Examining Preference Heterogeneity in Adoption of Emerging Transportation Technologies

A PhD Dissertation Submitted to the
Department of Engineering Systems and Environment
University of Virginia

In Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy

By
Wenjian Jia

Committee Members:
Chair: Dr. Brian L. Smith (ESE)
Advisor: Dr. T. Donna Chen (ESE)
Dr. Michael Albert (ESE)
Dr. John Miller (ESE)
Dr. Andrew Mondschein (Urban & Environmental Planning)
Dr. Wenwen Zhang (Planning & Public Policy, Rutgers University)

August 2021

Acknowledgement

I cannot believe it has been four years since I came to UVa. Now it is close to the end of this journey. Sitting in front of my laptop, lots of people come into mind. I would like to thank them for shining my four years in the U.S.

Thank Mid-Atlantic Transportation Sustainability University Center (MATS UTC) and Jeffress Trust for funding my PhD research. Thank my dissertation committee members, Dr. Michael Albert, Dr. John Miller, Dr. Andrew Mondschein, Dr. Brian L. Smith, and Dr. Wenwen Zhang for their feedback and comments on my dissertation.

I am deeply grateful to my advisor, Prof. T. Donna Chen. Prof. Chen gives me support on every aspect from work to life: word-by-word edits of my papers, suggestions on conference presentations, insights on research directions, encouragement when research fails, suggestions on future career, and care during both American and Chinese festivals. She is a role model for my incoming research and teaching career. Special thanks to Prof. Wenewen Zhang for her guidance in our collaborative research and her help in my job searching. It is a great experience to work with her and hope we can keep collaborating in the future.

Thank all research mates in Dr. Chen's group. Their friendships make my life in a foreign country less daunting. I appreciate all the experiences here: Thanksgiving dinner, Texas Roadhouse, Blue Ridge Mountain, Amtrak, and watching the 2020 presidential election, etc., all of these are new to me and help me understand this country.

Thank my parents. No words can describe how great they are. Hope they can stay healthy and happy. Thank my two sisters and two little nephews. All the tiredness disappears when seeing their cute faces. Thank the love of my life, Qian Wang. "海上升明月，天涯共此时". With her, the distance of the Pacific Ocean and 12-hour time lag become negligible.

Abstract

The transportation system is experiencing disruptive changes with the development of emerging transportation technologies like ride-hailing, electric vehicles (EVs), and automated vehicles (AVs). These technologies are double-edged swords, whose impacts on the transportation system, urban forms, energy consumption, and emissions are highly uncertain. To start to quantify the uncertainty, it is critical to study consumers' adoption preferences for these emerging technologies, which informs policy opportunities to support sustainable outcomes in deployment. This dissertation examines consumer preferences for the three innovative technologies in transportation (ride-hailing, EVs, and AVs), with a particular focus on preference heterogeneity.

Preference heterogeneity has been well studied in the research on the traditional transportation system. With the disruption of emerging technologies, recent studies highlight the importance of preference heterogeneity in the adoption of new technologies. This dissertation contributes to the existing literature in the following three aspects: 1) examine the disparity in ride-hailing usage under various spatial contexts, 2) examine the heterogeneous preferences for EVs using advanced discrete choice models which allow for random preference heterogeneity, and 3) examine the heterogeneity in AV mode choice preferences that can be linked to latent attitudinal constructs.

First, based on the 2017 National Household Travel Survey data, the ride-hailing analysis shows the disparities in ride-hailing usage. For example, seniors (compared to younger) and low-income (compared to high-income) travelers are less likely to use ride-hailing services. Such disparities between age groups and between income groups widen in urban and rural areas, respectively. Moreover, ride-hailing services are found to fill mobility gaps for non-vehicle owners from public transport desert communities. Findings provide insights for future ride-hailing research to consider the interplay between socio-economic characteristics and spatial contexts, rather than examining these two elements independently.

Second, the EV preference heterogeneity analysis develops mixed logit (MXL), latent class (LC), and latent class-mixed logit (LC-MXL) models based on stated choice experiments data collected in Virginia in 2018 (n = 837). Model results suggest that monetary incentives are the most effective in increasing EV market share, followed by deploying more charging infrastructure, while improvement in battery range is found to be least effective. Moreover, the comparison across the three statistical models shows that no one model is unanimously superior to the other models in uncovering consumer preference heterogeneity in EV adoption. Rather, altogether they provide a more comprehensive picture of the complex EV preference structure.

Third, AV mode choice preferences are examined using the integrated choice and latent variable (ICLV) model based on stated preference surveys distributed in the Seattle (n = 511) and Kansas City regions (n = 558) in 2020 and 2021. Model results suggest the importance of latent attitudes (e.g., attitudes towards AV technology, willingness to share travel with strangers) and mode-specific attributes (e.g., trip cost, trip time) in explaining AV mode choice outcomes.

Additionally, the sensitivity to in-vehicle travel time in private AVs can be significantly lower than in human-driven vehicles, suggesting the potential of induced vehicle miles traveled in the AV era.

In sum, this dissertation investigates three distinct aspects of preference heterogeneity (observed, unobserved, and latent attitudes) in the adoption of three emerging transportation technologies (ride-hailing, EVs, and AVs). Findings provide technology-specific policy insights for sustainable deployment. Moreover, as the three technologies are in different stages of adoption, insights from preferences for more mature technologies have implications for the newer technologies. For example, the findings of the ride-hailing analysis have implications for future studies on the deployment of AVs once real-world usage data is available. The potential spatial heterogeneity in using shared AVs and shared AVs with pooling modes should be explicitly considered in AV policy-making, ensuring that the benefits of new technologies can be shared by all groups of people. Lastly, considering that the transportation system keeps evolving with the introduction of new technologies, the study framework in this dissertation can also apply to future research on other emerging transportation technologies beyond ride-hailing, vehicle electrification, and automation.

Table of Contents

Acknowledgement 2

Abstract 3

Chapter 1. Introduction 10

1.1 Background 10

1.2 Trends of shard mobility, vehicle electrification, and vehicle automation 10

1.3 Modeling approaches for incorporating preference heterogeneity 11

1.4 Research objectives 12

1.5 Organization of dissertation 13

Chapter 2. Disparities in Ride-hailing Usage Under Different Spatial Contexts 15

2.1 Introduction 16

2.2 Literature review 16

2.2.1 Individual socio-economic characteristics 16

2.2.2 Spatial context variables 17

2.2.3 The interaction of individual socio-economic variables and spatial contexts 18

2.3 Data and methods 19

2.3.1 Data 19

2.3.2 Methods 22

2.4 Results and discussion 23

2.4.1 Age 23

2.4.2 Household income 25

2.4.3 Gender 27

2.4.4 Household vehicle ownership 29

2.4.5 Driving licensure 31

2.5 Conclusions and future work 32

Chapter 3. Investigating Heterogeneous Preferences for Electric Vehicles 44

3.1 Introduction 45

3.2 Literature review 46

3.3 Data and method 51

3.3.1 Survey 51

3.3.2 Method 55

3.4 Model results 56
3.4.1 MXL model 56
3.4.2 LC and LC-MXL model 60
3.5 Discussion 64
3.5.1 Goodness-of-fit and cross validation 64
3.5.2 Willingness-to-Pay (WTP) comparison 66
3.5.3 Policy scenario analysis 70
3.6 Conclusions 72
Chapter 4. Local Context Matters: Examining Mode Choice Preferences in the Autonomous Vehicle Era in Two U.S. Metropolitan Regions 78
4.1 Introduction 79
4.2 Literature review 80
4.2.1 Individual socio-economic characteristics 80
4.2.2 Attitudes 81
4.2.3 Mode-specific attributes 82
4.3 Data 83
4.3.1 Survey overview 83
4.3.2 Choice experiments design 84
4.3.3 Survey distribution 86
4.4 Method 91
4.4.1 Latent variable component 91
4.4.2 Discrete choice model component 92
4.4.3 Measurement model 93
4.4.4 Joint model likelihood function 93
4.5 Results and discussion 95
4.5.1 Mode-specific attributes 95
4.5.2 Attitudes 96
4.5.3 Individual socio-economic characteristics 96
4.5.4 Trip characteristics 97
4.6 Conclusion and limitations 100
Chapter 5. Conclusions 109

5.1 Summary of main findings.. 109
5.2 Contributions... 110
5.3 Limitations and future research 111
Publications and Presentations.. 113

List of Figures

Figure 2 - 1. Sample distribution 21
Figure 2 - 2. Ride-hailing usage odds ratio (senior vs. younger travelers) from different state samples.. 24
Figure 2 - 3. Ride-hailing usage odds ratio (low- vs. high-income travelers) 26
Figure 2 - 4. Ride-hailing usage odds ratio (female vs. male) 28
Figure 2 - 5. Ride-hailing usage odds ratio (travelers from households without vehicles vs. with vehicles) 30
Figure 2 - 6. Ride-hailing usage odds ratio (non-drivers vs. drivers) 31

Figure 3 - 1. Typical choice experiment for mid/full size cars 52
Figure 3 - 2. Cross validation results 65
Figure 3 - 3. Marginal WTP for battery range evaluated at 200 miles. 67
Figure 3 - 4. WTP for (a) DC fast charging stations spaced every 40 miles along interstate highways (relative to a 100-mile spacing). (b) workplace charging. (c) local public charging stations available at one in three local destinations (relative to no local public charging stations availability). (d) $1000 incentive 69
Figure 3 - 5. Fuel type market share simulation. (a) Overall market share based on different models. (b) market share for the three class from the LC model. (c) market share for the two classes in the LC-MXL model. S1: BEV battery range increases by 50%; S2: Greater access to charging infrastructure; S3: $7500 incentive for EVs. 71

Figure 4 - 1. An example of travel mode choice experiment.. 86
Figure 4 - 2. Spatial distribution of the two samples. 88
Figure 4 - 3. Distribution of responses to the attitudinal questions 90
Figure 4 - 4. ICLV model framework for modeling AV mode choice preferences.................... 94

List of Tables

Table 2 - 1. Descriptive statistics of the explanatory variables and response variables. 21

Table 3 - 1. EV Preference Studies Using Mixed Logit (MXL) or Latent Class (LC) Models ... 49
Table 3 - 2. Attribute levels for the mid/full sized car body type 53
Table 3 - 3. Comparison between survey sample and general population 54
Table 3 - 4. Mixed logit model (MXL) results 58
Table 3 - 5. Determining the number of classes for the LC model. 60
Table 3 - 6. Estimation results for LC and LC-MXL model 62

Table 4 - 1. Attribute levels for choice experiment design. 85
Table 4 - 2. Descriptive statistics of sample in the Seattle and Kansas City regions.................. 89
Table 4 - 3. Attitudinal statements 90
Table 4 - 4. ICLV model results for the Seattle region 98
Table 4 - 5. ICLV model results for the Kansas City region 99

Chapter 1. Introduction

1.1 Background

In the late 19 century, the invention of internal combustion engines, which are powered by fossil fuels, marked the paradigm shift in road transport from the carriage era to the automobile era. In 1908, the introduction of the Ford Model T made the automobile no longer a luxury, and thus the mass adoption of automobiles became a reality. Together with the construction of the Interstate Highway System, the U.S. became a nation on wheels, which shaped the current landscape of transportation systems and urban forms. However, the car-oriented culture also brings lots of negative impacts, such as traffic crashes, energy security challenges, greenhouse gas emissions, air pollution, traffic noise, etc.

Today we are at the cusp of another paradigm shift, as represented by the three revolutions in transportation (shared mobility, electric vehicles, and automated vehicles) (Sperling, 2018). The three innovative technologies bring both opportunities and challenges to the transportation system. The debate on the “dream” or “nightmare” outcomes of these technologies suggests high uncertainty of their impacts. Some predict “dream” outcomes where automated vehicles (AVs) will be shared and powered by electricity, mitigating traffic congestion, reducing parking demand, and achieving local and national emission reduction goals. In contrast, “nightmare” scenarios suggest that AVs will be mainly privately-owned and powered by fossil fuels, which will exacerbate congestions, energy use and emissions (Rouse, 2019). The uncertainty in impacts is largely dependent on whether and how consumers adopt the technologies. Thus, understanding consumer adoption preferences lays the foundation for better planning for sustainable outcomes in the deployment of these emerging transportation technologies.

1.2 Trends of shard mobility, vehicle electrification, and vehicle automation

Shared mobility services have rocketed in popularity with the rise of the shared economy and advances in information communication technologies. Taking Uber, for example, which had its first-ever trip in 2010 in San Francisco, has since expanded to more than 10,000 cities worldwide (Uber, 2021). According to the 2017 National Household Travel Survey in the U.S., about 7% of adult Americans have used ride-hailing services in the month prior to taking the survey. As an alternative travel mode, ride-hailing services have the potential to provide mobility options for the traditionally mobility-disadvantaged groups. Even more interest lies in growing the usage of pooled ride-hailing services, which allow travelers to share the cost, energy consumption, and emissions for riders headed in the same direction. Understanding the preferences for these ride-hailing services across different socio-economic groups informs policymaking for more inclusive shared mobility services for all travelers that yield significant reductions in energy and emissions, compared to private automobile travel.

In 2010, GM released the first commercially available plug-in hybrid EV “Chevy Volt”, and Nissan released the all-electric “LEAF”, respectively. As of 2020, there have been 83 EV models available in the U.S. market, representing significant growth for EVs, though there are still significantly more available models for gasoline-powered vehicles (U.S. DOE, 2020). Thanks to the environmental benefits of EVs, governments around the world have announced ambitious EV adoption targets. In Europe, Norway aims for a 100% EV market share by 2025, and Netherlands, Sweden, Denmark, Iceland, etc. also set targets for phasing out new passenger ICEV sales by 2030 (ICCT, 2020). China has announced a goal to achieve 25% EV market share by 2025 (State Council of the PRC, 2020). In North America, the State of California mandates that all light-duty vehicles on sale be EVs by 2035 (CA.GOV, 2020), and British Columbia in Canada by 2040 (GOV.BC.CA, 2020). In contrast to the ambitious EV adoption goals, the current EV market share is still modest in most of the world. Norway is the leading country in EV adoption with a 56% market share in 2019. Besides Norway, two other European countries achieved market shares greater than 10% in 2019: Netherlands at 15% and Sweden at 11%. The EV market shares in China, U.S., and Japan, the three largest auto markets, pale in comparison, at 5%, 2%, and 1%, respectively (IEA, 2020). Understanding consumers' preferences for EVs is the key to bridge the gap between the current low market share and the ambitious EV adoption targets.

The third "revolution" relates to vehicle automation. Society of Automobile Engineers (SAE) defines six levels of vehicle automation, from level 0 (no automation) to level 5 (full automation) (SAE, 2021). Currently, there are no commercially available vehicles with level 5 automation (referred to as AVs in this dissertation), which can drive everywhere in all conditions and will not require passengers to take over driving. Litman (2021) predicts that, by 2045, half of the new vehicle sales will be AVs, and 40% of vehicle travel will be fulfilled by AVs. Many experts envision three AV modes: private AVs (PAV), shared AVs (SAV), and shared AVs with pooling (SAVP). It is critical to study travelers' preferences for these different AV modes, providing policy insights for encouraging the adoption of more sustainable AV modes.

1.3 Modeling approaches for incorporating preference heterogeneity

It has been well acknowledged that preferences are heterogeneous across the population, as evidenced in many research areas, including transportation, energy, and health. Accounting for preference heterogeneity is important for two reasons. Practically, understanding heterogeneous preferences provides policy insights for more effective marketing strategies. Theoretically, ignoring heterogeneity can lead to biased coefficient estimates, which then leads to erroneous inferences and predictions (Hess, 2014). Over the years, three types of preference heterogeneity have been examined, including observed (systematic) heterogeneity, unobserved (random) heterogeneity, and heterogeneity related to latent attitudes. The transportation field (especially travel behavior research) serves as an important application domain and also contributes greatly to the methodology development to account for each type of preference heterogeneity.

Observed preference heterogeneity refers to the heterogeneity that can be linked to observed variables. For instance, in the travel mode choice context, the preferences for mode-specific attributes can be different across socio-demographic groups. By interacting the mode-specific attributes with individual socio-economic variables, the heterogeneous preferences for attributes can be related to the individual socio-economic characteristics. The commonly used multinomial logit (MNL) models or generalized extreme value (GEV) models, which have gained popularity since the 1970s, are readily convenient to capture such observed heterogeneity.

Unobserved preference heterogeneity is related to the heterogeneity that cannot be linked to observed variables but instead are purely random. Since the 2000s, the mixed logit (MXL) models have been used to capture unobserved heterogeneity. The MXL models assume the preferences for an attribute follow a random (continuous) distribution in a population. Alternatives to the MXL models to capture unobserved heterogeneity are the latent class (LC) models which allow the preference coefficients to follow a discrete distribution. Furthermore, Greene and Hensher (2013) proposed the LC-MXL model, which combines the LC and MXL models, to account for additional layers of unobserved heterogeneity.

In addition to socio-economic characteristic, the relationship between attitudes and preferences have gained interest in recent years. Early research practice uses several attitudinal statements to measure attitudes and subsequently enters the responses to attitudinal statements into utility functions directly. This approach, however, may result in endogeneity bias and measurement errors, because attitudes are difficult to directly measure (Daly et al., 2012). In contrast, the integrated choice and latent variable (ICLV) models treat the attitudes as latent variables (Ben-Akiva et al., 2002). The latent variables are then used to explain both discrete choice outcomes and responses to attitudinal statements. The ICLV models can relate the heterogeneous preferences to latent attitudes, which improves the explanatory power of choice models and informs policy opportunities for attitude interventions.

1.4 Research objectives

Preference heterogeneity has been well studied in transportation research over the past decades, particularly for vehicle purchase behavior, travel mode choice, route choice, etc. Various heterogeneity models described in section 1.3 have prominent application examples in traditional transportation systems. Today the transportation system is experiencing disruptive changes with the adoption of emerging technologies. However, little research has been dedicated to the application of these heterogeneity models to study consumer preferences of emerging transportation services and modes. This dissertation aims to fill this gap with the following three specific objectives:

- The first objective is to understand the disparities in ride-hailing usage for the mobility-disadvantaged groups, particularly how the disparities are related to spatial contexts.
- The second objective is to examine the heterogeneous preferences for EVs, and to compare various discrete choice models that captures random preference heterogeneity.
- The last objective is to explore travelers' mode choice preferences in the AV era while accounting for latent attitudinal constructs.

1.5 Organization of dissertation

The rest of the dissertation is organized in the following manner.

- Chapter 2: Disparities in Ride-hailing Usage Under Different Spatial Contexts.
    - This paper uses the systematic heterogeneity treatment approach to examine how the disparities in ride-hailing usage can be linked to spatial contexts (i.e., urban vs. rural, community public transport mode share, etc.), based on large-scale revealed preference data.
    - This chapter addresses research objective 1.
- Chapter 3: Investigating Heterogeneous Preferences for Electric Vehicles.
    - This paper compares three advanced discrete choice models that capture random heterogeneity in the application of studying EV preferences. The model fit, prediction performance, behavioral interpretation, and policy implications from the three models are discussed.
    - This paper addresses research objective 2.
- Chapter 4: Local Context Matters: Examining Mode Choice Preferences in the Autonomous Vehicle Era in Two U.S. Metropolitan Regions.
    - This paper explores how AV mode choice preferences are related to latent attitudinal constructs, using the ICLV modeling framework.
    - This paper addresses research objective 3.
- Chapter 5: Conclusions
    - Summarize the major findings of this dissertation
    - Discuss the contributions of this dissertation
    - Discuss the limitations of this dissertation and avenues for future work.

Reference

Ben-Akiva, M., McFadden, D., Train, K., Walker, J., Bhat, C., Bierlaire, M., Bolduc, D., Boersch-Supan, A., Brownstone, D., Bunch, D.S. and Daly, A., 2002. Hybrid choice models: Progress and challenges. Marketing Letters, 13(3), pp.163-175.

CA.GOV, 2020. Governor Newsom Announces California Will Phase Out Gasoline-Powered Cars & Drastically Reduce Demand for Fossil Fuel in California's Fight Against Climate Change. https://www.gov.ca.gov/2020/09/23/governor-newsom-announces. Last accessed Feb. 2021.

Daly, A., Hess, S., Patruni, B., Potoglou, D. and 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(2), pp.267-297.

GOV.BC.CA, 2020. Province puts in place rules for 100% electric-vehicle sales by 2040. https://archive.news.gov.bc.ca/releases/news_releases_2017-2021/2020EMPR0031-001416.htm. Last accessed Feb. 2021.

Greene, W.H. and Hensher, D.A., 2013. Revealing additional dimensions of preference heterogeneity in a latent class mixed multinomial logit model. Applied Economics, 45(14), pp.1897-1902.

Hess, S., 2014. Latent class structures: taste heterogeneity and beyond. In Handbook of choice modelling (pp. 311-329). Edward Elgar Publishing Cheltenham.

ICCT, 2020. Growing momentum: Global overview of government targets for phasing out sales of new internal combustion engine vehicles. https://theicct.org/blog/staff/global-ice-phaseout-nov2020 Last accessed Feb. 2021.

IEA, 2020. Global EV outlook 2020: Entering the decade of electric drive? International Energy Agency.

Litman, 2021. Autonomous Vehicle Implementation Predictions Implications for Transport Planning. Victoria Transport Policy Institute. https://www.vtpi.org/avip.pdf

Rouse, D., 2021. Commentary: what are the implications of autonomous vehicles for urban design practice?. Journal of Urban Design, 26(2), pp.178-183.

SAE, 2021. Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles J3016_202104. https://www.sae.org/standards/content/j3016_202104/

Sperling, D., 2018. Three revolutions: Steering automated, shared, and electric vehicles to a better future. Island Press.

State Council of the PRC, 2020. New Energy Automobile Industry Development Plan (2021-2035). http://www.gov.cn/zhengce/content/2020-11/02/content_5556716.htm Last accessed Feb. 2021.

Uber, 2021. The history of Uber. https://www.uber.com/newsroom/history/. Last accessed on June, 2021.

U.S. DOE, 2020. Find and Compare Cars. https://www.fueleconomy.gov/feg/findacar.shtml

Chapter 2. Disparities in Ride-hailing Usage Under Different Spatial Contexts

Wenjian Jia [a], T. Donna Chen [a *]
[a] Department of Engineering Systems & Environment, University of Virginia, PO Box 400742, Charlottesville, VA 22906, U.S.
[*] Corresponding author: tdchen@virginia.edu

Abstract
Ride-hailing services have become an important travel mode. However, existing studies show disparities in ride-hailing usage across socio-economic groups. This paper further examines how these disparities change with spatial contexts. Based on the add-on samples from the 2017 National Household Travel Survey, we develop logistic regression models with interaction terms between individual socio-economic variables and spatial contexts. Results suggest that seniors and low-income travelers are less likely to use ride-hailing, and these age and income disparities in ride-hailing usage are exacerbated in urban and rural settings, respectively. Interestingly, as the share of community seniors and community median income increase, the ride-hailing usage disparity between seniors and younger travelers, and between low-income and high-income travelers, decrease, respectively. The analysis also supports the idea that ride-hailing services are filling mobility needs in public transport deserts for non-vehicle owners, as the odds of ride-hailing usage for non-vehicle owners is 3-4 times of vehicle owners in communities with zero public transport mode share. With an increase in community public transport mode share, the importance of ride-hailing services to non-vehicle owners (relative to vehicle owners) reduces dramatically. Study results highlight the importance of considering the interaction effects between socio-economic characteristics and spatial contexts in examining ride-hailing usage, which informs policy-making for creating more inclusive ride-hailing services.

Key Words: Shared Mobility, Ride-hailing; Spatial Contexts; Transportation Equity, Logistic Regression

2.1 Introduction

With the rise of the shared economy and advances in information communication technologies, ride-hailing services, such as those provided by Uber and Lyft, have gained popularity around the world. These ride-hailing apps connect travelers and drivers via a real-time online platform which is equipped with advanced ride-matching algorithms and is often less expensive than traditional taxis. Furthermore, ride-hailing serves as a flexible mobility option for those belonging to traditionally mobility under-served groups (Brown, 2019).

Existing studies have examined two important questions related to ride-hailing usage patterns: 1) who (e.g., individual socio-economic attributes) are users of ride-hailing services; and 2) how spatial contexts (e.g., built environment characteristics) are related to the usage of ride-hailing services. Prior literature has shown disparities in ride-hailing usage between different socio-economic groups. For example, seniors (relative to younger travelers) and low-income (relative to high-income) travelers are found to be less likely to use ride-hailing services (Smith, 2016; Conway et al., 2018; Sikder, 2019). Many spatial context variables are also found to be important in explaining the usage of ride-hailing, such as neighborhood land use mix, population density, accessibility, etc. (Alemi et al., 2018a; Grahn et al., 2020; Malik et al., 2021).

However, the interplay between individual socio-economic characteristics and spatial contexts are not well understood. The interaction of these two elements helps improve the understanding of how spatial contexts are associated with exacerbating or dampening of disparities in ride-hailing usage across socio-economic groups, thereby providing policy insights to create more equitable and inclusive ride-hailing services. Motivated by this knowledge gap, this paper develops logistic regression models with various interaction terms, based on the add-on samples from seven different states from the 2017 National Household Travel Survey (NHTS). The research scope is confined to traditionally mobility-underserved groups, including seniors, low-income travelers, females, non-drivers, and travelers from zero-vehicle households. Related spatial contexts are incorporated, such as the community percent of seniors, median income, the share of females, public transport mode share, and urban/rural classification. Additionally, with the seven state samples, the analysis reveals the robustness of findings across different geographic regions.

2.2 Literature review

This section summarizes studies on the factors associated with the usage of ride-hailing services. The factors can be broadly divided into: 1) individual socio-economic characteristics, and 2) spatial contexts (such as built environment characteristics).

2.2.1 Individual socio-economic characteristics

Researchers often use individual travel surveys to examine how socio-economic characteristics are related to ride-hailing usage. The 2017 NHTS data is frequently used to examine ride-hailing usage patterns in the U.S. Consistently, studies showed that ride-hailing usage was negatively correlated with vehicle ownership (Conway et al., 2018; Sabouri et al., 2020a; Bansal et al., 2020) or households with more vehicles than workers (Sikder 2019). A positive relationship between ride-hailing usage and public transport use frequency is reported by Mitra et al. (2019) and Sikder (2019). Further, Sikder (2019) found that African Americans, seniors, and people from households with children were less likely to use ride-hailing services. Deka and Fei (2019) reported that ride-hailing usage frequency was lower for drivers than non-drivers.

Other studies are conducted at the U.S. national level but use self-administered surveys instead of 2017 NHTS. For example, Bansal et al. (2020) surveyed 11,902 respondents in the U.S. and found that younger individuals who had achieved higher education levels, from households with fewer vehicles, and belong to more affluent families were more likely to use ride-hailing services. Smith (2016) surveyed 4,787 U.S. adults and found that the propensity for ride-hailing was substantially higher among younger adults and people with high income and education levels. They also found that ride-hailing users were less likely to own cars and relied more heavily on public transit.

At the regional level, ride-hailing usage in California has been studied extensively (Alemi et al., 2018a; Alemi et al., 2018b, Alemi et al., 2019, and Malik et al., 2021). For example, Alemi et al. (2018a) examined factors associated with ride-hailing usage, based on the California Millennials Dataset (n = 1975). Results showed that younger, better educated, non-Hispanic individuals were more likely to use ride-hailing services. This ride-hailing usage was also correlated with frequency of plane travel, number of long-distance business trips, usage of smartphone transportation-related apps, and usage of taxi and car-sharing services. Rayle et al. (2016) compared ride-hailing with traditional taxi service usage in San Francisco, California. Study results suggested that these two types of mobility services differed significantly in terms of user characteristics and wait times, and the study found a latent demand from young and well-educated users for ride-hailing. Lavieri and Bhat (2019) revealed the role of latent attitudinal and lifestyle constructs in explaining ride-hailing usage based on survey data from 1607 commuters in Dallas-Fort Worth metropolitan area, Texas. Focusing on the city of Toronto, Canada, Young and Farber (2019) found that the wealthy younger travelers were more likely to be ride-hailing users.

In summary, the studies above using surveys conducted at various geographic scales all show various disparities in ride-hailing usage. Generally, senior, low-income, and female travelers are found to be less likely to use ride-hailing, while travelers from households without vehicles are more likely to use ride-hailing (Rayle et al., 2016; Dias et al., 2017; Conway et al., 2018, Sikder, 2019; Deka and Fei, 2019; Grahn et al., 2020; Bansal et al., 2020). But the evidence is also mixed. Smith (2016) found that men and women were equally likely to use ride-hailing services according to a national Pew Research Center survey of 4,787 American adults. Tirachini and del Río (2019) examined ride-hailing usage in Santiago De Chile and found that car availability was not a statistically significant factor in explaining ride-hailing usage.

2.2.2 Spatial context variables

The relationship between spatial contexts and ride-hailing usage patterns are often examined based on ride-hailing trip data released by service operators. Due to privacy concerns, such data lack travelers' socio-economic information. As a result, these studies often aggregate ride-hailing trips at a certain spatial scale (e.g., census tract), and explores determinants of the aggregated ride-hailing demand.

One of the popular datasets analyzed in previous literature is the "City of Chicago Transportation Network Providers trip database", which records the distance, duration, cost, and pick-up and drop-off points for each ride-hailing trip. Using this dataset, Barajas and Brown (2021) found lower levels of ride-hailing usage in low-income neighborhoods. Yan et al. (2020) used a random forest model to predict ride-hailing demand between census tract pairs in Chicago. They found that census tract socioeconomic and demographic characteristics were the most important predictor variables. Additionally, results showed that the percentage of workers earning $3333 or less per month in a census tract was negatively correlated with ride-hailing usage. Ghaffar et al. (2020) estimated random-effects negative binomial models and found that higher ride-hailing demand is associated with fewer parking spots and higher parking fees in a census tract.

The publicly available ride-hailing trip data from RideAustin, a transportation network company in Austin, Texas, has also been analyzed in several studies (Yu and Peng, 2019; Yu and Peng, 2020; and Dias et al., 2019). For example, Yu and Peng (2019) found that census blocks with greater land use mix are associated with higher ride-hailing demand. Based on data from 6.3 million Lyft trips in Los Angeles, California, Brown (2019) examined the relationship between Lyft demand and built environment and neighborhood socio-economic characteristics. The study found that Lyft trip demand was positively correlated with the percent of households without cars in a neighborhood.

In a more extensive study, Sabouri et al. (2020b) examined the impacts of built-environment on Uber usage based on trip data from 24 metropolitan regions in the U.S. Results showed that Uber trip demand was positively associated with the total population, employment, activity density, land use mix, and transit stop density in a census block group, and was negatively correlated with intersection density and destination accessibility by auto and transit.

In addition to these aggregate-level studies, a few studies based on individual travel surveys also explored the correlation between individual travelers' ride-hailing usage and spatial contextual variables. For example, Bansal et al. (2020) surveyed 11,902 respondents in the U.S and found that travelers living in metropolitan areas were more likely to use ride-hailing services. Alemi et al. (2018a)'s study in California showed that travelers from areas with greater land-use mix and regional car accessibility were more likely to use ride-hailing services.

2.2.3 The interaction of individual socio-economic variables and spatial contexts

Given the demonstrated disparities in ride-hailing usage between different socio-economic groups and between different spatial/neighborhood contexts, a natural next step is to examine how individual socio-economic characteristics' impact on ride-hailing usage varies across different spatial contexts. To the authors' knowledge, only two studies have investigated this issue. The first study is conducted by Dias et al. (2017) which used the 2014-2017 Puget Sound Regional travel survey data. By interacting the household vehicle ownership variable with a dummy variable representing a high-density neighborhood, the study found that households owning vehicles are less likely to use ride-hailing than households without vehicles when households reside in low-density neighborhoods (less than 5000 households per square mile). In contrast, the disparity of ride-hailing usage between households without and with vehicles disappears in high-density neighborhoods. However, the major focus of Dias et al. (2017) was on the main effects of various variables on ride-hailing usage, thereby only limited number of interaction effects were examined. On the other hand, Shirgaokar et al. (2021) explicitly examined the interaction effects of socio-demographics and spatial locations on the motivation to use ride-hailing based on 2,917 survey respondents from California. The study found that women from suburban or small town/rural areas were more open to ride-hailing than their male counterparts for reasons of non-reliance on others, not getting lost while driving, and getting help with carrying luggage. In contrast, women in urban areas were less likely to use ride-hailing for these reasons, compared to their male counterparts. However, Shirgaokar et al. (2021) focused on the stated motivation to use ride-hailing instead of the real-world ride-hailing usage, and the sample was limited to people aged 55 or older. Moreover, both studies focused only on a specific region, a more thorough and systematic examination of such interactions across multiple geographic regions is necessary to verify the robustness of findings.

This study fills the research gap on ride-hailing usage disparity by examining the intersection of individual socioeconomic characteristics and spatial contexts, based on large state add-on samples (CA, TX, GA, SC, NC, NY, and WI) from the 2017 NHTS survey dataset. For each state sample, this paper focuses on the ride-hailing usage disparities for mobility disadvantaged groups, and examine whether and how those disparities are related to spatial contexts.

2.3 Data and methods

2.3.1 Data

The 2017 NHTS dataset is the main data source used for this analysis. More specifically, this paper uses the add-on samples from state Departments of Transportation (DOTs), which are partners of 2017 NHTS administered by the Federal Highway Administration (FHWA). These add-on samples have large sample sizes, facilitating exploring the interaction effects between various socio-economic characteristics and spatial contexts. A total of nine state DOT partners participated in the 2017 NHTS data collection. Among the nine add-on samples, Arizona and Maryland samples show a small number of respondents who had used ride-hailing (237 and 187, respectively), and thus are excluded from the analysis, leaving seven add-on samples (CA, TX, GA, SC, NC, NY, and WI). The number of respondents and sample descriptive statistics for each state are shown in **Figure 2 - 1** and **Table 2 - 1**, respectively.

New to the 2017 iteration, the 2017 NHTS includes questions on ride-hailing usage for respondents who are at least 16 years old. Specifically, the survey asks, "In the past 30 days, how many times have you purchased a ride with a smartphone ride-share app (e.g., Uber, Lyft, Sidecar)?" This paper divides respondents into non-users and users depending on their responses to the ride-hailing usage question, leading to a binary dependent variable. The choice to reduce this response to binary classification is mainly due to the limited number of ride-hailing users in the samples. With more ride-hailing users in future iterations of the NHTS, it is possible to further classify users into more categories (e.g., infrequent users, frequent users, and heavy users).

The analysis focuses on socio-economic groups which are traditionally mobility-disadvantaged, including seniors, low-income people, females, non-drivers, and travelers from households without vehicles. To explore how the ride-hailing usage disparities across socio-economic groups are associated with the spatial contexts, more detailed geographic information of respondents is needed. However, the public version of 2017 NHTS contains respondents' geolocation information only at the state level. By signing a contract with FHWA, the authors obtained an additional dataset that recorded the residential census tract for each respondent. Thus, various community characteristics can be linked with each respondent by matching the respondents' census tract with the 2017 American Community Survey (ACS) data, including urban/rural classification, percent of females, percent of seniors, census tract median income, public transport mode share, etc. Lastly, the analysis includes a set of control variables that can be correlated with both socio-economic variables and ride-hailing usage, including respondents' educational attainment, smartphone availability, total daily trip count, and whether the respondents were born in the U.S.

Table 2 - 1. Descriptive statistics of the explanatory variables and response variables.

Variables	**CA** (n = 48,104)	**TX** (n = 45,266)	**GA** (n = 15,612)	**SC** (n = 12,507)	**NC** (n = 15,807)	**NY** (n = 31,192)	**WI** (n = 20,808)
Using ride-hailing at least once in the last month	12%	8%	8%	5%	4%	5%	4%
Female	53%	53%	55%	54%	55%	53%	52%
Age: 65 +	32%	28%	29%	34%	34%	32%	31%
Income lower than state median	47%	33%	40%	42%	42%	51%	36%
Non-driver	9%	7%	9%	8%	7%	11%	6%
Household without vehicles	3%	2%	4%	3%	3%	7%	2%
Control variables							
High school graduate or lower	21%	25%	29%	31%	29%	28%	30%
Bachelor's degree or higher	48%	46%	42%	39%	39%	45%	40%
Without smartphone	15%	12%	15%	19%	20%	23%	25%
Daily trip count (mean)	3.53	3.57	3.48	3.54	3.63	3.56	3.71
Born in the U.S.	83%	86%	92%	96%	95%	90%	96%

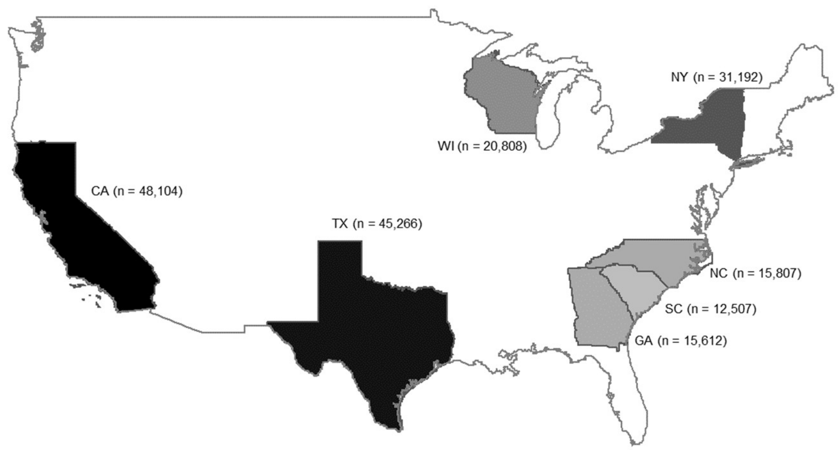


Figure 2 - 1. Sample distribution

2.3.2 Methods

Logistic regression models are developed for this analysis. The log odds of using ride-hailing are assumed to be a linear function of the predictor variables and their interactions:

$$logit(y) = \ln\left(\frac{P(y=1)}{1-P(y=1)}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 \times x_2 \quad (1)$$

Where, y represents whether a respondent has used ride-hailing in the past month. x_1 denotes individual socio-demographic variables and x_2 are community contextual variables. β are estimated coefficients.

The interpretation of logistic models often uses odds ratios. Taking the non-driver dummy variable for example, the odds ratio for this dummy variable denotes the odds of **non-drivers** using ride-hailing relative to the odds of **drivers** using ride-hailing. Due to the specification of an interaction term between the non-driver variable and community public transport mode share, the ride-hailing usage odds ratio for non-drivers (vs. drivers) depends on the community public transport mode share, as shown in the equation below. Thus, the authors can examine how the disparity in ride-hailing usage between non-drivers and drivers changes with the community public transport mode share. The odds ratio for other individual socio-demographic variables can be interpreted similarly.

$$OR_{nonDriver} = exp(\beta_{nonDriver} + \beta_{nonDriver \times PTShare} \times PTShare) \quad (2)$$

Where, $OR_{nonDriver}$ is the odds ratio of ride-hailing usage (non-drivers vs. drivers). $\beta_{nonDriver}$, $\beta_{nonDriver \times PTShare}$ are estimated coefficients. $PTShare$ denotes community public transport mode share.

2.4 Results and discussion

Logistic regression model results based on the seven samples are shown in the appendix (Table A.1 to Table A.7, respectively). Based on model coefficient estimates, the authors calculate the odds ratio of ride-hailing usage between different socio-economic groups, under various spatial contexts. Note that an odds ratio of one means that the odds of using ride-hailing is the same between two socio-economic groups. Next, the authors discuss the ride-hailing usage odds ratio between seniors and younger travelers, low-income and high-income travelers, females and males, non-drivers and drivers, and travelers living in households without vehicles and with vehicles.

2.4.1 Age

Figure 2 - 2 shows the ride-hailing usage odds ratio between seniors and younger travelers for the seven states. In general, seniors (aged 65 or more) show lower odds of using ride-hailing compared to younger travelers. This result is in line with previous findings (e.g., Smith, 2016; Conway et al., 2018; Sikder 2019; Deka and Fei, 2019; Grahn et al., 2020). Furthermore, the magnitude of the age disparity in ride-hailing usage is found to be correlated with the urban/rural classification, according to five out of the seven samples. Urban areas show a more severe age disparity in ride-hailing usage. Take the NC sample for example, in **rural areas** (where community seniors account for 20% of the population), the odds of seniors using ride-hailing is **48% lower** than their younger rural counterpart. In **urban areas** (where community seniors also account for 20% of the population), the age disparity is exacerbated: urban seniors show **80% lower** odds of ride-hailing usage compared to younger travelers. Note that the association between urban/rural classification and age disparity is not supported by all the samples, since the TX and SC samples show an insignificant coefficient for the interaction term.

The age disparity in ride-hailing usage is also found to be associated with the share of seniors in a community. With an increase in the share of community seniors, the ride-hailing usage odds ratio (seniors vs. younger travelers) increase. This finding is supported by four samples (CA, TX, SC, NY). For the CA and TX samples, when a senior is from a community with a low share of seniors, the odds of ride-hailing usage is much lower than a younger traveler from the same community. In contrast, when a senior is from a community with a high share of seniors, the age disparity shrinks. The SC and NY samples suggest that in senior-dominated communities, senior travelers show higher odds of using ride-hailing than younger travelers. Last, the authors note that the community senior share is found to be irrelevant to the age disparity in the other three samples (NC, GA, and WI), as shown by the insignificant coefficient for the interaction term between the senior dummy and share of community seniors.

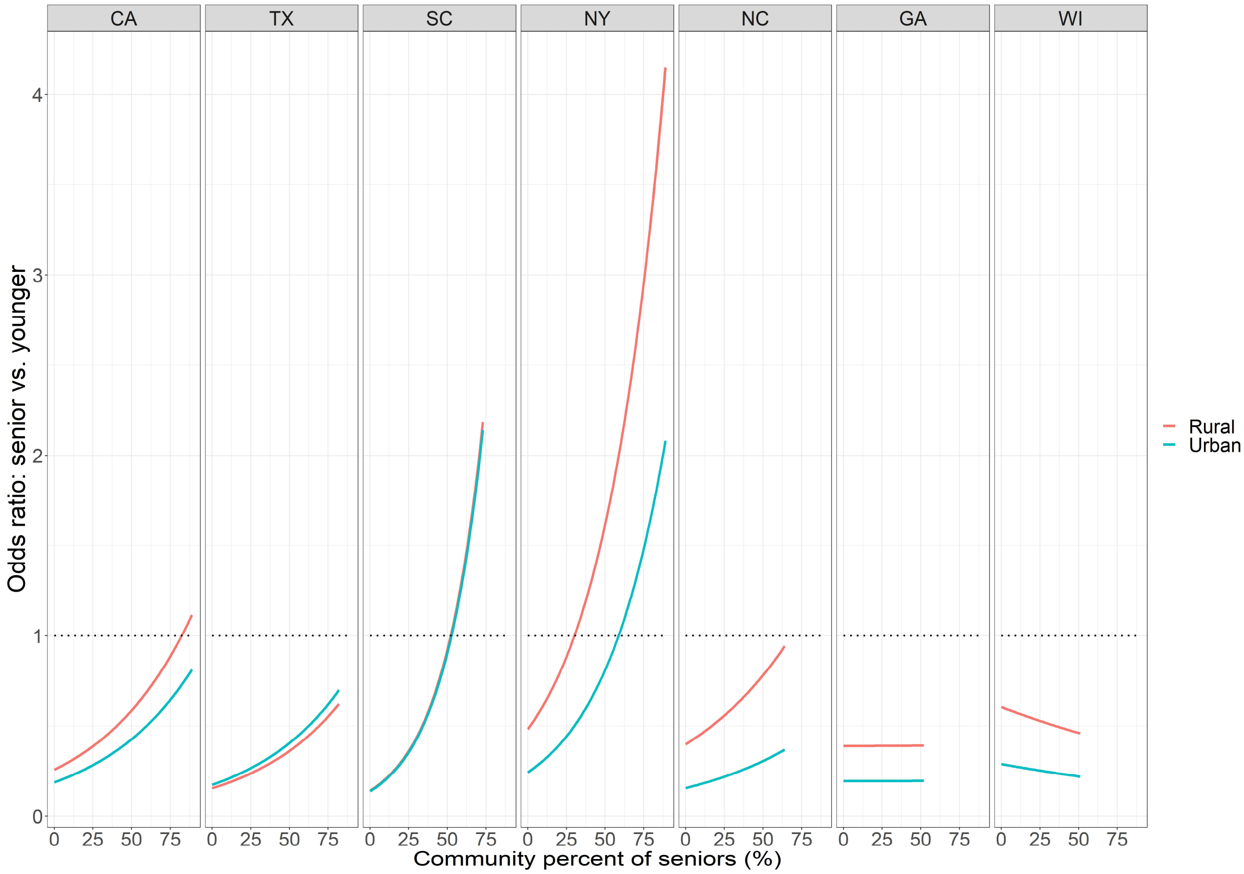


Figure 2 - 2. Ride-hailing usage odds ratio (senior vs. younger travelers) from different state samples

2.4.2 Household income

For each state sample, a respondent is classified to the low-income category when his/her reported household income is below the state median income. Generally, low-income travelers show lower odds of using ride-hailing than their high-income counterparts, as shown in **Figure 2 - 3.** Similar findings are reported by Smith (2016), Conway et al. (2018), Sikder (2019), Grahn et al. (2020), and Bansal et al. (2020), etc., indicating that the cost of ride-hailing services might be a barrier for low-income travelers.

Moreover, the ride-hailing usage disparity between low- and high-income travelers is found to be more severe in rural areas than in urban areas, based on the results of three samples (TX, NY, WI). For example, according to the TX sample, a low-income respondent from **an urban area** (with a community median income of $60,000) shows **35% lower** odds of using ride-hailing than his/her urban high-income counterpart. By contrast, a low-income respondent from **a rural area** (with a community median income of $60,000) shows **62% lower** odds of ride-hailing usage than its rural high-income counterpart. Another two samples (CA, SC) show similar disparity patterns related to the urban/rural classification, but with low statistical significance (p-value around 0.2). One potential explanation for such a disparity pattern is that travelers from rural areas generally have longer trips. As suggested by the 2017 NHTS data, the median trip distance for urban and rural travelers are 3.06 and 5.69 miles, respectively. A longer trip distance means higher cost for using ride-hailing services, which becomes less affordable for low-income travelers. Lastly, the association between urban/rural classification and the disparity in ride-hailing usage across income segments is not demonstrated by the remainder two samples (NC and GA).

Additionally, the community median income is also found to be relevant to the ride-hailing usage disparity based on the results of three samples (CA, TX, SC). According to the CA and TX samples, in poorer communities, the odds of using ride-hailing for a low-income traveler is much lower than a high-income traveler. Such a disparity is dampened in more wealthy communities. For the SC sample, in wealthy communities, low-income travelers show even higher odds of ride-hailing usage compared to their high-income counterparts. Similar trends are shown by the NC and NY samples, though with low statistical significance. These results suggest access to ride-hailing services may be an issue for low-income travelers from poor communities. Nonetheless, the underlying reason is not clear, due to the lack of transparency in service operators' strategies of supply distribution and pricing. Lastly, the authors note that opposite associations between community median income and ride-hailing usage disparities are found based on two samples (GA and WI), with low statistical significance.

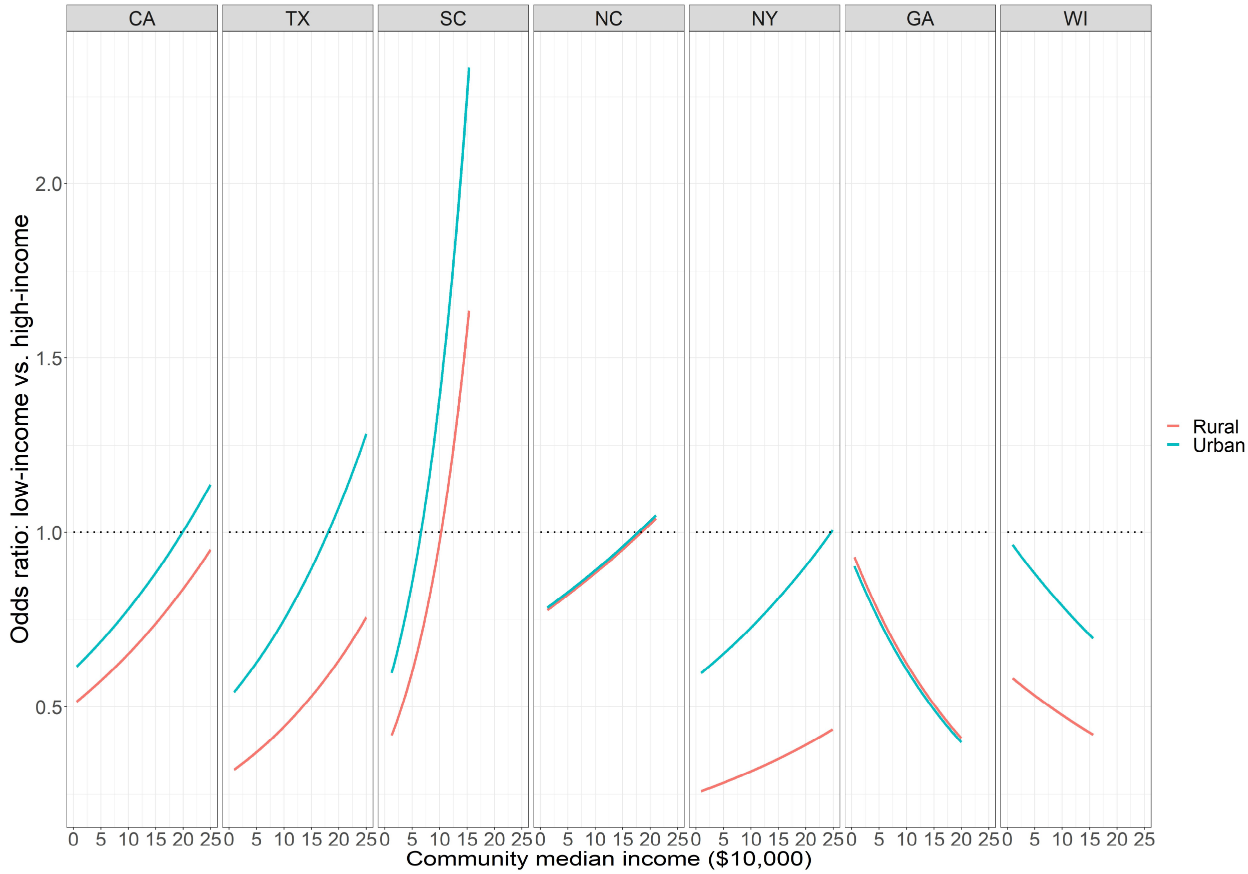


Figure 2 - 3. Ride-hailing usage odds ratio (low- vs. high-income travelers)

2.4.3 Gender

For all seven samples, the association between the gender disparity in ride-hailing usage and urban/rural classification appears to be minor, as represented by the low statistical significance for the interaction term between the female and urban dummy variables. A relevant study by Shirgaokar et al. (2021), however, demonstrates the role of spatial locations (urban vs. rural) in explaining gender differences in using ride-hailing. The study found that, in rural or suburbs areas, women were more likely than their male counterparts to use ride-hailing for reasons of being independent, not getting lost, and getting help with carrying bags. In contrast, in urban areas, women were less likely to use ride-hailing for these reasons. There are two possible explanations for the inconsistency in findings between this study and Shirgaokar et al. (2021). First, Shirgaokar et al. (2021) focused only on those travelers aged 55 or older, while this study focused on the general population. Second, Shirgaokar et al. (2021) examined the motivations to use ride-hailing while this study examined real-world ride-hailing usage.

Moreover, the percent of females in the community can be relevant to the gender disparity in ride-hailing usage, but the evidence is mixed among different state samples, as shown in **Figure 2 - 4**. According to the GA sample, the community female share is positively correlated with the ride-hailing usage odds ratio (females vs. males). In communities with fewer females, the odds of using ride-hailing for a female is much lower than for a male. In contrast, for communities dominated by females, a female shows higher odds of using ride-hailing than a male member in the same community. Interestingly, another sample (NY) supports opposite trends. Two other samples (WI and CA) show similar trends to the NY sample, but with low statistical significance (p-value around 0.2). Lastly, results from three samples (TX, SC, NC) suggest no association between the percent of females in a community and gender disparity in ride-hailing usage, as shown by the large p values (exceeding 0.8) for the coefficient of the interaction term between the female gender variable and community female share variable.

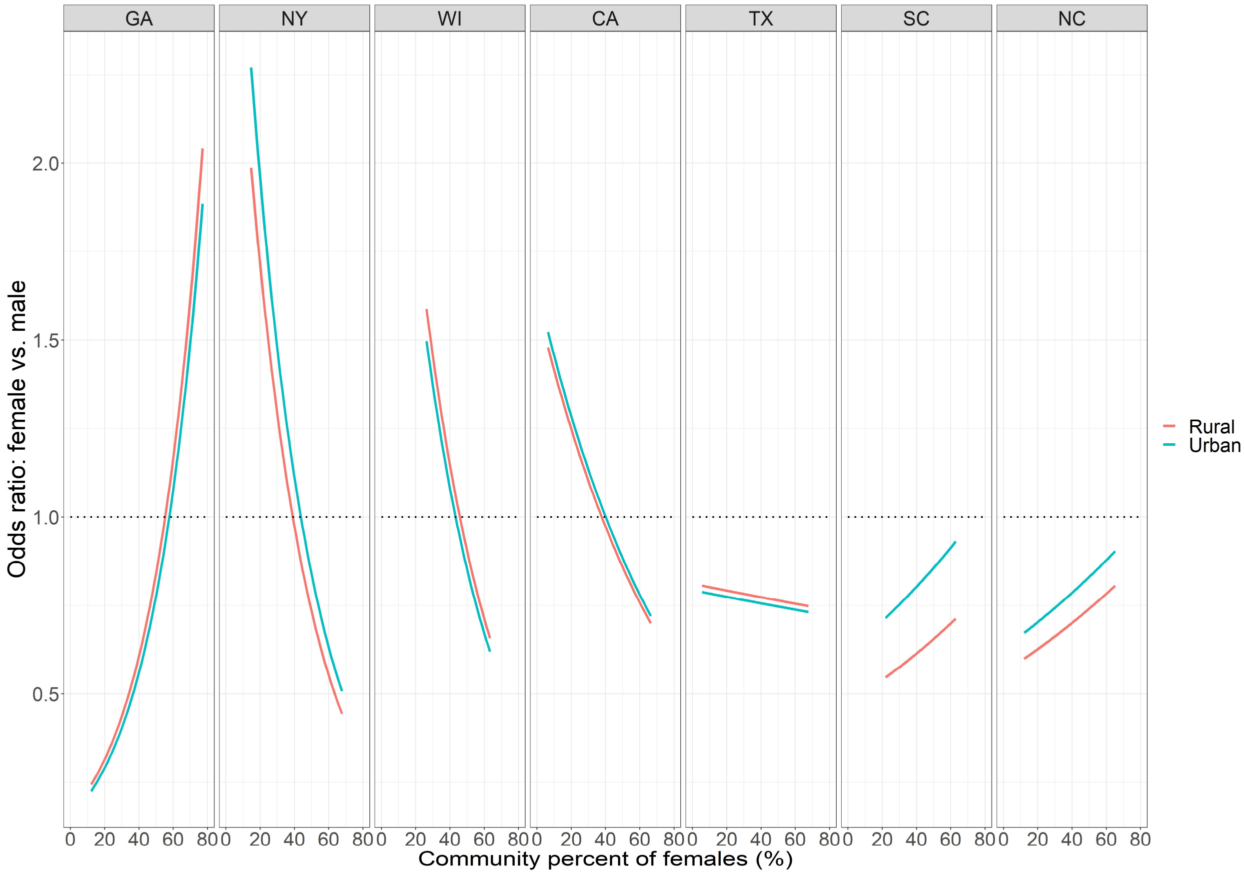


Figure 2 - 4. Ride-hailing usage odds ratio (female vs. male)

2.4.4 Household vehicle ownership

Figure 2 - 5 shows the odds ratio of ride-hailing usage between households without vehicles and with vehicles. In general, travelers from households without vehicles show higher odds of using ride-hailing than those from households owning vehicles. Similar findings are reported by Conway et al. (2018), Sikder (2019), Sabouri et al. (2020a), and Bansal et al. (2020). Furthermore, this study finds that the disparity in ride-hailing usage between households without vehicles and with vehicles is associated with the community public transport mode share, according to the results of three samples (CA, TX, NY). Taking the CA sample for example, in communities with zero public transport mode share, the odds for non-vehicle owners using ride-hailing is 3.1 times of those from households owning vehicles. With the increase in community public transport mode share, the importance of ride-hailing services to non-vehicle owners decreases. More specifically, in a community with a 58% public transport mode share (the highest public transport share among CA census tracts), the odds of a traveler from a zero-vehicle household using ride-hailing becomes even lower than those from households owning vehicles. Results of the remainder four samples (SC, NC, GA, WI) also show similar downward trends in odds ratios with the increase in community public transport mode share, but with low statistical significance. Lastly, note that in the SC sample, the coefficient for the interaction term between public transport mode share and household vehicle ownership is practically large, although statistically insignificant. As shown by the curve for the SC sample in **Figure 2 - 5**, the odds of using ride-hailing for travelers from zero-vehicle households in public transport desert communities is **184% higher** than those from households with vehicles. When community public transport mode share increase to 20% in SC, non-vehicle owners show **87% lower** odds of using ride-hailing than those owning vehicles. These results indicate that ride-hailing services increase mobility access for travelers with no access to private vehicles in transport desert communities.

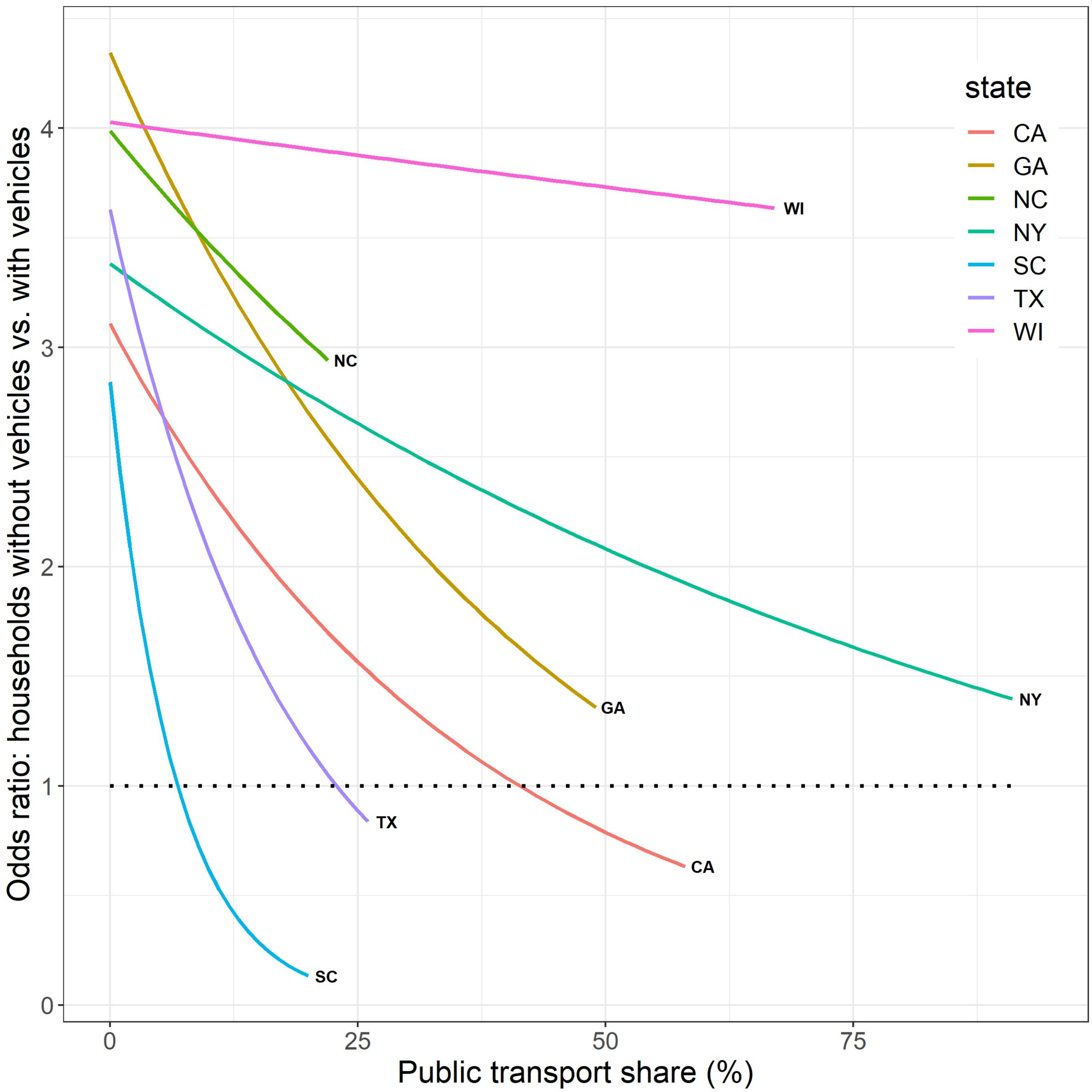


Figure 2 - 5. Ride-hailing usage odds ratio (travelers from households without vehicles vs. with vehicles)

2.4.5 Driving licensure

Figure 2 - 6 shows the ride-hailing usage odds ratio between non-drivers and drivers. Results of two samples (TX and GA) support the association between community public transport mode share and the disparity in ride-hailing usage between non-drivers and drivers. Taking the TX sample for example, in public transport desert areas, a non-driver shows **78% higher** odds of using ride-hailing compared to a driver. With the increase in community public transport mode share, the odds ratio of using ride-hailing (non-drivers vs. drivers) decreases. Specifically, in a community with 26% public transport share (the highest among TX census tracts), the odds of using ride-hailing for a non-driver is **67% lower** than that for a driver. The GA sample shows similar patterns. However, the results of the other five samples (CA, SC, NC, NY, WI) do not show statistically significant coefficients for the interaction terms between community public transport mode share and individual driver license status.

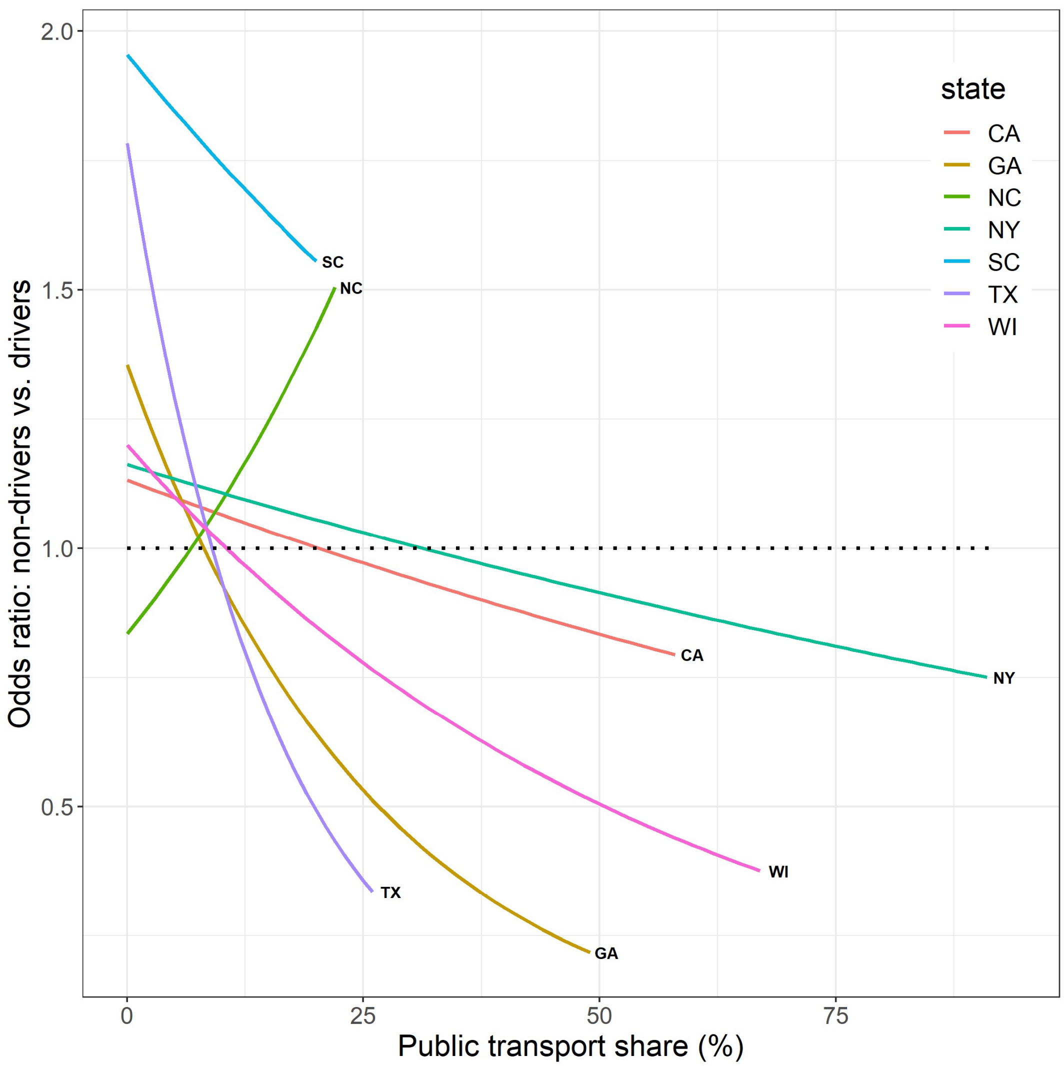


Figure 2 - 6. Ride-hailing usage odds ratio (non-drivers vs. drivers)

2.5 Conclusions and future work

This paper examines disparities in ride-hailing usage and how such disparities are associated with spatial contexts, based on seven add-on samples from the 2017 NHTS. Generally, results suggest that seniors (compared to younger) and low-income (compared to high-income) travelers are less likely to use ride-hailing. Furthermore, the ride-hailing usage disparity between seniors and younger travelers is exacerbated in urban areas, while the disparity between low-income and high-income travelers is exacerbated in rural areas. In addition to the urban/rural classification, the share of community seniors and community median income are also found to be relevant to the disparities in ride-hailing usage. The disadvantaged status for seniors using ride-hailing is improved with the increase in the share of community seniors. Wealthy communities see dampened disparities in ride-hailing usage between low-income and high-income travelers. Sometimes, the low-income travelers from more wealthy communities may show even higher odds of using ride-hailing than their high-income counterparts.

The gender disparities in ride-hailing usage do not show much correlation with the urban/rural classification in this analysis. The percent of community females can be relevant to gender-based gaps in ride-hailing usage, but with mixed evidence among the seven state samples. Results based on certain samples suggest that when a female is from a female-dominated community, the odds of using ride-hailing is higher than a male from the same community. The opposite trend is also supported by other samples: a female from a female-dominated community shows lower odds of using ride-hailing than its male counterpart. Either way, these results suggest that existing ride-hailing usage studies, which fail to consider the interaction effects between gender and spatial contexts, can mask the complexity of gender disparity in ride-hailing usage.

Moreover, results suggest that ride-hailing services appear to fill the mobility gap for people without cars living in public transport desert communities, as represented by their much higher odds of using ride-hailing compared to vehicle owners from the same community. With the increase in community public transport mode share, the odds of using ride-hailing for the two groups becomes closer, suggesting the diminishing role of ride-hailing services for travelers without vehicles where public transit is more available.

Finding of this paper highlight the need for future ride-hailing research to consider the interplay of socio-economic characteristics and spatial contexts, rather than examining these two elements independently as in most existing ride-hailing usage studies. Furthermore, the study results point to specific community contexts which place traditionally mobility underserved groups at an even greater disadvantage when it comes to ride-hailing usage, suggesting targeted areas for policy makers to improve equity in ride-hailing services.

The authors also note several limitations and potential future work. First, this analysis is based on observational cross-sectional data, therefore the findings can only suggest correlations while no causal relationship is warranted. Second, the 2017 NHTS survey was conducted between 2016 and 2017, and the number of respondents who have ride-hailing experiences are still limited, necessitating the use of a binary dependent variable (ride-hailing user and non-user) in this study For future studies with surveys conducted in more recent years (and greater market penetration of ride-hailing services), it is advisable to create multiple categories of ride-hailing users to capture nuances between non-users, infrequent users, and frequent users. Third, the NHTS survey does not differentiate between pooled and non-pooled ride-hailing trips. Considering the environmental and congestion alleviation benefits of pooling, future work should examine the usage patterns of pooled ride-hailing vs. single occupancy ride-hailing (Brown, 2020; Lazarus et al., 2021; Loa et a., 2021). Fourth, this paper focuses on how the disparities in ride-hailing usage differ across spatial contexts. The temporal evolution of such disparities is also important. As shown by Dias et al. (2021), there is a "democratization" of ride-hailing services over time. Lastly, this paper uses multiple state samples to check whether the findings are robust across different geographic regions, but the authors are unable to offer explanations on why findings based on a certain sample may differ from others. Comparative analysis between states could enrich our understanding of ride-hailing usage, accounting for state- and locality-specific ride-hailing regulations and policies.

Acknowledgements
The authors would like to thank Layla Sun and Stacey Bricka for providing details of the 2017 NHTS home location at Census Tract level. The authors also thank Shraddha Praharaj and Linda Lim for their help in editing this manuscript.

Author Contributions
The authors confirm contribution to the paper as follows: study conception and design: W. Jia, T.D. Chen; data collection: W. Jia; analysis and interpretation of results: W. Jia, T.D. Chen; draft manuscript preparation: W. Jia, T.D. Chen. All authors reviewed the results and approved the final version of the manuscript.

Appendix

Table A-1. Model results for the CA sample

	Est.	Std. err.	p value
Constant	-3.52	0.34	0.00
Female	0.47	0.46	0.31
Senior (age 65+)	-1.36	0.21	0.00
Low income (lower than state median income)	-0.68	0.18	0.00
Non-driver	0.12	0.09	0.15
Household without vehicles	1.13	0.12	0.00
High school graduate or lower	-0.68	0.06	0.00
Bachelor's degree or higher	0.64	0.04	0.00
Without smartphone	-1.96	0.11	0.00
Daily trip count	0.30	0.05	0.00
Born in the U.S.	0.35	0.04	0.00
Urban	0.91	0.12	0.00
Percent of females in census tract	0.00	0.01	0.59
Percent of seniors in census tract	-0.01	0.00	0.00
Median household income in census tract	0.04	0.00	0.00
Public transport share in census tract	0.06	0.00	0.00
Female × urban	0.03	0.15	0.85
Senior × urban	-0.32	0.19	0.10
Low income × urban	0.18	0.17	0.28
Female × Percent of females in census tract	-0.01	0.01	0.16
Senior × Percent of seniors in census tract	0.02	0.00	0.00
Low income × Median household income in census tract	0.03	0.01	0.01
Non-driver × Public transport share in census tract	-0.01	0.01	0.35
Household without vehicles × Public transport share in census tract	-0.03	0.01	0.00

Table A-2. Model results for the TX sample

	Est.	Std. err.	p value
Constant	-2.62	0.38	0.00
Female	-0.21	0.55	0.70
Senior (age 65+)	-1.86	0.32	0.00
Low income (lower than state median income)	-1.17	0.29	0.00
Non-driver	0.58	0.11	0.00
Household without vehicles	1.29	0.16	0.00
High school graduate or lower	-0.76	0.08	0.00
Bachelor's degree or higher	0.75	0.05	0.00
Without smartphone	-1.81	0.17	0.00
Daily trip count	0.28	0.07	0.00
Born in the U.S.	0.20	0.05	0.00
Urban	0.94	0.12	0.00
Percent of females in census tract	-0.03	0.01	0.00
Percent of seniors in census tract	-0.01	0.00	0.03
Median household income in census tract	0.06	0.01	0.00
Public transport share in census tract	0.13	0.01	0.00
Female × urban	-0.02	0.16	0.89
Senior × urban	0.11	0.30	0.70
Low income × urban	0.53	0.28	0.06
Female × Percent of females in census tract	0.00	0.01	0.91
Senior × Percent of seniors in census tract	0.02	0.01	0.03
Low income × Median household income in census tract	0.04	0.02	0.02
Non-driver × Public transport share in census tract	-0.06	0.03	0.02
Household without vehicles × Public transport share in census tract	-0.06	0.03	0.05

Table A-3. Model results for the GA sample

	Est.	Std. err.	p value
Constant	-2.88	0.62	0.00
Female	-1.81	0.83	0.03
Senior (age 65+)	-0.95	0.43	0.03
Low income (lower than state median income)	-0.05	0.32	0.87
Non-driver	0.30	0.18	0.10
Household without vehicles	1.47	0.21	0.00
High school graduate or lower	-0.66	0.13	0.00
Bachelor's degree or higher	0.85	0.08	0.00
Without smartphone	-1.59	0.23	0.00
Daily trip count	0.21	0.12	0.07
Born in the U.S.	-0.19	0.10	0.06
Urban	1.02	0.18	0.00
Percent of females in census tract	-0.02	0.01	0.12
Percent of seniors in census tract	-0.01	0.01	0.03
Median household income in census tract	0.09	0.01	0.00
Public transport share in census tract	0.09	0.01	0.00
Female × urban	-0.08	0.23	0.72
Senior × urban	-0.70	0.35	0.04
Low income × urban	-0.03	0.27	0.92
Female × Percent of females in census tract	0.03	0.02	0.04
Senior × Percent of seniors in census tract	0.00	0.02	0.99
Low income × Median household income in census tract	-0.04	0.03	0.19
Non-driver × Public transport share in census tract	-0.04	0.02	0.07
Household without vehicles × Public transport share in census tract	-0.02	0.02	0.26

Table A-4. Model results for the SC sample

	Est.	Std. err.	p value
Constant	-3.81	1.07	0.00
Female	-0.75	1.39	0.59
Senior (age 65+)	-1.96	0.49	0.00
Low income (lower than state median income)	-0.99	0.40	0.01
Non-driver	0.67	0.26	0.01
Household without vehicles	1.04	0.37	0.00
High school graduate or lower	-0.60	0.18	0.00
Bachelor's degree or higher	0.93	0.12	0.00
Without smartphone	-1.37	0.26	0.00
Daily trip count	0.49	0.16	0.00
Born in the U.S.	0.07	0.19	0.71
Urban	0.62	0.19	0.00
Percent of females in census tract	0.00	0.02	0.81
Percent of seniors in census tract	-0.03	0.01	0.00
Median household income in census tract	0.12	0.02	0.00
Public transport share in census tract	0.14	0.04	0.00
Female × urban	0.27	0.26	0.29
Senior × urban	-0.02	0.37	0.95
Low income × urban	0.35	0.31	0.25
Female × Percent of females in census tract	0.01	0.03	0.81
Senior × Percent of seniors in census tract	0.04	0.02	0.02
Low income × Median household income in census tract	0.10	0.05	0.07
Non-driver × Public transport share in census tract	-0.01	0.17	0.95
Household without vehicles × Public transport share in census tract	-0.15	0.19	0.43

Table A-5. Model results for the NC sample

	Est.	Std. err.	p value
Constant	-3.70	0.90	0.00
Female	-0.58	1.20	0.63
Senior (age 65+)	-0.92	0.40	0.02
Low income (lower than state median income)	-0.27	0.32	0.39
Non-driver	-0.18	0.27	0.50
Household without vehicles	1.38	0.29	0.00
High school graduate or lower	-0.47	0.16	0.00
Bachelor’s degree or higher	0.84	0.10	0.00
Without smartphone	-1.44	0.25	0.00
Daily trip count	0.28	0.14	0.05
Born in the U.S.	-0.11	0.14	0.45
Urban	1.11	0.19	0.00
Percent of females in census tract	-0.01	0.02	0.58
Percent of seniors in census tract	-0.01	0.01	0.10
Median household income in census tract	0.08	0.02	0.00
Public transport share in census tract	0.09	0.02	0.00
Female × urban	0.11	0.23	0.62
Senior × urban	-0.94	0.31	0.00
Low income × urban	0.01	0.26	0.97
Female × Percent of females in census tract	0.01	0.02	0.81
Senior × Percent of seniors in census tract	0.01	0.01	0.32
Low income × Median household income in census tract	0.01	0.04	0.72
Non-driver × Public transport share in census tract	0.03	0.07	0.71
Household without vehicles × Public transport share in census tract	-0.01	0.08	0.86

Table A-6. Model results for the NY sample

	Est.	Std. err.	p value
Constant	-3.99	0.58	0.00
Female	1.10	0.80	0.17
Senior (age 65+)	-0.73	0.30	0.02
Low income (lower than state median income)	-1.37	0.24	0.00
Non-driver	0.15	0.19	0.43
Household without vehicles	1.22	0.22	0.00
High school graduate or lower	-0.35	0.12	0.00
Bachelor's degree or higher	0.88	0.08	0.00
Without smartphone	-1.65	0.16	0.00
Daily trip count	0.30	0.11	0.01
Born in the U.S.	0.01	0.08	0.90
Urban	0.07	0.12	0.56
Percent of females in census tract	-0.01	0.01	0.53
Percent of seniors in census tract	0.01	0.01	0.27
Median household income in census tract	0.08	0.01	0.00
Public transport share in census tract	0.04	0.00	0.00
Female × urban	0.13	0.16	0.40
Senior × urban	-0.69	0.20	0.00
Low income × urban	0.84	0.21	0.00
Female × Percent of females in census tract	-0.03	0.02	0.07
Senior × Percent of seniors in census tract	0.02	0.01	0.06
Low income × Median household income in census tract	0.02	0.02	0.28
Non-driver × Public transport share in census tract	0.00	0.00	0.21
Household without vehicles × Public transport share in census tract	-0.01	0.00	0.02

Table A-7. Model results for the WI sample

	Est.	Std. err.	p value
Constant	-4.72	0.78	0.00
Female	1.09	1.07	0.31
Senior (age 65+)	-0.50	0.48	0.30
Low income (lower than state median income)	-0.52	0.41	0.21
Non-driver	0.18	0.25	0.47
Household without vehicles	1.39	0.30	0.00
High school graduate or lower	-1.00	0.15	0.00
Bachelor's degree or higher	0.78	0.09	0.00
Without smartphone	-1.65	0.19	0.00
Daily trip count	-0.09	0.13	0.47
Born in the U.S.	-0.47	0.14	0.00
Urban	0.69	0.16	0.00
Percent of females in census tract	0.03	0.02	0.03
Percent of seniors in census tract	-0.03	0.01	0.00
Median household income in census tract	0.05	0.02	0.00
Public transport share in census tract	0.06	0.01	0.00
Female × urban	-0.06	0.21	0.77
Senior × urban	-0.74	0.31	0.02
Low income × urban	0.51	0.33	0.12
Female × Percent of females in census tract	-0.02	0.02	0.27
Senior × Percent of seniors in census tract	-0.01	0.02	0.81
Low income × Median household income in census tract	-0.02	0.04	0.60
Non-driver × Public transport share in census tract	-0.02	0.02	0.39
Household without vehicles × Public transport share in census tract	0.00	0.02	0.95

Reference

Acheampong, R.A., Siiba, A., Okyere, D.K. and Tuffour, J.P., 2020. Mobility-on-demand: An empirical study of internet-based ride-hailing adoption factors, travel characteristics and mode substitution effects. Transportation Research Part C: Emerging Technologies, 115, p.102638.

Alemi, F., Circella, G., Handy, S. and Mokhtarian, P., 2018a. What influences travelers to use Uber? Exploring the factors affecting the adoption of on-demand ride services in California. Travel Behaviour and Society, 13, pp.88-104.

Alemi, F., Circella, G., Mokhtarian, P. and Handy, S., 2018b. Exploring the latent constructs behind the use of ridehailing in California. Journal of choice modelling, 29, pp.47-62.

Alemi, F., Circella, G., Mokhtarian, P. and Handy, S., 2019. What drives the use of ridehailing in California? Ordered probit models of the usage frequency of Uber and Lyft. Transportation Research Part C: Emerging Technologies, 102, pp.233-248.

Bansal, P., Sinha, A., Dua, R. and Daziano, R.A., 2020. Eliciting preferences of TNC users and drivers: evidence from the United States. Travel Behaviour and Society, 20, pp.225-236.

Barajas, J.M. and Brown, A., 2021. Not minding the gap: Does ride-hailing serve transit deserts?. Journal of Transport Geography, 90, p.102918.

Brown, A., 2019. Redefining car access: ride-hail travel and use in Los Angeles. Journal of the American Planning Association, 85(2), pp.83-95.

Brown, A.E., 2020. Who and where rideshares? Rideshare travel and use in Los Angeles. Transportation Research Part A: Policy and Practice, 136, pp.120-134.

Conway, M.W., Salon, D. and King, D.A., 2018. Trends in taxi use and the advent of ridehailing, 1995–2017: Evidence from the US National Household Travel Survey. Urban Science, 2(3), p.79.

Deka, D. and Fei, D., 2019. A comparison of the personal and neighborhood characteristics associated with ridesourcing, transit use, and driving with NHTS data. Journal of Transport Geography, 76, pp.24-33.

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

Dias, F.F., Lavieri, P.S., Kim, T., Bhat, C.R. and Pendyala, R.M., 2019. Fusing multiple sources of data to understand ride-hailing use. Transportation Research Record, 2673(6), pp.214-224.

Dias, F.F., Kim, T., Bhat, C.R., Pendyala, R.M., Lam, W.H., Pinjari, A.R., Srinivasan, K.K. and Ramadurai, G., 2021. Modeling the Evolution of Ride-Hailing Adoption and Usage: A Case Study of the Puget Sound Region. Transportation Research Record, 2675(3), pp.81-97.

Ghaffar, A., Mitra, S. and Hyland, M., 2020. Modeling determinants of ridesourcing usage: A census tract-level analysis of Chicago. Transportation Research Part C: Emerging Technologies, 119, p.102769.

Grahn, R., Harper, C.D., Hendrickson, C., Qian, Z, Matthews, H.S., 2020. Socioeconomic and usage characteristics of transportation network company (TNC) riders. Transportation, 47, pp. 3047–3067

Lavieri, P.S. and Bhat, C.R., 2019. Investigating objective and subjective factors influencing the adoption, frequency, and characteristics of ride-hailing trips. Transportation Research Part C: Emerging Technologies, 105, pp.100-125.

Lazarus, J.R., Caicedo, J.D., Bayen, A.M. and Shaheen, S.A., 2021. To Pool or Not to Pool? Understanding opportunities, challenges, and equity considerations to expanding the market for pooling. Transportation Research Part A: Policy and Practice, 148, pp.199-222.

Loa, P. and Habib, K.N., 2021. Examining the influence of attitudinal factors on the use of ride-hailing services in Toronto. Transportation Research Part A: Policy and Practice, 146, pp.13-28.

Malik, J., Bunch, D.S., Handy, S. and Circella, G., 2021. A deeper investigation into the effect of the built environment on the use of ridehailing for non-work travel. Journal of Transport Geography, 91, p.102952.

Mitra, S.K., Bae, Y. and Ritchie, S.G., 2019. Use of ride-hailing services among older adults in the United States. Transportation research record, 2673(3), pp.700-710.

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

Sabouri, S., Brewer, S. and Ewing, R., 2020a. Exploring the relationship between ride-sourcing services and vehicle ownership, using both inferential and machine learning approaches. Landscape and Urban Planning, 198, p.103797.

Sabouri, S., Park, K., Smith, A., Tian, G. and Ewing, R., 2020b. Exploring the influence of built environment on Uber demand. Transportation Research Part D: Transport and Environment, 81, p.102296.

Shirgaokar, M., Misra, A., Agrawal, A.W., Wachs, M. and Dobbs, B., 2021. Differences in ride-hailing adoption by older Californians among types of locations. Journal of Transport and Land Use, 14(1), pp.367-387.

Sikder, S., 2019. Who uses ride-hailing services in the United States? Transportation research record, 2673(12), pp.40-54.

Smith, A. (2016). Shared, collaborative, and on demand: The new digital economy. Pew Research Center. Retrieved from http://www.pewinternet.org/2016/05/19/the-new-digitaleconomy/

Tirachini, A. and del Río, M., 2019. Ride-hailing in Santiago de Chile: Users' characterisation and effects on travel behaviour. Transport Policy, 82, pp.46-57.

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

Yan, X., Liu, X. and Zhao, X., 2020. Using machine learning for direct demand modeling of ridesourcing services in Chicago. Journal of Transport Geography, 83, p.102661.

Young, M. and Farber, S., 2019. The who, why, and when of Uber and other ride-hailing trips: An examination of a large sample household travel survey. Transportation Research Part A: Policy and Practice, 119, pp.383-392.

Yu, H. and Peng, Z.R., 2019. Exploring the spatial variation of ridesourcing demand and its relationship to built environment and socioeconomic factors with the geographically weighted Poisson regression. Journal of Transport Geography, 75, pp.147-163.

Yu, H. and Peng, Z.R., 2020. The impacts of built environment on ridesourcing demand: A neighbourhood level analysis in Austin, Texas. Urban Studies, 57(1), pp.152-175.

Chapter 3. Investigating Heterogeneous Preferences for Electric Vehicles

Wenjian Jia [a], T. Donna Chen [a] *
[a] Department of Engineering Systems & Environment, University of Virginia, PO Box 400742, Charlottesville, VA 22906, U.S.
* Corresponding author: tdchen@virginia.edu

Abstract

Understanding consumer preference heterogeneity in adopting electric vehicles (EVs) is critical to inform EV policymaking. This paper develops mixed logit (MXL), latent class (LC), and latent class-mixed logit (LC-MXL) models based on stated choice experiments data collected in Virginia in 2018 (n = 837). Overall, all three models indicate that the monetary incentive is most effective in increasing EV market share, followed by deploying more charging infrastructure, while improvement in battery range is found to be least effective. Furthermore, the performance of these different models in examining EV preference heterogeneity is compared in terms of model fit, behavioral interpretation, and policy implications. Results show that no model is unanimously superior to the other models, rather, altogether they provide a more comprehensive picture of EV preference structure. Findings provide insights on the usefulness of each modeling framework for future EV preference research. Also, it informs policy-makers to be aware of alternative models which can provide a different perspective on policy implications.

Keywords: Electric Vehicles, Stated Preference, Preference Heterogeneity, Discrete Choice Models, Mixed Logit, Latent Class.

3.1 Introduction

Vehicle electrification represents a promising solution to improve energy security and reduce greenhouse gas (GHG) emissions. By the current global average carbon intensity of electricity generation, a global average mid-sized battery electric vehicle (BEV) and a plug-in hybrid electric vehicle (PHEV) emit similar amounts of GHG compared to a non-plug-in hybrid electric vehicle (HEV), and less GHG than a global average internal combustion engine vehicle (ICEV) using gasoline over each vehicle's life cycle (IEA, 2020). With further decarbonization in electricity generation, the emissions reduction benefits of EVs (including BEVs and PHEVs) over ICEVs on a life-cycle basis are expected to be more significant.

Ambitious EV adoption goals have been proposed by governments. For example, the state of California set a target of ending ICEV new sales by 2035 (CARB, 2021). The province of British Columbia in Canada passed the zero-emission vehicle (ZEV) act which set targets of 30% ZEV market share of new vehicle sales by 2030 and 100% by 2040 (GOV.BC.CA, 2019). In contrast to those ambitious goals, consumer demand for EVs is still low (though growing). As the leading state in adopting electric mobility in the US, California's new vehicle sales in 2019 saw only a 7.6% EV market share (CNCDA, 2021). At the country level, the 2019 EV market shares in the U.S. and Canada are 2.1% and 3.0%, respectively, lagging behind China and many European countries (IEA, 2020).

Understanding consumers' (heterogeneous) preferences is critical to inform EV support policies and marketing strategies to achieve mass EV adoption. Due to the lack of large-scale real-world EV purchase data, EV consumer preferences studies are usually based on stated preference (SP) survey data. In a SP survey, respondents are asked to select their preferred option from several alternatives in hypothetical scenarios, by conducting trade-offs among attributes associated with each alternative. Discrete choice models are often used to analyze the outcomes of stated choice data, where model coefficients represent marginal effects of attributes on the utility of an alternative (also represent consumers' tastes or preferences for the attributes). The model coefficients can then be used for eliciting various policy-related measures, such as calculating willingness-to-pay (WTP) for EV attributes (Hidure et al., 2011; Tanaka et al., 2014; Axsen et al., 2015; Hackbarth and Madlener, 2016; Noel et al., 2019), and simulating EV market shares under various incentive policy and charging infrastructure deployment scenarios (Valeri and Danielis, 2015; Kormos et al., 2019; Qian et al., 2019; Gong et al., 2020; Danielis et al., 2020).

To date, the mixed logit (MXL) and latent class (LC) are two predominant discrete choice models to study EV preferences, thanks to their capabilities to capture preference heterogeneity (especially unobserved heterogeneity). The two models have very different assumptions, where the MXL models assume preference coefficients follow continuous and unimodal distribution across the population while the LC models segment preference coefficients into a finite number of discrete classes. Neither of the assumptions can be warranted and an inappropriate assumption may bias the results and thus provide erroneous policy insights. However, existing studies lack in comparing the role of the two models in uncovering EV preferences. Furthermore, recent advances in choice modeling propose the latent class-mixed logit (LC-MXL) model, considering that both

MXL and LC models have their own limitations. The LC-MXL model combines the two approaches above, allowing preference coefficients to be randomly distributed within each latent class (Bujosa et al. 2010; Greene and Hensher, 2013), and shows an advantage in capturing complex preference heterogeneity structure (Keane and Wasi, 2013). Though has not been applied in the EV adoption domain, the LC-MXL may have the potential to study EV preferences since the choice experiments often involve many EV attributes and substantial preference heterogeneity is reported (Liao et al., 2017).

Motivated by these gaps, this study estimates the MXL, LC, and LC-MXL models to examine consumers' EV preferences, based on a stated vehicle fuel type choice survey in Virginia, U.S. The major contribution lies in discussing how the results of frequently used models are translating into EV policy implications. Specifically, we examined consumers' WTP for EV attributes and the effectiveness of various policy scenarios based on the three models above. Findings help to better understand the performance of different models in studying EV preferences, and provide insights for future research on model selection.

3.2 Literature review

Consumers' EV preference heterogeneity is widely studied in various regions using SP surveys. The preference heterogeneity represents the variation in tastes for an attribute across the population, which can be decomposed into observed heterogeneity (systematic heterogeneity) and unobserved heterogeneity (random heterogeneity). Early studies use multinomial logit (MNL) to capture systematic preference heterogeneity, by specifying utility functions with interaction terms between EV attributes and individual socio-economic characteristics. Often, analysts are unable to include all relevant factors that impact people's choices, which leads to the existence of unobserved heterogeneity. The MNL models, however, are unable to capture unobserved preference heterogeneity. When ignoring unobserved heterogeneity, the estimated model coefficients can be biased and inconsistent, and thus can yield erroneous inferences and policy implications (Hess, 2014; Mannering et al., 2016).

In more recent EV preference studies, it has become a common practice to allow for unobserved heterogeneity using more advanced discrete choice models (i.e., MXL or LC models), as shown in **Table 3 - 1**. In contrast to the fixed preference coefficients in MNL models, the MXL models specify coefficients to follow continuous and unimodal distributions (such as normal, lognormal, and triangular) across the population. The seminal work by Brownstone and Train (1998) illustrates an application of MXL models to study preferences for alternative fuel vehicles (AFVs). Based on a SP survey from 4654 respondents in California, Brownstone and Train developed MXL models with various random error components to allow for flexible substitution patterns and random taste heterogeneity. Recently, MXL models have been applied in studying EV preferences in Italy (Danielis et al., 2020), Japan (Khan et al., 2020), and South Korea (Kim et al., 2020, applied energy). Although MXL models are flexible and statistically powerful, the assumption that preference coefficients vary continuously and unimodally over respondents may not be warranted (Greene and Hensher, 2003). Also, the MXL model is inadequate when the sample consists of different segments of individuals with varied segment-specific preferences, or the sample show complex multimodal preferences distributions.

The LC models specify a finite number of discrete classes where each class has unique preference coefficients, and each respondent belongs to a class with a probability depending on the individual's socio-economic characteristics. Hidrue et al. (2011) is the first LC model application to study EV preferences, based on 3029 survey respondents in the U.S. The study identifies two latent classes (EV-oriented class and gasoline vehicle-oriented class) where the EV-oriented class is more sensitive to fuel cost and driving range than the gasoline vehicle-oriented class. Furthermore, younger and more educated respondents are more likely to be in the EV-oriented class, which informs the target population for marketing. More recently, Kormos et al. (2019) identified five unique EV preference classes (EV-enthusiast, PHEV-oriented, ZEV [zero-emission vehicle]-neutral, HEV-oriented, and ICEV-oriented), based on 2123 respondents in Canada. Gong et al. (2020) focused on Australia (n = 1076) and developed LC models with five classes to examine heterogeneous consumer preferences for EV attributes and incentive policies. Although the LC models reveal distinct preference profiles across different consumer segments, one limitation of the LC models is that the preference coefficient associated with each attribute is fixed within each class, making it less flexible than MXL models.

There has been consensus that the more advanced models (i.e., MXL and LC) are superior to standard MNL models in examining EV preferences. It is however unknown how the MXL and LC models compare with each other in uncovering EV preferences and eliciting policy-related indicators, since very few EV preference studies estimate multiple advanced models and contrast the findings. Exceptions are Sheldon et al. (2017), Guerra and Daziano (2020), and Li et al. (2020) which estimated both MXL and LC models. Findings from these studies suggest that both models provide valuable insights for preference heterogeneity and are complementary in understanding EV preferences. Beyond the EV adoption preferences domain, there are many studies that contrasts the LC models with the MXL models in application to various application contexts, including travel route choices (Greene and Hensher, 2003), travel mode choices (Shen, 2009; Dong and Koppelman, 2014), traffic crash severity modeling (Cerwick et al., 2014; Li et al., 2019), preferences for the quality of electricity supply (Julian, 2017), WTP for water quality improvement in the Caribbean coastal (Beharry-Borg and Scapra, 2010). Results from these comparison studies show substantial differences in model fit, prediction performance, and behavioral interpretation between MXL and LC models. For example, Dong and Koppelman (2014) found that the LC model was able to recover preference heterogeneity that is masked by the MXL model. Beharry-Borg and Scapra (2010) found that the MXL model better represents the unobserved heterogeneity for their non-snorkeler samples.

Considering the individual limitations of the MXL and LC models, recent advances in choice modeling propose the LC-MXL modeling framework by combining discrete and continuous heterogeneity representations of preferences. The LC-MXL model accounts for preference heterogeneity in two ways (Bujosa et al. 2010): 1) identifying different preference classes as a function of respondent-specific characteristics, and 2) allowing for within-class heterogeneity via random coefficients. The LC-MXL framework has been applied in various settings, including modeling revealed preferences for forest sites for recreation in Mallorca, Spain (Bujosa et al. 2010), preferences for land-use policies in the Netherlands (Boeri, 2011); SP freight trip choice in Australia (Greene and Hensher, 2013), preferences for renewable energy generation sources (Yoo and Ready, 2014), willingness-to-pay (WTP) for different levels of vehicle automation (Daziano et al. 2017), willingness to improve the resilience of New York City's transportation system (Wang et al. 2018), preference heterogeneity for treatment among people with Type 2 diabetes (Zhou and Bridges, 2019). All seven studies above conclude that the LC-MXL model outperforms the LC and MXL models in fitting their choice data (often based on AIC, BIC, or adjusted McFadden-R^2). In a more extensive study, Keane and Wasi (2013) compared the LC-MXL models to the MXL, LC, and other scale-adjusted logit models in analyzing 10 datasets covering a wide range of settings, including medical decision-making and choice of products such as pizza delivery, mobile phones, etc. The authors found the LC-MXL model was preferred when the preference heterogeneity structure was complex. However, the advantage of the LC-MXL model disappeared when the dataset had relatively simple heterogeneity patterns.

In summary, this paper identifies two major gaps in existing EV preference studies. First, considering the complex EV preferences structure, it is worthwhile to examine whether a hybrid model (i.e., LC-MXL) helps better understand EV preferences. Second, existing EV studies lack in discussing the role of these advanced modeling approaches in uncovering EV preferences and eliciting policy-related measures. This paper fills these gaps by estimating MXL, LC, and LC-MXL models based on a SP survey conducted in Virginia, 2018 (n = 837). Model fit, prediction accuracy, behavioral interpretation, and policy implications are contrasted across different models. Findings provide insights into the usefulness of each model framework within the context of EV preference, and are valuable for policy-makers and EV manufacturers to better understand consumer preferences to promote mass EV adoption.

Table 3 - 1. EV Preference Studies Using Mixed Logit (MXL) or Latent Class (LC) Models

Authors (year)	Study region	Survey year	Sample size	Model	Distribution of random parameters / error components	# of classes
Rotaris et al. (2021)	Italy and Slovenia	2018	1934	MXL [1]	Normal	-
Giansoldati et al. (2020)	Italy	2017	200	MXL [1]	Normal	-
Danielis et al. (2020)	Italy	2018	996	MXL	Normal	-
Giansoldati et al. (2018)	Italy	2017	318	MXL	Constrained triangular	-
Valeri and Danielis (2015)	Italy	2013	121	MXL	Constrained triangular	-
Liao et al. (2018)	Netherlands	2016	1003	LC	-	5
Rasouli and Timmermans (2016)	Netherlands	2012	726	MXL	Normal	-
Kim et al. (2014) [2]	Netherlands	2012	726	MXL [1]	Normal	-
Hoen and Koetse (2014)	Netherlands	2011	1903	MXL	Normal	-
Cherchi (2017)	Denmark	2014-2015	2363	MXL [1]	Normal	-
Jensen et al. (2013)	Denmark	2012	369	MXL [1]	Normal	-
Mabit and Fosgerau (2011)	Denmark	2007	2146	MXL	Normal	-
Hackbarth and Madlener (2016)	Germany	2011	711	LC	-	6
Hackbarth and Madlener (2013)	Germany	2011	711	MXL	Normal	-
Langbroek et al. (2016)	Sweden (Stockholm)	2014	294	MXL	Normal	-
Glerum et al. (2014)	Swiss	2011	593	MXL [1]	Normal	-
Noel et al. (2019)	Nordic region	2016-2017	4105	MXL	Normal	-
Jia et al. (2021)	US (Virginia)	2018	837	MXL	Normal	-
Guerra and Daziano (2020)	US (Philadelphia)	2018	1545	MXL, LC	Normal	2
Sheldon et al. (2017)	US	2013	1261	MXL, LC	Normal, lognormal	3
Cirillo et al. (2017)	US (Maryland)	2014	456	MXL	Normal, lognormal	-
Helveston et al. (2015)	US and China	2012-2013	832	MXL	Normal	-
Tanaka et al. (2014)	US and Japan	2012	8202	MXL	Normal	-
Maness and Cirillo (2012)	US (Maryland)	2010	141	MXL	Normal	-
Hidrue et al. (2011)	US	2008-2009	3029	LC	-	2
Brownstone and Train (1998)	US (California)	1993	4654	MXL	Normal, lognormal	-
Kormos et al. (2019)	Canada	2017	2123	LC	-	5
Abotalebi et al. (2019)	Canada	2015	11539	LC	-	7
Ferguson et al. (2018)	Canada	2015	17953	LC	-	4

Dimatulac et al. (2018)	Canada	2016	1000	LC	-	4
Axsen et al. (2015)	Canada	2013	1754	LC	-	5
Li et al. (2020)	China	2018	394	MXL, LC	NA	2
Qian et al. (2019)	China	2015	1076	MXL	Normal	-
Ma et al. (2019)	China	2017	1719	MXL	Normal	-
Nie et al. (2018)	China (Shanghai)	2014-2015	760	MXL	Normal	-
Wang et al. (2017)	China	2015	247	MXL	Normal	-
Kim et al. (2020a)	South Korea	NA	532	MXL	Normal, lognormal	-
Kim et al. (2020b)	South Korea	NA	665	MXL	Normal, lognormal	-
Kim et al. (2019)	South Korea	2017	1000	MXL	Normal, lognormal	-
Choi et al. (2018)	South Korea	2016	1002	MXL	Normal	-
Byun et al. (2018)	South Korea	2016	615	MXL	Normal	-
Khan et al. (2020)	Japan	NA	500	MXL	Normal	-
Ito et al. (2019)	Japan	2011	2408	MXL	Normal	-
Gong et al. (2020)	Australia (New South Wales) [3]	2018	1076	LC	-	5
Ghasri et al. (2019)	Australia (New South Wales)	2018	1076	MXL [1]	Gumbel	-

Note: [1] Studies use integrated choice and latent variable models. For the discrete choice model part, mixed logit models are specified.
NA: not explicitly mentioned by the authors.
[2] the dataset is the same as Rasouli and Timmermans (2016).
[3] The dataset is the same as Ghasri et al. (2019)

3.3 Data and method

3.3.1 Survey

This analysis is based on a statewide SP vehicle fuel type choice survey conducted in 2018 in Virginia. Detailed survey design and distribution is described in Jia et al. (2021). The pivot portion of the SP survey presents choice experiments where a respondent is required to select its most preferred alternative from four vehicle options (ICEV, HEV, PHEV, and BEV) with varied attributes, as shown in **Figure 3 - 1**. The choice experiments are customized for each respondent by asking for their preferred vehicle body type before choice experiments begin, so that attribute levels can be designed to be relevant for their selected vehicle body type (i.e., subcompact/compact car, mid/full size car, small/medium SUV, standard/large SUV or minivan, and pick-up truck). Taking a mid/full size car for example, the attribute levels for each attribute are shown in **Table 3 - 2**. Note that about 6% of respondents select pick-up truck as their preferred body type and these respondents are excluded from taking choice experiments as there are no available pick-up truck EVs for survey design reference.

The formal survey was distributed to a general respondent pool through Survey Sample International (SSI) and a targeted EV owner pool via advertisements on relevant listservs (such as Facebook groups of EV owners, Virginia Clean Cities alliance, etc.). A total of 837 (including 66 EV owners) complete responses were kept after removing those who are non-drivers or do not finish the choice experiments. Since each respondent is asked to take six choice experiments, a total of 5022 choice observations are used for subsequent model estimation. The sample descriptive statistics along with Virginia population characteristics are shown in **Table 3 - 3**. Note that this study aims not to predict the exact EV market share in Virginia (which requires representative sample). Rather, the oversampling of EV owners enables a more heterogenous sample and facilitates examining EV preference heterogeneity.

Choice scenario 2 of 6

You mentioned that you would most likely purchase a Mid/Full size car for your next vehicle and that you plan to drive 9,001 to 11,000 miles annually for this vehicle. Please carefully review each vehicle and all its attributes below. And then please select the one vehicle you would **MOST LIKELY** purchase.

	Internal Combustion Engine	Hybrid Electric	Plug-in Hybrid Electric	Battery Electric
Vehicle Technical Specs				
Fuel economy	30 mpg	45 mpg	45 mpg (gas-mode)	
Annual tailpipe CO_2 emissions	2.96 tonnes	1.98 tonnes	0.83 tonnes	0
Battery-only range	0	0	30 miles	200 miles
Charging Station Availability				
Long distance travel charging *(Charging rate is 60-80 miles of range per 20 minutes of charging)*				
DC fast charging station spacing along interstate highways			Fast charging station every 100 miles	Fast charging station every 100 miles
Local charging station availability *(Charging rate is 10-20 miles of range per 1 hour of charging)*				
Major destinations *(workplace/school)*			NO	NO
Other destinations *(shopping center, restaurant, etc.)*			30% (1 in 3 destinations)	30% (1 in 3 destinations)
Cost				
One-time purchase cost				
Purchase price	$22,000	$27,000	$22,000	$27,000
Annual cost				
Fuel/charging*	$833	$556	$400	$339
Maintenance/repair	$819	$737	$737	$573
Electric vehicle use fee			$100	$200
One-time incentives				
Federal tax credit			$4,500	$7,500
State rebates			$0	$0

** Gasoline price and electricity price are assumed to be $2.5/gallon and 11.08 cents/kWh*

Please select the ONE vehicle you would most likely purchase

Figure 3 - 1. Typical choice experiment for mid/full size cars

Table 3 - 2. Attribute levels for the mid/full sized car body type

Attributes	ICEV	HEV	PHEV	BEV
Battery range (mile)	-	-	(1) 15 (2) 30	(1) 100 (2) 200 (3) 300
Fuel economy (mpg)	(1) 30 for ICEV, 45 for HEV (2) 40 for ICEV, 60 for HEV		Same as HEV (hybrid mode)	-
Annual tailpipe CO_2 emissions (ton) [1]	Calculated	Calculated	Calculated	0
DC fast charging stations spacing along interstate highways	-	-	(1) 40 miles (2) 70 miles (3) 100 miles	
Local charging station at workplace/school	-	-	(1) Yes, (2) No	
Local public charging stations at other destinations (restaurant, shopping center, etc.)	-	-	(1) 0% (2) 15% (1 in 6 destinations) (3) 30% (1 in 3 destinations)	
Purchase price ($)	22,000 (Base)	27,000	(1) Base + 11,000, (2) Base + 5,500, (3) Base	(1) Base + 10,000, (2) Base + 5,000, (3) Base
Annual fuel/charging cost ($)[2]	Calculated	Calculated	Calculated	Calculated
Annual maintenance cost ($)	Base (calculated)	Base × 90%	(1) Base × 90% (2) Base × 80%	(1) Base × 70% (2) Base × 50%
Annual EV use fee ($)	-	-	(1) 0 (2) 100	(1) 100 (2) 200
Federal tax credit ($)	0	0	(1) 0 (2) 4,500/7,500 [3]	
State rebates ($)	0	0	(1) 0 (2) 10% of purchase price	

Notes: [1] The use of tailpipe emissions instead of well-to-wheel emissions enables respondents to be more aware of functionality difference between EVs and ICEVs. The tailpipe emissions are calculated based on fuel economy and respondent reported VMT.

[2] Fuel cost is calculated based on fuel economy, respondent reported VMT and fuel price which has three levels ($2/gallon, $2.5/gallon, and $3/gallon); charging cost is calculated based on energy efficiency, VMT, and electricity price (assumed at 11.08 cents/kWh). Energy efficiency for EVs are average values from US DOE EV models. Due to the relative stable nature of electricity prices compare to gasoline prices, we do not vary electricity price levels.

[3] The amount of federal tax credit depends on battery capacity.

Table 3 - 3. Comparison between survey sample and general population

Variable	Sample (count)	Sample (percent)	Virginia general population (Percent)
Gender			
Male	376	45%	49%
Female	461	55%	51%
Age			
Less than 18 years	0	0%	22%
18 to 24 years	44	5%	10%
25 to 34 years	188	22%	14%
35 to 44 years	156	19%	13%
45 to 54 years	117	14%	13%
55 to 64 years	163	19%	13%
65 +	169	20%	15%
Education			
Less than high school	4	0%	10%
High school graduate (includes equivalency)	73	9%	24%
Some college, no degree	162	19%	19%
Associate's degree	73	9%	8%
Bachelor's degree	278	33%	22%
Graduate or professional degree	247	30%	17%
Income			
Less than $25,000	71	8%	16%
$25,000 - $34,999	74	9%	8%
$35,000 - $49,999	110	13%	11%
$50,000 - $74,999	135	16%	16%
$75,000 - $99,999	141	17%	13%
$100,000 - $149,999	149	18%	17%
$150,000 - $199,999	86	10%	8%
$200,000 - $249,999	37	4%	11% 1
$250,000 - $299,999	17	2%	
$300,000 and over	17	2%	
Household size			
1	158	19%	28%
2	336	40%	35%
3 or more	343	41%	38%
Number of vehicles			
0	38	5%	6%
1	302	36%	30%
2	345	41%	38%
3 or more	152	18%	26%
Housing tenure			
Own	618	74%	66%
Rent	208	25%	34%
Other	11	1%	
EV owners	66	8%	-

Note: [1] $200,000 and over.

3.3.2 Method

In a standard multinomial logit (MNL) model, the probability of individual n in choice scenario t choosing alternative j_{nt}^*, conditional on a vector of estimated coefficients β, is given by **Equation 1.**

$$P_{nt}(j_{nt}^*|\beta) = \frac{\exp(V_{ntj_{nt}^*})}{\sum_{j=1}^{J}\exp(V_{ntj})} \quad (1)$$

Where J is the total number of alternatives (ICEV, HEV, PHEV, and BEV); $V_{ntj_{nt}^*}$ is the deterministic utility component given by $f(x_{ntj_{nt}^*}, \beta, z_n)$, where $x_{nt\ {}_{nt}^{*}}$ is the attributes of alternative j_{nt}^* as faced by individual n in choice scenario t, and z_n is the vector of individual-specific characteristics.

The MXL model extends the MNL model to allow for random coefficients (Train, 2009). Assuming β_n to be the taste coefficients for individual n and to be continuously distributed over individuals with density $f(\beta|\Omega)$ (Ω is a vector of parameters of this continuous distribution), the probability of the sequence of observed choices for individual n and the log-likelihood function for the entire observations are given in **Equation 2** and **Equation 3**, respectively:

$$P_n(\Omega) = \int_{\beta} \prod_{t=1}^{T} P_{nt}(j_{nt}^*|\beta) f(\beta|\Omega) d\beta \quad (2)$$

$$LL(\Omega) = \sum_{n=1}^{N} ln\ (P_n(\Omega)) \quad (3)$$

Where j_{nt}^* is the alternative chosen by individual n in choice scenario t; $P_{nt}(j_{nt}^*|\beta)$ is the standard logit probability of the observed choice for individual n in choice scenario t, conditional on β; T = 6 since each individual in this survey complete six choice tasks; and N= 837 representing the total number of respondents in this analysis.

For the LC model, it is assumed that individuals are implicitly assigned into a set of S discrete classes depending on their characteristics (e.g., socio-demographics). The preference coefficients differ across classes but are fixed within each class. Let $P_{nt}(j_{nt}^*|\beta_s)$ be the standard logit probability of the observed choice for individual n in choice scenario t, conditional on individual n belonging to class s, the likelihood of the sequence of observed choices for individual n is given by **Equation 4**:

$$L_n(\beta) = \sum_{s=1}^{S} \pi_{ns} [\prod_{t=1}^{T} P_{nt}(j_{nt}^*|\beta_s)] \quad (4)$$

Where β_s is taste coefficients associated with class s; π_{ns} is the class allocation probabilities which is given by **Equation 5**.

$$\pi_{ns} = \frac{\exp(\delta_s + g(\gamma_s, z_n))}{\sum_{q=1}^{S}\exp(\delta_q + g(\gamma_q, z_n))} \quad (5)$$

Where, δ_s are classes specific constants; γ_s is estimated coefficients representing the impacts of individual characteristics z_n on the class assignment probabilities, and $g(.)$ represents the utility function for the class assignment model.

Lastly, the LC-MXL model combines latent class with continuous random preference heterogeneity. The contribution of individual n to the likelihood function is obtained by integrating the within-class heterogeneity and then averaging across discrete classes, as represented in **Equation 6**. For the MXL and LC-MXL models which require simulation to estimate coefficients, maximum simulated likelihood estimation with 500 Halton draws was used to simulate the probability. All models were implemented in the apollo package in R (Hess and Palma, 2019).

$$L_n(\Omega) = \sum_{s=1}^{S} \pi_{ns} \int_{\beta_s} [\prod_{t=1}^{T} P_{nt}(j_{nt}^* | \beta_s)] f_S(\beta_s | \Omega_s) d\beta_s \tag{6}$$

3.4 Model results

3.4.1 MXL model

Mixed logit models specify the coefficients for EV-specific attributes (battery range, DC fast charging station availability, workplace charging availability, and local public charging station availability) to be random. Three versions of MXL models are estimated, as shown in **Table 3 - 4**. The MXL-N model specifies a normal distribution for random coefficients. The MXL-LN model allows the coefficient for battery range to follow lognormal distribution. The last MXL-LN-INT model explores systematic heterogeneity by interacting individuals' characteristics with EV-specific attributes. The sign of estimated coefficients in the three MXL models are consistent. On average, respondents prefer greater battery range, greater availability of different types of charging infrastructure, and more EV monetary incentives, while show negative preferences for higher purchase price and fuel cost.

In the MXL-N model, the estimated standard deviation of the battery range coefficient is statistically significant, indicating that preferences for battery range do indeed vary among respondents. The battery range coefficient is distributed with an estimated mean of 0.52 and an estimated standard deviation of 0.45, indicating that 87% of respondents positively value greater battery range. However, the reminder 13% of respondents place a negative coefficient on battery range, which is surprising since all respondents should theoretically prefer greater battery range. This may be due to the assumption of a normal distribution, which allows a coefficient distributed on both sides of zero. A one-sided distribution assumption (e.g., lognormal), which allows all consumers have the same sign for their taste coefficients, was also tested in this analysis (and discussed later in this section). Lastly, preferences for the three types of charging infrastructure are not found to vary, as shown by the highly insignificant standard deviations for the three charging infrastructure coefficients.

The MXL-LN model imposes a sign restriction on the battery range coefficient by assuming a log-normal distribution. As shown in **Table 3 - 4**, the MXL-LN model shows that the estimated mean (u) and standard deviation (s) for the logarithm of battery range coefficient are -

0.87 and 0.72, respectively. The mean and standard deviation of the battery range coefficient itself are then derived from the estimates of u and s, by $exp(u + s^2/2)$ and $\sqrt{exp(2u + s^2)[exp(s^2) - 1]}$, which yield 0.54 (t = 4.35) and 0.44 (t = 13.65), respectively. The AIC and BIC values of the MXL-LN model are 8051 and 8162, respectively, which are slightly lower than its counterpart model (MXL-N) with all normally distributed coefficients, possibly because the MXL-LN model avoids negative coefficients for battery range, a likely better reflection of reality.

The MXL-LN-INT model expands the MXL-LN model by including individual characteristics into the utility functions to explore systematic heterogeneity. Since the MXL-LN and MXL-LN-INT models are nested, the likelihood ratio test can be conducted and results reject the MXL-LN model. The interaction terms between alternative specific constants and individuals' socio-demographics show that younger males with higher educational attainment and higher household income looking for a subcompact/compact car are more likely to show interest in EVs. Additionally, the EV-specific attributes are interacted with the EV owner dummy. Results show that EV owners are more sensitive to the battery range and charging infrastructure availability than non-EV owners, though the two groups' differences in charging infrastructure preferences are statistically insignificant (but with practical significance).

Table 3 - 4. Mixed logit model (MXL) results

	Alt.	MXL-N		MXL-LN		MXL-LN-INT	
		Est.	Std. error	Est.	Std. error	Est.	Std. error
Means							
Alternative specific constant	ICEV	0.46	0.29	0.38	0.28	2.75	0.74
Alternative specific constant	HEV	-0.38	0.17	-0.38	0.17	0.08	0.31
Alternative specific constant	BEV	-4.37	0.74	-4.22	0.70	-5.16	0.83
Battery range	BEV	0.52	0.14	-0.87	0.32	-0.51	0.25
DC fast charging stations every 40 miles along interstate highways [1]	BEV	0.45	0.12	0.46	0.12	0.39	0.13
Workplace charging availability	BEV	0.30	0.11	0.29	0.11	0.24	0.12
Local public charging stations available at one in three destinations [2]	BEV	0.24	0.11	0.24	0.11	0.19	0.12
Purchase price (per $1,000)	ALL	-0.16	0.01	-0.16	0.01	-0.16	0.01
Fuel cost (per $1,000)	ALL	-1.30	0.29	-1.24	0.29	-1.14	0.30
Incentive (per $1,000)	ALL	0.14	0.02	0.14	0.02	0.14	0.02
Standard deviations							
Alternative specific constant	ICEV	6.13	0.45	6.11	0.40	5.59	0.39
Alternative specific constant	HEV	3.39	0.24	3.46	0.25	3.33	0.22
Alternative specific constant	BEV	1.81	0.17	1.87	0.23	2.57	0.23
Logarithm of battery range	BEV	-0.45	0.04	0.72	0.13	0.34	0.10
DC fast charging stations	BEV	0.06	0.15	-0.04	0.14	0.02	0.16
Workplace charging availability	BEV	-0.08	0.13	-0.05	0.10	-0.07	0.11
Local public charging stations	BEV	0.02	0.09	0.08	0.10	-0.07	0.12
Interaction terms							
Male	PHEV					0.66	0.30
Male	BEV					0.57	0.45
Age <= 35	Non-ICEV					1.28	0.69
Age >= 56	Non-ICEV					-1.50	0.67
Bachelor's degree or higher	HEV					2.50	0.80

Bachelor’s degree or higher	PHEV			2.51	0.69
Bachelor’s degree or higher	BEV			2.37	0.84
Household income > $100,000	HEV			0.75	0.68
Household income > $100,000	PHEV			1.13	0.61
Household income > $100,000	BEV			1.55	0.77
Subcompact/compact car	HEV			0.96	0.74
Subcompact/compact car	PHEV			1.28	0.61
Subcompact/compact car	BEV			1.10	0.70
EV owners × Battery range	BEV			0.68	0.14
EV owners × DC fast charging stations every 40 miles along interstate highways	BEV			0.49	0.38
EV owners × Workplace charging availability	BEV			0.43	0.37
EV owners × Local public charging stations available at one in three destinations	BEV			0.51	0.38
Model fit					
Number of individuals		837	837	837	
Number of observations		5022	5022	5022	
Number of coefficients		17	17	34	
Log likelihood at convergence		-4010	-4008	-3936	
Rho-square		0.4240	0.4242	0.4347	
Adj. Rho-square		0.4216	0.4218	0.4298	
AIC		8054	8051	7940	
BIC		8165	8162	8161	

[1] The base level is 100-mile spacing. DC fast charging stations every 70 miles is not statistically significant compared to the base level.
[2] The base level is no local public charging stations availability. Local public charging stations available at one in six destinations is not statistically significant compared to the base level. Akaike information criterion (AIC); Bayesian information criterion (BIC)
* p < 0.1. ** p < 0.05. *** p < 0.01

3.4.2 LC and LC-MXL model

For LC models, the number of classes is pre-specified by the analysts. As shown in **Table 3 - 5**, we run a series of LC models to determine the most suitable number of classes by comparing model fit (e.g., AIC, BIC) and behavioral interpretability (the sign and magnitude of the coefficient estimates). When the number of classes increase incrementally from 2 to 5, model fit performance improves accordingly. However, LC models with four or more classes show estimates with large standard errors, indicating potential overfitting issues. Thus, the final LC model specification chooses a three-class solution.

Table 3 - 5. Determining the number of classes for the LC model.

# of Classes	Log-likelihood	Rho-square	Adj. Rho-square	AIC	BIC	# of estimated parameters
2	-5028	0.278	0.274	10110	10286	27
3	**-4400**	**0.368**	**0.362**	**8887**	**9174**	**44**
4	-4113	0.409	0.400	8348	8746	61
5	-3985	0.428	0.416	8126	8634	78

The LC model in **Table 3 - 6** shows the preference coefficients and class assignment coefficients for each latent class. Battery range and purchase price coefficients are statistically significant across all three classes, though the magnitude of estimates varies greatly. The coefficients for other attributes, however, are significant for only one or two classes. These coefficients' variations in magnitude and statistical significance affirm heterogenous preferences for EVs.

As the base class to which the other two classes are compared, class 1 members are influenced by battery range and all monetary attributes (purchase price, fuel cost, and incentives) but exhibit invariance to the charging infrastructure provision, up to a class membership probability. Class 1 respondents could be termed "cost-conscious", in that they are the most sensitive to the three monetary attributes compared to the other two classes.

Class 2 members consider the least number of attributes in choice experiments. They express sensitivity to only battery range and purchase price. Charging infrastructure and EV incentives are highly insignificant, perhaps as a result of their low probability of selecting EVs. Notably, the class 2 segment is the most sensitive to battery range among all three classes, which suggest that this segment can be coined the "high range anxiety" class. According to class assignment model results, respondents who are younger and have greater educational attainment are more likely to be assigned to this "high range anxiety" class.

Class 3 could be thought of as the "charging infrastructure-aware" class, given that it is the only class which shows sensitivity to charging infrastructure availability, with significantly positive preferences for greater availability of DC fast charging stations, workplace charging, and local public charging stations. Members from this class consider EVs seriously in vehicle fuel type choice decision making, since they are motivated by all of the EV-specific attributes. Compared to the "cost-conscious" class (the base class), the probability of being a member of "charging infrastructure-aware" class increases with being male, younger age, greater educational attainment, higher income, and purchase intent of buying a small car (relative to large vehicles).

The LC model assumes fixed coefficients within each latent class. This paper further estimates a LC-MXL model which allows taste coefficients within each class to vary. As shown in **Table 3 - 6**, the LC-MXL model identifies two latent classes[1] (Non-EV-oriented class and EV-oriented class) with normally distributed coefficients for battery range and alternative specific constants within each class. The significant estimates of standard deviations indicate that there exists preference variation which remains unexplained by the LC models or MXL models alone, but can be captured by the LC-MXL model. Note that the classes across the LC and LC-MXL models are not equivalent and cannot be compared directly.

The LC-MXL model results show that respondents from non-EV-oriented class are indifferent to battery range and charging infrastructure variables, possibly because they are unlikely to consider EVs in vehicle fuel type choices. Notably, members of the non-EV-oriented class are very sensitive to monetary attributes (purchase price, fuel cost, and monetary incentives). For example, the purchase price coefficient for the non-EV-oriented class is more than four times of that in the EV-oriented class.

In contrast, respondents in the EV-oriented class are motivated by EV-specific attributes, with positive preferences for greater battery range and greater availability of charging infrastructure. Compared to the non-EV-oriented class, the probability of belonging to the EV-oriented class increases with being male, younger age, greater educational attainment, higher income, and purchase intent of a small car (as opposed to a mid/full size car, SUV, or van).

[1] A LC-MXL model with three latent classes did not prove to be stable, as represented by large standard errors and inflated coefficient estimates.

Table 3 - 6. Estimation results for LC and LC-MXL model

	Alternative	LC						LC-MXL			
		Class 1 "Cost-Conscious"		Class 2 "High Range Anxiety"		Class 3 "Charging Infrastructure-Aware"		Class a "Non-EV-oriented"		Class b "EV-oriented"	
		Est.	Std. err	Est.	Std. err	Est.	Std. err	Est.	Std. err	Est.	Std. err
Class-specific Preference Coefficients											
Alternative specific constant	ICEV	3.44	0.32	-0.51	0.39	-1.74	0.24	5.61	0.94	-2.16	0.38
Alternative specific constant	HEV	-0.37	0.48	1.80	0.27	-1.13	0.16	1.53	0.57	-0.87	0.17
Alternative specific constant	BEV	-5.47	2.10	-9.45	4.22	-2.30	0.54	-4.15	2.27	-6.29	1.30
Battery range	BEV	0.84	0.40	1.47	0.74	0.39	0.11	0.58	0.44	0.90	0.24
DC fast charging stations every 40 miles along interstate highways [1]	BEV	-0.79	0.46	0.02	0.64	0.33	0.08	0.49	0.34	0.40	0.18
Workplace charging availability	BEV	0.45	0.37	-0.36	0.64	0.19	0.07	0.49	0.29	0.25	0.16
Local public charging stations available at one in three destinations [2]	BEV	-0.24	0.40	-0.81	1.31	0.22	0.08	-0.38	0.29	0.47	0.16
Purchase price (per $1,000)	ALL	-0.18	0.03	-0.10	0.02	-0.10	0.01	-0.41	0.05	-0.10	0.02
Annual fuel cost (per $1,000)	ALL	-1.15	0.57	0.40	0.85	-0.39	0.34	-1.53	0.87	-1.06	0.46
Incentive (per $1,000)	ALL	0.15	0.04	0.04	0.04	0.12	0.02	0.30	0.05	0.10	0.03
Class-specific Preference Coefficients											
Alternative specific constant (σ) [3]	ICEV							7.74	1.06	2.32	0.41
Alternative specific constant (σ)	HEV							6.80	0.88	1.47	0.20
Alternative specific constant (σ)	BEV							1.76	0.47	3.37	0.46
Battery range (σ)	BEV							0.18	0.07	-0.23	0.20
Class assignment coefficients											
Constant				-0.93	0.26	-0.87	0.22			-1.11	0.28
Male				0.13	0.21	0.34	0.18			0.42	0.21
Age <= 35				0.68	0.27	0.73	0.23			0.67	0.24
Age >= 56				-0.26	0.24	-0.56	0.21			-0.69	0.23
Bachelor's degree or higher				0.47	0.22	0.81	0.19			0.40	0.23
Household income > $100,000				0.21	0.24	0.58	0.20			0.62	0.23
Subcompact/compact car				0.08	0.25	0.42	0.22			0.65	0.25
Class probabilities		37%		23%		40%		58%		42%	
Model fit											
Number of individuals		837						837			
Number of observations		5022						5022			
Number of estimated coefficients		44						35			
Log likelihood at convergence		-4400						-3909			

Rho-square	0.3681	0.4385
Adj. Rho-square	0.3617	0.4335
Akaike information criterion (AIC)	8887	7888
Bayesian information criterion (BIC)	9174	8117

[1] The base level is 100-mile spacing. DC fast charging stations every 70 miles is not statistically significant compared to the base level.
[2] The base level is no local public charging stations availability. Local public charging stations available at one in six destinations is not statistically significant compared to the base level.
[3] σ represents standard deviation estimates for random coefficients.
* $p < 0.1$. ** $p < 0.05$. *** $p < 0.01$.

3.5 Discussion

3.5.1 Goodness-of-fit and cross validation

The goodness-of-fit of different models developed in this paper is measured by log-likelihood, AIC, and BIC values. The three MXL models, though with different specifications, show quite a similar model fit. Contrary to the findings by Greene and Hensher (2003) and Shen (2009), but consistent with Keane and Wasi (2013), the LC model's fitting performance ranks last when compared to the MXL or LC-MXL models in this study. While its model fit performance is relatively poor, the LC model is useful in gaining an intuition for the structure of preference heterogeneity. Depending on the number of classes and the unique preferences exhibited by each class, the analyst can determine whether a more complex model (e.g., LC-MXL) is needed (Keane and Wasi, 2013). The LC-MXL shows the best model fit, as shown by the lowest AIC and BIC values among all models estimated in this paper. The model fit gain from the LC-MXL model is possibly due to the complex EV preference heterogeneity pattern among respondents that neither the MXL nor the LC model could capture well, as noted by Keane and Wasi (2013).

In addition to the goodness-of-fit, five-fold cross-validation is conducted to compare the out-of-sample prediction performance across different models. The Fitting factor and Brier score (Parady et al., 2021) are calculated for each model, following equation 7 and equation 8, respectively. The Fitting factor represents the average predicted probability for the chosen alternative, and has a range between zero and one. The larger the fitting factor, the better model predicts. The Brier score is a more comprehensive measure than the fitting factor, as the former considers not only the predicted probability of the chosen alternative, but also the non-chosen alternatives. The Brier score ranges from zero to two, and a lower Brier score represents better prediction performance. The cross-validation results are shown in **Figure 3 - 2**. Across the three MXL models, MXL-LN-INT model, which incorporates systematic heterogeneity, shows better prediction performance than the other two models. Furthermore, the MXL-LN-INT also predicts slightly better than the LC or LC-MXL models.

$$Fitting\ factor = \frac{1}{NT}\sum_{n=1}^{N}\sum_{t=1}^{T}\sum_{j=1}^{J}\widehat{P_{njt}} \times d_{njt} \tag{7}$$

$$Brier\ score = \frac{1}{NT}\sum_{n=1}^{N}\sum_{t=1}^{T}\sum_{j=1}^{J}(\widehat{P_{njt}} - d_{njt})^2 \tag{8}$$

Where, $\widehat{P_{njt}}$ is the model predicted probability for choosing alternative j for respondent n in scenario t; d_{njt} equals to 1 if the observed choice outcome for respondent n in scenario t is alternative j.

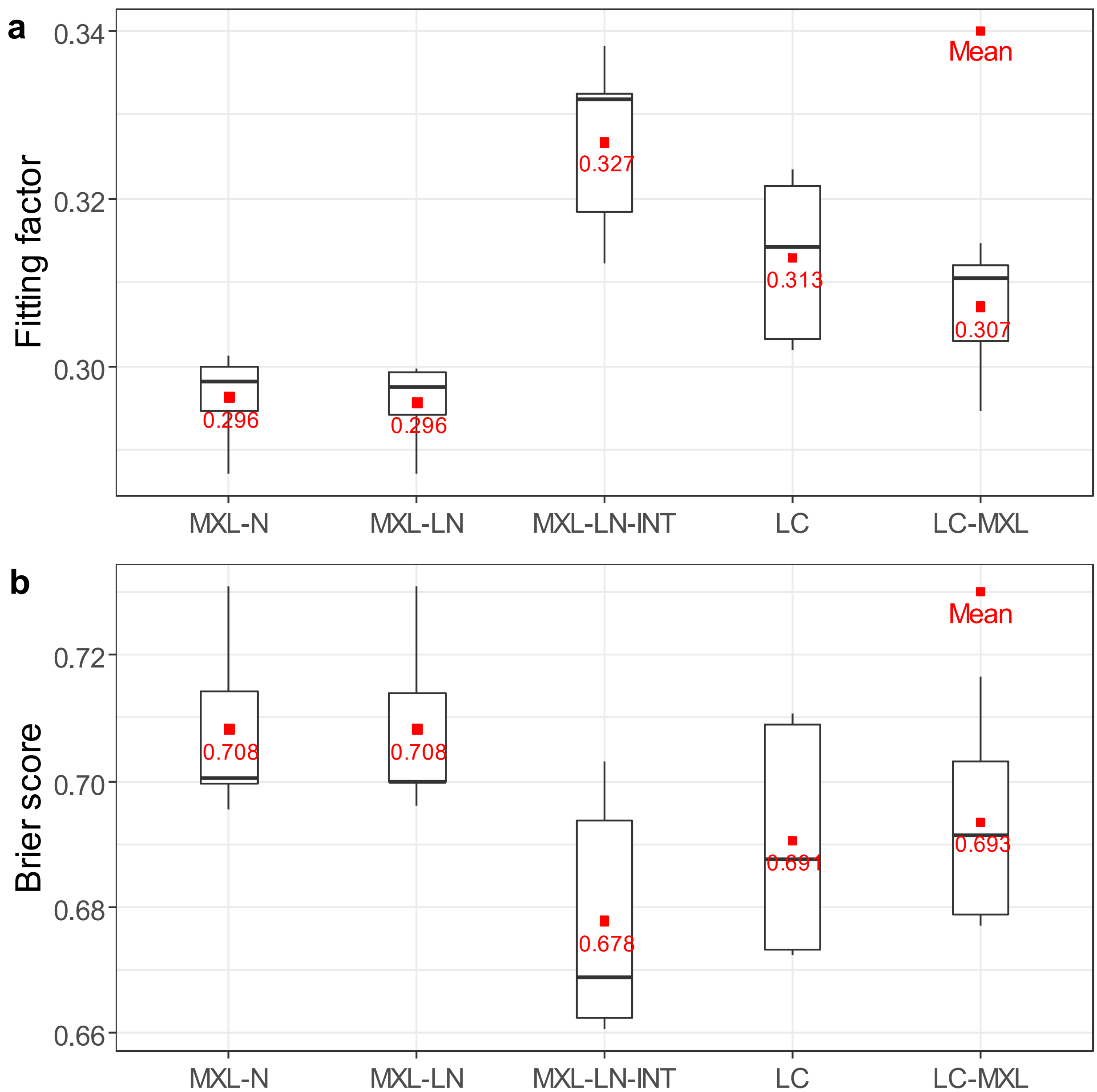


Figure 3 - 2. Cross validation results

3.5.2 Willingness-to-Pay (WTP) comparison

Moving beyond model fit, this section discusses behavioral interpretation based on different models. Specifically, the WTPs for EV attributes are compared across MXL-LN-INT, LC, and LC-MXL models. The battery range is specified using a logarithmic transformation in all models, under the assumption of a decreasing marginal WTP for battery range (Dimitropoulos et al., 2013; Hackbarth and Madlener, 2016; Daziano et al. 2017). Thus, the marginal WTP for battery range is dependent on the level of battery range. In **Figure 3 - 3**, the marginal WTP is evaluated at 200 miles. **Figure 3 - 3 (a)** shows the WTP calculated from the MXL-LN-INT model, which exhibits a log-normal distribution. The calculated mean and median of WTP for battery range are $21 and $19, respectively. The LC model results suggest that the “high range anxiety” class (accounts for 23% of respondents) has large marginal WTP ($77 for a one-mile increase in range, calculated at 200 miles), almost four times of the other two classes, as shown in **Figure 3 - 3 (b)**. Finally, the LC-MXL models, which combine the characteristics of LC and MXL models, suggest bimodality of preferences for battery range, as shown in **Figure 3 - 3 (c)**. Overall, the WTP for battery range uncovered in this study match results from a meta-analysis paper by Dimitropoulos et al. (2013) which reported significant valuation of battery range, with mean WTP for an additional mile of battery range varying from $20 to $200.

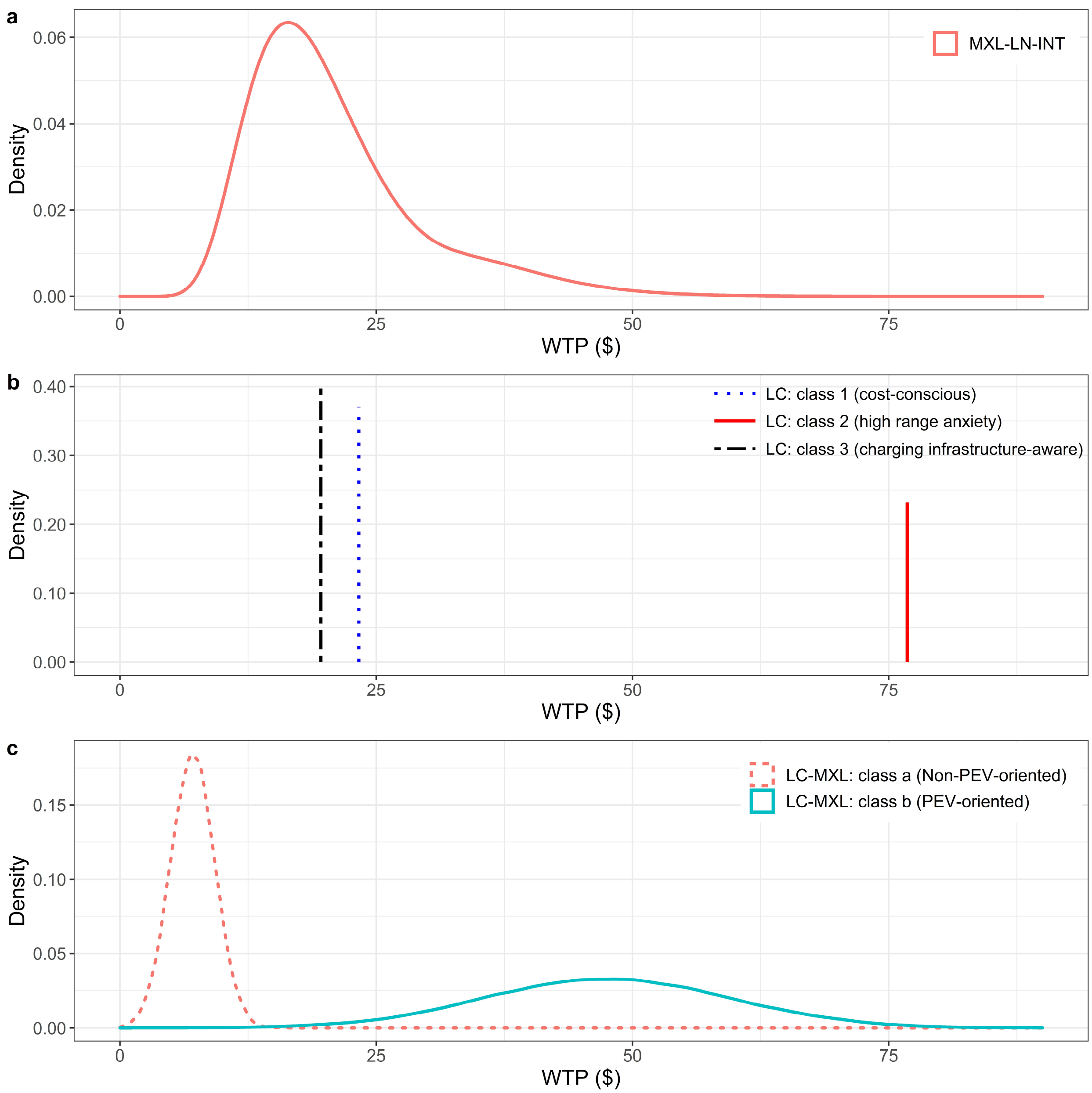


Figure 3 - 3. Marginal WTP for battery range evaluated at 200 miles.
(a) MXL-LN-INT. (b) LC. (c) LC-MXL

Figure 3 - 4 (a) to (c) shows the WTP for the availability of DC fast-charging stations, workplace charging, and local public charging stations, respectively. For MXL models, the heterogeneity in WTP for charging infrastructure is not significant, so that only mean WTP is discussed. Specifically, the MXL-LN-INT model shows that, on average, respondents are willing to pay approximately $2621 for DC fast-charging stations spaced every 40 miles along interstate highways (relative to a 100-mile spacing), about $1663 for workplace charging, and about $1393 for local public charging stations available at one in three local destinations (relative to no local public charging stations availability). In contrast, the LC model results show heterogeneous preferences for charging infrastructure as only a certain segment of respondents show significant preferences for charging infrastructure. Respondents of "charging infrastructure-aware" class (40% of the sample) would pay a statistically significant amount of about $3345, $1912, and $2191 for the same provision of DC fast-charging stations, workplace charging, and local public charging stations, respectively. For respondents from the other two classes ("cost-consciousness" and "high range anxiety") in the LC model, WTP for the three types of charging infrastructure is hardly significant. Similarly, the LC-MXL model results show that the EV-oriented class (42% of the sample) has significant WTP for charging infrastructure while the non-EV-oriented class does not. The authors note that the WTP for charging infrastructure availability should be interpreted cautiously as these results are related to the design of attribute levels. All model results indicate that DC fast charging station spacing of 70 miles and 100 miles are not valued differently by respondents, and that preferences between local public charging stations at one in six destinations and no local public charging stations do not show a significant difference. Future research should design the choice experiments covering finer tiers of charging infrastructure provision.

Respondents' monetary evaluation for a $1000 incentive is shown in **Figure 3 - 4 (d)**. Results of the MXL-LN-INT model show that respondents, on average, discount EV monetary incentives, with a WTP of approximately $885 for a $1000 incentive. On the other hand, the LC model results show that WTP for incentives varies greatly among consumer segments. Respondents in class 1 ("cost-conscious" class, 37% of the sample) evaluate the $1000 incentive at $846 while members in class 2 ("high range anxiety" class, 23% of the sample) do not show significant WTP for monetary incentives. One promising finding is that class 3 ("charging infrastructure-aware" class, 40% of the sample) in the LC model assigns a higher value ($1245) for a $1000 incentive. Perhaps individuals in this class interpret the EV incentive not only for monetary value but also as a sign of government support for EVs. Similarly, latent class model results in Ferguson et al. (2018) and Kormos et al. (2019) show that respondents of some classes place a greater dollar value for EV monetary incentives while the other classes discount incentives. Lastly, the LC-MXL model results indicate that EV-oriented members, on average, value a $1000 incentive at $1011 while the non-EV-oriented class at $722.

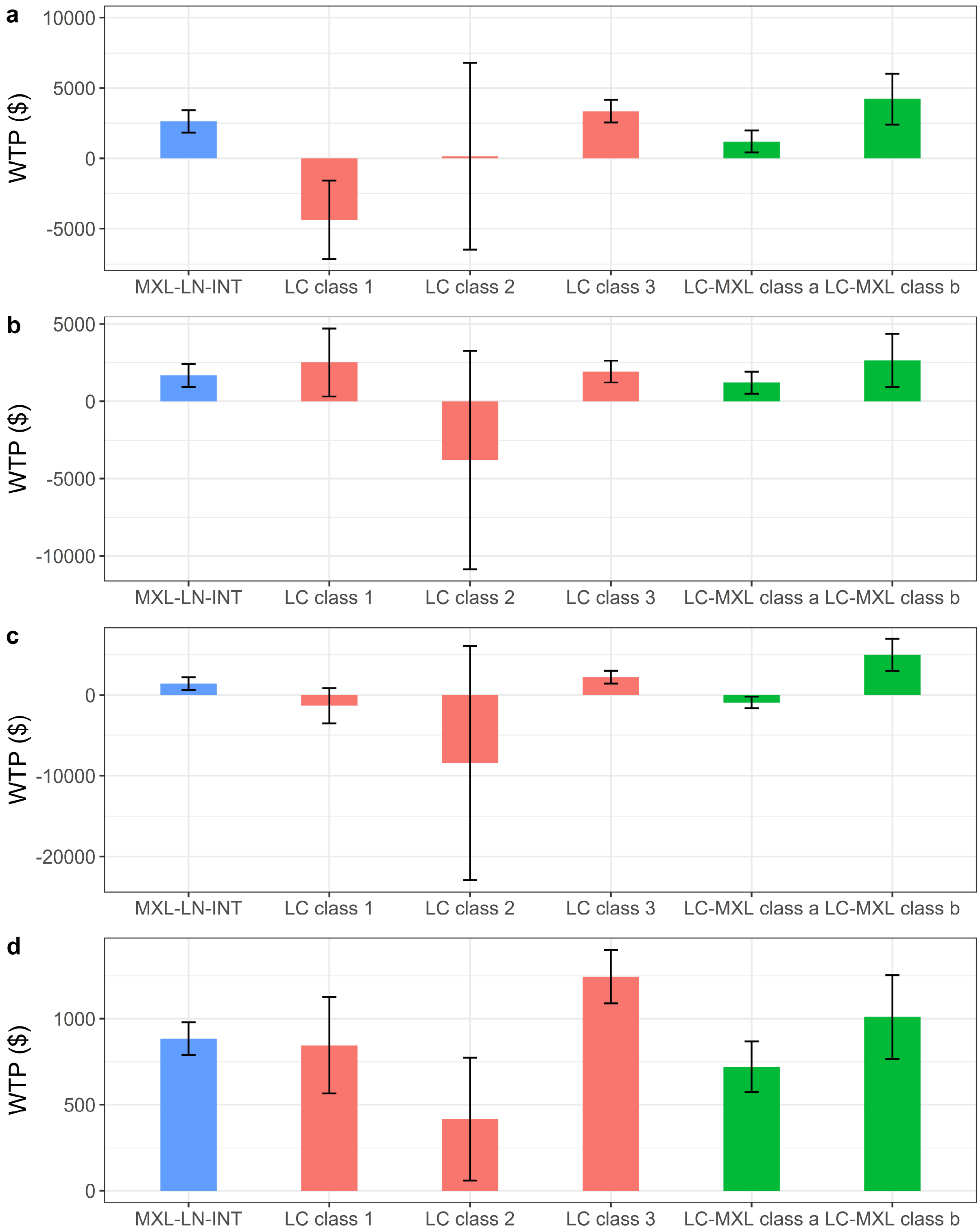


Figure 3 - 4. WTP for (a) DC fast charging stations spaced every 40 miles along interstate highways (relative to a 100-mile spacing). (b) workplace charging. (c) local public charging stations available at one in three local destinations (relative to no local public charging stations availability). (d) $1000 incentive

3.5.3 Policy scenario analysis

This section conducts market share simulations under various policy scenarios. We first develop a base scenario, which is defined as follows: 100-mile DC fast charging station spacing along interstate highways, no workplace charging available, no local public charging stations available, no EV monetary incentives, and all other attributes keeping the same as those presented to respondents in choice experiments. The first scenario (S1) represents "range improvement" where the BEV battery range increases by 50%. In the second scenario (S2), charging infrastructure is widely deployed, with 40-mile DC fast charging station spacing along interstate highways, access to workplace charging, and one in three local destinations having public charging stations. The third scenario (S3) provides a $7,500 incentive for EVs. **Figure 3 - 5** shows simulated market share under each policy scenario. Before discussing the detailed results, the authors note that this exercise aims not to predict the exact EV market share, but to show how different models inform policy implications.

Figure 3 - 5 (a) shows the market share results based on MXL-LN-INT, LC, and LC-MXL, respectively. For the LC and LC-MXL models, the market share is aggregated over all classes. Overall, the simulation results from different models are similar. According to the MXL-LN-INT model, both the BEV battery range improvement scenario and charging infrastructure deployment scenario show an increase in BEV share, with a decrease in PHEV share, relative to the base scenario. Merging PHEV with BEV, EV shares for the two scenarios increase from 31% (base scenario) to 32% and 34%, respectively. By contrast, the incentive scenario appears to be more effective. The $7500 EV incentive is associated with a ten percentage point increase in EV share, increase from 31% (base scenario) to 41%.

The LC model allows to examine policy effectiveness for each class, as shown in **Figure 3** - 5 **(b)**. For the "cost-conscious" class, the effects of battery range increase and greater charging infrastructure availability are minor, while providing monetary incentives increase EV share from 3% (base scenario) to 9%. For the "high range anxiety" class (with dominating preferences for HEVs and strongly dislike BEVs), none of the three scenarios show strong effects. The "charging infrastructure-aware" class is most responsive to the three policy scenarios, where EV share increases from 73% (base scenario) to 75%, 80%, and 87%, respectively. Particularly, for the charging infrastructure deployment scenario, BEV share alone increases by 16 percent point, from 31% (base scenario) to 47%, although with a nine percentage point decrease in PHEV share.

Lastly, as shown in **Figure 3 - 5 (c),** for the two latent classes from the LC-MXL models, the "non-EV-oriented" class is indifferent to the range increase and charging infrastructure deployment scenarios, while providing monetary incentives is associated with a seven percentage point increase in EV share, from 11% (base scenario) to 18%. For the "EV-oriented" class, battery range increase and charging infrastructure deployment scenarios increase BEV share by three and nine percent point, respectively, with a decrease in PHEV shares. When providing incentives, BEV share increase from 27% (base scenario) to 29%, and PHEV share increase from 36% (base scenario) to 44%.

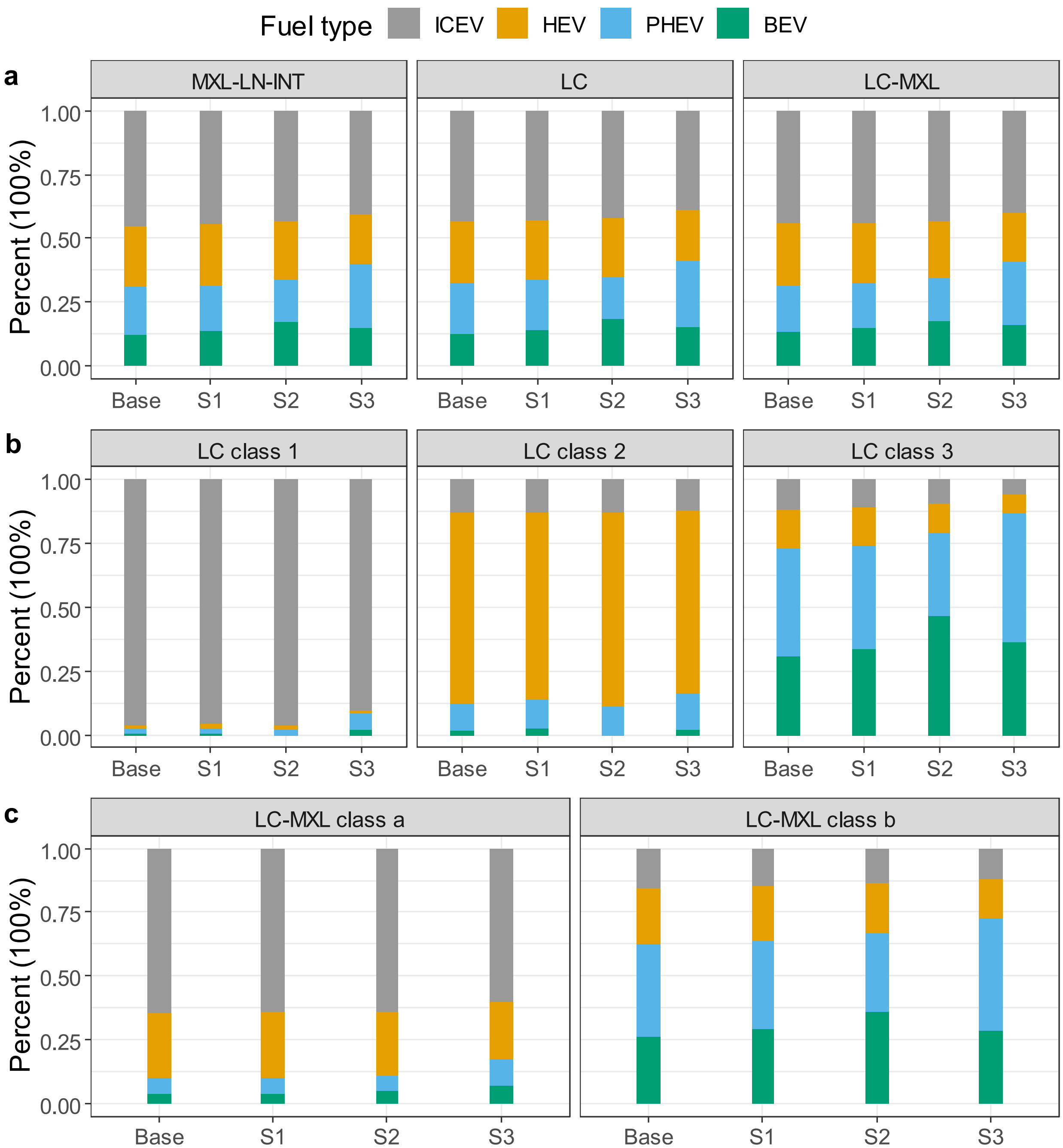


Figure 3 - 5. Fuel type market share simulation. (a) Overall market share based on different models. (b) market share for the three class from the LC model. (c) market share for the two classes in the LC-MXL model. S1: BEV battery range increases by 50%; S2: Greater access to charging infrastructure; S3: $7500 incentive for EVs.

3.6 Conclusions

This study investigates heterogeneous EV preferences based on stated choice experiment data. MXL, LC, and LC-MXL models are developed and model results are contrasted in terms of model fit, behavioral interpretation, and policy implications. Consistently, all models suggest that monetary incentives are most effective in improving EV market share, followed by greater charging infrastructure deployment, while the effectiveness of battery range improvement appears to be minor. Additionally, it is not reasonable to conclude a single model specification is unambiguously preferred. Rather, each model shows its strengths and limitations in uncovering EV preference profiles. The combined modeling analysis provides a more comprehensive picture of respondents' preferences for EVs which could inform more effective EV policy design and implementation.

The MXL models have the best prediction performance, and show an advantage in exploring systematic preference heterogeneity, since it can directly interact an EV attribute with a specific individual characteristic. For example, the MXL-LN-INT model results suggest that EV owners are more sensitive to battery range than non-EV owners. In contrast, the LC and LC-MXL models uncover the sources of preference heterogeneity in a less deterministic way. The preference profile of each class is linked to a series of individual characteristics, up to a probability.

The advantage of LC and LC-MXL models lies in capturing heterogeneous preferences for charging infrastructure, as the models can identify classes that are sensitive to the availability of the three types of charging infrastructure and classes that are not. Specifically, the "charging infrastructure-aware" class (40% of the sample) from the LC model and the "EV-oriented" class (42% of the sample) from the LC-MXL model express high sensitivity to charging infrastructure availability, suggesting the effectiveness of charging infrastructure deployment for these classes. Relative to the base scenario, greater availability of charging infrastructure increases the "charging infrastructure-aware" class' BEV market share from 31% to 47%, and the "EV-oriented" class from 27% to 36%.

The LC model identifies a "high range anxiety" class with the highest marginal WTP ($77) for battery range, which is approximately four times of the other classes. However, none of the three tested policy scenarios are found to be effective for this class. Specifically, a 50% increase in the BEV battery range is associated with only one percentage point increases in BEV market share, from 2% (base scenario) to 3%. This is possibly because of this class's strongly negative preferences for BEVs. More research is needed to study the fuel type choice motivations of this class, and informs policy for the electrification of this class.

The comparison of different models suggests that no model can unambiguously be superior to others in exploring EV preference heterogeneity. Although this result might be dataset-specific, it informs future research on EV preferences to discuss more on model selection, especially when delivering policy-related measures. For more prudent EV policy-making, policymakers can benefit from knowing that there are alternative models which may show different perspectives on policy implications.

In terms of study limitations, the authors note the hypothetical bias limitation of SP survey data. People's stated choice behavior may not represent real preferences particularly given their lack of direct EV experience, and thus the findings should be interpreted with caution. Secondly, vehicle choice is a complex consumer choice not completely captured by the variables considered in this analysis. Latent attitudes towards different fuel types, perception of brand image (popularity, reliability, etc.), availability of financing mechanisms for purchase or lease, crash safety ratings, risk of theft, insurance premiums, etc. are not included in this study targeted at analyzing generalized EV preferences.

Acknowledgements

This study is funded by Mid-Atlantic Transportation Sustainability University Center (MATS UTC) Grant No: 146221.

Author Contributions

The authors confirm contribution to the paper as follows: study conception and design: W. Jia, T.D. Chen; data collection: W. Jia, T.D. Chen; analysis and interpretation of results: W. Jia, T.D. Chen; draft manuscript preparation: W. Jia, T.D. Chen. All authors reviewed the results and approved the final version of the manuscript.

References

CARB, 2021. https://ww2.arb.ca.gov/resources/fact-sheets/governor-newsoms-zero-emission-2035-executive-order-n-79-20

Abotalebi, E., Scott, D.M. and Ferguson, M.R., 2019. Why is electric vehicle uptake low in Atlantic Canada? A comparison to leading adoption provinces. Journal of Transport Geography, 74, pp.289-298.

Axsen, J., Bailey, J. and Castro, M.A., 2015. Preference and lifestyle heterogeneity among potential plug-in electric vehicle buyers. Energy Economics, 50, pp.190-201.

Beck, M.J., Rose, J.M. and Hensher, D.A., 2013. Environmental attitudes and emissions charging: An example of policy implications for vehicle choice. Transportation Research Part A: Policy and Practice, 50, pp.171-182.

Beharry-Borg, N. and Scarpa, R., 2010. Valuing quality changes in Caribbean coastal waters for heterogeneous beach visitors. Ecological Economics, 69(5), pp.1124-1139.

Bujosa, A., Riera, A. and Hicks, R.L., 2010. Combining discrete and continuous representations of preference heterogeneity: a latent class approach. Environmental and Resource Economics, 47(4), pp.477-493.

Byun, H., Shin, J. and Lee, C.Y., 2018. Using a discrete choice experiment to predict the penetration possibility of environmentally friendly vehicles. Energy, 144, pp.312-321.

BC Gov News. 2018. *Provincial government puts B.C. on path to 100% zero-emission vehicle sales by 2040*. https://news.gov.bc.ca/releases/2018PREM0082-002226

GOV.BC.CA. 2019. Zero-Emission Vehicles Act. https://www2.gov.bc.ca/gov/content/industry/electricity-alternative-energy/transportation-energies/clean-transportation-policies-programs/zero-emission-vehicles-act

Boeri, M., 2011. Advances in Stated Preference Methods: Discrete and Continuous Mixing Distributions in Logit Models for Representing Variance and Taste Heterogeneity. Ph.D. Thesis. Queen's University of Belfast, School of Biological Sciences – Gibson Institute for Land, Food and the Environment.

Brownstone, D. and Train, K., 1998. Forecasting new product penetration with flexible substitution patterns. Journal of econometrics, 89(1-2), pp.109-129.

CA. GOV, 2018. *Governor Brown Takes Action to Increase Zero-Emission Vehicles, Fund New Climate Investments*. https://www.ca.gov/archive/gov39/2018/01/26/governor-brown-takes-action-to-increase-zero-emission-vehicles-fund-new-climate-investments/index.html

California New Car Dealers Association (CNCDA) (February 2021). "California Auto Outlook Covering Fourth Quarter 2020: State New Light Vehicle Registrations Predicted to Approach 1.8 Million Units in 2021" (PDF). *CNCDA*. Retrieved 2021-02-07. A total of 132,742 plug-in electric passenger cars (101,628 BEVs and 31,114 PHEVs) and were registered in 2020. Registration figures between 2016 and 2019 were revised, resulting in significantly higher estimates than previously reported thorugh 2019.

Cerwick, D.M., Gkritza, K., Shaheed, M.S. and Hans, Z., 2014. A comparison of the mixed logit and latent class methods for crash severity analysis. Analytic Methods in Accident Research, 3, pp.11-27.

Chen, T.D. and Paleti, R., 2018. Would You Consider a "Green" Vehicle? Anticipating Electric Vehicle Adoption Patterns and Emissions Impacts in Virginia. Mid-Atlantic Transportation Sustainability University Transportation Center University of Virginia.

Cherchi, E., 2017. A stated choice experiment to measure the effect of informational and normative conformity in the preference for electric vehicles. Transportation Research Part A: Policy and Practice, 100, pp.88-104.

Choi, H., Shin, J. and Woo, J., 2018. Effect of electricity generation mix on battery electric vehicle adoption and its environmental impact. Energy policy, 121, pp.13-24.

Cirillo, C., Liu, Y. and Maness, M., 2017. A time-dependent stated preference approach to measuring vehicle type preferences and market elasticity of conventional and green vehicles. Transportation Research Part A: Policy and Practice, 100, pp.294-310.

Daly, A., Hess, S. and de Jong, G., 2012. Calculating errors for measures derived from choice modelling estimates. Transportation Research Part B: Methodological, 46(2), pp.333-341.

Danielis, R., Rotaris, L., Giansoldati, M. and Scorrano, M., 2020. Drivers' preferences for electric cars in Italy. Evidence from a country with limited but growing electric car uptake. Transportation Research Part A: Policy and Practice, 137, pp.79-94.

Daziano, R.A., Sarrias, M. and Leard, B., 2017. Are consumers willing to pay to let cars drive for them? Analyzing response to autonomous vehicles. Transportation Research Part C: Emerging Technologies, 78, pp.150-164.

Dimatulac, T., Maoh, H., Khan, S. and Ferguson, M., 2018. Modeling the demand for electric mobility in the Canadian rental vehicle market. Transportation Research Part D: Transport and Environment, 65, pp.138-150.

Dimitropoulos, A., Rietveld, P. and Van Ommeren, J.N., 2013. Consumer valuation of changes in driving range: A meta-analysis. Transportation Research Part A: Policy and Practice, 55, pp.27-45.

EVAdoption, 2019. *EV Market Share - California*. https://evadoption.com/ev-market-share/ev-market-share-california/

Ferguson, M., Mohamed, M., Higgins, C.D., Abotalebi, E. and Kanaroglou, P., 2018. How open are Canadian households to electric vehicles? A national latent class choice analysis with willingness-to-pay and metropolitan characterization. Transportation Research Part D: Transport and Environment, 58, pp.208-224.

Giansoldati, M., Rotaris, L., Scorrano, M. and Danielis, R., 2020. Does electric car knowledge influence car choice? Evidence from a hybrid choice model. Research in Transportation Economics, 80, p.100826.

Glerum, A., Stankovikj, L., Thémans, M. and Bierlaire, M., 2014. Forecasting the demand for electric vehicles: accounting for attitudes and perceptions. Transportation Science, 48(4), pp.483-499.

Gong, S., Ardeshiri, A. and Rashidi, T.H., 2020. Impact of government incentives on the market penetration of electric vehicles in Australia. Transportation Research Part D: Transport and Environment, 83, p.102353.

Greene, W.H. and Hensher, D.A., 2003. A latent class model for discrete choice analysis: contrasts with mixed logit. Transportation Research Part B: Methodological, 37(8), pp.681-698.

Greene, W.H. and Hensher, D.A., 2013. Revealing additional dimensions of preference heterogeneity in a latent class mixed multinomial logit model. Applied Economics, 45(14), pp.1897-1902.

Hackbarth, A. and Madlener, R., 2013. Consumer preferences for alternative fuel vehicles: A discrete choice analysis. Transportation Research Part D: Transport and Environment, 25, pp.5-17.

Hackbarth, A. and Madlener, R., 2016. Willingness-to-pay for alternative fuel vehicle characteristics: A stated choice study for Germany. Transportation Research Part A: Policy and Practice, 85, pp.89-111.

Helveston, J.P., Liu, Y., Feit, E.M., Fuchs, E., Klampfl, E. and Michalek, J.J., 2015. Will subsidies drive electric vehicle adoption? Measuring consumer preferences in the US and China. Transportation Research Part A: Policy and Practice, 73, pp.96-112.

Hess, S., 2014. Latent class structures: taste heterogeneity and beyond. In Handbook of choice modelling (pp. 311-329). Edward Elgar Publishing Cheltenham.

Hess, S. and Palma, D., 2019. Apollo: a flexible, powerful and customisable freeware package for choice model estimation and application. Journal of Choice Modelling, p.100170.

Hidrue, M.K., Parsons, G.R., Kempton, W. and Gardner, M.P., 2011. Willingness to pay for electric vehicles and their attributes. Resource and energy economics, 33(3), pp.686-705.

Hoen, A. and Koetse, M.J., 2014. A choice experiment on alternative fuel vehicle preferences of private car owners in the Netherlands. Transportation Research Part A: Policy and Practice, 61, pp.199-215.

IEA, 2019. Global EV outlook 2019. International Energy Agency.

Jia, W. and Chen, T.D., 2021. Are Individuals' stated preferences for electric vehicles (EVs) consistent with real-world EV ownership patterns? Transportation Research Part D: Transport and Environment, 93, p.102728.

Keane, M. and Wasi, N., 2013. Comparing alternative models of heterogeneity in consumer choice behavior. Journal of Applied Econometrics, 28(6), pp.1018-1045.

Kim, J.H., Kim, H.J. and Yoo, S.H., 2019. Willingness to pay for fuel-cell electric vehicles in South Korea. Energy, 174, pp.497-502.

Kim, J., Seung, H., Lee, J. and Ahn, J., 2020a. Asymmetric preference and loss aversion for electric vehicles: The reference-dependent choice model capturing different preference directions. Energy Economics, 86, p.104666.

Kim, I., Kim, J. and Lee, J., 2020b. Dynamic analysis of well-to-wheel electric and hydrogen vehicles greenhouse gas emissions: Focusing on consumer preferences and power mix changes in South Korea. Applied Energy, 260, p.114281.

Kormos, C., Axsen, J., Long, Z. and Goldberg, S., 2019. Latent demand for zero-emissions vehicles in Canada (Part 2): Insights from a stated choice experiment. Transportation Research Part D: Transport and Environment, 67, pp.685-702.

Li, L., Wang, Z., Chen, L. and Wang, Z., 2020. Consumer preferences for battery electric vehicles: a choice experimental survey in china. Transportation Research Part D: Transport and Environment, 78, p.102185.

Liao, F., Molin, E. and van Wee, B., 2017. Consumer preferences for electric vehicles: a literature review. Transport Reviews, 37(3), pp.252-275.

Liao, F., Molin, E., Timmermans, H. and van Wee, B., 2018. The impact of business models on electric vehicle adoption: A latent transition analysis approach. Transportation Research Part A: Policy and Practice, 116, pp.531-546.

Ma, S.C., Xu, J.H. and Fan, Y., 2019. Willingness to pay and preferences for alternative incentives to EV purchase subsidies: An empirical study in China. Energy Economics, 81, pp.197-215.

Mabit, S.L. and Fosgerau, M., 2011. Demand for alternative-fuel vehicles when registration taxes are high. Transportation Research Part D: Transport and Environment, 16(3), pp.225-231.

Mannering, F.L., Shankar, V. and Bhat, C.R., 2016. Unobserved heterogeneity and the statistical analysis of highway accident data. Analytic methods in accident research, 11, pp.1-16.

Nie, Y., Wang, E., Guo, Q. and Shen, J., 2018. Examining Shanghai Consumer Preferences for Electric Vehicles and Their Attributes. Sustainability, 10(6), p.2036.

Noel, L., Carrone, A.P., Jensen, A.F., de Rubens, G.Z., Kester, J. and Sovacool, B.K., 2019. Willingness to pay for electric vehicles and vehicle-to-grid applications: A Nordic choice experiment. Energy Economics, 78, pp.525-534.

Parady, G., Ory, D. and Walker, J., 2021. The overreliance on statistical goodness-of-fit and under-reliance on model validation in discrete choice models: A review of validation practices in the transportation academic literature. Journal of Choice Modelling, 38, p.100257.

Qian, L., Grisolía, J.M. and Soopramanien, D., 2019. The impact of service and government-policy attributes on consumer preferences for electric vehicles in China. Transportation Research Part A: Policy and Practice, 122, pp.70-84.

Rasouli, S. and Timmermans, H., 2016. Influence of social networks on latent choice of electric cars: a mixed logit specification using experimental design data. Networks and Spatial Economics, 16(1), pp.99-130.

Revelt, D. and Train, K., 1998. Mixed logit with repeated choices: households' choices of appliance efficiency level. Review of economics and statistics, 80(4), pp.647-657.

Rotaris, L., Giansoldati, M. and Scorrano, M., 2021. The slow uptake of electric cars in Italy and Slovenia. Evidence from a stated-preference survey and the role of knowledge and environmental awareness. Transportation Research Part A: Policy and Practice, 144, pp.1-18.

Sheldon, T.L., DeShazo, J.R. and Carson, R.T., 2017. Electric and plug-in hybrid vehicle demand: lessons for an emerging market. Economic Inquiry, 55(2), pp.695-713.

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

Tanaka, M., Ida, T., Murakami, K. and Friedman, L., 2014. Consumers' willingness to pay for alternative fuel vehicles: A comparative discrete choice analysis between the US and Japan. Transportation Research Part A: Policy and Practice, 70, pp.194-209.

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

United States Environmental Protection Agency, 2019. Inventory of U.S. Greenhouse Gas Emissions and Sinks 1990 – 2017. EPA 430-P-19-001.

Valeri, E. and Danielis, R., 2015. Simulating the market penetration of cars with alternative fuel powertrain technologies in Italy. Transport Policy, 37, pp.44-56.

Wang, N., Tang, L. and Pan, H., 2017. Effectiveness of policy incentives on electric vehicle acceptance in China: A discrete choice analysis. Transportation Research Part A: Policy and Practice, 105, pp.210-218.

Wang, C., Sun, J., Russell Jr, R. and Daziano, R.A., 2018. Analyzing willingness to improve the resilience of New York City's transportation system. Transport Policy, 69, pp.10-19.

Yoo, J. and Ready, R.C., 2014. Preference heterogeneity for renewable energy technology. Energy Economics, 42, pp.101-114.

Zhou, M. and Bridges, J.F., 2019. Explore preference heterogeneity for treatment among people with Type 2 diabetes: A comparison of random-parameters and latent-class estimation techniques. Journal of choice modelling, 30, pp.38-49.

Chapter 4. Local Context Matters: Examining Mode Choice Preferences in the Autonomous Vehicle Era in Two U.S. Metropolitan Regions

Wenjian Jia [a], T. Donna Chen [a *], Wenwen Zhang [b], Kaidi Wang [c]
[a] Department of Engineering Systems & Environment, University of Virginia, PO Box 400742, Charlottesville, VA 22906, U.S.
[b] Edward J. Bloustein School of Planning and Public Policy, Rutgers University, 33 Livingston Avenue, New Brunswick, NJ 08901
[c] School of Public and International Affairs, Virginia Tech, 140 Otey St. Blacksburg, VA 24060

[*] Corresponding author: tdchen@virginia.edu

Abstract
Private autonomous vehicles (PAVs) and shared autonomous vehicles (SAVs) are two potential travel modes in the AV era. This paper examines travelers' preferences for these modes using integrated choice and latent variable (ICLV) models based on stated preference surveys distributed in the Seattle (n = 511) and Kansas City (n = 558) metropolitan regions. Model results for both regions suggest the importance of latent attitudes (e.g., attitudes towards the AV technology, willingness to share travel with strangers) and trip-related characteristics (e.g., trip purpose, trip congestion level) in explaining mode choice. First, socio-economic characteristics are found to be related to latent attitudes, but the relationships are inconsistent across the two study regions. Second, travelers are found to prefer to use PAVs for commute and congested trips, but the two study regions show inconsistent results regarding the sensitivity to in-vehicle travel time (IVTT). Seattle travelers penalize the IVTT in PAVs 14% less than in human-driven vehicles, whereas such a difference is not statistically significant for the Kansas City region. The inconsistent finding suggests the importance of local contexts in evaluating AV impacts, which should be explicitly considered in AV planning.

Keywords: Autonomous Vehicles, Travel Mode Choice, Stated Preference, Discrete Choice Models, Latent Attitudes.

4.1 Introduction

Autonomous vehicles (AVs) are expected to bring disruptive changes to the transportation system, urban landscape, energy system, and greenhouse gas emissions. Studies have shown that AVs have the potential to greatly reduce traffic crashes, decrease vehicle ownership, save parking spaces, and provide mobility options for traditionally under-served populations (Papadoulis et al., 2019; Zhang et al., 2018; Zhang et al., 2015; Harper et al., 2016; Zhang & Wang, 2020). However, researchers are also concerned that AVs may induce more vehicle miles traveled (VMT), exacerbate traffic congestion and urban sprawl, and produce more greenhouse gas emissions, especially in sprawled regions (Wadud et al., 2016; Wang & Zhang, 2021). The uncertainty of AV impacts largely depends on the adoption of different AV modes: private AVs (PAVs) or shared AVs (SAVs). The dream AV future is often described as one where all trips are served by SAVs with dynamic ride-sharing. This scenario, certainly, requires travelers' willingness to adopt AVs and to rely exclusively on shared mobility. It is, therefore, important to study travelers' preferences for these different modes and AV ownership (Zhang et al., 2020) to provide policy insights for encouraging the adoption of more sustainable AV modes.

Several critical questions emerge when studying AV mode preferences. First, how do different socio-economic groups' AV preferences vary? Answers to this question help ensure that all groups can enjoy the benefits of AV technology. The second question relates to the role of travelers' attitudes. Understanding the relationship between various attitudinal constructs and AV mode preferences informs policy interventions to encourage specific AV modes. Lastly, how are mode-specific attributes (e.g., in-vehicle travel time) perceived differently in AVs compared to human-driven vehicles (HVs)? Since the value of travel time (VOTT) is an important parameter in regional travel demand models, exploring the changes in VOTT in the AV era can help metropolitan planning organizations (MPOs) better assess the local impacts of AVs.

The above research questions have been examined in recent years in regions ranging from Europe to North America to China. As detailed in the literature review section, however, the current understanding of AV mode preferences is still limited, and existing findings are often mixed in terms of different socio-economic groups' AV preferences and the sensitivity to AV's in-vehicle travel time. This paper examines travelers' AV mode choice preferences with the following contributions. First, the study distributes the same survey instrument in two metropolitan areas in the U.S. (Seattle and Kansas City) with distinctly contrasting population densities and travel mode shares. Findings help MPOs understand how local contexts shape AV mode preferences. Second, unlike the other studies reviewed here, a best-worst stated choice experiment design is used, which helps elicit more information about travelers' trade-offs among different modes compared to the traditional "pick-one" choice experiments. This approach is particularly useful when there is a dominating alternative among the choice set. Lastly, this paper applies the integrated choice and latent variable (ICLV) modeling framework to the best-worst choice data. The ICLV models treat attitudes as latent variables, allowing for potential endogeneity and measurement errors of attitudes (Ben-Akiva et al., 2002; Daly et al., 2012).

4.2 Literature review

Because AVs and AV services are not widely available in the real world, the SP survey approach is most often used to study AV preferences. Some studies directly ask survey respondents about their adoption intentions of AVs, including surveys conducted in Texas (Bansal et al. 2016; Bansal and Kockelman 2018; Sener et al. 2019), Puget Sound Region (Lavieri et al. 2017; Wang and Akar 2019; Nazari et al., 2018; Asgari and Jin 2019), and across entire nations like the U.S. (Nodjomian and Kockelman 2019; Barbour et al. 2019; Wang et al. 2020), Australia (Cunningham et al. 2019), France (Payre et al. 2014), Spain (Montoro et al. 2019), China (Jing et al. 2019), and international contexts (Kyriakidis et al. 2015). These studies uncover a variety of variables (i.e., individual socio-demographic characteristics, land use, attitudes, etc.) associated with AV adoption intention. However, these studies generally do not provide survey respondents with specific mode choice contexts in eliciting their AV preferences.

Another branch of studies design SP surveys with travel mode choice experiments. The choice experiments are often designed based on a specific trip reported by respondents. Respondents are asked to imagine a future with AV modes, and then make mode choice decisions under the reference trip context. The mode choice set for each respondent often includes PAVs, SAVs, and/or conventional (human-operated) travel modes. The attributes describing each mode include trip time and cost, waiting time, etc. Discrete choice models are then used to explore influential factors related to mode choice outcomes, such as individual socio-economic characteristics, attitudes, and mode-specific attributes. Results of these influential factors are summarized as follows.

4.2.1 Individual socio-economic characteristics

Krueger et al. (2016) examined stated mode choices among the current mode, SAV, and SAV with pooling (SAVP), based on 435 respondents from major metropolitan areas of Australia. They found that young people and individuals with multimodal travel patterns were more likely to adopt SAVs. Also in Australia, Zhou et al. (2020) examined heterogeneous preferences for car-sharing and SAV modes, based on 1,500 survey responses across major Australian cities. The authors concluded that females, non-drivers, and seniors showed negative preferences for SAVs, even though these groups are the most likely to benefit from SAVs. In an SP survey (n = 663) conducted in the Netherlands, Ashkrof et al. (2019) found that middle-aged males taking longer-distance leisure trips were more likely to adopt ADTS (automated driving transport service). Tian et al. (2021) conducted an SP survey in China with 542 respondents and concluded that respondents, in general, were more likely to keep their current vehicles or to buy an AV than to use an SAV. Also, younger people were more likely to be interested in PAVs or SAVs. In the UK, Wadud et al. (2021) quantified the convenience value of owning a PAV relative to SAV ride-hailing services via examining preferences for PAV, SAV, and SAVP. With 800 respondents from London and Manchester, the study found that women showed significant WTP for ownership (£2020 per year) relative to the SAV mode, whereas men were found to be indifferent between owning AVs and using ride-hailing services. While these studies show heterogeneous AV preferences across different socio-economic groups, other studies suggest that socio-economic characteristics only play a minor role in AV mode choice preference. For example, Steck et al. (2018) found no socio-economic variables that were relevant to AV mode choice preferences based on 172 commuters from Germany. Also in Germany, Kolarova et al. (2019) found that age and gender appeared to be irrelevant to the AV preferences (n = 485).

A few studies designed choice experiments that specifically focused on the pooling behavior in an SAV. König and Grippenkoven (2020) examined travelers' willingness to share rides in Germany (n = 150) and found that more monetary compensation was needed for seniors and females to accept the SAVP mode. Based on survey data from the Dallas-Fort Worth metropolitan area (n = 1607), Lavieri and Bhat (2019) found that travelers showed lower sensitivities to the presence of strangers for commute trips than for leisure trips.

4.2.2 Attitudes

Existing studies have also shown the importance of attitudinal variables in explaining AV mode choice preferences, including trust in AVs, positive attitude toward AV efficiency, environmental consciousness (Ashkrof et al. 2019), perception of the convenience of AVs (Correia et al., 2019), pro-AV sentiments, enjoyment of driving (Haboucha et al. 2017), attitude towards public transit and organized time style (Etzioni et al. 2021), driving control, mobility control, safety concerns, tech-savviness (Asmussen et al. 2020), and attitude toward work productivity in AVs (Yap et al., 2016; Lavieri and Bhat, 2019).

In studies that consider attitudinal variables, the latent nature of attitudes has been well acknowledged. Two modeling approaches are often used to incorporate the latent attitudes. The first approach is a sequential estimation approach (see, e.g., Yap et al., 2016; Haboucha et al., 2017; Ashkrof et al., 2019). Factor analysis is first conducted to identify latent attitudes among a series of attitudinal measurement indicators. Then the latent attitudes enter the utility functions of the discrete choice models. The two steps are conducted in a sequential manner. Therefore, the estimation is inefficient since the latent attitudes are derived only from the attitudinal indicator information, without using the information from the discrete choice component (Daly et al., 2012). Additionally, more complex relationships among socio-economic characteristics and latent attitudinal constructs cannot be examined using this approach (Ben-Akiva et al., 2002).

Another approach uses the ICLV modeling framework (see, e.g., Lavieri and Bhat, 2019; Asmussen et al., 2020; Correia et al., 2019). The ICLV model allows the use of information from the discrete choice component to inform the estimation of latent attitudinal constructs (Daly et al., 2012). More importantly, the latent attitudinal constructs can be linked to exogenous socio-demographic variables. Therefore, the direct and indirect effects of socio-demographic variables on AV mode preferences can be identified. For example, Asmussen et al. (2020)'s study in Austin, Texas (n = 1,021) showed a direct age effect, where older adults were less likely to adopt AVs. Furthermore, the study also showed indirect age effects on AV adoption preferences through latent attitudes, such as driving control, safety concerns, and tech-savviness. These attitudes are related to AV preferences, and vary across age segments.

Beyond the AV mode choice studies, the ICLV model has also been used in various transportation research, such as traditional travel mode choice without AV modes (Atasoy et al., 2013; Hess et al., 2018), route choice (Alizadeh et al., 2019), rail travel preferences (Daly et al., 2012; Hess et al., 2013), walking and cycling preferences (Kamargianni and Polydoropoulou, 2013; Motoaki and Daziano, 2015), electric vehicle (EV) purchase preferences (Jensen et al., 2013; Kim et al., 2014; Giansoldati te al., 2020; Rotaris et al., 2021), and EV charging behavior (Pan et al., 2019).

4.2.3 Mode-specific attributes

Travel cost, travel time, and wait time are all found to be important influential factors of AV mode choices (Krueger et al., 2016). The sensitivity to in-vehicle time in AVs is of particular interest to researchers and transportation planners since the VOTT is a critical parameter to evaluate AV impacts in travel demand models. However, there exists mixed evidence on whether the in-vehicle time is perceived more or less negatively in AV modes compared to in HVs.

Steck et al. (2018) focused on commute trips in Germany (n = 172). The study found that PAVs (in autonomous mode) were associated with a 31% reduction in VOTT compared to the PAVs in manually-driven mode. Differently, Yap et al. (2016) reported that in-vehicle time in an AV was perceived more negatively than the in-vehicle time in a HV, based on 761 survey respondents from the Netherlands. Krueger et al. (2019) jointly modeled mode and housing choice based on survey responses from 512 commuters in the Sydney metropolitan area in Australia. Results suggested no significant change in VOTT in PAVs compared to HVs. Gao et al. (2019) used existing ride-hailing services as an analogy for AVs considering both modes free travelers from the effort of driving. Based on survey responses of 520 commuters in the U.S., the study found that, when the choice experiments displayed information that ride-hailing services were driverless, the VOTT in that driverless vehicle was 15% higher than in a HV. In contrast, when respondents were reminded about the multitasking capabilities in human-driven ride-hailing services, the VOTT in the ride-hailing services was about half of a HV.

Some studies report more nuanced results regarding VOTT shifts in the AV era, which can be related to trip purpose, distance, or vehicle interior design. With 485 respondents from Germany, Kolarova et al. (2019) found that PAVs were associated with a 41% reduction in VOTT compared to HVs for commute trips. In contrast, for leisure or shopping trips, the VOTT did not change significantly. In the Netherlands, Ashkrof et al. (2019) found that in-vehicle time in an ADTS was more pleasant than in a HV for short-distance commute trips. For long-distance commute trips, however, no difference in travel time sensitivity was found. Also in the Netherlands, Correia et al. (2019) examined the impacts of AV (with office or leisure interior) on the VOTT for commute trips. They found that the VOTT for an office-interior AV was 26% lower than that for a HV, whereas no significant difference in VOTT was found between leisure-interior AVs and HVs.

In summary, the influential factors associated with AV mode choice preferences are not yet well understood, especially the role of socio-economic characteristics and in-vehicle travel time. Mixed findings are often reported across existing studies, which are conducted in different regions with various modeling frameworks. In this paper, travel mode choice preferences in the AV era are examined in two distinct metropolitan areas in the U.S., using the same survey instrument and modeling approach.

4.3 Data

4.3.1 Survey overview

The data used for this analysis comes from an SP survey which consists of five parts: 1) eliciting a trip that respondents take on a typical week; 2) education materials to familiarize respondents to the concept of AVs; 3) travel mode choice experiments; 4) respondents' socio-economic characteristic and household attributes; and 5) respondents' attitudes toward AVs, willingness to share travel with strangers, trip location data privacy sensitivity, enjoyment of driving, and environmental consciousness. Responses to the attitudinal statements use a five-point Likert scale, such as ranging from "strongly disagree", "disagree", "Neither disagree nor agree", "agree" to "strongly agree".

The survey did not ask respondents to report their most frequent trips since doing so made it difficult to obtain adequate responses with trip purposes other than commuting. Instead, respondents were asked to report all trip purposes (including work/work-related business, school/daycare/religious activity, medical/dental services, shopping/errands, meals, social/recreational, and transporting someone else) that they undertook during a typical week before the COVID-19 outbreak [2]. To avoid excessive cognitive burden, the survey asked respondents to report detailed trip characteristics for only one trip, which was randomly chosen from the reported trip purposes. The trip characteristics included departure and arrival time (which also served as an attention check: those who reported the arrival time before the departure time were disqualified from continuing the survey), trip time (in-vehicle time and waiting time), trip distance (total distance and walk access distance), trip cost (fare and parking cost), trip mode, trip frequency, whether traveling alone or with other people, whether bring luggage/cargo, dwelling time at destination, duration of parking search, and overall congestion level rating for the trip. This elicited revealed preference trip served as the basis around which the AV mode choice experiments were designed, as described in section 3.2.

Before beginning the choice experiments, respondents were presented with educational materials, as detailed Appendix A. A 2-minute Waymo video (Waymo, 2016) was embedded in the educational material section to give respondents a general introduction to AVs. After watching the video, respondents were informed that new travel modes would become available: PAVs and SAVs. The SAV is described as a driverless taxi or an Uber/Lyft car without a driver. In addition, respondents were told they could pool with someone else when using a SAV.

[2] Since the survey was distributed amid the COVID-19 pandemic, as detailed in section 3.3, respondents were asked to recall their typical trips before the COVID-19 outbreak.

4.3.2 Choice experiments design

Instead of the typically used "pick-one" choice experiment design, a best-worst choice experiment format was used in this study, where respondents are asked to select their most and least preferred alternatives. The "best-worst" choice experiment design allows analysts to obtain more information than the "pick-one" format with the same number of choice scenarios, facilitating the evaluation of trade-offs across alternatives.

The stated choice experiments are tailor-made for each respondent based on his/her reported trip characteristics. Each respondent is asked to respond to six choice tasks. Across the six choice tasks, attributes of the HV alternative remain the same as respondents' reported values, but attributes of the three new modes (PAV, SAV, SAVP) vary. For each choice task, respondents are asked to select their most and least preferred modes among the four options: HV, PAV, SAV, and SAVP. The attributes describing the alternatives include in-vehicle time, waiting time, walking distance, travel cost, fares, and parking cost. Specific attribute levels for experiment design are shown in **Table 4 - 1**.

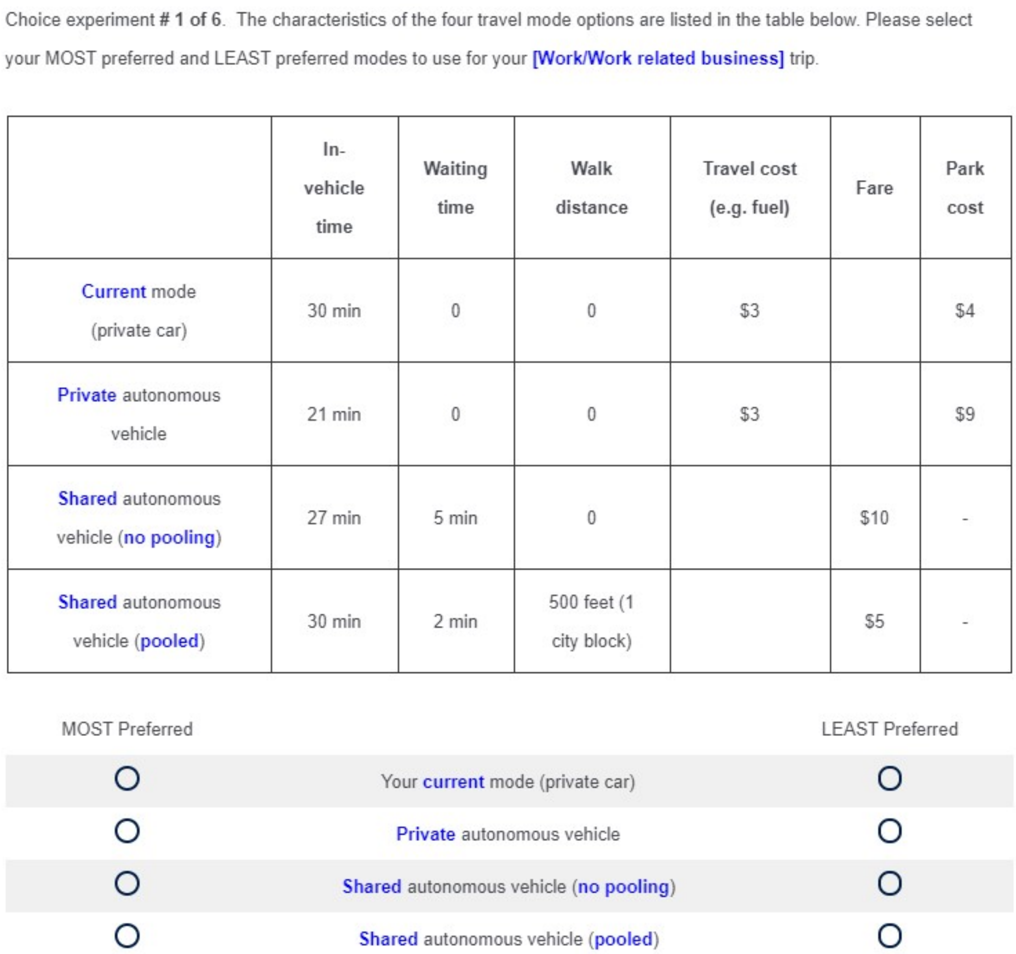
Note that the waiting time attribute is applied only to SAV and SAVP. Literature has shown that the waiting time of existing ride-hailing services is about 4.8 minutes to 5.6 minutes in Austin, Texas (Yang et al., 2021), under 3 minutes in dense urban cores and over 10 minutes in low-density distant suburbs in Chicago (Thebault-Spieker et al., 2017), 3 minutes to 10 minutes for UberX and 3 minutes to 13 minutes for UberBlack in Atlanta (Wang and Mu, 2018), 5.5 minutes for Lyft and 6.1 minutes for Uber, on average, in Los Angeles (Brown, 2019). Although no relevant literature showing the ride-hailing services' waiting time in the Seattle and Kansas City regions was identified, existing literature cited above covers a wide variety of regions across the U.S. Therefore, the waiting time attribute is designed with three levels: 2, 5 and 10 minutes. The levels of walking distance are designed by referencing the Uber Express Pool services, which allow riders to walk up to five minutes (about a quarter-mile) to a pickup spot for a lower fare (Gehrke et al., 2021). Three levels of walking distance are thus designed for the choice experiments: zero, 500 feet, and a quarter mile.

The travel cost for the HV alternative is calculated based on the reported trip distance and cost per mile of an average sedan (18.45 cents/mile, AAA, 2019). Here, only the operational costs (fuel and maintenance) are considered, while the ownership cost components are ignored. The PAV's travel cost is set to be 30% lower, 10% lower, or 20% greater than the HV alternative (both positive and negative levels capture the uncertainty of AVs' operational costs). For the SAV mode, three fare levels are set at 50, 100, or 150 cents/mile, while the SAVP mode's fare is set at 10%, 30%, or 50% lower than the SAV mode. Lastly, the PAV parking cost attribute has three levels:10% lower than the reported parking cost, 50% higher than a $2 plus the reported parking cost, and 100% higher than a $2 plus the reported parking cost. Note that a constant of $2 is added to the respondents' self-reported parking costs to ensure variations across different levels for the PAV alternative's parking cost, considering that some respondents report zero parking cost.

Combining all levels for all attributes, a full factorial design gives 3^12 = 531441 possible combinations. We then use SAS "MktEx macro" function to generate a fractional factorial design with 36 choice scenarios. These 36 choice scenarios are further divided into 6 blocks. Each respondent is randomly assigned one block which contains six choice tasks. An example of a choice task is shown in **Figure 4 - 1**.

Table 4 - 1. Attribute levels for choice experiment design.

	Your current mode (private car)	Private autonomous vehicle	Shared Autonomous vehicle (no pooling)	Shared autonomous vehicle (pooled)
In-vehicle time (min)	Base (Reported)	Base – 30% Base – 10% Base + 20%	Base – 30% Base – 10% Base + 20%	SAV-alone + 10% SAV-alone + 20% SAV-alone + 40%
Waiting time (min)	0	0	2 5 10	2 5 10
Walking distance (mile)	Reported	0 500 feet (1 city block) A quarter mile (2-3 city blocks)	0 500 feet (1 city block) A quarter mile (2-3 city blocks)	0 500 feet (1 city block) A quarter mile (2-3 city blocks)
Travel cost ($, fuel and maintenance)	Base (calculated)	Base - 40% Base - 10% Base + 20%	-	-
Fares ($, calculated based on cents/mile)	-	-	50 cents/ mile 100 cents/mile 150 cents/mile	SAV-alone - 10% SAV-alone - 30% SAV-alone - 50%
Parking cost ($)	Reported	Reported – 10% (Reported + $2) + 50% (Reported + $2) + 100%	-	-



Choice experiment # 1 of 6. The characteristics of the four travel mode options are listed in the table below. Please select your MOST preferred and LEAST preferred modes to use for your [Work/Work related business] trip.

	In-vehicle time	Waiting time	Walk distance	Travel cost (e.g. fuel)	Fare	Park cost
Current mode (private car)	30 min	0	0	$3		$4
Private autonomous vehicle	21 min	0	0	$3		$9
Shared autonomous vehicle (no pooling)	27 min	5 min	0		$10	-
Shared autonomous vehicle (pooled)	30 min	2 min	500 feet (1 city block)		$5	-

MOST Preferred		LEAST Preferred
○	Your current mode (private car)	○
○	Private autonomous vehicle	○
○	Shared autonomous vehicle (no pooling)	○
○	Shared autonomous vehicle (pooled)	○

Figure 4 - 1. An example of travel mode choice experiment

4.3.3 Survey distribution

The survey is distributed in Seattle and Kansas City metropolitan areas, which represent two typical but distinct urban forms in the U.S. According to the U.S. Census Bureau (2019), the Seattle region's population density is 678 people per square mile, which is more than twice the density of the Kansas City region (297 people per square mile). For commuting, the Seattle region shows a high public transit mode share of 11% and 77% private car share, while Kansas City is more car-oriented with only a 1% public transit share and 91% private car share. The mean commuting time in the Seattle region is 31.6 minutes, which is about 32% longer than the Kansas City region (23.9 minutes). The Seattle region also shows a greater median household income ($94,027) and a greater percent of people with a bachelor's degree or higher level of educational attainment (44%), as opposed to $70,215 median income and 38% bachelor's degree or higher educational attainment in the Kansas City region. Lastly, the parking cost in Seattle is also higher than Kansas City, as the highest 24-hour downtown garage parking cost is about $30 in Seattle and $20 in Kansas City (Parkopedia, 2020).

The pilot survey was conducted in June 2020. More than 30 responses were collected, including general respondents and metropolitan planning organization staff from both regions. The survey was then revised based on pilot survey feedback. For the formal survey, panels were recruited through Qualtrics' sampling services, and respondents were compensated for completing the survey. At the beginning of the survey, several criteria are set to exclude disqualified respondents, including 1) younger than 18 years of age, 2) reported trip distance greater than 50 miles (long-distance travel is beyond the scope of this analysis), 3) reported trip arrival time ahead of the departure time, which suggests that the respondents did not take the survey seriously, and 4) fail to pass the attention checks at least twice (in the last section of the survey capturing attitudes, three attention checks are embedded by explicitly asking respondents to select a specified answer [e.g., "Please select 'Strongly agree'"] to the five-Likert scale question, as detailed Jia (2021)). The first wave of the formal survey was conducted from July to September 2020. A total of 612 and 174 qualified respondents who completed the survey were obtained from the Seattle and Kansas City regions, respectively. Due to the small sample size of the Kansas City region, we conducted a second wave survey for this region from February to April 2021 and obtained additional 384 qualified responses. For this paper, we focus only on travelers using the private car mode, since the number of respondents who reported other modes are too small to conduct a rigorous statistical modeling. After removing responses with non-private car modes, 516 and 558 respondents are left for the Seattle and Kansas City regions, giving 3096 and 3348 choice observations, respectively.

The raw survey response data underwent further cleaning. A few respondents reported very large parking costs (which consequently created unrealistically high parking costs in the choice experiments). To ensure that the choice experiments are realistic to respondents, we remove those choice experiments with parking costs greater than \$30 and \$20 from the Seattle sample and Kansas City sample, respectively. These two thresholds are determined by the highest 24-hour garage parking fees in the downtown area in respective regions (Parkopedia, 2020). Finally, a total of 3038 and 3336 choice observations are kept for the Seattle and Kansas City regions, respectively, which represent 98.1% and 99.6% of the original choice observations. The final samples' geographic distribution by zipcode is shown in **Figure 4 - 2**.

Table 4 - 2 shows the summary socio-demographic characteristics of samples from both regions and compares them to the general population for both regions and the national population. Since the responses used in this analysis are limited to travelers using the private car mode, their demographics may not be aligned with the general population. Overall, the distribution of respondents' demographics is expected. One exception is that females (at 68%) are overrepresented in the Kansas City sample. Lastly, responses to the attitudinal questions are presented in **Figure 4 - 3**, and **Table 4 - 3** shows the specific wording of each attitudinal statement.

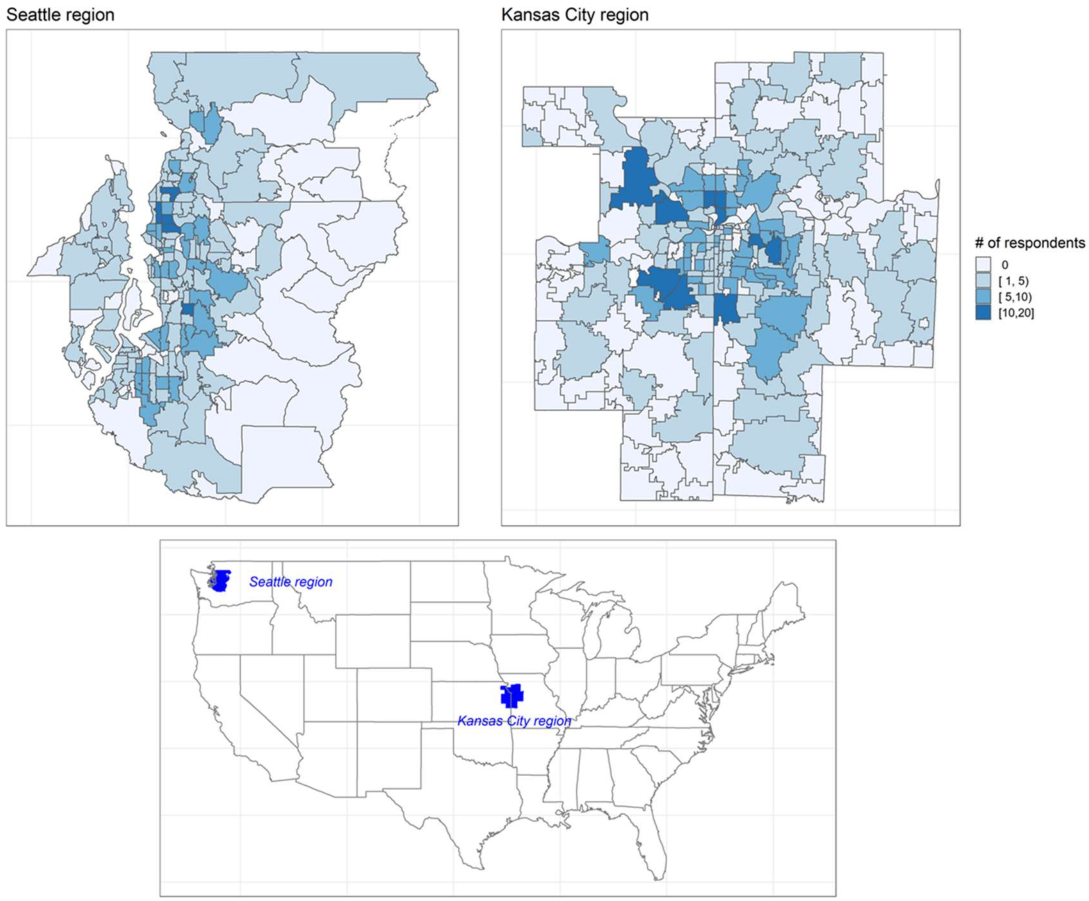


Figure 4 - 2. Spatial distribution of the two samples.

Table 4 - 2. Descriptive statistics of sample in the Seattle and Kansas City regions

	Seattle: sample (n = 511)	Seattle: Census population	Kansas City: sample (n = 558)	Kanas City: Census population	National: Census population
Female	50%	50%	68%	51%	51%
Age: 18 – 35	41%	32%	35%	29%	29%
Age: 35 – 65	49%	51%	38%	51%	49%
Age: 65 +	9%	17%	27%	20%	22%
High school graduate or lower	15%	26%	23%	34%	38%
Some college / Associate's degree	35%	30%	32%	28%	29%
Bachelor's degree	33%	27%	30%	24%	20%
Graduate or professional degree	17%	17%	15%	14%	13%
Drive license	95%	-	94%	-	-
Household income: Less than $35,000	-	-	26%	25%	26%
Household income: $35,000 - $100,000	-	-	47%	44%	42%
Household income: $100,000 or more	-	-	23%	31%	31%
Household income: Less than $50,000	25%	25%	-	-	38%
Household income: $50,000 - $150,000	55%	48%	-	-	46%
Household income: $150,000 or more	15%	27%	-	-	16%
Have pre-school children	13%	-	11%	-	-
Household size (mean)	2.85	-	2.45	-	-
Vehicle count per adult (mean)	1.06	-	1.10	-	-

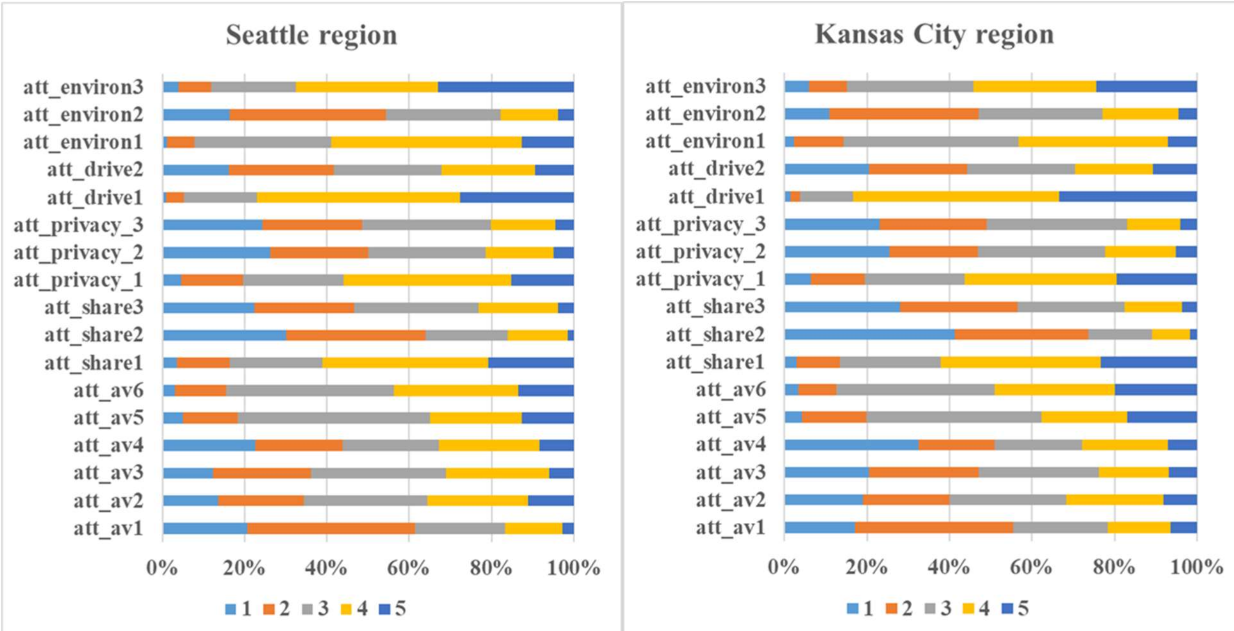


Figure 4 - 3. Distribution of responses to the attitudinal questions

Table 4 - 3. Attitudinal statements

Variables	Attitudinal statements
att_av1	Learning how to use new technologies is frustrating for me
att_av2	I like the idea of using a vehicle that is fully autonomous.
att_av3	I like to be among the first people to have the latest technology.
att_av4	How likely are you to purchase a self-driving car once the technology is fully developed?
att_av5	In general, how concerned are you about self-driving vehicles not driving as well as human drivers?
att_av6	How safe would you feel being a passenger in a self-driving car?
att_share1	I avoid being in close proximity to unfamiliar people.
att_share2	I feel comfortable with the idea of sharing a ride with strangers in a self-driving vehicle if it means having a reduced fee.
att_share3	I would share a self-driving vehicle with fellow travelers who have a similar route as me.
att_privacy_1	I'm worried that technology is invading our privacy too much.
att_privacy_2	How comfortable are you in allowing trip-location data usage for general surveillance?
att_privacy_3	How comfortable are you in allowing trip-location data usage to facilitate directed advertising?

att_drive1	I like the idea of driving as a means of transportation.
att_drive2	I would rather have someone else do the driving.
att_environ1	I am committed to having an environmentally friendly lifestyle.
att_environ2	I hardly ever considered the impact on the environment in making my daily choices.
att_environ3	Greenhouse gases from human activities are creating serious problems.

Note:

- att_av1, att_av2, att_av3, att_share1, att_share2, att_share3, att_privacy_1, att_drive1, att_drive2, att_environ1, att_environ2, att_environ3: response from 1 to 5 represent "Strongly disagree", "Disagree", "Neither disagree nor agree", "Agree", "Strongly agree", respectively.
- att_av4: response from 1 to 5 represent "Extremely unlikely", "Somewhat unlikely", "Neither unlikely nor likely", "Somewhat likely", "Extremely likely", respectively.
- att_av5: response from 1 to 5 represent "Not at all concerned", "Not so concerned", "Somewhat concerned", "Very concerned", "Extremely concerned", respectively.
- att_av6: response from 1 to 5 represent "Extremely safe", "Very safe", "Somewhat safe", "Not so safe", "Not at all safe", respectively.
- att_privacy_2, att_privacy3: response from 1 to 5 represent "Very uncomfortable", "Somewhat uncomfortable", "Unsure", "Somewhat comfortable", "Very comfortable", respectively.

4.4 Method

This paper uses the integrated choice and latent variable (ICLV) models to study AV mode choice preferences. The ICLV models are readily convenient to incorporate the impacts of attitudes, via treating the attitudes as latent variables. As shown in **Figure 4 - 4**, the ICLV model consists of three major components: latent variable component, discrete choice model component, and measurement model component.

4.4.1 Latent variable component

We first define several latent variables which represent respondents' underlying attitudinal constructs, including attitudes toward AVs, willingness to share travel with strangers, attitudes towards trip location data privacy, enjoyment of driving, and environmental consciousness. The latent variable α_{ln} is a function of individuals' socio-demographic characteristics plus a random disturbance.

$$\alpha_{ln} = \gamma_l z_n + \eta_{ln}$$

Where,

- $l = 1,2, .. L$, and $n = 1,2, ... N$. L and N denotes the total number of latent variables and the total number of respondents, respectively.
- z_n is a vector of socio-demographic characteristics of the respondent n. γ_l is a vector of to be estimated coefficients associated with z_n.
- η_{ln} is an error term that follows standard Normal distribution across respondents.

4.4.2 Discrete choice model component

The discrete choice model component is specified as an underlying random utility model. The utility function for alternative j, respondent n, in choice scenario t is given by:

$$U_{njt} = V_{njt} + \varepsilon_{njt}$$

$$V_{njt} = ASC_j + \sum_{l=1}^{L} \lambda_{lj}\alpha_{ln} + \delta_j z_n + \beta_j x_{njt} + \tau_{nj}$$

Where,

- ASC_j is the estimated alternative specific constant for alternative j (j = 1, 2, 3, 4).
- λ_{lj} is an estimated coefficient representing the impacts of latent attitudes α_{ln} on the utility of alternative j.
- δ_j is a vector of estimated coefficients that capture the direct effects of individual socio-economic characteristics on the utility of alternative j.
- x_{njt} is a vector of alternative attributes and β_j are estimated coefficients associated with x_{njt}.
- τ_{nj} is an error component that is normally distributed across respondents with mean zero and standard deviation θ_j.
- ε_{njt} is an error term that follows identically and independently extreme value distribution.

Conditional on α_{ln} and τ_{nj}, the likelihood of individual n choosing its most and least preferred alternatives in choice scenario t is given by the equation below:

$$L_{nt}^{Choice}(\alpha, \tau) = \frac{e^{V_{best_{nt}}}}{\sum_j e^{V_{njt}}} \times \frac{e^{-\mu V_{worst_{nt}}}}{\sum_{j \neq bes\ nt} e^{-\mu V_{njt}}}$$

Where,

- $best_{nt}$ and $worst_{nt}$ refer to the most and least preferred alternatives by respondent n in choice scenario t, respectively.
- μ is a scale parameter that represents the scale difference between the stage of choosing the best alternative and the stage of choosing the worst alternative. The model assumes that respondents select the best alternative first, and then the worst option is selected from the reminder set of alternatives. A negative sign is applied to the utility function in the stage of choosing the worst alternative, since the alternative with the lowest utility is most likely to be chosen as the least preferred option.

Since each respondent take six choice tasks (T = 6), the likelihood (conditional on α, τ) of an individual n makes his or her sequence of choices is the product of $P_{nt}(\alpha, \tau)$ over his or her multiple choice tasks:

$$L_n^{Choice}(\alpha, \tau) = \prod_{t=1}^{T} L_{nt}^{Choice}(\alpha, \tau)$$

4.4.3 Measurement model

In the measurement model component, the latent variables are used to explain respondents' answers to attitudinal questions. A continuous measurement model is specified, as shown in the equation below.

$$L_n^{Indicator}(\alpha) = \prod_{i=1}^{I} \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{(I_{ni}-\bar{I}_i-\zeta_{li}\alpha_{ln})^2}{2\sigma_i^2}}$$

Where,

- I is the total number of attitudinal indicators.
- ζ_{li} is an estimated coefficient denoting the impact of the latent variable α_{ln} on the attitudinal indicator I_i.
- σ_i is an estimated standard deviation.

4.4.4 Joint model likelihood function

Lastly, the three components are modeled jointly. The combined log-likelihood function (LL) is shown below. Model coefficients are estimated using maximum simulated likelihood estimation with 1000 mlhs draws (Hess, 2005).

$$LL = \sum_{n=1}^{N} log\,(\int_a \int_\tau L_n^{Choic}\;(\alpha, \tau) L_n^{Indicator}(a)\, f(\tau) f(\alpha) d\tau da)$$

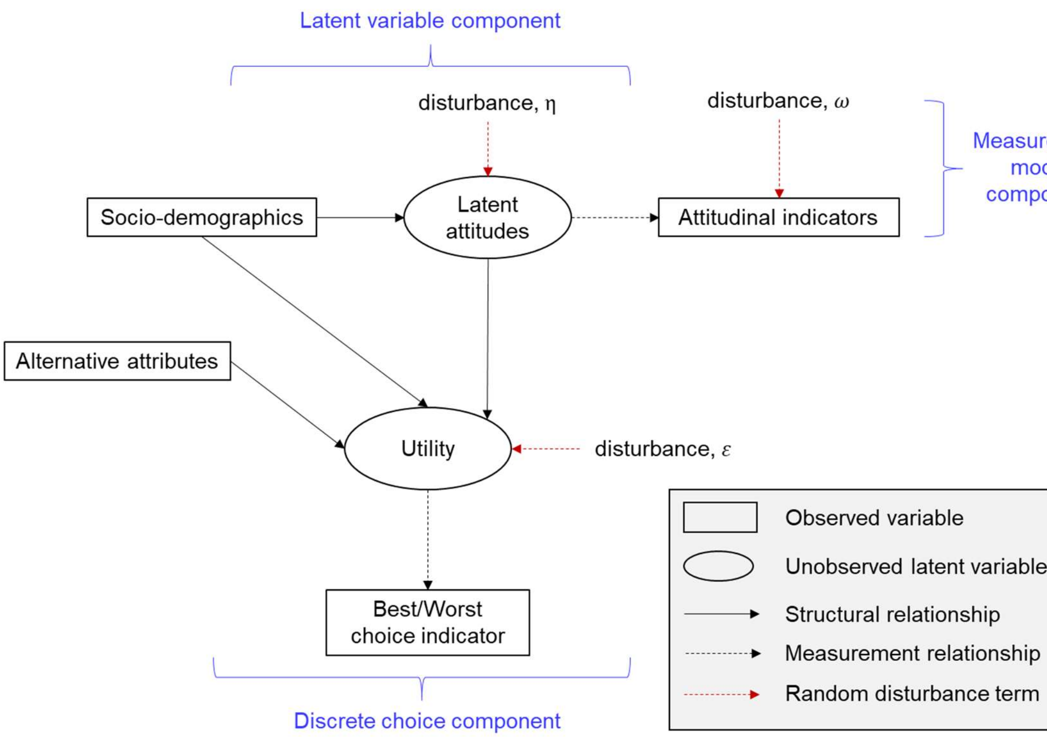


Figure 4 - 4. ICLV model framework for modeling AV mode choice preferences

4.5 Results and discussion

Table 4 - 4 and **Table 4 - 5** show the ICLV model estimation results for the Seattle and Kansas City regions, respectively. In these two tables, only the results of discrete choice component and the latent variable component are presented. The discrete choice component shows the impacts of mode-specific attributes, trip-specific attributes, and latent attitudinal constructs on the AV mode choice preferences. The latent variable component shows how socio-economic characteristics are related to the latent attitudinal constructs. Results of the measurement model component, which show the relevant attitudinal statements used to measure the latent attitudinal constructs, are attached in the Appendix Table B-1 and B-2.

4.5.1 Mode-specific attributes

For both regions, the in-vehicle travel time, wait time, walking distance, and trip costs are found to significantly impact mode choices. This result is consistent with previous findings (e.g., Krueger et al., 2016; Kolarova et al., 2019; Correia et al., 2019). Travelers are found to be indifferent to reductions in wait time from 5 to 2 minutes, while a 10-minute wait time significantly reduces the utility of the respective mode. As expected, increases in walk distance are associated with lower utilities for the respective mode. Lastly, trip costs (fare, fuel, and parking cost) are negatively correlated with the utilities for respective modes.

The coefficients for in-vehicle travel time are specified to be specific to each mode alternative, and the coefficient differences across alternatives are examined. First, we compare the in-vehicle time sensitivity between HVs and PAVs. For the Seattle sample, travelers' sensitivity to in-vehicle time in a PAV is 14% lower than in a HV ($t = -2.2$). In contrast, in the Kansas City region, the in-vehicle time coefficient for PAVs is close to that for HVs, without statistical significance ($t = 0.6$). Second, the in-vehicle time sensitivity between the SAV and SAVP modes is compared. Travelers from the Seattle region perceive the SAVP in-vehicle time 14% more negative than the SAV, with low statistical significance ($t = 1.3$). For the Kansas City region, the difference is more significant where travelers' sensitivity to the SAVP in-vehicle time is 23% higher than the SAV mode ($t = 1.7$).

Lastly, for both regions, we find that the sensitivity to the in-vehicle time of PAVs is higher than the SAVs, possibly because people attach some depreciation costs associated with using PAVs. Note that for the Kansas City sample, in the second wave of the survey deployment, we designed two versions of choice experiments. The first version does not present the vehicle ownership cost information (as shown in **Figure 4 - 1**). The second version adds a reminder statement, "the private car and private autonomous vehicle modes have roughly \$0.3 - \$0.6/mile ownership costs (e.g., depreciation, insurance, license, taxes, etc.), though not shown in the table below". Respondents are randomly presented with one of the two versions of choice experiments. Results show that the presentation of vehicle ownership cost information neither impacts the mode preferences nor the sensitivity to in-vehicle time.

4.5.2 Attitudes

For both regions, two latent attitudinal constructs are found to be significantly correlated with AV mode preferences. First, respondents with more positive attitudes toward AVs show greater preferences for the three AV modes, especially for the PAV mode. The latent attitude toward AVs is measured by the level of interest in AV technology, the level of AV safety concern, and general perception of new technologies. The second latent attitudinal construct, which is also important for both study regions, is the willingness to share travel with strangers. As expected, travelers show a higher probability of choosing the SAVP mode when they are more willing to share travel with strangers.

Some latent attitudinal constructs are also found to be important, but only for a specific region. For the Kansas City sample, those who are less concerned about trip-location data privacy are more likely to select the SAV and SAVP modes, particularly for the SAV mode. For the Seattle sample, travelers who do not enjoy driving are found to show greater preferences for the PAV mode, although the impact of this "enjoyment of driving" latent variable shows low statistical significance ($t = 1.4$). Lastly, we find that the latent environmental consciousness construct is not significant for the AV mode choice preferences for either study region. This is not surprising since the AV modes in our choice experiments are not described with a specific powertrain fuel type and the public may not be aware of the environmental footprints of different AV modes.

4.5.3 Individual socio-economic characteristics

The individual socio-economic characteristics do not show direct effects on the utility of AV modes, but show indirect effects when linking to the latent attitudes. As shown in the latent variable model component in **Table 4 - 4**, for the Seattle sample, females are less likely to show a positive attitude toward AVs and are less willing to share travel with strangers. Moreover, seniors (compared to those younger than 65) in Seattle are found to be more likely to enjoy driving. Other socio-economic variables, such as educational attainment, household income, presence of children in the household, household size, and vehicle ownership, etc. are tested but show no significant associations with the latent attitudinal constructs among Seattle respondents.

For the Kansas City sample, the association between socio-economic characteristics and latent attitudes is very different from the Seattle sample. As shown in the latent variable component in **Table 4 - 5**, no specific socio-economic characteristics are found to be associated with the latent AV attitude. In other words, different socio-economic groups from the Kansas City sample do not show significant variation in their attitudes toward AVs. Moreover, unlike the Seattle region, gender does not appear to impact the willingness to share travel with strangers. Interestingly, seniors from the Kansas City region are found to be less willing to share rides with strangers. Lastly, females in Kansas City appear to be less sensitive to trip location privacy concerns while seniors are more likely to be concerned.

4.5.4 Trip characteristics

Several trip-related characteristics appear to be significant for both study regions, including trip purpose, self-rated trip congestion level, and trip distance. First, commute trips are associated with a greater probability of choosing PAVs, compared to other trip purposes. Second, travelers who reported a higher level of congestion during the trip are more likely to select the PAV mode, relative to those who reported less congested trips. These two results suggest that travelers may prefer AVs when desiring to improve productivity during commute and to avoid stressful congestion. Lastly, compared to short trips, travelers with longer trip distances are more likely to pool with others when using SAVs. This result is expected since the longer the trip, the more expensive the fare, and travelers, thus, have greater motivation to pool with other travelers to save cost.

Table 4 - 4. ICLV model results for the Seattle region

	Alternative	Estimate	Std. Err.	T value
Discrete choice model component				
Alternative specific constant	hv	7.65	0.89	8.56
Alternative specific constant	pav	3.39	0.68	4.95
Alternative specific constant	sav	-	-	-
Alternative specific constant	savp	-1.62	0.53	-3.04
Trip purpose: work	pav	0.50	0.31	1.64
Trip distance	savp	0.46	0.16	2.81
Self-rated trip congestion level	pav	0.16	0.09	1.89
In-vehicle time	hv	-2.92	0.30	-9.72
In-vehicle time	pav	-2.50	0.24	-10.30
In-vehicle time	sav	-2.02	0.26	-7.67
In-vehicle time	savp	-2.31	0.28	-8.24
Waiting time (10 min)	sav, savp	-0.27	0.09	-3.20
Waiting time (2 or 5 min)	sav, savp	-	-	-
Walking distance (0)	all	0.47	0.08	6.01
Walking distance (500 feet)	all	-	-	-
Walking distance (a quarter mile)	all	-0.38	0.08	-4.73
Total cost [1]	all	-1.03	0.09	-10.88
Scale parameter [2]	all	0.86	0.07	2.00
Error component	hv	2.48	0.16	15.78
Error component	pav	0.14	0.24	0.60
Error component	savp	2.42	0.21	11.67
Error component	hv, pav	2.51	0.22	11.48
Positive AV attitudes	pav	2.03	0.18	11.19
Positive AV attitudes	sav	1.99	0.22	9.11
Positive AV attitudes	savp	1.48	0.23	6.53
Willingness to share travel with strangers	savp	0.87	0.16	5.45
Not enjoy driving	pav	0.23	0.17	1.36
Latent variable component				
Positive AV attitudes: female		-0.15	0.06	-2.53
Positive AV attitudes: age 65 +		-0.14	0.15	-0.97
Willingness to share: female		-0.27	0.07	-3.74
Not enjoy driving: female		0.09	0.10	0.91
Not enjoy driving: age 65 +		-0.51	0.23	-2.17

Note: for HV and PAV, the total cost refers to operational cost plus the parking cost; for SAV and SAVP, the total cost refers to fares.

Table 4 - 5. ICLV model results for the Kansas City region

	Alternative	Estimate	Std. Err.	T value
Discrete choice model component				
Alternative specific constant	hv	6.80	0.65	10.44
Alternative specific constant	pav	2.55	0.53	4.77
Alternative specific constant	sav	-	-	-
Alternative specific constant	savp	-1.84	0.44	-4.21
Trip purpose: work	pav	0.48	0.30	1.62
Trip distance	savp	0.63	0.20	3.19
Self-rated trip congestion level	pav	0.35	0.10	3.51
In-vehicle time	hv	-2.23	0.26	-8.69
In-vehicle time	pav	-2.11	0.22	-9.75
In-vehicle time	sav	-1.68	0.23	-7.29
In-vehicle time	savp	-2.07	0.24	-8.46
Waiting time (10 min)	sav, savp	-0.22	0.08	-2.87
Waiting time (2 or 5 min)	sav, savp	-	-	-
Walking distance (0)	all	0.39	0.07	5.39
Walking distance (500 feet)	all	-	-	-
Walking distance (a quarter mile)	all	-0.58	0.08	-7.26
Total cost [1]	all	-0.70	0.09	-8.07
Scale parameter [2]	all	0.90	0.07	1.43
Error component	hv	1.47	0.13	11.60
Error component	pav	1.56	0.14	10.90
Error component	savp	2.25	0.18	12.40
Error component	hv, pav	2.08	0.18	11.65
Positive AV attitudes	pav	2.43	0.17	14.55
Positive AV attitudes	sav	1.52	0.21	7.33
Positive AV attitudes	savp	1.70	0.20	8.33
Unwillingness to share travel with strangers	savp	-0.62	0.15	-4.25
Less concern about trip-location data privacy	sav	0.80	0.17	4.84
Less concern about trip-location data privacy	savp	0.45	0.18	2.55
Latent variable model component				
Unwillingness to share: age 65+		0.25	0.10	2.50
Less concern about data privacy: female		0.20	0.06	3.41
Less concern about data privacy: age 65+		-0.33	0.09	-3.57

Note: for HV and PAV, the total cost refers to operational cost plus the parking cost; for SAV and SAVP, the total cost refers to fares.

4.6 Conclusion and limitations

This paper examines travelers' mode choice preferences in the AV era based on SP surveys conducted in two distinct metropolitan regions (Seattle and Kansas City) in the U.S. ICLV models are developed to incorporate the impacts of latent attitudinal constructs on AV mode preferences. Results uncover several important latent attitudes, including attitudes toward AVs, willingness to share travel with strangers, sensitivity to trip location data privacy, and enjoyment of driving. Although socio-economic characteristics do not show direct impacts on the AV mode choices, their indirect effects are identified through latent attitudes. Such a linkage between socio-economic characteristics and latent attitudes indicates policy opportunities to encourage certain AV modes among specific socio-economic groups. For example, according to the Seattle sample, females show more negative attitudes toward AVs, possibly because of concern about AV safety. Targeted marketing campaigns on the safety benefits of AVs relative to HVs may change women's AV mode preferences. Qualitative study approaches, such as semi-structured interviews, are needed for a better understanding of women's hesitation to adopt AVs, so that policy-makers and AV manufactures can develop tailored solutions to address such concerns. Moreover, females are found to be less willing to share travel with strangers, which in turn impedes selecting the SAVP mode. Women's aversion to pooling is also evident in existing ride-hailing services (Kang et al., 2021). Moving to the AV era, efforts by SAV operators and regulators are needed to alleviate women's concerns about pooling to ensure a more sustainable AV future.

The mode-specific attributes (such as travel cost, in-vehicle time, etc.) are found to be important in mode choice decisions. Policies that increase the cost of using private modes while decrease the cost for SAV and SAVP modes have the potential to increase average vehicle occupancy in the AV era. Future work can examine the impacts of various policies (such as occupancy-based VMT fees, parking fees, etc.) on regional travel mode share. Additionally, according to the Seattle sample, travelers are 14% less sensitive to the in-vehicle time in PAVs than in HVs. Such a significant reduction in sensitivity to in-vehicle time suggests potential induced VMT due to the AV adoption. Meanwhile, the Kansas City sample shows no significant difference in sensitivity to in-vehicle time between PAVs and HVs. The inconsistent findings between the two study regions echo the mixed evidence in existing literature. With more future studies conducted in various regions, we can better understand the role of local contexts in shaping AV mode choice preferences.

We note several limitations of this study. First, the SP survey approach inherently has hypothetical bias. None of the respondents have experiences with AVs. Therefore, this study aims not to predict the exact travel mode share in the AV era, but rather to understand influential factors of travelers' AV mode preferences. Second, the SP survey was conducted during the Covid-19 pandemic; therefore, study findings might be attached with some specific "pandemic effects". For example, we find that travelers from both regions evaluate the in-vehicle time in SAVP more negatively than in SAV, which may be partly due to the social distancing norms of the pandemic.

In the post-pandemic period, whether this trend (or the magnitude of the trend) will still hold warrant further investigation. Third, this study focuses only on current car travelers. The impacts of AVs on other traditional modes (e.g., walk, bike, and transit) represent an avenue for future work. Lastly, we note a methodology limitation associated with the ICLV model. In the measurement model component, we use a continuous model specification for the dependent variables (i.e., responses to the attitudinal questions). Considering the ordinal nature of those responses, an ordered logit/probit specification may be more appropriate. However, the ordered specification would greatly increase the number of estimated coefficients compared to a continuous specification, therefore making it more difficult for coefficient estimation.

Acknowledgements
This study is funded by Jeffress Memorial Trust. The authors would like to thank Suzanne Childress for her help in distributing pilot survey in the Puget Sound Region, and Eileen Yang in distributing pilot survey in the Kansas City. The authors also thank Linda Lim and Mirla Abi Aad for their help in implementing the survey in Qualtrics.

Author Contributions
The authors confirm contribution to the paper as follows: study conception and design: T.D. Chen, W. Zhang, W. Jia, K. Wang; data collection: W. Jia; analysis and interpretation of results: W. Jia, T.D. Chen, W. Zhang, K. Wang; draft manuscript preparation: W. Jia, T.D. Chen, W. Zhang, K. Wang.

References

AAA, Your driving cost. https://exchange.aaa.com/wp-content/uploads/2019/09/AAA-Your-Driving-Costs-2019.pdf

Asgari, H. and Jin, X., 2019. Incorporating attitudinal factors to examine adoption of and willingness to pay for autonomous vehicles. Transportation Research Record, 2673(8), pp.418-429.

Ashkrof, P., Homem de Almeida Correia, G., Cats, O. and van Arem, B., 2019. Impact of automated vehicles on travel mode preference for different trip purposes and distances. Transportation Research Record, 2673(5), pp.607-616.

Asmussen, K.E., Mondal, A. and Bhat, C.R., 2020. A socio-technical model of autonomous vehicle adoption using ranked choice stated preference data. Transportation Research Part C: Emerging Technologies, 121, p.102835.

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

Bansal, P. and Kockelman, K.M., 2018. Are we ready to embrace connected and self-driving vehicles? A case study of Texans. Transportation, 45(2), pp.641-675.

Barbour, N., Menon, N., Zhang, Y. and Mannering, F., 2019. Shared automated vehicles: A statistical analysis of consumer use likelihoods and concerns. Transport Policy, 80, pp.86-93.

Ben-Akiva, M., McFadden, D., Train, K., Walker, J., Bhat, C., Bierlaire, M., Bolduc, D., Boersch-Supan, A., Brownstone, D., Bunch, D.S. and Daly, A., 2002. Hybrid choice models: Progress and challenges. Marketing Letters, 13(3), pp.163-175.

Brown, A.E., 2019. Prevalence and mechanisms of discrimination: Evidence from the ride-hail and taxi industries. Journal of Planning Education and Research, p.0739456X19871687.

Cunningham, M.L., Regan, M.A., Horberry, T., Weeratunga, K. and Dixit, V., 2019. Public opinion about automated vehicles in Australia: Results from a large-scale national survey. Transportation research part A: policy and practice, 129, pp.1-18.

Daly, A., Hess, S., Patruni, B., Potoglou, D. and 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(2), pp.267-297.

de Almeida Correia, G.H., Looff, E., van Cranenburgh, S., Snelder, M. and van Arem, B., 2019. On the impact of vehicle automation on the value of travel time while performing work and leisure activities in a car: Theoretical insights and results from a stated preference survey. Transportation Research Part A: Policy and Practice, 119, pp.359-382.

Etzioni, S., Daziano, R.A., Ben-Elia, E. and Shiftan, Y., 2021. Preferences for shared automated vehicles: A hybrid latent class modeling approach. Transportation Research Part C: Emerging Technologies, 125, p.103013.

Gao, J., Ranjbari, A. and MacKenzie, D., 2019. Would being driven by others affect the value of travel time? Ridehailing as an analogy for automated vehicles. Transportation, 46(6), pp.2103-2116.

Gehrke, S.R., Huff, M.P. and Reardon, T.G., 2021. Social and Trip-Level Predictors of Pooled Ride-Hailing Service Adoption in the Greater Boston Region. Case Studies on Transport Policy.

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

Harper, C.D., Hendrickson, C.T., Mangones, S. and Samaras, C., 2016. Estimating potential increases in travel with autonomous vehicles for the non-driving, elderly and people with travel-restrictive medical conditions. Transportation research part C: emerging technologies, 72, pp.1-9.

Hess, S., Train, K.E. and Polak, J.W., 2006. On the use of a Modified Latin Hypercube Sampling (MLHS) method in the estimation of a Mixed Logit model for vehicle choice. Transportation Research Part B: Methodological, 40(2), pp.147-163.

Jia, 2021. Examining Preference Heterogeneity in Adoption of Emerging Transportation Technologies. PhD dissertation, University of Virginia.

Jing, P., Huang, H., Ran, B., Zhan, F. and Shi, Y., 2019. Exploring the factors affecting mode choice Intention of autonomous vehicle based on an extended theory of planned behavior—A case study in China. Sustainability, 11(4), p.1155.

Kang, S., Mondal, A., Bhat, A.C. and Bhat, C.R., 2021. Pooled versus private ride-hailing: A joint revealed and stated preference analysis recognizing psycho-social factors. Transportation Research Part C: Emerging Technologies, 124, p.102906.

Kolarova, V., Steck, F. and Bahamonde-Birke, F.J., 2019. Assessing the effect of autonomous driving on value of travel time savings: A comparison between current and future preferences. Transportation Research Part A: Policy and Practice, 129, pp.155-169.

König, A. and Grippenkoven, J., 2020. Travellers' willingness to share rides in autonomous mobility on demand systems depending on travel distance and detour. Travel Behaviour and Society, 21, pp.188-202.

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

Krueger, R., Rashidi, T.H. and Dixit, V.V., 2019. Autonomous driving and residential location preferences: Evidence from a stated choice survey. Transportation Research Part C: Emerging Technologies, 108, pp.255-268.

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

Lavieri, P.S., Garikapati, V.M., Bhat, C.R., Pendyala, R.M., Astroza, S. and Dias, F.F., 2017. Modeling individual preferences for ownership and sharing of autonomous vehicle technologies. Transportation research record, 2665(1), pp.1-10.

Lavieri, P.S. and Bhat, C.R., 2019. Modeling individuals' willingness to share trips with strangers in an autonomous vehicle future. Transportation research part A: policy and practice, 124, pp.242-261.

Montoro, L., Useche, S.A., Alonso, F., Lijarcio, I., Bosó-Seguí, P. and Martí-Belda, A., 2019. Perceived safety and attributed value as predictors of the intention to use autonomous vehicles: A national study with Spanish drivers. Safety Science, 120, pp.865-876.

Nazari, F., Noruzoliaee, M. and Mohammadian, A.K., 2018. Shared versus private mobility: Modeling public interest in autonomous vehicles accounting for latent attitudes. Transportation Research Part C: Emerging Technologies, 97, pp.456-477.

Nodjomian, A.T. and Kockelman, K., 2019. How does the built environment affect interest in the ownership and use of self-driving vehicles?. Journal of Transport Geography, 78, pp.115-134.

Papadoulis, A., Quddus, M. and Imprialou, M., 2019. Evaluating the safety impact of connected and autonomous vehicles on motorways. Accident Analysis & Prevention, 124, pp.12-22.

Parkopedia, 2020. https://en.parkopedia.com/. Accessed on October, 2020.

Payre, W., Cestac, J. and Delhomme, P., 2014. Intention to use a fully automated car: Attitudes and a priori acceptability. Transportation research part F: traffic psychology and behaviour, 27, pp.252-263.

Sener, I.N., Zmud, J. and Williams, T., 2019. Measures of baseline intent to use automated vehicles: A case study of Texas cities. Transportation research part F: traffic psychology and behaviour, 62, pp.66-77.

Steck, F., Kolarova, V., Bahamonde-Birke, F., Trommer, S. and Lenz, B., 2018. How autonomous driving may affect the value of travel time savings for commuting. Transportation research record, 2672(46), pp.11-20.

Stoiber, T., Schubert, I., Hoerler, R. and Burger, P., 2019. Will consumers prefer shared and pooled-use autonomous vehicles? A stated choice experiment with Swiss households. Transportation Research Part D: Transport and Environment, 71, pp.265-282.

Sweet, M.N., 2021. User interest in on-demand, shared, and driverless mobility: Evidence from stated preference choice experiments in Southern Ontario. Travel Behaviour and Society, 23, pp.120-133.

Thebault-Spieker, J., Terveen, L. and Hecht, B., 2017. Toward a geographic understanding of the sharing economy: Systemic biases in UberX and TaskRabbit. ACM Transactions on Computer-Human Interaction (TOCHI), 24(3), pp.1-40.

Tian, Z., Feng, T., Timmermans, H.J. and Yao, B., 2021. Using autonomous vehicles or shared cars? Results of a stated choice experiment. Transportation Research Part C: Emerging Technologies, 128, p.103117.

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

Wadud, Z. and Chintakayala, P.K., 2021. To own or not to own–That is the question: The value of owning a (fully automated) vehicle. Transportation Research Part C: Emerging Technologies, 123, p.102978.

Wadud, Z., MacKenzie, D. and Leiby, P., 2016. Help or hindrance? The travel, energy and carbon impacts of highly automated vehicles. Transportation Research Part A: Policy and Practice, 86, pp.1-18.

Wang, K. and Akar, G., 2019. Factors affecting the adoption of autonomous vehicles for commute trips: an analysis with the 2015 and 2017 Puget Sound Travel Surveys. Transportation Research Record, 2673(2), pp.13-25.

Wang, K., & Zhang, W. (2021). The role of urban form in the performance of shared automated vehicles. Transportation Research Part D: Transport and Environment, 93, 102744.

Wang, S., Jiang, Z., Noland, R.B. and Mondschein, A.S., 2020. Attitudes towards privately-owned and shared autonomous vehicles. Transportation research part F: traffic psychology and behaviour, 72, pp.297-306.

Wang, M. and Mu, L., 2018. Spatial disparities of Uber accessibility: An exploratory analysis in Atlanta, USA. Computers, Environment and Urban Systems, 67, pp.169-175.

Waymo, 2016. Say Hello to Waymo. https://www.youtube.com/watch?v=uHbMt6WDhQ8&feature=emb_title

Yang, H., Liang, Y. and Yang, L., 2021. Equitable? Exploring ridesourcing waiting time and its determinants. Transportation Research Part D: Transport and Environment, 93, p.102774.

Yap, M.D., Correia, G. and Van Arem, B., 2016. Preferences of travellers for using automated vehicles as last mile public transport of multimodal train trips. Transportation research part a: policy and practice, 94, pp.1-16.

Zhang, W., Guhathakurta, S., Fang, J. and Zhang, G., 2015. Exploring the impact of shared autonomous vehicles on urban parking demand: An agent-based simulation approach. Sustainable Cities and Society, 19, pp.34-45.

Zhang, W., Guhathakurta, S. and Khalil, E.B., 2018. The impact of private autonomous vehicles on vehicle ownership and unoccupied VMT generation. Transportation Research Part C: Emerging Technologies, 90, pp.156-165.

Zhang, W., & Wang, K., 2020. Parking futures: Shared automated vehicles and parking demand reduction trajectories in Atlanta. Land Use Policy, 91, 103963.

Zhang, W., Wang, K., Wang, S., Jiang, Z., Mondschein, A., & Noland, R. B. (2020). Synthesizing neighborhood preferences for automated vehicles. Transportation Research Part C: Emerging Technologies, 120, 102774.

Zhou, F., Zheng, Z., Whitehead, J., Washington, S., Perrons, R.K. and Page, L., 2020. Preference heterogeneity in mode choice for car-sharing and shared automated vehicles. Transportation Research Part A: Policy and Practice, 132, pp.633-650.

Appendix A

Educational material was presented to respondents before choice experiments. Respondents were told that the survey focused on the fully autonomous vehicles which were described using the following sentences:

- *"You do not have to be in control at all during travel".*
- *"The autonomous vehicle can handle any road and weather condition as an experienced human driver. All you need to do is tell the vehicle your destination, and it will deliver you to your destination".*
- *"The autonomous vehicle will also have emergency features and safety protocols".*
- *"No commercial production of a fully autonomous vehicle exists, but companies such as Zoox, Google's Waymo, and many others are working towards this goal."*

Appendix B

Table B-1. ICLV model results (measurement model component) for the Seattle region

	Estimate	Std. Err.	T value
AV attitude indicator 1: ζ	-0.29	0.05	-5.78
AV attitude indicator 2: ζ	1.01	0.05	20.81
AV attitude indicator 3: ζ	0.55	0.05	10.89
AV attitudes indicator 4: ζ	1.07	0.05	21.22
AV attitudes indicator 5: ζ	-0.57	0.05	-12.43
AV attitudes indicator 6: ζ	-0.79	0.04	-20.47
AV attitudes indicator 1: σ	1.00	0.03	31.61
AV attitudes indicator 2: σ	0.71	0.03	22.96
AV attitudes indicator 3: σ	0.96	0.03	30.55
AV attitudes indicator 4: σ	0.76	0.03	23.42
AV attitudes indicator 5: σ	0.83	0.03	29.84
AV attitudes indicator 6: σ	0.60	0.02	24.18
Willingness to share indicator 1: ζ	-0.19	0.05	-3.66
Willingness to share indicator 2: ζ	0.84	0.05	16.42
Willingness to share indicator 3: ζ	0.76	0.05	13.91
Willingness to share indicator 1: σ	1.04	0.03	31.58
Willingness to share indicator 2: σ	0.66	0.05	13.17
Willingness to share indicator 3: σ	0.83	0.04	19.90
Enjoyment of driving indicator 1: ζ	-0.54	0.09	-6.13
Enjoyment of driving indicator 2: ζ	0.46	0.09	5.33
Enjoyment of driving indicator 1: σ	0.64	0.07	9.22
Enjoyment of driving indicator 2: σ	1.12	0.05	24.27

Table B-2. ICLV model results (measurement model component) for the Kansas City region

	Estimate	Std. Err.	T value
AV attitudes indicator 1: ζ	-0.31	0.05	-6.01
AV attitudes indicator 2: ζ	1.05	0.04	24.47
AV attitudes indicator 3: ζ	0.61	0.05	11.86
AV attitudes indicator 4: ζ	1.14	0.05	24.26
AV attitudes indicator 5: ζ	-0.79	0.04	-19.41
AV attitudes indicator 6: ζ	-0.88	0.04	-24.35
AV attitudes indicator 1: σ	1.09	0.03	33.11
AV attitudes indicator 2: σ	0.70	0.03	25.25
AV attitudes indicator 3: σ	1.03	0.03	32.14
AV attitudes indicator 4: σ	0.76	0.03	24.97
AV attitudes indicator 5: σ	0.75	0.03	28.90
AV attitudes indicator 6: σ	0.58	0.02	24.96
Willingness to share indicator 1: ζ	0.18	0.05	3.36
Willingness to share indicator 2: ζ	-0.78	0.05	-14.31
Willingness to share indicator 3: ζ	-0.70	0.06	-12.43
Willingness to share indicator 1: σ	1.02	0.03	32.95
Willingness to share indicator 2: σ	0.70	0.05	14.36
Willingness to share indicator 3: σ	0.89	0.04	22.47
Concern about data privacy indicator 1: ζ	-0.50	0.05	-9.22
Concern about data privacy indicator 2: ζ	0.88	0.05	16.75
Concern about data privacy indicator 3: ζ	0.84	0.05	15.77
Concern about data privacy indicator 1: σ	1.02	0.03	30.69
Concern about data privacy indicator 2: σ	0.82	0.04	19.91
Concern about data privacy indicator 3: σ	0.74	0.04	18.00

Chapter 5. Conclusions

This dissertation investigates consumer preference heterogeneity in the adoption of emerging technologies. In this closing chapter, we first reiterate the main findings, and then discuss the contributions to existing literature. Lastly, research limitations are discussed and future research directions are identified.

5.1 Summary of main findings

This research discusses three preference heterogeneity issues in the adoption of three emerging technologies in transportation, including ride-hailing services, EVs, and AVs. Due to the different adoption stages of the three technologies, different kinds of data (and suitable modeling approaches) are used. Major findings are detailed below.

Using large-scale revealed preference (RP) survey data, Chapter 2 shows the observed disparities in ride-hailing usage across different socio-demographic groups, and also shows how various spatial contexts are related to such disparities. Specifically, seniors (compared to younger) and low-income (compared to high-income) travelers are found to be less likely to use ride-hailing. Furthermore, the ride-hailing usage disparity between seniors and younger travelers is exacerbated in urban areas, while the disparity between low-income and high-income travelers is exacerbated in rural areas. In addition to the urban/rural classification, the share of community seniors and community median income are also found to be relevant to the disparity patterns. Lastly, ride-hailing services appear to fill the mobility gap for non-vehicle owners in public transport desert communities, as represented by their much higher odds of using ride-hailing services compared to those owning vehicles from the same community. With the increase in community public transport mode share, the odds of using ride-hailing for the two groups become close, suggesting the diminishing role of ride-hailing services for travelers without private vehicles in transit-rich communities.

Chapter 3 demonstrates the unobserved heterogeneous preference for EVs based on stated choice experiments. MXL, LC, and LC-MXL models are developed and model results are contrasted in terms of model fit, behavioral interpretation, and policy implications. Consistently, all models suggest that monetary incentives are most effective in improving future EV market share, followed by greater charging infrastructure deployment, while the effectiveness of battery range improvement appears to be minor. Additionally, it is not reasonable to conclude a single model specification is unambiguously preferred. Rather, each model shows its strengths and limitations in uncovering EV preference profiles. The combined modeling analysis provides a more comprehensive picture of respondents' preferences for EVs, which could inform more effective EV policy design and implementation.

Chapter 4 presents the impacts of latent attitudinal constructs on AV mode choice preferences, using an ICLV model based on stated preference choice experiments. Results suggest that attitudes toward AVs, willingness to share travel with strangers, sensitivity to data privacy, and enjoyment of driving are important latent attitudinal constructs in explaining AV mode choice preferences. Furthermore, although the socio-demographic characteristics do not show direct effects on AV mode preferences, they are found to be related to the latent attitudinal constructs, thereby indirectly impacting AV mode choices. For example, females, on average, are found to show more negative attitudes towards AVs and less willing to share travel with strangers, which translates to their lower probability of selecting AVs and sharing modes compared to an average male. Lastly, this study shows the potential reduction in sensitivity to in-vehicle travel time in the AV era (up to 14%), possibly because AVs enable greater productivity during travel and free travelers from tedious driving tasks.

5.2 Contributions

Overall, the dissertation contributes to the existing literature by systematically examining consumer preference heterogeneity in the adoption of emerging transportation technologies. From the practical/empirical perspective, findings of this dissertation provide policy insights for sustainable outcomes in the deployment of new transportation technologies. From the methodology perspective, the three papers use various datasets and heterogeneity models, which inform future research on the adoption of new transportation technologies. Specific contributions of the three papers are listed as follows.

The ride-hailing analysis in Chapter 2 is motivated by the knowledge gap about how spatial contexts are associated with the exacerbating or dampening of disparities in ride-hailing usage across socio-economic groups. These results suggest the target areas for policy-makers to improve equity in ride-hailing usage. This study also demonstrates the necessity to consider the interplay of socio-economic characteristics and spatial contexts for ride-hailing usage research, instead of examining these two aspects independently as in most existing studies. Furthermore, the observed heterogeneity modeling approach by interacting spatial contextual variables and socio-economic variables can be applied to investigate the spatial heterogeneity issue for the adoption of other emerging technologies (e.g., EVs, AVs) when large-scale RP data are available.

The EV preference heterogeneity analysis in Chapter 3 contributes to existing EV adoption studies by comparing different models which capture random preference heterogeneity. Practically, findings provide insights for policymakers to deploy charging infrastructure and evaluate the effectiveness of EV incentive policies. Methodologically, our results demonstrate the strength and limitations of various advanced discrete choice models in understanding EV preference heterogeneity. Future research on EV preferences should discuss more on model selection, especially when delivering policy-related measures. For more prudent EV policy-making, policymakers can benefit from knowing that there are alternative models which may show different results and interpretations.

Lastly, the AV mode choice preferences analysis in Chapter 4 applies an ICLV modeling framework to the best-worst choice experiment data, uncovering the interrelationships among socio-demographic characteristics, latent attitudes, and AV mode choice preferences. Results indicate policy opportunities to encourage certain AV modes among specific socio-economic groups, such as via marketing campaigns to intervene in their latent attitudes. Additionally, the study distributes the same survey instrument in two metropolitan areas in the U.S. (Seattle and Kansas City) with distinct population densities and existing travel mode shares. Findings help metropolitan planning organizations (MPOs) understand how local contexts matter in shaping AV mode preferences, which should be explicitly considered in AV planning.

5.3 Limitations and future research

We first note the data limitation of this research. Unlike ride-hailing services, EVs and AVs have limited real-world RP data for preference analysis. Thus, the preference analysis for EVs and AVs is based on SP surveys, which suffer from hypothetical bias. People's stated choice behavior may not represent real preferences, particularly given their lack of direct EV or AV experiences, and thus the findings should be interpreted with caution. With the gradual deployment of vehicle electrification and automation, the joint RP and SP data serve as a potential avenue for future work to study preferences for EVs and AVs (e.g., Jia et al., 2021).

The second limitation relates to the modeling approaches. Discrete choice models are used for analyzing the consumers' choice preferences in this dissertation. The discrete choice models have a long tradition in travel behavior research and show well-established records in eliciting behavioral interpretation and policy implications. However, it has limitations in prediction accuracy and handling non-linear and threshold effects compared with the emerging machine learning approaches (Hillel et al., 2021). Combining the discrete choice modeling and machine learning approaches represents a promising methodology frontier for studying consumer preferences.

Despite the aforementioned limitations, the results of this dissertation, which account for consumers' preference heterogeneity, can be integrated with agent-based models to evaluate the impact of emerging transportation technologies on regional vehicle-mile-traveled (VMT), mode share, energy consumption, and emissions. The consumer heterogeneous preferences fueled simulations can also be used to assess the effectiveness of various policies (e.g., EV purchase incentives, VMT use fee, parking pricing, etc.) in reducing VMT and emissions. Moving beyond the realm of the "three revolutions", the research framework can be used to study the adoption of future emerging technologies, as the transportation system keep evolving.

Reference

Hillel, T., Bierlaire, M., Elshafie, M. and Jin, Y., 2021. A systematic review of machine learning classification methodologies for modelling passenger mode choice. Journal of Choice Modelling, 38, p.100221.

Jia, W. and Chen, T.D., 2021. Are Individuals' stated preferences for electric vehicles (EVs) consistent with real-world EV ownership patterns?. Transportation Research Part D: Transport and Environment, 93, p.102728.

Publications and Presentations

Publication Plan

- Jia, W., Chen, T.D. Disparities in Ride-hailing Usage Under Different Spatial Contexts. (To be submitted to Journal of Transport Geography)
- Jia, W., Chen, T.D. Investigating Heterogeneous Preferences for Electric Vehicles. (To be submitted to Transportation Research Part A)
- Jia, W., Chen, T.D., Zhang, W., Wang, K. Local Context Matters: Examining Mode Choice Preferences in the Autonomous Vehicle Era in Two U.S. Metropolitan Regions. (To be submitted to Transportation Research Part C)

Under Review

- Jia, W., Chen, T.D. 2021. Beyond Adoption: Examining Zero-Emission Vehicle Miles Traveled (zVMT) in Households with Zero-emission Vehicles. (Under review in Transportation Research Record)

Published Papers

- Jia, W. and Chen, T.D., 2021. Are Individuals' stated preferences for electric vehicles (EVs) consistent with real-world EV ownership patterns? Transportation Research Part D: Transport and Environment, 93, p.102728.
- Jia, W., Jiang, Z., Chen, T.D. and Paleti, R., 2019. Evaluating Fuel Tax Revenue Impacts of Electric Vehicle Adoption in Virginia Counties: Application of a Bivariate Linear Mixed Count Model. Transportation Research Record, 2673(9), pp.548-561.

Presentations

- Jia, W., Chen, T.D. "Beyond Adoption: Examining Zero-Emission Vehicle Miles Traveled (zVMT) in Households with Zero-Emission Vehicles". TRB 100th Annual Meeting, Washington, D.C., 2021/01.
- Jia, W., Chen, T.D., Zhang, W., Lim, L., Aad, M. "Willingness-to-Relocate: Analyzing Travelers' Parking Preferences for Private Automated Vehicles". TRB 100th Annual Meeting, Washington, D.C., 2021/01.
- Jia, W., Chen, T.D. "Are individuals Stated Preferences for Electric Vehicles (EVs) Consistent with Real-world EV adoption patterns? A Case Study in Virginia, U.S.". TRB 99th Annual Meeting, Washington, D.C., 2020/01.
- Jia, W., Chen, T.D. "Investigating Virginians' Heterogeneous Preferences for Electric Vehicles: Uncovering Policy Implications via a Comparison of Choice Modeling Methods". TRB 99th Annual Meeting, Washington, D.C., 2020/01.

- Jia, W., Jiang, Z., Chen, T.D. and Paleti, R. "Evaluating the Fuel Tax Revenue Impacts of Vehicle Electrification in Virginia Counties: Application of a Bivariate Linear Mixed Count Model." ASCE International Conference on Transportation & Development. Alexandria, Virginia, 2019/06.
- Jia, W., Chen, T.D. "Are Individuals' Stated Preferences for Electric Vehicles (EVs) Consistent with Real-world EV Adoption Patterns?" Engineering Sustainability 2019, April 2019, Pittsburgh, Pennsylvania, 2019/04.
- Jia, W., Chen, T.D. "Evaluating the Fuel Tax Revenue Impacts of Vehicle Electrification in Virginia Counties: Application of a Bivariate Linear Mixed Count Model". TRB 98th Annual Meeting, Washington, D.C., 2019/01.
- Jia, W., Jiang, Z., Chen, T.D. and Paleti, R. "Evaluating Fuel Tax Revenue Impacts of Electric Vehicle Adoption: A County-Based Virginia Case Study." International Transportation Economics Association Annual Meeting 2018, Hong Kong, China, 2018/06.