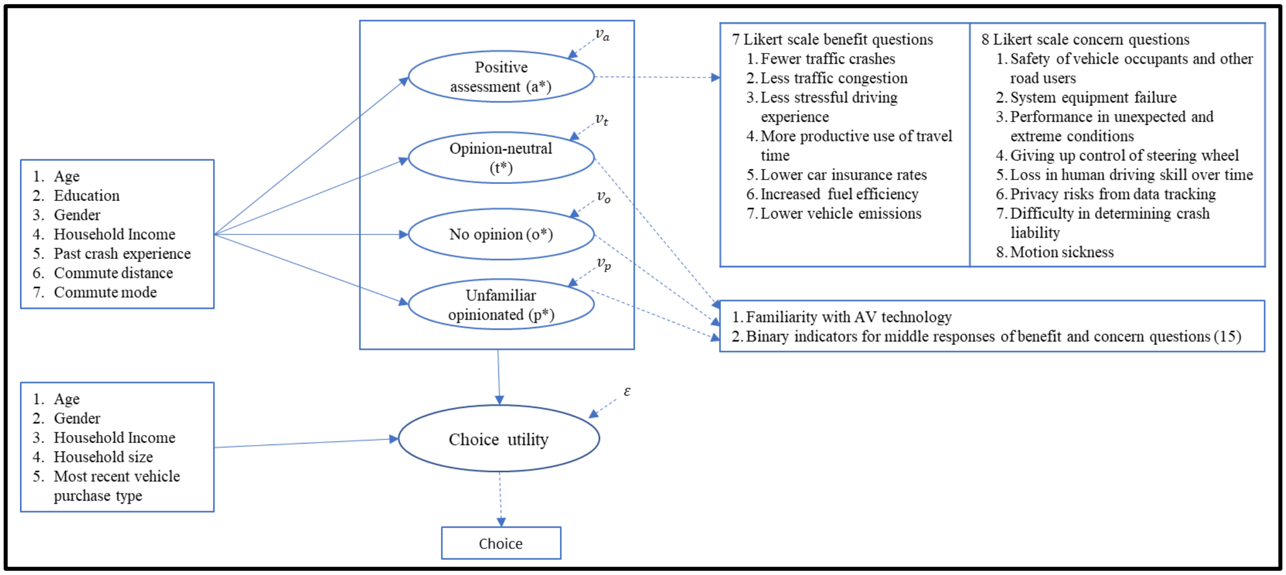

Figure 6.2. ICLV Model Path Diagram

The complete model structure is shown via a path diagram in **Error! Reference source not found.**. Overall, the framework is an integrated choice and latent variable model with a mixed logit base. The joint probability of observing the intended usage choice y_n and the indicators is (conditioning on respondent characteristics and model parameters omitted for clarity):

$$P(y_n, i_{a,n}, i_{t,n}, i_{o,n}, i_{p,n})$$
$$= \int_{a_n^*, t_n^*, o_n^*, p_n^*} \left\{ \begin{array}{c} \dfrac{\exp\left(U_{n,i} | a_n^*, t_n^*, o_n^*, p_n^*\right)}{\sum_j \exp\left(U_{n,j} | a_n^*, t_n^*, o_n^*, p_n^*\right)} \cdot \left[f_a\left(i_{a,n} | a_n^*, D_a\right) g_a(a_n^*) da_n^*\right] \cdot \left[f_t\left(i_{t,n} | t_n^*\right) g_t(t_n^*) dt_n^*\right] \cdot \\ \left[f_o\left(i_{o,n} | o_n^*\right) g_o(o_n^*) do_n^*\right] \cdot \left[f_p\left(i_{p,n} | p_n^*\right) g_p(p_n^*) dp_n^*\right] \end{array} \right\}$$

In this study, the model is estimated simultaneously using Bayesian inference with diffuse priors (Train, 2009). The burn-in period uses 50,000 draws and the following 60,000 draws are used for estimation. The models are estimated using the RSGHB package in R (Dumont et al., 2009).

6.4. Results and Analysis

The results of the ICLV model estimation are presented over three tables. Table 6.2 presents the structural and measurement model results for the three neutrality-familiarity type propensity latent variables. Table 6.4 presents the structural and measurement model results for the positive assessment latent variable. The intended AV usage choice model is presented in Table 6.5. Only the full model is presented and no comparison models are presented. This is due to the study's focus on differentiating the neutrality-familiarity types. The case study is a proof of concept for the ICLV framework. Additionally, loglikelihood values are presented for the complete model and choice model only. These are provided for completeness but have little practical value since ICLV models are optimized on the choice of not only the intended AV usage but also the indicator questions as well.

The values of the parameters measurement equations for the neutrality-familiarity type propensity latent variables have limited practical meaning. Their directionality was enforced in the likelihood function specification and given signs according to a prior behavioral assumption as described in the previous section. The models were normalized but setting the familiarity indicator to magnitude of 1 or -1 depending on the expected familiarity direction for the given neutrality-familiarity type. The no-opinion and unfamiliar opinionated types have more likely to have low self-reported familiarity with AV technology.

The structural model for the neutrality-familiarity types provides information on the demographic characteristics of each type. Both opinion neutral and no-opinion types tended to be younger and have lower incomes. Although both groups tended to have less formal education than opinionated respondents, the no-opinion respondents have a greater propensity to have less education than the neutral respondents. This results follows from previous research that found that no-opinion respondents tended to be less formally educated. Older respondents, respondents with higher education, and higher income respondent tend to be in the uninformed opinionated group.

Table 6.2. Latent Variable Models for Neutrality-Familiarity Type Propensities

	Opinion Neutral Propensity	No-Opinion Propensity	Unfamiliar Opinionated Propensity
Parameter	Mean	Mean	Mean
Structural Model			
ln(Respondent's Age)	-0.07*	-0.08*	0.08*
College Degree and Older than 24 Years	-0.10*	-0.21*	0.12*
Woman Respondent	0.00	0.10*	-0.04
Household Income / $25,000	-0.02*	-0.03*	0.01*
Respondent Involved in Auto Crash	-0.03	-0.12*	-0.01
Respondent's Commute Mode is Auto	-0.05*	-0.06	0.08*
Respondent's Commute Distance	-0.00	-0.00	0.00
Error Term Standard Deviation $\sigma_t, \sigma_o, \sigma_p$	0.41*	0.54*	0.42*
Measurement Model			
Fewer crashes & increased safety	1.57	1.23	-1.54
Less traffic congestion	1.30	0.93	-1.22
Less stressful driving experience	1.82	1.40	-1.83
More productive use of travel time	1.65	1.19	-1.65
Lower car insurance rates	1.20	0.96	-1.04
Increased fuel efficiency	0.99	0.83	-1.09
Lower vehicle emissions	0.75	0.53	-0.61
Safety of occupants & road users	2.05	1.43	-1.95
System failure or hacking	2.01	1.27	-1.75
Performance in environment	1.96	1.38	-2.01
Motion sickness	1.06	0.76	-0.97
Giving up control of steering wheel	1.86	1.42	-1.76
Loss in human driving skill	2.46	1.72	-2.27
Privacy risks from data tracking	1.97	1.39	-1.92
Difficulty in liability	1.41	1.06	-1.28
Familiarity with AV Technology (fixed)	1.00	-1.00	-1.00

Note:
* denotes statistically significant variables in the structural model.

All variables in the measurement model are statistically significant and are not denoted with an * in order to reduce clutter.

Opinion neutral respondents tended away from auto commutes while unfamiliar opinionated respondents tended to be auto commuters. This result fits a hypothesized effect of opinion formation on AVs tended towards a preferences of auto usage. Considering the familiar opinionated types as having an effect size of zero, both opinionated, non-neutral types have a greater likelihood of having auto commuters than the non-opinionated, neutral types. Table 6.3 summarizes the demographic characteristics of the neutrality-familiarity types. Lastly, the variance of the LVs in the structural equations are all similar. This is an encouraging sign as the scale of these LVs will also be similar which makes interpretation in the choice model more straightforward.

Table 6.3. Comparative characteristics for Neutrality-Familiarity Types

	Familiar	**Unfamiliar**
Neutral	*Opinion Neutral Respondents*	*No-Opinion Respondents*
	Younger Less college Lower incomes Non-auto commuters	Younger Less college Women Lower incomes Less likely in auto crash
Non-Neutral	*Familiar Opinionated Respondents*	*Uninformed Opinionated Respondents*
	Used as Base Group	Older More college Higher incomes Auto commuters

The structural model of the positive assessment latent variable suggests that older respondents and respondents with a crash experience in the past tend have more favorable assessment of AVs compared to their respective counter groups. On the other hand, female respondents and respondents with lower commute distance tend assess AVs less favorably. The measurement model results for the positive assessment latent variable are data-driven and directionality is left unrestricted. It was hypothesized that respondents with more positive assessments would tend to find AV benefits to be more likely and to be not concerned about possible AV concerns. This is confirmed through the directionality of the measurement equation parameters. Most parameters are significant with expected signs. The insignificant parameters are for three concern indicators with non-positive means estimates.

Table 6.4. Positive Assessment Latent Variable Model

Parameter	Mean		Confidence Interval 2.5%	97.5%
Structural Model				
ln(Respondent's Age)	0.22	*	0.19	0.26
College Degree and Older than 24 Years	-0.02		-0.11	0.03
Woman Respondent	-0.16	*	-0.21	-0.09
Respondent Involved in Auto Crash	0.25	*	0.20	0.35
Respondent's Commute Distance	-0.04	*	-0.07	-0.02
Error Term Standard Deviation σ_a	2.16		1.98	2.34
Measurement Model				
Fewer crashes & increased safety (fixed)	1.00		fixed	fixed
Less traffic congestion	0.47	*	0.43	0.51
Less stressful driving experience	1.13	*	1.07	1.18
More productive use of travel time	0.88	*	0.85	0.90
Lower car insurance rates	0.45	*	0.41	0.49
Increased fuel efficiency	0.70	*	0.66	0.74
Lower vehicle emissions	0.55	*	0.51	0.57
Safety of occupants & road users	-0.02		-0.04	0.01
System failure or hacking	-0.03		-0.05	0.00
Performance in environment	0.00		-0.03	0.02
Motion sickness	-0.19	*	-0.23	-0.17
Giving up control of steering wheel	-0.13	*	-0.16	-0.11
Loss in human driving skill	-0.11	*	-0.13	-0.08
Privacy risks from data tracking	-0.09	*	-0.12	-0.06
Difficulty in liability	-0.11	*	-0.14	-0.08

* denotes statistically significant variables in the structural model.

Table 6.5 presents the estimation results of the choice model for the respondent's most preferred way of using AVs when they become readily available, i.e. owning, renting/use it as a service and not all using AVs. Here, the *"not at all using an AV"* is considered as the base category. Women and older respondents are less inclined to use AVs either in owning form or renting form. On the other hand, respondents with higher income are more likely to use AVs in either form. In addition, people whose recent vehicle purchased was new were found to be more likely to own and less likely to rent AVs as compared to not using AVs. As expected, people who are in favor of AVs tend to use AVs either in own form or in service form with a stronger preference for ownership. Interestingly, people with a propensity towards opinion neutrality are more likely to use them in the form of service rather than owning them. Whereas people who do not possess any opinion are less likely to use them in the form of service. Finally, uninformed opinionated type people intend to own AVs rather than renting or using them as a service.

6.5. Discussion

This study proposed a first-step in developing an ICLV framework to account for differences between opinion neutrality and unsure/unfamiliar (no-opinion) survey respondents. The framework identifies these individual through their prevalence of choosing middle options over a series of Likert scale questions related to a topic. Additionally, the framework requires a way of measuring familiarity with such a topic to differentiate these groups. Differentiating these group has implications in measuring the preferences of individuals who choose the middle option in Likert scale survey questions. This group is heterogeneous in their motivations to choose middle responses which may impact forecasting efforts.

Table 6.5. Choice Model Results

Parameter	Mean		Confidence Interval 2.5%	97.5%
Alternative Specific Constant				
– Own Alternative	-0.12	*	-0.21	-0.03
– Rent/Service Alternative	0.20	*	0.14	0.26
Woman Respondent				
– Own Alternative	-0.51	*	-0.60	-0.36
– Rent/Service Alternative	-0.47	*	-0.55	-0.39
Household Size				
– Own Alternative	-0.01		-0.08	0.04
– Rent/Service Alternative	-0.07	*	-0.17	-0.01
ln(Respondent's Age)				
– Own Alternative	-0.09	*	-0.17	-0.03
– Rent/Service Alternative	-0.20	*	-0.25	-0.15
Household income / $25k				
– Own Alternative	0.05	*	0.00	0.10
– Rent/Service Alternative	0.06	*	0.03	0.10
Most Recent Vehicle Purchased was New				
– Own Alternative	0.24	*	0.15	0.32
– Rent/Service Alternative	-0.13	*	-0.20	-0.03
Assessment Latent Variable (LV)				
– Own Alternative	1.20	*	1.15	1.27
– Rent/Service Alternative	0.85	*	0.78	0.93
Neutral Respondent LV				
– Own Alternative	-0.18	*	-0.25	-0.12
– Rent/Service Alternative	0.09	*	0.02	0.15
No-Opinion Respondent LV				
– Own Alternative	0.04		-0.09	0.12
– Rent/Service Alternative	-0.22	*	-0.27	-0.16
Uninformed Opinionated Respondent LV				
– Own Alternative	0.20	*	0.16	0.25
– Rent/Service Alternative	-0.35	*	-0.42	-0.26
Number of Observations			1245	
Model Log-likelihood at Posterior Means			-68512.6	
Choice Model Null Log-likelihood			-1367.8	
Choice Model Log-likelihood at Posterior Means			-1267.1	

* denotes statistically significant variables in the structural model.

The autonomous vehicles case study presented found evidence of differing intended usage between these groups. Specifically, it was found that opinion neutral respondents were less receptive to owning an AV than no-opinion respondents. But opinion neutral respondents were more receptive to renting and using AVs as a service versus the no-opinion respondents. The characteristics of the neutral opinion and no-opinion respondents was found to have similarities and variation in this study. Both groups tended to be younger and have lower incomes. Although both groups tended to have less formal education than opinionated respondents, the no-opinion respondents has a greater propensity to have less education than the neutral respondents. This results follows from previous research that found that no-opinion respondents tended to be less formally educated.

6.5.1 Model Enhancements

The formulation in this case study is limited in its linear-in-parameter formulation. These latent variables can be used to analyze heterogeneity in individuals' preferences through interactions with respondent characteristics. For example, the model in the case study could interact the positive assessment latent variable with each respondent type propensity to analyze heterogeneity in assessment by familiarity and neutrality. Additionally, the latent variable could be used to create heterogeneity in preference through systematic taste variation or heterogeneity-in-means model formulations.

Additionally, the model can be enhanced by transforming the neutrality-familiarity type propensity latent variables. A logit transformation could control the scale of the latent variables. In this way, they could be seen as probabilities of type membership and used to create weighted preferences for various model variables. This naturally leads to the possibility of using a latent class structure where the propensities are mapped to class memberships. This 4-class latent class model with indicators structure (Ben-Akiva and Bruno, 1995) would allow for random preference heterogeneity between the neutrality-familiarity types.

6.5.2 Likert Scale Changes

The framework can be adapted to handle variations in the asking of Likert scale questions. For an even-numbered scale, two most middle options could be used as a factor correlating with unsure respondents. The latent variable measurement equation indicators would be changed such that the indicator function corresponds to choosing either of the two middle options. Additionally, the framework can also handle the inclusion of an explicit "No opinion" option. The latent variable measurement equation indicators would be changed such that the indicator function corresponds to choosing the "No opinion" option. In this setup, a familiarity question is unnecessary to differentiate neutral respondents from no-opinion respondents – the groups are naturally differentiated by the additional choice option. The framework still brings value in that situation since it provides a way to identify the likely population of no-opinion respondents and to incorporate their preferences into the choice model. Asymmetric scales have seen limited usage in ICLV models (Bahamonde Birke et al., 2016) but present a challenge for this framework. Specifically, there is limited research on which option(s) respondents would choose if they had no opinion on the matter. Additional research is needed on respondent behavior with asymmetric Likert scales.

6.5.3 Context-Demographics Link

Additional research is needed to explore the context-demographics link in familiarity and neutrality. The lack of clarity in the linkage is an open question and may be because familiarity in topics varies greatly. The formal education link seems most clear, and this may be due to biases from societal norms and expectations. Specifically, we found that more educated respondents were more comfortable making more polar opinionated statements even if ill-informed about autonomous vehicles. As attitudinal questions are incorporated into travel, activity, and energy research more, there may be concerns that more attention will need to focus on ascertaining the opinions of potentially vulnerable groups such those with lower incomes and less educations. This also raises questions about how forecasting with these groups is performed when these groups have quite heterogeneous preference although they are recorded with similar opinions and attitudes.

6.5.4 Other Reasons for Choosing the Middle Option

This study concentrated on only two motivations for choosing middle options in Likert scale questionnaires. Additional reason have been shown to exist and it is important to also account for these possibilities in future study design and analysis. Apart from true neutrality, Nowlis et al. (2002) have identified respondents may select the middle option when they have ambivalent feelings toward the object thus indicate indifference. Furthermore, Kulas et al. (2008) demonstrated that the Likert-type middle response category is at least sometimes used as a dumping ground – selected when a more appropriate alternative (such as it depends, or I can’t decide) is not available. This could occur if the respondent has beliefs simultaneously at both ends of the attitudinal scale.

Chapter 7: Improving the spatial transferability of travel demand forecasting models: An empirical assessment of the impact of incorporating attitudes on model transferability

7.1. Introduction

Spatial transferability of travel demand forecasting models, i.e. the ability to use a travel demand forecasting model developed in one region for travel demand forecasting in another region, is of considerable interest due to a variety of reasons (Atherton and Ben-Akiva, 1976; Sikder et al., 2014, 2013). This is so, because the ability to transfer models between regions can save significant cost and time for regions that cannot afford to build a model from scratch. The issue of spatial transferability is relevant to not just small/mid-sized regions in the United States, who are generally short of funds to conduct an extensive data-collection. It is also relevant to planning agencies in many developing countries, which generally have a meager budget for transportation planning (Santoso and Tsunokawa, 2005).

Ben-Akiva (1981) and Hansen (1981) suggested four different levels at which transferability (spatial, temporal or cross-cultural) must be considered from a theoretical standpoint: 1) underlying theory of travel behavior, which involves transferability of broad behavioral postulates, such as the random utility maximization decision rule; 2) model structure, which involves transferability of mathematical model structure, such as logit, nested logit, mixed logit and probit models of discrete choice; 3) empirical specification, which involves transferability of explanatory variables in the model specification; and 4) parameter values, which involves transferability of parameter estimates across contexts. Ideally, a forecasting model is considered perfectly transferable between contexts, if the model is transferable from the above mentioned four standpoints. However, perfect transferability is an unreasonable expectation due to a variety of reasons. First, there is increasing evidence of violations of homo economicus human behavior and inability of present statistical/behavioral models to account for these variations make them less transferable. Second, choice of a particular model structure from a variety of plausible ones can also introduce approximations and reduce transferability. Third, variations in model specifications – including omission of certain explanatory variables, neglecting observed and unobserved heterogeneity, and sampling and measurement errors – can also amplify the issue and make it harder to achieve perfect transferability. Lerman (1981) pointed out that as models are only abstractions of reality, the expectation of perfect transferability is too overly restrictive and a reasonable expectation would be whether models from different contexts are close enough to being substitutable for some pre-defined purpose. Further, Koppelman and Wilmot (1986) pointed out that transferability should not be viewed as a dichotomous property. Rather, transferability assessment should talk about degree of transferability. This degree of transferability can be measured in various ways and one possible method is to use the locally estimated model as a yardstick against which a transferred model is assessed. For practical purposes, if a transferred model performs better than (or as good as) a model estimated using locally available data, then the transferred model can be used at the local context.

Over the years, researchers have argued that a positive relationship exists between model specification and model transferability (Atherton and Ben-Akiva, 1976; Koppelman and Wilmot, 1986; Lerman, 1981; Tardiff, 1979). That hypothesis is that since transferability is based on the assumption of behavioral regularity across contexts, well specified models should be able to capture this behavioral regularity better than naïve models and hence are expected to be more transferable (Koppelman and Wilmot, 1986). In this context, it has been speculated that inclusion of so-called “soft factors” (or latent variables)—attitudes, perception, norms, and beliefs—which greatly influence an individual’s decisions, might produce models which have higher transferability than traditionally estimated models with only observable socio-economic characteristics (Louviere, 1981).

This study is aimed at testing the hypothesis that travel demand forecasting models with observable as well as latent variables are better than traditionally estimated models with just observable explanatory variables. Specifically, we compare the spatial transferability of traditionally estimated multinomial logit (MNL) models with the spatial of transferability of integrated choice and latent variable (ICLV) models, which are used to incorporate “soft factors” (or latent variables) in traditional discrete choice models. ICLV models can offer greater insights into the decision-making process by including additional information through measurement equations for the latent variables. It is also believed that ICLVs produces more efficient model outputs (i.e. with less variation), such as demand elasticities and market predictions (Vij and Walker, 2016). Even though improved spatial transferability from incorporation of latent factors has long been hypothesized, there is limited empirical evidence in the literature that this hypothesis is true. Further, although existing work has examined whether ICLV models are more behaviorally sound and offer better predictions, but the spatial transferability of ICLV models has not been fully explored. Does the increased behavioral realism and predictability extend beyond the original spatial context to other areas? We hypothesize that well-specified ICLV models – particularly, ones that better capture the observed/unobserved heterogeneity in the data – will perform better (in terms of spatial transferability) than traditionally estimated choice models. We speculate that this will be due to how ICLV models produce non-linearity in the impact of exogenous variables on the choice outcome.

To test the above-mentioned hypothesis, this paper makes use of data from a survey administered among 811 respondents in the state of Florida and Michigan in the United States to conduct an empirical study. The available data is used to model an individual’s intention to use AVs using traditionally estimated multinomial logit models (without any latent variables) and integrated choice and latent variable (ICLV) models. Then, a spatial transferability assessment is performed using various assessment techniques and metrics available in the literature.

The structure of this study is as follows. Section 2 describes the econometric model structures, transferability assessment techniques, and transferability assessment metrics used in this study. Section 3 describes the empirical setting in this study and details of the estimated models. Section 4 present the transferability assessment procedure and results. Finally, section 5 summarizes and concludes the study.

7.2. Econometric Models and Transferability Assessment Techniques

In the forthcoming empirical analysis, we study the spatial transferability of two econometric model structures: 1) multinomial logit, and 2) integrated choice and latent variable models. This section describes these two model structures in detail and the transferability assessment technique and measures.

7.2.1 Multinomial Logit (MNL) Model

One of the most popular econometric model structures for modeling discrete choice outcomes is the multinomial logit (MNL) model. Its popularity is largely due to its closed form choice outcome probability expression and easy interpretability (Train, 2009). Consistent with random utility maximization theory, the multinomial logit model can be easily represented by the following two equations:

$$U_n = Bx_n + \varepsilon_n \tag{7.1}$$

$$y_{nj} = \begin{cases} 1 \; if \; U_{nj} > U_{nj'} \; \forall j' \in \{1, \ldots, \ldots, \ldots J\} \\ \quad\quad\quad 0 \; otherwise \end{cases} \tag{7.2}$$

where U_n is a $(J \times 1)$ vector of utilities of each of J alternatives, as perceived by decision maker n, x_n is the $(K \times 1)$ vector of observable explanatory variables, B is a $(J \times K)$ matrix of model parameters denoting sensitivities to the observable variables, ε_n is the $(J \times 1)$ vector denoting the random component of the utility specification, which is independent and identically distributed (IID) extreme value, and y_{nj} is the choice indicator, equal to one if decision-maker chose alternative j, zero otherwise.

The probability that a decision maker n chooses alternative j has the following functional form:

$$P(y_{nj} = 1 | x_n; B) = \frac{exp\,(\beta_{j*} x_n)}{\sum_{j'=1}^{J} exp\,(\beta_{j'*} x_n)} \tag{7.3}$$

where β_{j*} is a $(1 \times K)$ vector corresponding to the j^{th} row of B. Equation 3 maybe combined over all alternatives to yield following probability of observing the vector of choices y_n for decision maker n:

$$f_y(y_n | x_n; B) = \prod_{j'=1}^{J} [P(y_{nj'} = 1 | x_n; B)]^{y_{nj'}} \tag{7.4}$$

The parameter estimation in the multinomial logit model is done using maximum likelihood estimation of Equation 7.4.

7.2.2 Integrated Choice and Latent Variable (ICLV) Model

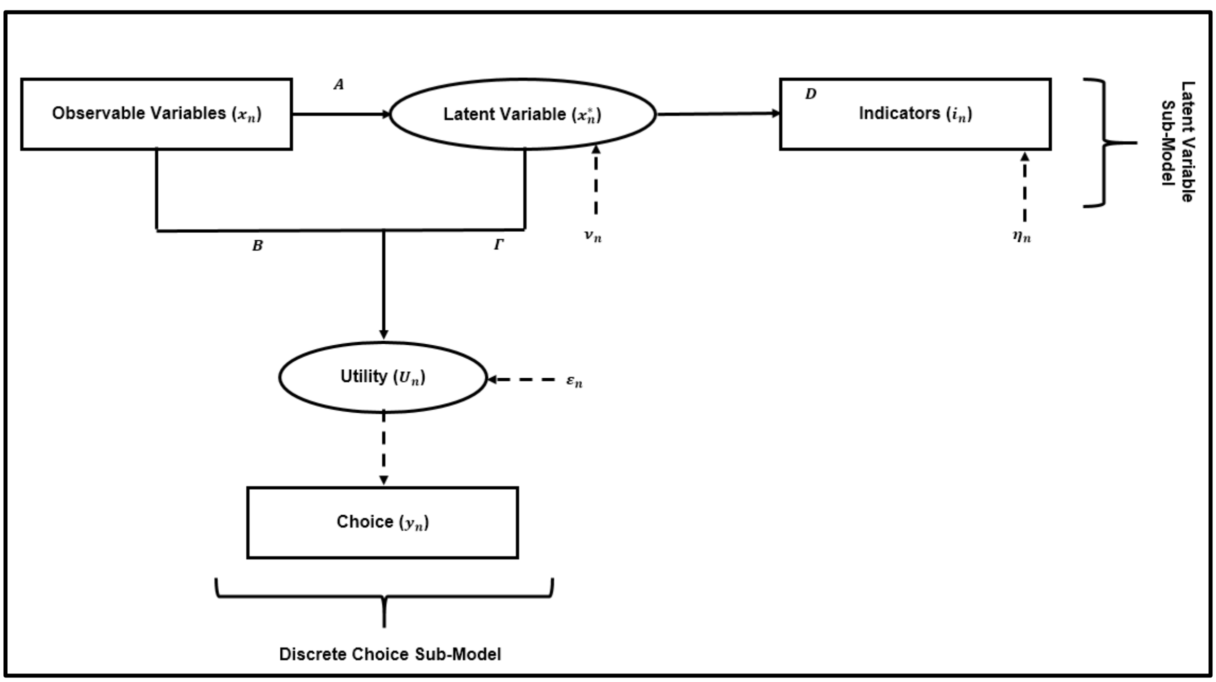

Figure 7.1. Integrated Choice and Latent Variable Framework (Vij and Walker, 2016)

Increasing emphasis on incorporation of psychological factors like attitudes, norms, perception, and beliefs in discrete choice models led to the development of integrated choice and latent variable models. The idea was that incorporation of these psychological factors will lead to more behaviorally realistic representation of choice processes and such models will have better explanatory power than traditional models without latent variables. Seminal papers on ICLV models by McFadden (1986), Train et al. (1987), Ashok et al. (2002), and Ben-Akiva et al. (2002) greatly popularized this model structure among the members of the travel behavior research community. Further, recent papers by Bolduc et al. (2005), Daly et al. (2012), Bhat and Dubey (2014), Vij and Walker (2016) have explored the benefits and limitations of ICLV models.

Figure 7.1 illustrates the ICLV model structure framework, which consists of two sub-components: a multinomial discrete choice model and a latent variable model. Each sub-component consists of a structural and measurement equation. Mathematically, the ICLV model is expressed using following four equations:

$$U_n = Bx_n + \Gamma x_n^* + \varepsilon_n \tag{7.5}$$

$$x_n^* = Ax_n + \upsilon_n \tag{7.6}$$

$$i_n = Dx_n^* + \eta_n \tag{7.7}$$

$$y_{nj} = \begin{cases} 1 \; if \; u_{nj} > u_{nj'} \; \forall j' \in \{1, \dots, \dots, \dots J\} \\ \qquad 0 \; otherwise \end{cases} \tag{7.8}$$

where U_n is a $(J \times 1)$ vector of utilities of each of J alternatives, as perceived by decision maker n, x_n is the $(K \times 1)$ vector of observable explanatory variables and x_n^* is the $(M \times 1)$ vector of latent explanatory variables, B and Γ are the $(J \times K)$ and $(J \times M)$ matrices of model parameters denoting sensitivities to the observable and latent variables, respectively, and ε_n is the $(J \times 1)$ vector denoting the random component of the utility specification; A is the $(M \times K)$ matrix of model parameters denoting the structural relationship between the latent and observable variables, and υ_n is the $(M \times 1)$ vector denoting the random component of that relationship; i_n is the $(R \times 1)$ vector of indicators used to measure the latent variables, assumed to represent deviations from the mean, D is the $(R \times M)$ matrix of model parameters denoting the sensitivities of the measurement equation, η_n is the $(R \times 1)$ vector denoting the random component of the measurement equation; and y_{nj} is the choice indicator, equal to one if decision-maker n chose alternative j, zero otherwise. The random components ε_n, υ_n, and η_n are assumed to be mutually independent.

The most popular form of the ICLV model in the literature is the *logit kernel*, where each element of ε_n, denoted by ε_{nj}, is IID gumbel across alternatives and decision makers with location parameter one zero and scale parameter one. Conditional on the latent variable, the probability that a decision-maker n chooses alternative j has the following functional form:

$$P\left(y_{nj} = 1 \middle| x_n, x_n^*; B, \Gamma\right) = \frac{exp\,(\beta_{j*}x_n + \gamma_{j*}x_n^*)}{\sum_{j'=1}^{J} exp\,(\beta_{j'*}x_n + \gamma_{j'*}x_n^*)} \tag{7.9}$$

where β_{j*} and γ_{j*} are $(1 \times K)$ and $(1 \times M)$ vectors corresponding to the j^{th} row of B and Γ, respectively. Equation 7.9 may be combined over all alternatives to yield following conditional probability of observing the vector of choices y_n for decision maker n:

$$f_y(y_n | x_n, x_n^*; B, \Gamma) = \prod_{j'=1}^{J} [P\left(y_{nj'} = 1 \middle| x_n, x_n^*; B, \Gamma\right)]^{y_{nj'}} \tag{7.10}$$

With regards to the measurement indicators, in this study, we assumed that the indicators represent Likert-scale type ordered response variable as in Daly et al. (2012). This need not always be the case as many studies in the literature have represented indicators as both continuous and unordered response variables (Ben-Akiva et al., 2002; Bhat, 2015; Bolduc et al., 2005). In the ordered representation of the indicators, the measurement equation ($i_n = Dx_n^* + \eta_n$) is assumed to be a propensity function driving choice of ordered response from L possible outcomes for each indicator. Assuming that η_n is independently (need not be identically) distributed normally across R indicators and N decision makers, the probability of the decision-maker n choosing ordered choice outcome l in the r^{th} measurement equation is written as:

$$P(w_{nrl} = 1|x_n, x_n^*; D, S) = \Phi\left[\frac{\psi_l^r - \delta_r x_n^*}{s_r}\right] - \Phi\left[\frac{\psi_{l-1}^r - \delta_r x_n^*}{s_r}\right] \quad (7.11)$$

where w_{nrl} is the ordered choice indicator, which is equal to one if the decision maker n chooses l^{th} ordered outcome for the r^{th} indicator, $\Phi[\cdot]$ is the cumulative distribution function of standard normal distribution, ψ_l^r is the l^{th} threshold dividing the propensity function for the r^{th} indicator, and S is the $(R \times 1)$ vector of scale parameters of η_n and s_r is the r^{th} element of S. Equation 7.12 can also be combined over L ordered outcomes and R indicators to yield following probability distribution function:

$$f_w(w_n|x_n x_n^*; D, S) = \prod_{r=1}^{R} \prod_{l=1}^{L} [P(w_{nrl} = 1|x_n, x_n^*; D, S)]^{w_{nrl}} \quad (7.12)$$

With regards to the structural equation, the latent variables are represented using a linear-in-parameter formulation, where v_n is assumed to be distributed normally with a mean vector of zeros and covariance matrix Ω. The probability distribution associated with latent variables are expressed as:

$$f_{x_n^*}(x_n^*|x_n; A, \Omega) = (2\pi)^{-\frac{M}{2}} |\Omega|^{-\frac{1}{2}} exp\left(-\frac{1}{2}(x_n^* - Ax_n)^T \Omega^{-1} (x_n^* - Ax_n)\right) \quad (7.13)$$

The joint unconditional probability distribution function for the choice and measurement indicators is written as:

$$f_{y,w}(y_n, w_n|x_n, x_n^*; B, \Gamma, D, S, A, \Omega) =$$
$$\int_{x_n^*} f_y(y_n|x_n, x_n^*; B, \Gamma) f_w(w_n|x_n x_n^*; D, S) f_{x_n^*}(x_n^*|x_n; A, \Omega) dx_n^* \quad (7.14)$$

The unknown parameters in equation 7.14 are estimated using maximum simulated likelihood estimation.

Three important points need to be noted here. First, as standard practice in ordered response models, $\psi_0^r = -\infty$ and $\psi_{L+1}^r = \infty$. Second, all elements of S were fixed to one. Third, for identification reasons, the on-diagonal and off diagonal elements of Ω were fixed to one and zero respectively.

How are ICLV models used for forecasting?

It can be considered that in the ICLV model, the measurement equation is only an auxiliary in the estimation of the structural equation. Typically, the measurement equation is not used for forecasting purposes[26]. From a forecasting standpoint, as we are normally interested in changes in the utility functions and their repercussion on choice probabilities, any changes in the observable explanatory variables will induces changes in the latent variables, and these updated latent variables can be used to calculate updated choice probability. When forecasting, the choice model probabilities in an ICLV model are calculated by marginalizing Equation 7.10 over the distribution of latent variable x_n^* and is written as:

$$f_y(y_n|x_n, x_n^*; B, \Gamma, A, \Omega) = \int_{x_n^*} f_y(y_n|x_n, x_n^*; B, \Gamma) f_{x_n^*}(x_n^*|x_n; A, \Omega) dx_n^* \quad (7.15)$$

7.2.3 Assessment Techniques

[26] This is a good thing as if the measurement equation were required for forecasting from the choice model, it would require the analyst to first generate forecasts for the indicators. But forecasted future values for indicators are generally not available.

There are two popular approaches to assess spatial transferability of travel demand forecasting models: (a) the application-based approach, and (b) the estimation-based approach. In the application-based approach, model parameters are estimated using data from one region (the base context) and applied to data in another region (the application context) to assess how well the model in the base context predicts in the application context. This approach tests the transferability of a model as a whole, without allowing an examination of which specific parameters are transferable. In the estimation-based approach, also known as joint-context estimation, data from the base and application contexts are combined to estimate a single model while recognizing potential differences between the two contexts. This is done by estimating context specific difference parameters. Simple t-tests on these difference parameters can shed light on whether the parameter estimates are different between the two contexts. Advantage of this approach is that one can test whether each (and every) parameter in a model is transferable. But the predominantly used application-based approach is used in this study to examine the transferability of MNL and ICLV models.

7.2.4 Assessment Metrics

The assessment measures can be classified into absolute and relative measures, where absolute measures assess how well the transferred model represents the observed behavior in the application context and relative measures assess the performance of transferred model relative to the application context model. Relative Aggregate Transfer Error (RATE) measure is used in this study for the transferability assessment. Let *i* and *j* represent the indices for the estimated and transferred models in the study region, respectively. Similarly, PS_k and OS_k represents the predicted and observed shares for the choice alternative *k*. Relative Error Measure for each alternative (REM_k) is defined by $\frac{PS_k - OS_k}{OS_k}$ and Root Mean Square Error for a model (β) applied to a dataset *m* is represented by $RMSE_m(\beta)$ and defined by $(\frac{\sum_k PS_k \times REM_k^2}{\sum_k PS_k})^{1/2}$. Finally, the RATE measure is the ratio of RMSE values of the transferred model (β_j) and the estimated model (β_i) in a region, i.e. $\frac{RMSE_i(\beta_j)}{RMSE_i(\beta_i)}$. It is important to note here that the constants of transferring model are adjusted using the procedure demonstrated in (Train, 2009) before calculating predicted shares in the study region.

7.3. Study Setting

In this study, we make use of data collected from a survey conducted among 414 and 397 respondents from Florida and Michigan states respectively. Respondents were asked about their preferred way of using autonomous vehicles (AVs), when they become readily available. Available options for the respondent to choose included:

- C1. Own AVs and use them only for personal use or use by family members
- C2. Own an AV and earn extra income on the side by making it available to other drivers when not needed
- C3. Own an AV and earn extra income on the side by providing rides for fellow passengers when you use it
- C4. Rent an AV as the need arises
- C5. Use AVs in the form of transportation (taxi, or public transit) provided by a service provider
- C6. Neither interested in investing in an AV nor using AVs as a transportation service.

Due to inadequate number of responses for some choices, they were grouped into three choice alternatives, namely, owning an AV (C1+C2+C3), using it as a shared vehicle (C4+C5) and not using an AV at all (C6). Apart from the response on preferred way of using an AV, the survey collected information on personal and household demographics, information on current travel behavior, perceptions of various attributes of AV technologies and opinion on familiarity with the technology, and perceptions of the benefits and concerns with AVs. Table 7.1 presents a comparison of demographic characteristics between Florida and Michigan. The comparison suggests that Florida and Michigan respondents have similar demographic characteristics and spatial transferability of models between these two regions can be explored further. With all the available information, MNL and ICLV models are developed and discussed in detail in the next two sub-sections.

7.3.1 MNL Models

Table 7.2 shows the estimation results of MNL models for Florida and Michigan data respectively, obtained after an extensive model specification testing. For identification reasons, *not using an AV* is considered as the base alternative. In the Florida model, people older than 60 years or people with a household containing at least one child younger than 16 years are more likely to intend to use AVs either in own or sharing forms. On the other hand, Floridians with household income greater than $100,000 are more inclined to use AVs in either forms. Moreover, male Floridians and white Floridians are less inclined to use AVs in owning and sharing forms respectively. Finally, Floridians who live in a single household or whose one-way commute

Table 7.1. Comparison of demographic characteristics between Florida and Michigan

Variable name	Mean value (Florida)	Mean value (Michigan)
Age indicator (1 if age is less than 40years, 0 otherwise)	0.08	0.05
Middle age indicator (1 if age is greater than 40 years and less than 60 years, 0 otherwise)	0.34	0.39
Old age indicator (1 if the respondent is more than 60 years old, 0 otherwise)	0.58	0.56
Male indicator (1 if the respondent is male, 0 otherwise)	0.59	0.65
White ethnicity indicator (1 if the respondent's ethnicity is white, 0 otherwise)	0.87	0.90
High household income indicator (1 if the respondent's household income is more than $100,000, 0 otherwise)	0.37	0.49
Single person household indicator (1 if the respondent lives in a single person household, 0 otherwise)	0.22	0.15
Short commute indicator (1 if the respondent's typical one-way commute distance is less than 5 miles, 0 otherwise)	0.25	0.20
Household with child indicator (1 if the household has at least one child with age less than 16 years, 0 otherwise)	0.12	0.15
Number of cars indicator (1 if the household has more than 3 cars, 0 otherwise)	0.05	0.15
Crash indicator (1 if the respondent had a crash in the past, 0 otherwise)	0.95	0.83
Inability indicator (1 if the household has people with physical or cognitive constraints, 0 otherwise)	0.10	0.08
Worker indicator (1 if the respondent is a worker; 0 otherwise)	0.50	0.59
Education indicator (1 if respondent holds a bachelor's degree or above, 0 otherwise)	0.67	0.67

Table 7.2. Estimation results of MNL models for Florida and Michigan (Not using an AV is the base alternative)

Variable name	Florida model parameter estimate (t-stat)		Michigan model parameter estimate (t-stat)	
	Own an AV for personal use	Share an AV	Own an AV for personal use	Share an AV
Constant	0.0198 (0.08)	-0.0107 (-0.03)	-0.262 (-1.69)	-1.27 (-5.81)
Old age indicator (1 if the respondent is more than 60 years old, 0 otherwise)	-0.529 (-2.09)	-0.770 (-2.36)		
Male indicator (1 if the respondent is male, 0 otherwise)	0.312 (1.45)			
White ethnicity indicator (1 if the respondent's ethnicity is white, 0 otherwise)		-1.00 (-2.91)		
High household income indicator (1 if the respondent's household income is more than $100,000, 0 otherwise)	0.454 (1.93)	0.532 (1.70)	0.721 (3.10)	0.884 (2.90)
Single person household indicator (1 if the respondent lives in a single person household, 0 otherwise)		0.571 (1.77)		
Short commute indicator (1 if the respondent's typical one-way commute distance is less than 5 miles, 0 otherwise)		0.535 (1.87)		
Household with child indicator (1 if the household has at least one child with age less than 16 years, 0 otherwise)	-1.06 (-2.71)	-0.622 (-1.37)		
Number of cars indicator (1 if the household has more than 3 cars, 0 otherwise)			-1.16 (-3.40)	-0.910 (-2.10)
No. of observations	414		397	
Log-likelihood at convergence	-412.05		-395.277	
Log-likelihood for constants only model	-429.823		-406.722	

distance is less than 5 miles tended to choose AVs in sharing form. Similar to Floridians, Michiganians whose household income is greater than $100,000 tend to use AVs in either form. However, Michiganians with more than 3 cars in the household tend not to use AVs in either forms.

7.3.2 ICLV Models

Table 7.3a presents the structural models for the Florida and Michigan regions. Results suggest that people in both Florida and Michigan who are younger than 40 years tend to be more favorable towards AVs. Similarly, Floridians who work and Michiganians with at least a bachelor's degree are more likely to be in favor of AVs. However, Floridians with children or with physically disabled people in their household are less likely to be in favor of AVs.

The choice model specification for each ICLV model is adopted from its corresponding MNL model to allow for a fairer spatial transferability comparison between the different models. As one may expect, all the adopted variables in the ICLV's choice model might not be statistically significant at 95% confidence interval and the same can be seen in Table 7.3b. However, both the Florida and Michigan model results suggest that people who are in favor of AVs tend to use AVs either by owning or sharing. Other interesting findings are that Floridians with shorter commute distance tend to use AVs in sharing form and white Floridians tend not to use AVs in sharing form.

Finally, Table 7.3c presents the measurement model results. As expected, all the coefficients of assessment latent variable are positive for the benefit indicators such as AVs helps in fewer traffic crashes, less traffic congestion and lower insurance rates etc.; and negative for the concern indicators such as AVs can have system and equipment failure, create motion sickness and poor performance in unexpected weather conditions etc.

Table 3a: Structural models of ICLV models for Florida and Michigan regions

Variable name	Florida model parameter estimate (t-stat)	Michigan model parameter estimate (t-stat)
Age indicator (1 if age less than 40years, 0 otherwise)	0.483 (3.12)	0.780 (3.33)
Household with child indicator (1 if the household has at least one child with age less than 16 years, 0 otherwise)	-0.765 (-2.96)	
Crash indicator (1 if the respondent had a crash in the past, 0 otherwise)	0.0792 (0.92)	
Inability indicator (1 if the household has people with physical or cognitive constraints, 0 otherwise)	-0.441 (-3.42)	
Worker indicator (1 if the respondent is a worker; 0 otherwise)	0.214 (2.65)	
Education indicator (1 if respondent holds a bachelor's degree or above, 0 otherwise)		0.209 (1.87)

Table 3b: Choice models of ICLV models for Florida and Michigan regions

Variable name	Florida model parameter estimate (t-stat)	Michigan model parameter estimate (t-stat)
Variables in the "owning an AV" utility of choice model		
Constant	0.0141 (0.02)	-1.16 (-2.32)
Old age indicator (1 if the respondent is more than 60 years old, 0 otherwise)	-0.0541 (-0.14)	
Male indicator (1 if the respondent is male, 0 otherwise)	0.186 (0.68)	
High household income indicator (1 if the respondent's household income is more than $100,000, 0 otherwise)	0.247 (0.70)	0.816 (1.72)
Household with child indicator (1 if the household has at least one child with age less than 16 years, 0 otherwise)	-1.82 (-1.50)	0.445 (0.55)
Assessment latent variable	1.49 (2.10)	2.48 (8.24)

Variables in the "sharing an AV" utility of choice model		
Constant	0.331 (0.55)	-1.53 (-2.86)
Old age indicator (1 if the respondent is more than 60 years old, 0 otherwise)	-0.430 (-1.01)	
White ethnicity indicator (1 if the respondent's ethnicity is white, 0 otherwise)	-1.07 (-3.12)	
High household income indicator (1 if the respondent's household income is more than $100,000, 0 otherwise)	0.416 (1.03)	0.503 (0.96)
Single household indicator (1 if the respondent lives in single person household, 0 otherwise)	0.365 (1.13)	
Short commute indicator (1 if the respondent's typical one-way commute distance is less than 5 miles, 0 otherwise)	0.530 (1.80)	
Household with child indicator (1 if the household has at least one child with age less than 16 years, 0 otherwise)	-1.51 (-1.17)	0.412 (0.46)
Assessment latent variable	1.32 (1.79)	2.14 (6.84)

Table 3c: Measurement models of ICLV models for Florida and Michigan regions

Indicator description	Florida model parameter estimate (t-stat)	Michigan model parameter estimate (t-stat)
Fewer traffic crashes	1.74 (11.98)	1.83 (12.04)
Less traffic congestion	1.37 (12.65)	1.34 (12.60)
Less stressful driving experience	2.29 (10.59)	2.13 (11.37)
More productive (than driving) use of travel time	1.49 (12.30)	1.42 (12.54)
Lower car insurance rates	1.08 (12.10)	1.04 (11.73)
Increased fuel efficiency	1.05 (11.94)	1.04 (11.73)
System/equipment failure or AV system hacking	-0.425 (-6.90)	-0.539 (-8.33)

Performance in (or response to) unexpected traffic situations	-0.462 (-7.37)	-0.597 (-8.79)
Motion sickness	-0.486 (-7.70)	-0.474 (-7.53)
Giving up my control of the steering wheel to the vehicle	-0.262 (-4.50)	-0.566 (-8.63)

7.4 Transferability Assessment

For the transferability assessment, each of the Florida and Michigan datasets were randomly divided into 8 sets of estimation and validation datasets with an 80-20 proportion of observations respectively. Using the specifications from section 3.1 and section 3.2, MNL and ICLV models are estimated for all the 16 datasets. Table 7.4a and Table 7.4b presents the estimation results of MNL models for each of the 8 datasets in Florida and Michigan respectively. Similarly, Table 7.4c and Table 7.4d presents the estimation results of ICLV models for Florida and Michigan respectively.

The transferred models for both the MNL and ICLV models were adjusted between the estimation and validation steps. The alternative specific constants in each model were iteratively adjusted (Train, 2009) until they matched the observed shares in the transferring region's estimation dataset. For each of a study area's eight dataset partitions, the constants adjusted models are then applied to each validation dataset to get the predicted percentage shares.

Table 7.5a and Table 7.6a present the calculated RATE values for Florida's MNL and ICLV models transferred to Michigan, respectively. Similarly, Table 7.5b and Table 7.6b present the calculated RATE values for Michigan's MNL and ICLV models transferred to Florida, respectively. In Tables 7.5a, 7.5b, 7.6a, and 7.6b, the RMSE1 row refers to the validation of the locally estimated model while RMSE2 refers to the validation of the transferred model. The models seem fairly transferrable between the regions after constant adjustment, as the four median RATE values (1.031, 1.268, 0.94 and 0.883) are close to 1. This suggests that both the transferred MNL and ICLV models are performing as good as to their corresponding local models in terms of predictions.

Table 4a: Estimation results of ICLV models for 8 estimation datasets of Michigan

Variable description	Full dataset	D1	D2	D3	D4	D5	D6	D7	D8
Variables in the structural equation									
Age indicator (1 if age less than 40years, 0 otherwise)	0.780	0.693	0.792	0.879	0.720	0.727	0.884	0.891	0.732
Education indicator (1 if respondent holds a bachelor's degree or above, 0 otherwise)	0.209	0.0933	0.187	0.256	0.165	0.261	0.244	0.201	0.225
Variables in the "owning an AV" utility of choice model									
Constant	-1.16	-1.02	-1.13	-1.43	-1.11	-1.21	-0.96	-1.27	-1.16
Household with child indicator (1 if the household has at least one child with age less than 16 years, 0 otherwise)	0.445	0.434	0.439	1.04	0.245	0.629	-0.06	0.376	0.631
High household income indicator (1 if the respondent's household income is more than $100,000, 0 otherwise)	0.816	0.852	0.905	1.05	0.751	0.832	0.441	1.01	0.707
Assessment latent variable	2.48	2.54	2.44	2.60	2.60	2.42	2.44	2.30	2.62
Variables in the "sharing an AV" utility of choice model									
Constant	-1.53	-1.51	-1.82	-1.55	-1.42	-1.50	-1.49	-1.56	-1.48
Household with child indicator (1 if the household has at least one child with age less than 16 years, 0 otherwise)	0.412	0.417	0.700	0.630	0.529	0.336	-0.42	0.332	0.808
High household income indicator (1 if the respondent's household income is more than $100,000, 0 otherwise)	0.503	0.689	0.710	0.431	0.342	0.534	0.339	0.698	0.370
Assessment latent variable	2.14	2.25	2.17	2.20	2.26	1.99	2.20	1.99	2.19

Table 4b: Estimation results of ICLV models for 8 estimation datasets of Florida

Variable description	Full dataset	D1	D2	D3	D4	D5	D6	D7	D8
Variables in the structural equation									
Age indicator (1 if age less than 40years, 0 otherwise)	0.483	0.110	0.253	0.374	0.373	0.327	0.507	-0.167	0.284
Household with child indicator (1 if the household has at least one child with age less than 16 years, 0 otherwise)	-0.765	-0.457	-0.536	-0.736	-0.333	-0.635	-0.686	1.47	-0.586
Crash indicator (1 if the respondent had a crash in the past, 0 otherwise)	0.0792	0.128	0.203	0.098	0.168	0.155	0.121	-1.09	0.175
Inability indicator (1 if the household has people with physical or cognitive constraints, 0 otherwise)	-0.441	-0.196	-0.236	-0.205	-0.169	-0.253	-0.206	0.980	-0.166
Worker indicator (1 if the respondent is a worker; 0 otherwise)	0.214	0.124	0.107	0.126	0.078	0.157	0.094	0.936	0.105
Variables in the "owning an AV" utility of choice model									
Constant	0.014	0.376	0.299	0.502	0.492	0.779	0.383	0.564	0.501
Old age indicator (1 if the respondent is more than 60 years old, 0 otherwise)	-0.054	0.069	0.054	0.150	0.168	-0.165	0.143	0.082	0.049
Male indicator (1 if the respondent is male, 0 otherwise)	0.186	0.407	0.201	0.242	0.193	0.213	0.468	0.800	0.156
High household income indicator (1 if the respondent's household income is more than $100,000, 0 otherwise)	0.247	0.384	0.514	-0.007	0.137	0.269	0.329	0.700	0.302
Household with child indicator (1 if the household has at least one child with age less than 16 years, 0 otherwise)	-1.82	-1.39	-1.94	-1.58	-1.34	-1.69	-1.56	-0.220	-1.28
Assessment latent variable	1.49	1.79	2.15	1.85	2.08	1.92	2.10	2.79	2.23

Variables in the "sharing an AV" utility of choice model									
Constant	0.331	0.651	0.817	0.877	0.971	1.18	1.06	0.974	1.07
Old age indicator (1 if the respondent is more than 60 years old, 0 otherwise)	-0.430	-0.344	-0.333	-0.125	-0.536	-0.505	-0.368	-0.467	-0.268
White ethnicity indicator (1 if the respondent's ethnicity is white, 0 otherwise)	-1.07	-0.986	-0.997	-1.10	-0.759	-1.23	-1.03	-1.19	-1.21
High household income indicator (1 if the respondent's household income is more than $100,000, 0 otherwise)	0.416	0.845	0.549	0.282	0.068	0.445	0.606	0.347	0.291
Single person household indicator (1 if the respondent lives in a single person household, 0 otherwise)	0.365	0.726	0.228	0.350	0.194	0.469	0.243	0.423	0.378
Short commute indicator (1 if the respondent's typical one-way commute distance is less than 5 miles, 0 otherwise)	0.530	0.432	0.586	0.518	0.587	0.547	0.612	0.845	0.465
Household with child indicator (1 if the household has at least one child with age less than 16 years, 0 otherwise)	-1.51	-1.18	-1.21	-0.726	-14.1	-1.47	-1.46	-0.040	-1.10
Assessment latent variable	1.32	1.40	1.65	1.47	1.72	1.60	1.61	0.749	1.78

Table 4c: Estimation results of MNL models for 8 estimation datasets of Michigan

Variable description	Full dataset	D1	D2	D3	D4	D5	D6	D7	D8
Variables in the "owning an AV" utility									
Constant	-0.262	-0.290	-0.196	-0.247	-0.263	-0.291	-0.245	-0.249	-0.321
Number of cars indicator (1 if the household has more than 3 cars, 0 otherwise)	-1.16	-1.02	-1.27	-1.23	-1.21	-0.988	-1.26	-0.997	-1.15
High household income indicator (1 if the respondent's household income is more than $100,000, 0 otherwise)	0.721	0.702	0.726	0.797	0.682	0.771	-1.47	0.682	0.803
Variables in the "sharing an AV" utility									
Constant	-1.27	-1.31	-1.34	-1.32	-1.30	-1.25	0.611	-1.13	-1.25
Number of cars indicator (1 if the household has more than 3 cars, 0 otherwise)	-0.910	-0.795	-1.02	-1.15	-0.886	-0.735	-0.750	-1.01	-0.986
High household income indicator (1 if the respondent's household income is more than $100,000, 0 otherwise)	0.884	0.990	0.865	0.964	0.904	0.874	0.829	0.848	0.794

Table 4d: Estimation results of MNL models for 8 estimation datasets of Florida

Variable description	Full dataset	D1	D2	D3	D4	D5	D6	D7	D8
Variables in the "owning an AV" utility									
Constant	0.019	-0.197	0.0341	-0.919	0.129	0.194	-0.034	-0.104	-0.025
Old age indicator (1 if the respondent is more than 60 years old, 0 otherwise)	-0.529	-0.470	-0.643	-0.684	-0.490	-0.572	-0.449	-0.435	-0.518
Male indicator (1 if the respondent is male, 0 otherwise)	0.312	0.456	0.275	0.270	0.208	0.224	0.336	0.377	0.356
High household income indicator (1 if the respondent's household income is more than $100,000, 0 otherwise)	0.454	0.561	0.538	0.345	0.331	0.416	0.515	0.508	0.429
Household with child indicator (1 if the household has at least one child with age less than 16 years, 0 otherwise)	-1.06	-0.979	-1.20	-1.20	-1.08	-1.17	-1.09	-0.831	-0.932
Variables in the "sharing an AV" utility									
Constant	-0.011	-0.403	0.005	-0.020	0.0267	0.079	-0.005	0.069	0.129
Old age indicator (1 if the respondent is more than 60 years old, 0 otherwise)	-0.770	-0.663	-0.843	-0.715	-1.01	-0.716	-0.838	-0.813	-0.593
White ethnicity indicator (1 if the respondent's ethnicity is white, 0 otherwise)	-1.00	-0.962	-0.919	-1.09	-0.621	-1.15	-0.947	-1.07	-1.22
High household income indicator (1 if the respondent's household income is more than $100,000, 0 otherwise)	0.532	0.720	0.464	0.600	0.122	0.562	0.722	0.620	0.435
Single person household indicator (1 if the respondent lives in a single person household, 0 otherwise)	0.571	1.00	0.429	0.678	0.363	0.656	0.466	0.430	0.527
Short commute indicator (1 if the respondent's typical one-way commute distance is less than 5 miles, 0 otherwise)	0.535	0.451	0.544	0.539	0.579	0.561	0.608	0.513	0.476
Household with child indicator (1 if the household has at least one child with age less than 16 years, 0 otherwise)	-0.622	-0.200	-0.722	-0.529	-0.611	-0.710	-0.858	-0.735	-0.608

Additionally, when comparing the RMSE of the localized models to the transferred ICLV models, predictive strength is comparable. In the case of the Florida to Michigan transfer, the median RATE between the localized MNL and transferred ICLV is 1.02. For the Michigan to Florida transfer, the median RATE between the localized MNL and transferred ICLV is 0.99. This result suggests that ICLV model are not improving on fit, but are at least replicating the prediction of market shares. The implications of this are developed in the discussion in section 7.5.

7.5 Summary and Conclusions

This chapter explored the spatial transferability of ICLV models in search of providing empirically support for the hypothesis that ICLV may improve the transferability of travel behavior and travel demand models. Through a case study, the intention to use autonomous vehicles were modeled with ICLV models using demographic data and autonomous vehicle opinions. The models found support for the use of attitudinal data in modeling the choice of intended use. A positive assessment latent variable was modeled using the opinion data in a series of measurement equations and linking the correlation and directionality of those responses to demographic variables via a structural equation.

The case study explored the transference of models estimated in Florida and Michigan – states located in different regions of the United States. This effort found that transferred ICLV models tended to outperform their locally estimated counterpart – in contrast to the MNL models used. Additionally, when comparing performance between locally estimated MNL models to transferred ICLV models, predictive accuracy was very similar with nearly equivalent median RMSE values observed. This result does not provide support for the hypothesis of improved

Table 5a: RATE values for the Florida MNL model transferred to Michigan

FL to MI	Dataset 1	Dataset 2	Dataset 3	Dataset 4	Dataset 5	Dataset 6	Dataset 7	Dataset 8	
RMSE1	0.275	0.378	0.232	0.103	0.165	0.218	0.811	0.109	Median
RMSE2	0.242	0.389	0.218	0.094	0.102	0.26	0.809	0.119	
RATE	1.134	0.971	1.06	1.087	1.624	0.84	1.002	0.909	1.031

Table 5b: RATE values for the Michigan MNL model transferred to Florida

MI to FL	Dataset 1	Dataset 2	Dataset 3	Dataset 4	Dataset 5	Dataset 6	Dataset 7	Dataset 8	
RMSE1	0.257	0.233	0.152	0.078	0.281	0.191	0.1	0.068	Median
RMSE2	0.251	0.276	0.111	0.114	0.235	0.143	0.064	0.047	
RATE	1.024	0.844	1.366	0.683	1.2	1.337	1.553	1.444	1.268

Table 6a: RATE values for the Florida ICLV model transferred to Michigan

FL to MI	Dataset 1	Dataset 2	Dataset 3	Dataset 4	Dataset 5	Dataset 6	Dataset 7	Dataset 8	
RMSE1	0.288	0.378	0.209	0.109	0.152	0.21	0.767	0.128	Median
RMSE2	0.281	0.365	0.224	0.163	0.134	0.261	0.808	0.144	
RATE	1.026	1.035	0.929	0.668	1.133	0.806	0.95	0.89	0.94

Table 6b. RATE values for the Michigan ICLV model transferred to Florida

MI to FL	Dataset 1	Dataset 2	Dataset 3	Dataset 4	Dataset 5	Dataset 6	Dataset 7	Dataset 8	
RMSE1	0.21	0.226	0.132	0.09	0.232	0.188	0.08	0.047	Median
RMSE2	0.196	0.262	0.15	0.065	0.285	0.212	0.16	0.009	
RATE	1.071	0.862	0.88	1.385	0.816	0.887	0.497	4.976	0.883

transferability in this case. Further study is needed to understand whether this is specific to this case study or a more general finding. In particular, this study used small samples for the transference and it is known that ICLV need larger samples to obtain efficiencies and for estimation convergence. Additionally, this is a question about whether opinions about new technology are transferable between regions. If these opinions are not transferable, then an ICLV is unlikely to transfer well since the latent variable equations likely would result in overfitting.

Although support for improved transferability was not found, there were still encouraging results that support continued research and possible application. In particular, equivalent transferability to a local estimated MNL suggests that the non-linearity added, through the ICLV structural equation, is not overfitting the data. This suggests that it may be possible to use the transferred ICLV to conduct analysis that could be done with the MNL model solely.

Additionally, a limitation of this study for practical usage is that equivalent parameter sets for the models were tested between the ICLV and MNL. The MNL model was used for the basis of inclusion or exclusion of variable through the statistical significance of parameter estimates[27]. An ICLV model estimated in another region could suggest the retention of some variables through the use of the structural equation. There is no guarantee that a variable that is insignificant in an MNL estimation will produce insignificant estimates in a corresponding ICLV model. The combined non-linear effect from the latent variable structural model and the choice model may have a mean not significantly different than zero, but a component(s) of that non-linearity may be statistically significant when separated. Thus an ICLV model may be able to guide itself to a reduced form choice model specification that may have been unconsidered by the analyst (when developing from a multinomial logit base only) and possibly better fitted (Vij and Walker, 2016). If this is the case, then it will be beneficial for regions that generally lack detailed attitudinal/perception data to borrow ICLV model from regions that have access to such detailed attitudinal/perception data to support the use of particular non-linear formulations.

[27] This practice is often frowned upon by academics, but it is commonly still used in practice for variable selection. As the case study is more about imitating practiced techniques, statistical significance was used.

REFERENCES

Abay, K., & Mannering, F. (2016). An empirical analysis of risk-taking in car driving and other aspects of life. *Accident Analysis and Prevention, 97*, 57-68.

Abdelwahab, H., & Abdel-Aty, M. (2001). Development of artificial neural network models to predict driver injury severity in traffic accidents at signalized intersections. *Transportation Research Record, 1746*, 6-13.

Ajzen, I. (2001). Nature and operation of attitudes. *Annual Review of Psychology, 52*(1), 27-58.

Alsnih, R., & Hensher, D. (2003). The mobility and accessibility expectations of seniors in an aging population. *Transportation Research Part A, 37*(10), 903-916.

American Association of State Highway and Transportation Officials. (2010). Highway Safety Manual, Washington, DC.

American Automobile Association. (2017). Organization Overview. Retrieved from: https://autoclubsouth.aaa.com/About/organization-overview.aspx (June 8, 2017)

Anastasopoulos, P., & Mannering, F. (2009). A note on modeling vehicle accident frequencies with random-parameters count models. *Accident Analysis and Prevention, 41(*1), 153–159.

Anderberg, M. (1973). Cluster analysis for applications. Academic Press. New York, 1973.

Anderson, M. (2015). The demographics of device ownership. Retrieved from: http://www. pewinternet.org/2015/10/29/the-demographics-of-device-ownership/ (August 18, 2016)

Anderson, J., Kalra, N., Stanley, K., Sorensen, P., Samaras, C., Oluwota, O., 2014. Autonomous Vehicle Technology – A Guide for Policymakers. RAND Corporation, Santa Monica, CA.

APTA. (2013). Millennials and Mobility: Understanding the Millennial Mindset, Report, 2013. Retrieved from: http://www.apta.com/resources/reportsandpublications/Documents/APTA-Millennials-and-Mobility.pdf (October 21, 2014).

Ashok, K., Dillon, W.R., & Yuan, S. (2002). Extending discrete choice models to incorporate attitudinal and other latent variables. *Journal of Marketing Research, 39*, 31–46.

Atherton, T., & Ben-Akiva, M. (1976). Transferability and updating of disaggregate travel demand models. *Transportation Research Record, 610*, 12–18.

Bahamonde-Birke, F.J., & Hanappi, T., (2016). The potential of electromobility in Austria: Evidence from hybrid choice models under the presence of unreported information. *Transportation Research Part A: Policy and Practice, 83,* pp.30-41.

Baka, A., Figgou, L., & Triga, V. (2012). 'Neither agree, nor disagree': a critical analysis of the middle answer category in Voting Advice Applications. *International Journal of Electronic Governance, 5*(3-4), 244-263.

Bansal P., & Kockelman, K. (2017). Forecasting Americans' long-term adoption of connected and autonomous vehicle technologies. *Transportation Research Part A, 95,* 49-63.

Bansal, P., Kockelman, K.M., & Singh, A. (2016). Assessing public opinions of and interest in new vehicle technologies: an Austin perspective. *Transportation Research Part C, 67*, 1-14.

Behnood, A., & Mannering, F. (2015). The temporal stability of factors affecting driver-injury severities in single-vehicle crashes: Some empirical evidence. *Analytic Methods in Accident Research, 8*, 7-32.

Behnood, A., & Mannering, F. (2016). An empirical assessment of the effects of economic recessions on pedestrian-injury crashes using mixed and latent-class models. *Analytic Methods in Accident Research, 12*, 1-17.

Behnood, A., & Mannering, F. (2017a). The effect of passengers on driver-injury severities in single-vehicle crashes: A random parameters heterogeneity-in-means approach. *Analytic Methods in Accident Research, 14*, 41-53.

Behnood, A., & Mannering, F. (2017b). Determinants of bicyclist injury severities in bicycle-vehicle crashes: A random parameters approach with heterogeneity in means and variances. Under review in Analytic Methods in Accident Research.

Beige, S. & Axhausen, K. (2012). Interdependencies between turning points in life and long-term mobility decisions. *Transportation, 39*(4), 857-872.

Ben-Akiva, M.E. (1981). Issues in Transferring and Updating Travel-Behavior Models, in: New Horizons in Travel-Behavior Research. Lexington Books, pp. 665–682.

Ben-Akiva, M., & Boccara, B. (1995). Discrete choice models with latent choice sets. *International Journal of Research in Marketing, 12*(1), pp.9-24.

Ben-Akiva, M., Walker, J., Bernardino, A.T., Gopinath, D.A., Morikawa, T., & Polydoropoulou, A. (2002). Integration of Choice and Latent Variable Models, in: In Perpetual Motion. Elsevier, pp. 431–470.

Benoit, J.-P., & Dubra, J. (2011). Apparent overconfidence. *Econometrica, 79*(5), 1591-1625.

Benoit, J.-P., & Dubra, J. (2017). When do populations polarize? An explanation. Working Paper. London Business School, London, UK.

Berube, A., Deakin, E, & Raphael, S. (2006). Socioeconomic differences in household automobile ownership rules: Implications for evaluation policy. Report. Retrieved from: http://socrates.berkeley.edu/~raphael/BerubeDeakenRaphael.pdf (August 18, 2016)

Bhat, C. (2003). Simulation estimation of mixed discrete choice models using randomized and scrambled Halton sequences. *Transportation Research Part B, 37*(1), 837–855.

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

Bhat, C., & Dubey, S. (2014). A new estimation approach to integrate latent psychological constructs in choice modeling. *Transportation Research Part B, 67*, 68-85.

Bickel, W., Miller, M., Yi, R., Kowal, B., Lindquist, D., & Pitcock, J. (2007). Behavioral and neuroeconomics of drug addiction: Competing neural systems and temporal discounting processes. *Drug and Alcohol Dependence, 90S,* S85-91.

Bills, T., & Walkers, J. (2017). Looking beyond the mean for equity analysis: Examining distributional impacts of transportation improvements. *Transport Policy, 54*, 61-69.

Bliss, L. (2016). 3 Bumps on the Road Ahead for Shared Autonomous Vehicles. Retrieved from: https://www.citylab.com/transportation/2016/12/regulating-shared-autonomous-vehicles-on-city-streets/511544/ (November 10, 2017)

Blumenberg, E., Taylor, B.D., Smart, M.J., Ralph, K., Wander, M., & Brumbaugh, S. (2012). What's youth got to do with it? Exploring the travel behavior of teens and young adults (No. UCTC-FR-2012–2014). University of California Transportation Center, Berkeley, CA.

Bolduc, D., Ben-Akiva, M., Walker, J., & Michaud, A. (2005). Hybrid choice models with logit kernel: Applicability to large scale models. In: Lee-Gosselin, M., Doherty, S. (eds.) Integrated Land-Use and Transportation Models: Behavioural Foundations, Elsevier, Oxford, 275-302.

Boston. (2017). Autonomous Vehicles: Boston’s Approach . Retrieved from: https://www.boston.gov/departments/new-urban-mechanics/autonomous-vehicles-bostons-approach (November 10, 2017)

Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1997). Beyond bipolar conceptualizations and measures: The case of attitudes and evaluative space. *Personality and Social Psychology Review, 1(*1), 3-25.

Camerer, C., Loewenstein, G., & Prelec, D. (2005). Neuroeconomics: How neuroscience can inform economics. *Journal of Economic Literature, 43*(1), 9-64.

Carlsson, F. (2010). Design of stated preference surveys: Is there more to learn from behavioral economics? *Environmental Resource Economics, 46*(2), 167-177.

Chang, G., T. Lin, & Lindely, J. (1992). Understanding Suburban Commuting Characteristics: An Empirical Study in Suburban Dallas. *Transportation Planning and Technology, 16*(3), 167-193.

Chang, J., Lusk, J., & Norwood, F. (2009). How closely do hypothetical surveys and laboratory experiments predict field behavior. *American Journal of Agricultural Economics, 91*(2), 518-534.

Chang, L.-Y. (2005). Analysis of freeway accident frequencies: Negative binomial regression versus artificial neural network. *Safety Science, 43*(8), 2005, 541-557.

Chatterjee, K., Sherwin, H. & Jain, J. (2013). Triggers for changes in cycling: the role of life events and modifications to the external environment. *Journal of Transport Geography, 30*, 183-193.

Chong, M., Abraham, A., & Paprzycki, M. (2005). Traffic accident analysis using machine learning paradigms. *Informatica, 29*(1), 89-98.

Chorus, C.G., Arentze, T.A., & Timmermans, H.J.P. (2008). A Random Regret-Minimization model of travel choice. *Transportation Research Part B, 42*, 1–18.

Chui, T., Fang, D., Chen, J., Wang, Y., & Jeris, C. (2001). A robust and scalable clustering algorithm for mixed type of attributes in large database environment. In Proceedings of Seventh ACM SIGKDD International Conference on knowledge discovery and data mining, 263-268.

Circella, G., Fulton, L., Alemi, F., Berliner, R. M., Tiederman, K., Moktharian, P., & Handy, S. (2016). What affects millennials' mobility? Part I: Investigating the environmental concerns, lifestyles, mobility-related attitudes and adoption of technology of young adults in California, National Center for Sustainable Transportation Report, May 2016.

Clark, B., Chatterjee, K., & Melia, S. (2015). Changes in level of household car ownership: the role of life events and spatial context. Transportation 43(4), 565-599.

Constantinou, E., Panayiotou, G., Konstantinou, N., Loutsiou-Ladd, A., & Kapardis, A. (2011). Risky and aggressive driving in young adults: Personality matters. *Accident Analysis and Prevention, 43*(1), 1323-1331.

Daly, A., Hess, S., Patruni, B., Potoglou, D., & Rohr, C. (2012). Using ordered attitudinal indicators in a latent variable choice model: a study of the impact of security on rail travel behaviour. *Transportation, 39*, 267–297.

Dargay, J., & Hanly, M. (2007). Volatility of car ownership, commuting mode and time in the UK. *Transportation Research Part A, 41*(1), 934-948.

Dia, H., & Javanshour, F. (2017). Automated shared mobility-on-demand: Melbourne pilot simulation study simulation study. *Transportation Research Procedia, 22*, 285-296.

Delen, D., Sharda, R., & Bessonov, M. (2006). Identifying significant predictors of injury severity in traffic accidents using a series of artificial neural networks. *Accident Analysis and Prevention, 38*(3), 434-444.

Dumont, J., Keller, J., & Carpenter, C. (2015). RSGHB: functions for hierarchical Bayesian estimation: a flexible approach. R package version, 1(2).

Edison, S., & Geissler, G. (2003). Measuring attitudes towards general technology: Antecedents, hypotheses and scale development. *Journal of Targeting, Measurement and Analysis for Marketing*, *12*(2), 137-156.

Elliot, M., & Shaheen, S. (2011). The impact of carsharing on household vehicle ownership. Access 38 (Spring), 22-27.

Fagnant, D., & Kockelman, K. (2014). The travel and environmental implications of shared autonomous vehicles, using agent-based model scenarios. *Transportation Research Part C, 40*, 1-13.

Fagnant, D., & Kockelman, K. (2015a). Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations. *Transportation Research Part A, 77*, 167-181.

Fagnant, D., & Kockelman, K. (2015b). Dynamic ride-sharing and optimal fleet sizing for a system of shared autonomous vehicles. In Transportation Research Board 94th Annual Meeting Compendium of Papers, paper no. 15-1962.

Fagnant, D., Kockelman, D., & Bansal, P. (2015). Operations of a shared automated vehicle fleet for Austin, Texas. *Transportation Research Record, 2536*, 98-106.

Festinger, L. (1962). *Cognitive dissonance. Scientific American, 207*(4), 93-107.

Firnkorn, J., & Muller, M. (2012). Selling mobility instead of cars: new business strategies of automakers and the impact on private vehicle holding. *Business Strategy and the Environment, 21(*4), 264–280.

Firnkorn, J., & Muller, M. (2015). Free-floating electric car sharing-fleets in smart cities: the dawning of a post-private car era in urban environments? *Environmental Science and Policy, 45*, 30–40.

Fisher, F., & Shell, K. (1969). Taste and quality change in the pure theory of the true cost-of-living index. In “Value, Capital and Growth”, J. N. Wolfe (Ed.), Edinburgh University Press, Edinburgh, UK.

Fishhoff, B., Bostrom, A., & Quadrel, M. (1993). Risk perception and communication. *Annual Review of Public Health, 14,* 183-203.

Fry, R. (2016). Millennials overtake baby boomers as America’s largest generation. Retrieved from: http://www.pewresearch.org/fact-tank/2016/04/25/millennials-overtake-baby-boomers/ (August 18, 2016)

Glasman, L., & Albarracib, D. (2006). Forming attitudes that predict future behavior: A meta-analysis of the attitude-behavior relation. *Psychological Bulletin*, *132*(5), 778-822.

Green, L., & Myerson, J., (2004). A discounting framework for choice with delayed and probabilistic awards. *Psychological Bulletin, 130*(5), 769-792.

Greenblatt J., & Saxena, S. (2015). Automated taxis could greatly reduce greenhouse gas emissions of U.S. light-duty vehicles. Nature Climate Change, Letters, July.

Greenblatt, J., & Shaheen, S. (2015). Automated Vehicles, On-Demand Mobility, and Environmental Impacts. *Current Sustainable Renewable Energy Reports, 2*, 74–81.

Greene, W. (1997). Econometric Analysis, third edition Macmillan, New York. NY.

Gorman, W. (1967). Tastes, habits, and choices. *International Economic Review, 8*(2), 218-222.

Guo, Y., Peeta, S., & Mannering, F. (2016). Rail-truck multimodal freight collaboration: a statistical analysis of freight-shipper perspectives. *Transportation Planning and Technology, 39(*5), 484-506.

Guiso, L., Sapienza, P., & Luigi, G. (2013). Time varying risk aversion. Working paper 19284, National Bureau of Economic Research, Cambridge, MA.

Haboucha, C.J., Ishaq, R., & Shiftan, Y. (2015). User preferences regarding autonomous vehicles: giving up your private car. In: IATBR 2015 – WINDSOR. Presented at the IATBR 2015, Windsor, U.K.

Haboucha, C., Ishaq, R., & Shiftan, Y. (2017). User preferences regarding autonomous vehicles. *Transportation Research Part C, 78*, 37-49.

Halton, J. (1960). On the efficiency of evaluating certain quasi-random sequences of points in evaluating multi-dimensional integrals. *Numerische Mathematik*, *2*(1), 84-90.

Hansen, S. (1981). In Favor of Cross-Cultural Transferability of Travel-Demand Models, in: New Horizons in Travel-Behavior Research. Lexington Books, pp. 637–648.

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

Hawkins, A.J. (2017). Las Vegas is expanding its self-driving shuttle experiment. Retrieved from: https://www.theverge.com/2017/11/6/16614388/las-vegas-self-driving-shuttle-navya-keolis-aaa (November 10, 2017)

Heffner, R., Kurani, K., Turrentine, T. (2007). Symbolism in California's early market for hybrid electric vehicles. *Transportation Research Part D, 12*(6), 396-413.

Hess, S., Stathopoulos, A., Daly, A. (2012). Allowing for heterogeneous decision rules in discrete choice models: an approach and four case studies. *Transportation, 39*, 565–591.

Hildebrand, E. (2003). Dimensions in elderly travel behavior: A simplified activity-based model using lifestyle clusters. *Transportation, 30*(3), 285-306.

Jaynes, N. (2016). How driverless cars will change car ownership forever. Retrieved from: http://mashable.com/2016/05/30/car-ownership-autonomy-column/#ZhXSbHx_WqqW (August 29, 2016)

Kahneman, D., Slovic, P., & Tversky, A. (editors) (1982). Judgment under uncertainty: Heuritics and biases. Cambridge University Press, New York, NY.

Kalton, G., Roberts, J., & Holt, D. (1980). The effects of offering a middle response option with opinion questions. The Statistician, 65-78.

Kodransky, M., & Lewenstein M. (2014). Connecting low-income people to opportunity with shared mobility. Institute for Transportation and Development Policy.

Koppelman, F.S., & Wilmot, C.G. (1986). The effect of omission of variables on choice model transferability. *Transportation Research Part B, 20*, 205–213. https://doi.org/10.1016/0191-2615(86)90017-2

Kosoff, M. (2016). Lyft's president says car ownership will "all but end" by 2025. Retrieved from: http://www.vanityfair.com/news/2016/09/lyfts-president-says-car-ownership-will-all-but-end-by-2025 (September 22, 2016)

Krosnick, J.A., Holbrook, A.L., Berent, M.K., Carson, R.T., Hanemann, W.M., Kopp, R.J., Mitchell, R.C., Presser, S., Ruud, P.A., Smith, V.K., & Moody, W.R. (2002). The impact of "no opinion" response options on data quality: non-attitude reduction or an invitation to satisfice?. *Public Opinion Quarterly*, *66*(3), pp.371-403.

Krueger, R., Rashidi, T., & Rose, J. (2016). Preferences for shared autonomous vehicles. *Transportation Research Part C, 69*, 343-355.

Kuhnimhof, T., Zumkeller, D., & Chlond, B. (2013). Who made peak car, and how? A breakdown of trends over four decades in four countries. *Transport Reviews*, *33*(3), 325-342.

Kulas, J.T., Stachowski, A.A., & Haynes, B.A. (2008). Middle response functioning in Likert-responses to personality items. *Journal of Business and Psychology*, *22*(3), pp.251-259.

Lavieri, P., Garikapati, V., Bhat, C., Pendyala, R., Astroza, S., & Dias, F. (2017). Modeling individual preferences for ownership and sharing of automated vehicle technologies. Submitted to Transportations Research Board (No. 17-05843).

Lerman, S. (1981). A Comment on Interspatial, Intraspatial, and Temporal Transferability, in: New Horizons in Travel-Behavior Research. pp. 628–632.

Le Vine, S., Zolfaghari, A., & Polak, J. (2015). Autonomous cars: The tension between occupant experience and intersection capacity. *Transportation Research Part C, 52*, 1-14.

Li, X., Lord, D., Zhang, Y., & Xie, Y. (2008). Predicting motor vehicle crashes using support vector machine models. *Accident Analysis and Prevention, 40*(4), 1611-1618.

Lieberman, M., Gaunt, R., Gilbert, D., & Trope, Y. (2002). Reflection and reflexion: A social cognitive neuroscience approach to attributional inference. In "Advances in experimental psychology" M. Zanna (Ed.), Academic Press, New York, NY, 199-249.

Litman, T. (2018). Automated Vehicle Implementation Predictions: Implications for Transport Planning. Victoria Transport Policy Institute, Victoria, BC.

Lluch, C. (1974). Expenditure, savings, and habit formation. *International Economic Review, 15*(3), 786-797.

Lord, C., Ross, L., & Lepper, M. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology, 37*(11), 2098-2109.

Lord, D., & Mannering, F. (2010). The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. *Transportation Research Part A, 44*(5), 291-305.

Louviere, J.J. (1981). Some comments on premature expectations regarding spatial, temporal, and cultural transferability of travel-choice models, in: New Horizons in Travel-Behavior Research. Lexington Books, pp. 653–663.

Maheshri, V., & Winston, C. (2016). Did the great depression keep bad drivers off the road? *Journal of Risk and Uncertainty, 52*(3), 255-280.

Malmendier, U., & Nagel, S. (2011). Depression babies: Do macroeconomic experiences affect risk-taking? *Quarterly Journal of Economics, 126*(1), 373-416.

Malyshkina, N., & Mannering, F. (2008). Effect of increases in speed limit on severities of injuries in accidents. *Transportation Research Record, 2083*, 122-127.

Malyshkina, N., & Mannering, F. (2009). Markov switching multinomial logit model: An application to accident-injury severities. *Accident Analysis and Prevention, 41*(4), 829-838.

Malyshkina, N., & Mannering, F. (2010). Zero-state Markov switching count-data models: An empirical assessment. *Accident Analysis and Prevention, 42*(1), 122-130.

Malyshkina, N., Mannering, F., & Tarko, A. (2009). Markov switching negative binomial models: An application to vehicle accident frequencies. *Accident Analysis and Prevention, 41*(2), 217-226.

Maness, M., & Cirillo, C. (2016). An indirect latent informational conformity social influence choice model: Formulation and case study. *Transportation Research Part B, 93*, 75-101.

Maness, M., Cirillo, C., & Dugundji, E. (2015). Generalizable behavioral framework for choice models of social influence: Behavioral and data concerns in travel behavior. *Journal of Transport Geography, 46*, 137-150.

Mannering, F. (2017). Cross-sectional modelling. Forthcoming in "Safe mobility: Methodology and crash data analysis", D. Lord, S. Washington (Eds.), Emerald Group Publishing, Binley, West Yorkshire, UK.

Mannering, F. (2018). Temporal instability and the analysis of highway accident data. *Analytic Methods in Accident Research, 17,* 1-13.

Mannering, F., & Bhat, C. (2014). Analytic methods in accident research: Methodological frontier and future directions. *Analytic Methods in Accident Research, 1*, 1-22.

Mannering, F., Murakami, E., & Kim, S.-G. (1994). Temporal stability of travelers' activity choice and home-stay duration: Some empirical evidence. *Transportation, 21*(4), 371-392.

Mannering, F., Shankar, V., & Bhat, C. (2016). Unobserved heterogeneity and the statistical analysis of highway accident data. *Analytic Methods in Accident Research, 11,* 1-16.

Mannering, F., & Winston, C. (1985). A dynamic empirical analysis of household vehicle ownership and utilization. *Rand Journal of Economics, 16*(2), 215-236.

Mannering, F., & Winston C. (1995). Automobile air bags in the 1990's: Market failure or market efficiency? *Journal of Law and Economics, 38*(2), 265-279.

Mannering, F., Winston, C., & Starkey, W. (2002). An exploratory analysis of automobile leasing by US households. *Journal of Urban Economics*, *52*(1), 154-176.

Martin E., & Shaheen S. (2011). Greenhouse gas emissions impacts of carsharing in North America. *IEEE Transactions on Intelligent Transportation Systems, 12*(4),1074-1086.

Martin E., & Shaheen S. (2011). The impact of carsharing on public transit and non-motorized travel: an exploration of North American carsharing survey data. *Energies*, *4*(11), 2094 - 2114.

Martin, E., Shaheen, S.A., & Lidicker, J. (2010). Impact of carsharing on household vehicle holdings. *Transportation Research Record, the Journal of the Transportation Research Board, 2143,* 150-158.

McClure, S., Laibson, D., Loewenstein, G., & Cohen, J. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science*, *306*(5695), 503-507.

McDonald, N. (2015). Are millennials really the "go-nowhere" generation? *Journal of the American Planning Association, 81*(2), 90-103.

McFadden, D. (1981). Econometric Models for Probabilistic Choice. Structural Analysis of Discrete Data Using Econometric Applicatios. MIT Press, Cambridge, MA.

McFadden, D. (1986). The choice theory approach to market research. *Marketing Science, 5*, 275–297.

McFadden, D. (2013). The new science of pleasure: Consumer choice behavior and the measurement of well-Being, Handbook of Choice Modelling, S. Hess and A. J. Daly, (eds.), Edward Elgar. http://elsa.berkeley.edu/wp/mcfadden122812.pdf

McFadden, D., & Ruud, P. (1994). Estimation by simulation. *The Review of Economics and Statistics, 76*(4), 591-608.

McFadden, D., & Train, K. (2000). Mixed MNL models for discrete response. *Journal of Applied Econometrics, 15*, 447–470.

McKenzie, B. (2015). Who drives to work? Commuting by automobile in the United States: 2013. Retrieved from: http://www.census.gov/hhes/commuting/files/2014/acs-32.pdf (August 18, 2016)

Menon, N., Barbour, N., Zhang, Y., Pinjari, A.R., & Mannering, F. (2018). Shared autonomous vehicles and their potential impacts on household vehicle ownership: An exploratory empirical assessment. *International Journal of Sustainable Transportation,* pp.1-12.

Menon N., Pinjari, A.R., Zhang Y., & Zou L. (2016, January). Consumer perception and intended adoption of autonomous vehicle technology: Findings from a university population survey. Presented at the Transportation Research Board's 95th Annual Meeting, Washington DC.

Metz, D. (2013). Peak car and beyond: the fourth era of travel. Transport Reviews 33(3), 255-270.

Meyer, G., & Shaheen, S. (2017). Disrupting Mobility: Impacts of Sharing Economy and Innovative Transportation on Cities. Springer. ISBN 978-3-319-51602-8 (e-Book)

Millard-Ball, A., & Schipper, L. (2011). Are we reaching peak travel? Trends in passenger transport in eight industrialized countries. *Transport Reviews, 31*(3), 357-378.

Milton, J., Shankar, V., & Mannering, F. (2008). Highway accident severities and the mixed logit model: An exploratory empirical analysis. *Accident Analysis and Prevention, 40*(1), 260-266.

Mokhtarian, P. (2002). Telecommunications and travel: the case for complementarity. *Journal of Industrial Ecology, 6(*2), 43-57.

Mokhtarian, P. (2009). If telecommunication is such a good substitute for travel, why does congestion continue to get worse? *Transportation Letters, 1*(1), 1-17.

Moore, G. (2002). Crossing the chasm: Marketing and selling disruptive products to mainstream customers. New York: Harper Business.

Motion Digest. (2017). EasyMile Receives NHTSA Approval to Test Shared Autonomous Vehicles on Public Roads. Retrieved from: https://motiondigest.com/2017/10/13/easymile-receives-nhtsa-approval-to-test-shared-autonomous-vehicles-on-public-roads/ (November 10, 2017)

National Academies. (2013). Frontiers in massive data analysis. The National Academies Press, Washington, DC.

Ng, L., Barfield, W. & Mannering, F. (1998). Analysis of private drivers' commuting and commercial drivers' work-related travel behavior. *Transportation Research Record, 1621*, 50-60.

Nobis, C. (2006). Carsharing as key contribution to multimodal and sustainable mobility behavior. *Transportation Research Record, the Journal of the Transportation Research Board, 1986*, 89-97.

Nowlis, S.M., Kahn, B.E., & Dhar, R. (2002). Coping with ambivalence: The effect of removing a neutral option on consumer attitude and preference judgments. *Journal of Consumer Research, 29*(3), 319-334.

Oakil, A., Ettema, D., Arentze, T., & Timmermans, H., (2014). Changing household car ownership level and life cycle events: An action in anticipation or an action on occurrence. *Transportation, 41(*4), 889-904.

Ohta, H., Fujii, S., Nishimura, Y., & Kozuka, M. (2013). Analysis of the acceptance of carsharing and eco-cars in Japan. *International Journal of Sustainable Transportation, 7(*6), 449-467.

Persaud, B., Lan, B., Lyon, C., & Bhim, R. (2010). Comparison of empirical Bayes and full Bayes approaches for before–after road safety evaluations. *Accident Analysis and Prevention, 42(*1), 38-43.

Peston, M. (1967). Changing utility functions. In "Essays in mathematical economics", M. Shubik (Ed.), Princeton University Press, Princeton, NJ.

Peterman, D. (2013). Federal traffic safety programs: An overview, congressional research service report for congress. Washington, DC.

Peters, J., & Buchel, C. (2011). The neural mechanisms of intertemporal decision-making: Understanding variability. *Trends in Cognitive Science, 15*(5), 227-239.

Phlips, L. (1972). A dynamic version of the linear expenditure model. *The Review of Economics and Statistics, 54*(4), 450-458.

Pinjari, A., Eluru, N., Bhat, C., Pendyala, R.M., & Spissu, E. (2008). Joint model of choice of residential neighborhood and bicycle ownership: accounting for self-selection and unobserved heterogeneity. *Transportation Research Record: Journal of the Transportation Research Board, 2082*, 17-26.

Pollak, R. (1970). Habit formation and dynamic demand functions. *Journal of Political Economy, 78(*4), 745-763.

Pollak, R., & Wales, T. (1969). Estimation of the linear expenditure system. *Econometrica, 37*(4), 611-628.

Polzin, S., Chu, X., & Godfrey, J. (2014). The impact of millennials' travel behavior on future personal vehicle travel. *Energy Strategy Reviews, 5,* 59-65.

Rashidi, T., Mohammadian, A., & Koppelman, F. (2011). Modeling interdependencies between vehicle transaction, residential relocation and job change. *Transportation, 38*, 909-932.

Riviere, C., Lauret, P., Ramsamy, J., & Page, Y. (2006). A Bayesian neural network approach to estimating the energy equivalent speed. *Accident Analysis and Prevention, 38*(2), 248-259.\

Rosenbloom, S. (2001). Sustainability and automobility among the elderly: An international assessment. *Transportation*, *28*(4), 375-408.

Rossi, T., & Bhat, C. (2014). Guide for travel model transfer. Federal Highway Administration, United States Department of Transportation, FHWA-HEP-15-006, Washington, DC.

Ruhm, C. (2000). Are recessions good for your health? *Quarterly Journal of Economics, 115*(2), 617-650.

Santoso, D.S., & Tsunokawa, K. (2005). Spatial transferability and updating analysis of mode choice models in developing countries. *Transportation Planning and Technology, 28,* 341–358.

Savolainen, P., Mannering, F., Lord, D., & Quddus, M. (2011). The statistical analysis of crash-injury severities: A review and assessment of methodological alternatives. *Accident Analysis and Prevention, 43*(5), 1666-1676.

Semega, J., Fontenot, K., & Kollar, M. (2017). Income and Poverty in the United States: 2016. Current Population Report. U. S. Department of Commerce. Retrieved from: https://www.census.gov/ (March, 24th, 2018)

Seraneeprakarn, P., Huang, S., Shankar, V., Mannering, F., Venkataraman, N., & Milton, J. (2017). Occupant injury severities in hybrid-vehicle involved crashes: A random parameters approach with heterogeneity in means and variances. *Analytic Methods in Accident Research, 15*, 41-55.

Schoettle, B., & Sivak, M. (2014). A survey of public opinion about autonomous and self-driving vehicles in the US, the UK, and Australia. University of Michigan, Technical Report No.UMTRI-2014-21. Retrieved from: http://deepblue.lib.umich.edu/bitstream/handle/2027.42/108384/103024.pdf?sequence=1andisAllowed=y (August 13, 2014).

Seraneeprakarn, P., Huang, S., Shankar, V., Mannering, F., Venkataraman, N., & Milton, J. (2017). Occupant injury severities in hybrid-vehicle involved crashes: A random parameters approach with heterogeneity in means and variances. *Analytic Methods in Accident Research, 15*, 41-55.

Shaheen, S., & Cohen, A., (2013). Carsharing and personal vehicle services: worldwide market developments and emerging trends. *International Journal of Sustainable Transportation, 7*(1), 5-34.

Shaheen S., Cohen A., & Zohdy, I. (2016). Shared Mobility: Current Practices and Guiding Principles. U.S. Department of Transportation. Report No. FHWA-HOP-16-022

Shaheen, S., Cano, L., & Camel, M. (2016). Exploring electric vehicle carsharing as a mobility option for older adults: A case study of a senior adult community in the San Francisco Bay Area. *International Journal of Sustainable Transportation, 10*(5), 406-417.

Shaheen, S., Rodier, C., Murray, G., Cohen, A., & Martin, E. (2010). Carsharing and public parking policies: assessing benefits, costs, and best practices in North America. Report CA-MTI-10-2612, Mineta Transportation Institute.

Shen, J. (2009). Latent class model or mixed logit model? A comparison by transport mode choice data. *Applied Economics, 41* (22), 2915-2924.

Sheela, P., & Mannering, F. (2018). The effect of information on changing opinions toward autonomous vehicle adoption: An exploratory analysis. Working paper. College of Engineering, University of South Florida, Tampa, FL.

Sikder, S., Augustin, B., Pinjari, A., & Eluru, N. (2014). Spatial transferability of tour-based time-of-day choice models. *Transportation Research Record: Journal of the Transportation Research Board*, *2429*, 99–109.

Sikder, S., & Pinjari, A.R.(2012). Immobility levels and mobility preferences among elderly in the United States: Evidence from the 2009 national household travel survey. *Transportation Research Record: Journal of the Transportation Research Board*, *2318*, 137-147.

Smith, A. (2011). Trends in cell phone usage and ownership. Retrieved from: http://www.pewinternet.org/2011/04/28/trends-in-cell-phone-usage-and-ownership/ (August 18, 2016)

Smith, A. (2013). Smartphone ownership 2013. Retrieved from: http://www.pewinternet.org/2013/06/05/smartphone-ownership-2013/ (August 18, 2016)

Spieser, K., Ballantyne, K., Treleaven, K., Zhang, R., Frazzoli, E., Morton, D., & Pavone, M. (2014). Toward a systematic approach to the design and evaluation of automated mobility-on-demand systems: A case study in Singapore. In G. Meyer and S. Beiker (Eds.), Road Vehicle Automation: Lecture Notes in Mobility (pp. 229-245). Switzerland: Springer International Publishing.

Stocker, A., & Shaheen, S. (2017). Shared Automated Vehicles: Review of Business Models. Transportation Sustainability Research Center, University of California, Berkeley, Discussion Paper 9.

Sturgis, P., Roberts, C., & Smith, P. (2014). Middle alternatives revisited: how the neither/nor response acts as a way of saying "I don't know"?. *Sociological Methods and Research, 43*(1), 15-38.

Sumer, N. (2003). Personality and behavioral predictors of traffic accidents: testing a contextual mediated model. *Accident Analysis and Prevention*, *35*(1), 949-964.

Svenson, O. (1981). Are we all less risky and more skillful than our fellow drivers? *Acta Psychologica, 94*(2), 143-148.

Tardiff, T.J. (1979). Specification analysis for quantal choice models. *Transportation Science, 13*, 179–190.

Train, K.E. (2009). Discrete choice methods with simulation. Cambridge university press.

Train, K.E., McFadden, D.L., & Goett, A.A. (1987). Consumer attitudes and voluntary rate schedules for public utilities. *The Review of Economics and Statistics*, *69*, 383–383.

Transportation Research Board. (2013). Critical issues in transportation. Retrieved from: http://onlinepubs.trb.org/Onlinepubs/general/criticalissues13.pdf (March 21, 2017)

Tversky, A. (1969). Intransitivity of preferences. *Psychological Review*, *76*, 31–48.

Tversky, A. (1972). Choice by elimination. *Journal of Mathematical Psychology, 9*, 341–367.

Tversky, A., & Kahneman, D. (1974). Judgement under uncertainty: Heuristics and biases. *Science, 185(*4157), 1124-1131.

Ulleberg, P., & Rundmo, T. (2003). Personality, attitudes and risk perception as predictors of risky driving behavior among young drivers. *Safety Science, 41*(5), 427-443.

van Wee, B. (2015). Peak car: The first signs of a shift towards ICT-based activities replacing travel? A discussion paper. *Transport Policy, 42*, 1-3.

Vij, A., & Walker, J. L. (2016). How, when and why integrated choice and latent variable models are latently useful. *Transportation Research Part B, 90*, 192-217.

Wang, F., Yang, M., & Yang, R. (2006). Dynamic fleet management for cybercars. Proceedings of the IEEE Intelligent Transportation Systems Conference, Toronto, Canada.

Washington, S., Karlaftis, M., & Mannering, F. (2011). Statistical and econometric methods for transportation data analysis. Second edition. Chapman and Hall/CRC, Boca Raton, FL.

Winston, C., Maheshri, V., & Mannering, F. (2006). An exploration of the offset hypothesis using disaggregate data: The case of airbags and antilock brakes. *Journal of Risk and Uncertainty. 32*(2), 83-99.

World Health Organization. (2015). Global status report on road safety 2015. World Health Organization, Geneva, Switzerland.

Xie, Y., Lord, D., & Zhang, Y. (2007). Predicting motor vehicle collisions using Bayesian neural networks: An empirical analysis. *Accident Analysis and Prevention, 39*(5), 922-933.

Xiong, Y., Tobias, J., & Mannering, F. (2014). The analysis of vehicle crash injury-severity data: A Markov switching approach with road-segment heterogeneity. *Transportation Research Part B, 67*, 109-128.

Yu, R., & Abdel-Aty, M. (2013). Utilizing support vector machine in real-time crash evaluation. *Accident Analysis and Prevention, 51*, 252-259.

Zhang, J., Yu, B., & Chikaraishi, M. (2014). Interdependences between household residential and car ownership behaviour: a life history analysis. *Journal of Transport Geography, 34*, 165-174.

Zhang, W., & Guhathakurta, S. (2017). Parking spaces in the age of shared automated vehicles: How much parking will we need and where? Transportation Research Board Conference Paper, Washington, DC.