Title
RISE OF THE MACHINES? A characterization of users of Hyundai’s Robotaxi Pilot Program

Author
Halog, Janelle

Publication Date
2021

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,
IRVINE

RISE OF THE MACHINES?
A characterization of users of Hyundai's Robotaxi Pilot Program

THESIS

submitted in partial satisfaction of the requirements
for the degree of

MASTER OF SCIENCE

in Civil Engineering

by

Janelle M. Halog

Thesis Committee:
Professor Jean-Daniel Saphores, Chair
Professor Michael Hyland
Professor Wenlong Jin

2021

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

ACKNOWLEDGEMENTS

Firstly, I would like thank my advisor, Professor Jean-Daniel Saphores, for his kindness and advice during my research. I am truly grateful for his guidance and the genuine enthusiasm he has shown in my research.

Thank you as well to the members of my committee, Professor Michael Hyland and Professor Wenlong Jin, for your valuable feedback and time in ensuring the quality of this thesis.

I would like to thank my fellow coworkers at Advanced Civil Technologies for their patience and encouragement while I balanced work and school. I especially would like to thank Karen Cohoe, Jamal Salman, and Isaac Alonso Rice for their understanding and words of motivation.

I would also like to thank my significant other, Kevin Nguyen, for supporting my well-being and helping me stay determined. Thank you for always being there for me through thick and thin.

Lastly, I would like to thank my family, friends, and community for cheering me on and believing in me throughout my educational career and endeavors. Your kind gestures do not go unforgotten.

To those that have ever provided moral support or checked in on me, thank you.

ABSTRACT OF THE THESIS

RISE OF THE MACHINES?

A characterization of users of Hyundai's Robotaxi Pilot Program

by

Janelle M. Halog

Master of Science in Civil Engineering

University of California, Irvine, 2021

Professor Jean-Daniel Saphores, Chair

On November 4th, 2019, in cooperation with Pony.ai and VIA, Hyundai launched in Irvine, California, BotRide, an innovative pilot program and robotaxi service. With BotRide, participants could hail a ride in an autonomous vehicle to various destinations in Irvine for free through a mobile application. After the program ended, a survey was sent in February 2020 to BotRide participants to collect socioeconomic data and information about user experience. BotRide is unique, since, in addition to having ridehailing and ridesharing elements, it is a robotaxi service. As robotaxis are just emerging, robotaxi research is still in its infancy, so my thesis is one of the first studies to contribute to the robotaxi literature. The purpose of this thesis is to analyze and investigate how various socioeconomic characteristics, such as occupation, gender, or ethnicity, may influence a user's likelihood to use Hyundai's BotRide service. To achieve this goal, I analyzed the February 2020 survey data provided by Hyundai and estimated a binary logit model (to characterize BotRide users) and a multinomial logit model (to contrast light and heavy BotRide users) using Stata. Results

indicate that household income, the convenience or functionality of using BotRide, working fulltime in Irvine, or being an Irvine resident may reduce the likelihood of using BotRide. In contrast, having some level of familiarity with and knowledge about autonomous vehicles increases the likelihood of using BotRide. Hence, Hyundai may consider targeting a certain group of users, encouraging ridership by providing more information on BotRide itself, and exploring ways to make BotRide a more convenient and functional option for users.

CHAPTER 1: INTRODUCTION

1.1 BACKGROUND

On November 4th, 2019, Hyundai in cooperation with Pony.ai and VIA launched in Irvine (CA) BotRide, an innovative pilot program. With BotRide, users were able to hail a ride in an autonomous vehicle to a list of approved destinations for free through a mobile application. Each trip in this BotRide ridesharing service had a maximum of two riders in the vehicle. After the program ended, a survey was sent in February 2020 to BotRide participants to collect socioeconomic data and information about their experience. BotRide is unique since in addition to having ridehailing and ridesharing elements, its novelty stems from the fact that it is a robotaxi service. Since robotaxi systems are just emerging, my thesis is one of the first to characterize BotRide users to understand their motivation to use robotaxis.

1.2 RESEARCH OBJECTIVE

The focus of this research is to investigate how various socioeconomic characteristics may influence a user's tendency to use the BotRide service through the development and use of binary logit and multinomial logit models in Stata.

1.3 THESIS OUTLINE

The remainder of this thesis is organized as follows.

In Chapter 2, I review selected recent papers dealing with ridehailing, ridesharing, and robotaxis and discuss the relevance of each field of research to BotRide.

In Chapter 3, I introduce the survey data I received from Hyundai and explain how I created variables used in the development of the two models estimated in this thesis. After commenting on some data limitations, I present my logit and multinomial logit models.

Chapter 4 discusses and analyzes the results of the logit model and multinomial logit model. For each model, specification tests are also briefly discussed.

Finally, Chapter 5 summarizes the findings of this thesis and provides suggestions for future areas of study.

CHAPTER 2: LITERATURE REVIEW

To inform my modeling choices and especially what variables to consider, I reviewed selected papers dealing with ridehailing, ridesharing, and self-driving taxi services (robotaxis). From my review of the literature, the following variables were found to be statistically significant in the ridehailing, ridesharing, and robotaxi literature: age/generation, race, gender, work status, household income (HHI), children or elderly persons in the household, occupation, vehicle ownership, and household location. Table 1 summarizes key features and relevant findings of the papers I reviewed for this study.

2.1 RIDEHAILING

A number of recent papers have started to characterize users of ridehailing services (such as Uber and Lyft), in which a user hails a vehicle to take them to a destination, and their travel behavior [1][2]. I found four studies that analyze users of ridehailing at the national level, and one at the international level. Several papers in this strand of the literature have focused on specific groups of users of ridehailing services, such as Millennials and members of Generation X, or older riders [1][2][3][4][5]. Other studies have explored the potential impacts of shared mobility on vehicle ownership and vehicle preferences as well as its impact on the use of public transportation [6][7]. Overall, these studies suggest that sociodemographic variables play an important role in predicting the adoption of ridehailing.

Table 1 - Reviewed Papers and Key Features

Authors (Year)	Key Features			
	Main Question	Data	Method	Main Findings
Alemi, Circella, Mokhtarian, and Handy (2019)	What drives the use of ridehailing in California?	California Millennials Dataset, collected in fall 2015 through an online survey	Ordered probit model with sample selection and a zero-inflated ordered probit model with correlated error terms	Sociodemographic variables are important predictors of service adoption but do not explain much of the variation in usage frequency. Individuals who use apps for other aspects of travel, longer-distance travelers, and those more willing to pay for reduced travel time are more likely to use ridehailing. Those who prefer to own a vehicle and with stronger safety concerns are less likely to use ridehailing.
Circella, Alemi, Tiedeman, Handy, and Mokhtarian (2018)	How are shared mobility services (including carsharing, ridehailing and bikesharing) used in California? What factors drive the use of ridehailing? Under what circumstances are individuals more likely to use Uber and Lyft? How frequently do Californians use ridehailing, and how does that frequency vary with sociodemographics, built environment characteristics, individual lifestyles and attitudes?	Fall 2015 online survey with more than 2000 respondents, including millennials and members of Generation X	Surveys, model estimation (binary logit model, latent - class adoption model)	Better-educated individuals who live in predominantly urban areas are more likely to use ridehailing services. Increased land-use mix and regional auto accessibility increase the likelihood of using ridehailing. Adoptions of on-demand ride services is higher among individuals who make more long-distance trips and those who travel more by plane. Individuals with stronger preferences to own a vehicle are less likely to be frequent users of Uber and Lyft.

Authors (Year)	Key Features			
	Main Question	Data	Method	Main Findings
	What limits/encourages the use of these services?			
Payyanadan and Lee (2018)	To understand the current practices and barriers to ridesharing among older adults in rural and urban settings and to translate findings into a web-based ridesharing tool customized to fit the needs of older adults	39 drivers 65 years and older from urban and rural areas from a Midwestern US state, vehicle instrumentation to record ridesharing trips, post-drive interviews	Contextual Design approach (a framework for developing front-end design such that the user data drives the overall system design and development)	Results from the Contextual Design showed that older adults faced four main challenges, which are limited social network, efficient communication of trip details and needs, and establishing trip reliability and privacy.
Dias, Lavieri, Kim, Bhat, and Pendyala (2019)	To better understand the use of ride-hailing services	RideAustin Data, Austin Zoning and Parcel Data, 2016 American Community Survey	Multivariate ordered probit model, data fusion methodology	Socio-demographic variables are important determinants of trip frequencies. Data fusion allows the possibility to answer questions on why and for what purpose ridehailing services are used, where ridehailing users reside, and how many ridehailing trips will be undertaken in different time periods.
Tang, Li, Yu, and Wei (2020)	To examine how app-based ride hailing would impact passengers' choice of travel mode and change car-purchasing behaviors	China is taken as the empirical contex t and data are from a large-scale app-based survey conducted vias the largest app-based ride-hailing platform in the world (DIDI Chuxing platform)	Travel behavior models (binary logit), choice model for alternative travel mode, and choice model for car-purchasing behavior	App-based ridehailing services attract people who pursue fast, affordable, and around 10 to 30-minute point-to-point transport. App-based ride hailing also helps mitigate the inconvenience of limits on private car usage. More than 35% of app-based ridehailing users would have taken traditional taxi services, reflecting the competitive relationship between app-based ride hailing and traditional taxi services. 37% of app-based ridehailing users would have taken public transportation, which shows the necessity of improving current public transport systems and achieving a balance between app-based ridehailing and public transportation. Factors that influence passengers to choose app-based ridehailing are related to their household income and household type.

Authors (Year)	Key Features			
	Main Question	Data	Method	Main Findings
				App-based ridehailing is likely to change future willingness to purchase cars. More than half of app-based ridehailing users showed a change in attitude toward car purchases, and 6.6% of the respondents affirmed they would not purchase new private cars if app-based ridehailing services were permanently available.
Sikder (2019)	Who uses ride-hailing services in the United States?	2017 National Household Travel Survey	Descriptive analysis and ordered logit model (OLM)	Race, work status, children, elderly persons, income, and vehicle household influence the use of ridehailing services. These services also seem to be complementary of public transit.
Sadowsky and Nelson (2017)	How has ridehailing services impacted public transportation use?	Monthly public transportation ridership data at the urbanized area level provided by the Federal Transit Authority (FTA)	Discontinuity regression analysis	There is speculation that the introduction of Uber complemented public transportation by solving the “last mile” problem and by offering a safer option at night when public transportation services are reduced. It is speculated that the entrance of Lyft, the second ridehailing service, increased competition and supply in ridehailing, making an entire trip with a ridehail service more cost-effective and convenient than splitting a trip between a rideshare company and public transportation.
Merat, Madigan, and Nordhoff (2017)	To give an overview of the social-psychological factors that are likely to influence the trust and acceptance of shared Society of Automotive Engineers (SAE) Level 4 Automated Vehicle (AVs)	Studies of human robot interaction (HRI), results from real-world surveys and online surveys	HRI and surveys	It is recommended that the pathway to adoption and acceptance of AVs should be incremental as well as iterative, while providing users with hands-on experience at every stage of the system. Manufacturers should use new technologies, social networks, and crowdsourcing methods for consumer feedback and input to encourage the adoption and acceptance of shared AVs.
Morales Sarriera, Escovar Álvarez, Blynn, Alesbury,	To investigate whether people perceive dynamic ridesharing as having positive or negative utility with respect to its social aspects, what	A survey of transportation network company (TNC) users conducted through	Survey built on an online survey development service called Qualtric, interviews	Dynamic ridesharing users reported social interactions were relevant to mode choice, but not as much as traditional factors such as time and cost. The possibility of having a negative social interaction was more of a deterrent to use of dynamic ridesharing than potentially having a positive social interaction. A substantial number

Authors (Year)	Key Features			
	Main Question	Data	Method	Main Findings
Scully, and Zhao (2017)	influences those perceptions, and how they compare with traditional factors like time and cost	Mechanical Turk, which is a task distribution company, in June and July of 2016, which had 997 respondents across the United States, personal open interviews conducted with individuals that used dynamic ridesharing	assessed the impact of social factors on the perception and use of dynamic ridesharing services	of riders felt prejudice toward passengers of different social class and race, and these passengers were more likely to prefer having more information about potential future passengers. Most dynamic ridesharing users were motivated by ease and speed, in comparison with walking and public transportation. Safety in dynamic ridesharing was an important issue, especially for women, many of whom reported feeling unsafe and preferred to be matched with passengers of the same sex.
Zhang and Zhang (2018)	To better understand individuals' ridesharing behaviors and the interdependencies between vehicle ownership and ridesharing usage	2017 National Household Travel Survey	Zero-inflated negative binomial regression models	Results indicated that a one-vehicle reduction was associated with a 7.9% increase in ridesharing usage frequency as well as a 23.0% increase in the probability of ridesharing usage. Vehicle ownership affects rideshare frequency more in higher density populations than areas with lower density. Young people, men, those that cannot drive, those with high household income, and those that live in areas with rail service or higher population density, have a tendency to use ridesharing more.
Lee, Yoo, Kim, Kim, and Kang (2020)	How do robotaxi service experience and demographics, and positive and negative emotions of robotaxis service experience affect user acceptance? How do main evaluation factors for the quality of the	User experience data, post- and pre-surveys and interviews, 71 participants	Structural equation modeling (SEM) and path analysis to analyze factors affecting user acceptance	User experience had a significant effect on user acceptance. Service quality in the traveling stage had the largest effect on user experience, and overall satisfaction had the largest effect on user acceptance whereas a user's willingness to pay had a relatively low effect. "Cutting-edge" was selected as the typical emotion that had a positive relationship with user acceptance. "Apprehensive" was the typical emotion that had a negative relationship. Service quality in the traveling and

Authors (Year)	Key Features			
	Main Question	Data	Method	Main Findings
	robotaxi service by stage affect robotaxi satisfaction?			drop-off stages were found to have a significant effect on overall satisfaction. Reliability, speed, and kindness were found to be crucial factors in the traveling stage while accessibility, information, and communication were found to be important factors in the drop-off stage.
Meurer, Pakusch, Stevens, Randall, and Wulf (2020)	To study passengers' robotaxi service experiences in real-life settings	10 participants, 33 rides recorded on video; pre- and post-interviews	Wizard of Oz Study, which allowed the user to interact with an interface (simulated robotaxi with a hidden driver in this case) without knowing responses were produced by a human	It is recommended that travelers should be able to appropriate the space of the robotaxi during their trip. It is also suggested that passengers should feel they are in control of their situation and that designing passenger-centric information systems should resolve this issue. The customer's trip should be designed to be a coherent and consistent experience.
Sanguinetti, Ferguson, Oka, Alston-Stepnitz, and Kurani (2021)	To articulate potential design solutions to promote pooling	12 different shared automated vehicle (SAV) designs	Reviews current SAV designs and literature	To mitigate the perceived risks of ride-pooling, it is recommended to increase personal space, defensible space, and perceived control. The benefits of ride-pooling were defined as restorative environments, which are sites that provide stress relief through aesthetic design as well as social capital.
Vosooghi, Kamel, Puchinger, Leblond, and Jankovic (2019)	To explore the impact of user trust and willingness-to-use on robotaxi fleet size	Transportation system data of the Rouen-Normandie metropolitan area in France, local survey for explore variation of user trust and	Multi-agent simulation (MATSim) of the transportation system of the Rouen-Normandie metropolitan, synthetic	Modal share increases proportionally to the fleet size. Mode shifts toward robotaxi come from public transport, car, and walk. However, the use of public transport decreases significantly relative to other ones. Women and elderly people are less likely to use a robotaxi while students and persons younger than 14 use robotaxis more significantly. Robotaxi usage for employed people is minor and fluctuated for unemployed people. Maximum

Authors (Year)	Key Features			
	Main Question	Data	Method	Main Findings
		their willingness-to-use future robotaxis	population generation	fleet usage happens with a fleet size between 2000 and 3000.
Yoo, Lee, Kim, Kim, Hwangbo, and Kang (2020)	To derive various internal/external factors that contribute to the anxieties of robotaxi passengers, and to propose a human-machine interface concept to resolve such factors	28 subjects in the central area of Seoul, user experience surveys and interviews	Wizard of Oz methodology, which used a remote system for safe testing of the robotaxi, used in two field tests	The major anxiety factors in riding robotaxis were "cut-in, turning, pedestrian, illegal parking, alley, accident occurrence alarm, reckless driving (external vehicle), horn sound (external vehicle), speed, and protruding vehicle." It was also found that people preferred flexibility in robotaxi driving and outside interaction.

In a 2019 study by Alemi *et al.*, a dataset created from a 2015 online survey distributed to Californian Millennials and members of Generation X was analyzed. This study found that individuals who have a tendency to use smartphone applications to manage other traveling aspects, such as choosing a route or checking traffic, have a higher likelihood of using ridehailing. Factors like preferences towards owning a personal vehicle or having safety concerns, however, appear to discourage the use of ridehailing. Other statistically significant variables included age, household income, presence of children in the household, neighborhood type, and general attitudes [2].

Circella *et al.* (2018) found similar results. Variables like age were also significant and indicated the use of ridehailing was the greatest in higher-educated Californian Millennials compared to other groups [1]. This differs in comparison to older adults who prefer the comfort of riding in a personal vehicle, which may result from challenges in socializing outside of their network, lack in efficient communication of trip requirements, and concerns related to privacy and reliability as suggested by Payyanadan and Lee in their 2018 paper [3].

Income appears significantly relevant to an individual's tendency to use ridehailing. According to a study done by Dias *et al.* (2019), those with lower income use ridehailing for practical purposes while those with higher incomes resort to ridehailing for airport trips and recreational activities [4]. An empirical study in China by Tang *et al.* (2020) also supports the influence of income on user frequency while recognizing that household types may have an influence as well [6]. In line with the idea of household types having influence, Sikder

(2019) indicates that employment schedules and the presence of elders, children, or both matter to explain frequency of use [5].

Ideally, ridehailing should complement public transportation, which can be the case according Sikder's 2019 study [5]. However, in urban areas at the very least, Sadowsky and Nelson (2017) found that the staggered introduction of services like Uber and Lyft resulted in a decrease in public transportation ridership [7]. As such, collaboration between public transportation agencies and ridehailing services should occur to create a more integrated transportation system. Hence, understanding why users opt to ride BotRide and drawing connections between user characteristics and usage may bring to light how the public travels at least in Irvine, California.

Since BotRide is a form of ridehailing, many of the socioeconomic factors found in the literature to impact ridership helped shape the variables used in the models I estimated to analyze BotRide usage. While many previous studies cover ridehailing in larger areas, BotRide is restricted by design to Irvine, California.

2.2 RIDESHARING

I also reviewed selected recent papers dealing with attitudes towards ridesharing services, and especially shared autonomous vehicles (AVs) [8][9]. I found three studies that analyze users of ridesharing, in which users carpool in a vehicle. Two of these studies are at the national level and one at the international level. Of the three papers reviewed, one focuses mainly on the relationship between vehicle ownership and ridesharing while the other two focus more on rideshare acceptance [8][9][10]. These studies suggest that mentality and attitudes toward ridesharing and autonomous vehicles (AVs) play an important role in predicting the willingness to use those services.

According to Merat *et al.* (2017), based on studies mostly done in Europe, shared AVs systems should have the following characteristics in order to increase user acceptance: reliability, consistent availability, safety, comfort, privacy, and accessibility to children and elders [8]. Additionally, Sarriera *et al.* (2017) conducted in 2016 a survey of transportation network companies users through Mechanical Turk and their findings highlight that safety in ridesharing is important, particularly for woman due to reports of many feeling unsafe in their shared AV experiences [9].

Vehicle ownership is also tied to the use of ridesharing services. Based on their zero-inflated negative binomial models estimated on data from the 2017 National Household Travel Survey, Zhang and Zhang (2018) found that owning one less vehicle results in a 7.9% increase in rideshare frequency with a 23.0% increase in probability in using a rideshare service. Furthermore, younger generations, men, those unable to operate a vehicle, people with higher household incomes, and people living in high population density areas or in regions with rail service are more inclined to use ridesharing services [10].

Although the focus of this thesis is BotRide as a robotaxi, I also reviewed selected papers related to ridesharing. Even though it is an older concept, the factors underlying the use of ridesharing are relevant to BotRide. For example, safety is a factor that has potential influence on the use of both services. As such, BotRide has similarities with ridesharing, so I considered in my models the variables found to be statistically significant in previous ridesharing studies, when they were available.

2.3 ROBOTAXIS

As robotaxis are becoming more commonplace, some researchers have begun to analyze the experience and demographics of robotaxi service users [11][12][13][14][15]. I found four studies that analyze users of robotaxi services as well as their experiences at the international level, with two in Europe [12][14] and two in South Korea [11][15]. Moreover, I found another study at the national level that mainly looked at how to promote robotaxi design to attract ridership [13]. These studies suggest that experience and emotional attitudes play an important role in predicting and encouraging the use of robotaxi services.

Currently, there is a lack of research surrounding data obtained directly from robotaxi user experience. One exception is Lee *et al.* (2020), who analyzed user experience data from a robotaxi service in downtown Seoul and Daejeon (South Korea) to better understand the relationship between user experience and user acceptance, emotional factors, and satisfaction. Overall, they found that reliability, speed, kindness, accessibility, and communications are crucial factors in the willingness to use robotaxi services [11].

In their 2020 study, Meurer *et al.* simulated an autonomous robotaxi service where in fact the driver of an electric vehicle was hidden from the passenger's view. Interviews were then conducted before and after the service was implemented to gain feedback from participants. Ultimately, that study found that users need to be able to appropriate the interior robotaxi space, feel in control, and have consistent experiences [12]. Although there are many perceivable benefits to robotaxis such as potential to reduce traffic, emissions, and energy use, user acceptance needs encouragement to help robotaxis become more commonplace. As discussed in a recent paper by Sanguinetti *et. al.* (2021), design suggestions to improve acceptance include personal and defensible space, perceived control, a restorative environment, and social capital [13].

From the above studies, safety is an important concern among users. For example, according to Vosooghi *et al.* (2019), women and elderly persons, which include homemakers and retired individuals, may be less likely to utilize robotaxis for fear of being placed in uncomfortable situations [14]. Other anxiety factors related to safety concerns, which include turns and risky driving maneuvers, may also hinder robotaxi use according to Yoo *et al.* [15].

Since many worries surrounding the use of robotaxis are due to safety, the BotRide data was also analyzed using variables potentially linked to safety and welfare such as personality, ethnicity, attitude and familiarity with AVs, and gender. Furthermore, as mentioned earlier, there is a lack in research that analyzes user experience. Hence, this study analyzes BotRide user data to address this gap and add to the emerging literature on robotaxis.

CHAPTER 3: DATA AND MODELS

3.1 DATA AND MODEL VARIABLES

My first model developed for this thesis is a logit model. A logit model explains a binary dependent variable assumed to be tied to a latent variable explained linearly by a set of explanatory variables [16]. Here my binary dependent variable equals 1 for respondents who tried Hyundai's BotRide robotaxi service and 0 otherwise.

My second model is a multinomial logit model. This type of model is similar to a logit model except that it considers more than two outcomes [16]. Here my dependent variable equals 1 if the respondent tried BotRide once or twice, 2 if the respondent tried BotRide three or more times, and 0 otherwise.

As mentioned in previous chapters, the data analyzed in this thesis come from a survey of BotRide users that was conducted in February 2020 by Hyundai. Unfortunately, I was not involved in the development of the survey and could not add questions that would have been useful to this study. As a result, I had to work with information that was provided by Hyundai's survey results.

From the BotRide dataset, I created a number of binary variables based on my literature review (see Chapter 2). The original dataset had a sample size of 1,279, however, 442 users of that sample provided an incomplete response. To avoid missing data issues, the original sample size was reduced to 837 to ensure there were responses for all the variables of interest. If a response was missing for any one of the variables in Table 2, the user was removed from the dataset.

Table 2 - Variables Used in Models

Variable Category	Variable Name/Description	Definition
Dependent Variables		
Usage (Logit)	BotRide usage	1 if user took 1 ride or more, 0 otherwise
Usage (Multinomial Logit)	BotRide usage	0 if the user took no rides 1 if the user took 1-2 rides 2 if the user took 3 or more rides
Binary Explanatory Variables (* indicates mutual exclusivity)		
Personality	Risktaker	1 if user is a risktaker, 0 otherwise
	Spontaneous	1 if user is spontaneous, 0 otherwise
	Saver	1 if user is a saver, 0 otherwise
	Pessimist	1 if user is a pessimist, 0 otherwise
Occupation*	UCI Student/Faculty	1 if user is a UCI Student/Faculty, 0 otherwise
	Irvine Resident	1 if user is an Irvine resident, 0 otherwise
	Works fulltime in Irvine	1 if user works fulltime in Irvine, 0 otherwise
	Other	1 if user is either an HCA employee, City of Irvine Employee, in journalism or media, or is an automotive or self-driving related employee; 0 otherwise
Live With	Roommate(s)	1 if user lives with roommates, 0 otherwise
	Parents	1 if user lives with parents, 0 otherwise
	Spouse/Significant Other	1 if user lives with a spouse or significant other, 0 otherwise
	Children under 18	1 if user lives with children under the age of 18, 0 otherwise
	Children 18+	1 if user lives with children of age 18+, 0 otherwise
	Living Alone	1 if user lives alone, 0 otherwise
Ethnicity*	White/Caucasian	1 if user is White/Caucasian, 0 otherwise
	African American/Black	1 if user is African American/Black, 0 otherwise
	Asian American	1 if user is Asian American, 0 otherwise
	Other	1 if user is an ethnicity not listed, 0 otherwise
	Prefer not to say	1 if user prefers not to say their ethnicity, 0 otherwise
	Mixed	1 if user is of mixed ethnicity, 0 otherwise
	Latinx	1 if user is Latinx, 0 otherwise
Gender	Female	If user is female, 0 otherwise

Variable Category	Variable Name/Description	Definition
HHI*	Under $35,000	1 if user has an HHI under $35,000, 0 otherwise
	$35,000 - $74,999	1 if user has an HHI between $35,000 - $74,999, 0 otherwise
	$75,000 - $99,999	1 if user has an HHI between $75,000 - $99,999, 0 otherwise
	$100,000 - $149,999	1 if user has an HHI between $100,000 - $149,999, 0 otherwise
	$150,000 - $199,999	1 if user has an HHI between $150,000 - $199,999, 0 otherwise
	$200,000 or more	1 if user has an HHI of $200,000 or more, 0 otherwise
	Prefer not the say	1 if user prefers not to say their HHI, 0 otherwise
Personal Vehicle	Does not have regular access to a personal vehicle	1 if user has regular access to a personal vehicle, 0 otherwise
	Has a personal vehicle	1 if user has a personal vehicle, 0 otherwise
	Does not have a personal vehicle, but has regular access to someone else's	1 if user does not have a personal vehicle, but has regular access to someone else's; 0 otherwise
Attitude Towards AVs	Very comfortable with the idea of riding in a self-driving vehicle/AV	1 if user is very comfortable, 0 otherwise
Public Transit Experience	Has a lot of experience	1 if user has a lot of experience, 0 otherwise
	Has some experience	1 if user has some experience, 0 otherwise
	Does not have much experience	1 if user does not have much experience, 0 otherwise
Familiarity with AVs Based on Experience	Has extensive knowledge	1 if user has extensive knowledge, 0 otherwise
	Has solid knowledge	1 if user has solid knowledge, 0 otherwise
	Has a little bit of knowledge	1 if user has a little bit of knowledge, 0 otherwise
	Does not really know anything	1 if user does not really know anything, 0 otherwise
Convenience vs. Free	BotRide is more about convenience than saving	1 if to the user, BotRide is about convenience in the user's routine or schedule; 0 otherwise
Function vs. Fun	Botride is more about function than fun	1 if to the user, BotRide is about function and getting the user from point A to B; 0 otherwise

For the dataset, variables for the following categories were developed, which cover (at least partially) the significant variables previously mentioned in the literature review: personality, occupation, living situation, race/ethnicity, gender, household income (HHI), and access to a personal vehicle. Variables for each of the above-mentioned categories are shown in Table 2.

I used Stata to modify variables and to compile the final dataset for the models. The development for the variables in each of the categories is detailed below.

3.1.1 DEPENDENT VARIABLE (BOTRIDE USAGE)

The variable "BotRide usage" was developed to contrast users and non-users of BotRide services. The original BotRide data for usage had the following possible categories:

- User did not take a ride with BotRide
- User took 1-2 rides
- User took 3-5 rides
- User took 6-9 rides
- User took 10+ rides

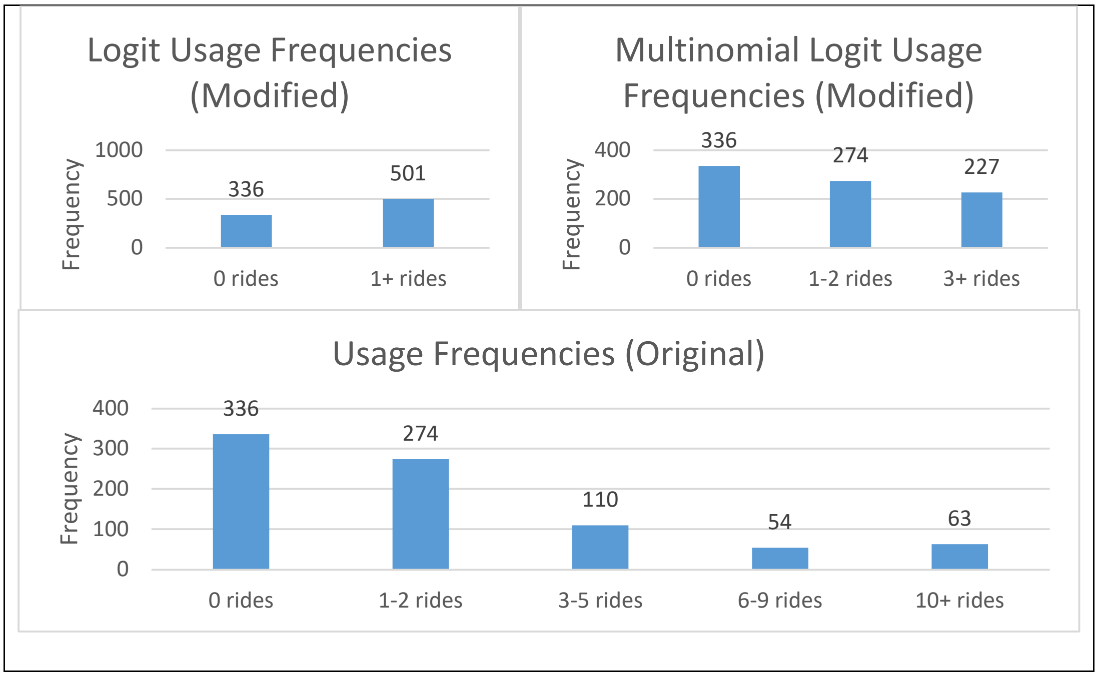
To convert this variable to a binary variable, if a user took a ride, it was coded as 1, and 0 otherwise. For use in the multinomial model, "BotRide usage" was split into the following categories:

- 0 if the user took no rides

- 1 if the user took 1-2 rides
- 2 if the user took 3 or more rides

Figure 1 shows the modified usage data frequencies used in the logit and the multinomial logit models, as well as the original data.

Figure 1 - Modified and Original Usage Data



3.1.2 BINARY EXPLANATORY VARIABLES

<u>Personality</u>

The personality variable consists of the categories "Risktaker", "Spontaneous", "Saver", and "Pessimist". In the original dataset, users self-selected their personality from ranking personality traits on a scale of 1 to 5, where a rank of 5 meant the user strongly associated with the personality trait and 0 indicated the reverse. To convert each personality category into a binary variable, I recoded a ranking of 1 to 3 as "0" and a ranking of 4 to 5 as "1". For example, for “Risktaker”, users coded as 0 were defined to be cautious, while users coded as 1 were defined as a risktaker. For “Spontaneous”, 0 was defined as planner and 1 was defined as spontaneous. For “Saver”, 0 was defined as spender and 1 was defined as saver. For “Pessimist”, 0 was defined as optimist, and 1 was defined as pessimist.

<u>Occupation</u>

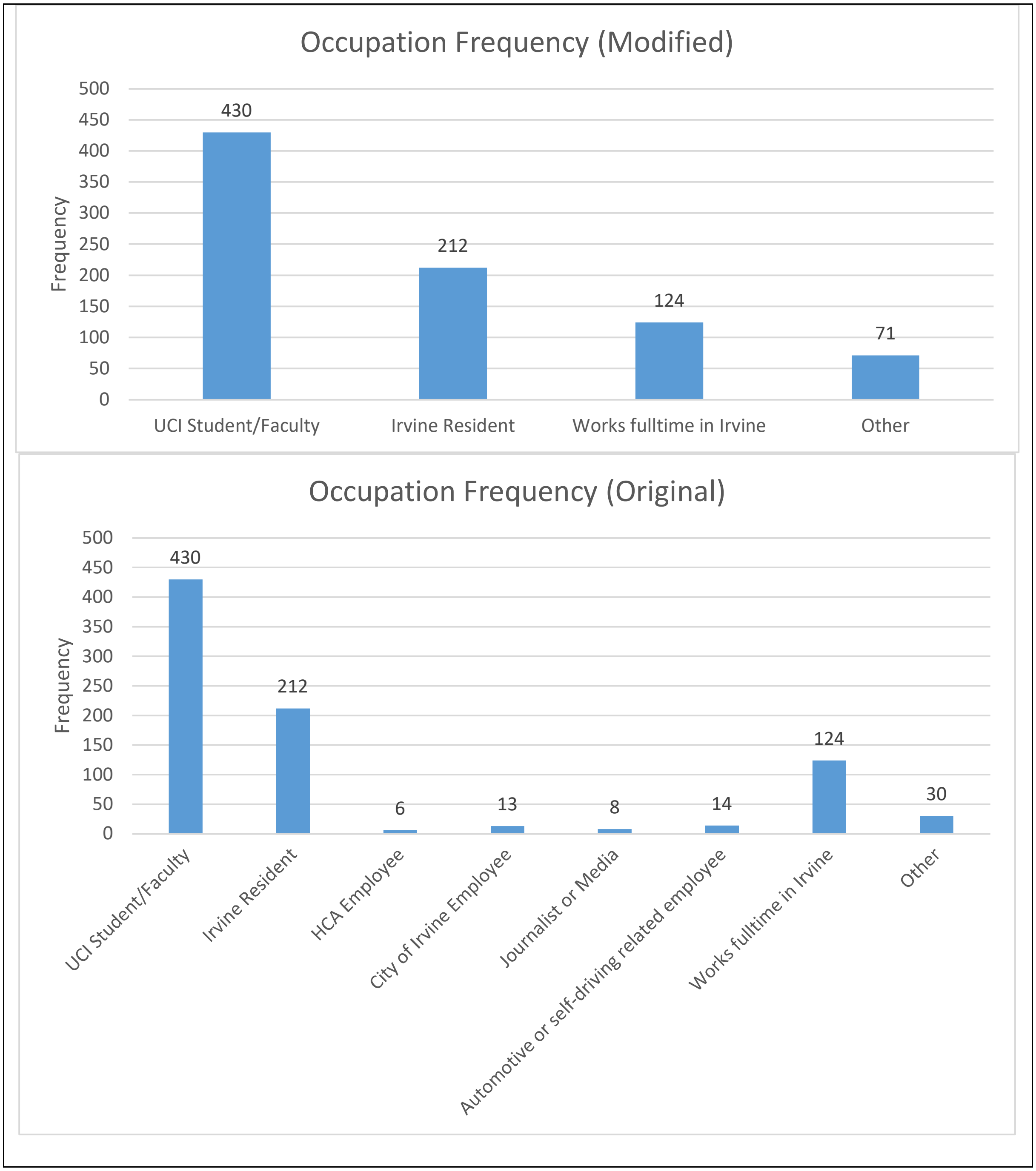
In the “Occupation” category, the original dataset gave users the following choices:

- UCI Student/Faculty *
- Irvine Resident *
- HCA (Healthcare) Employee
- City of Irvine Employee
- Journalist or Media
- Automotive or self-driving related Employee
- Works fulltime in Irvine *

Of the above options, those labeled with “*” had the largest responses compared to the other remaining choices. As such, the remaining categories were combined into a new variable called “Other”.

Figure 2 below shows the modified occupation data frequencies used in the models as well as the original data.

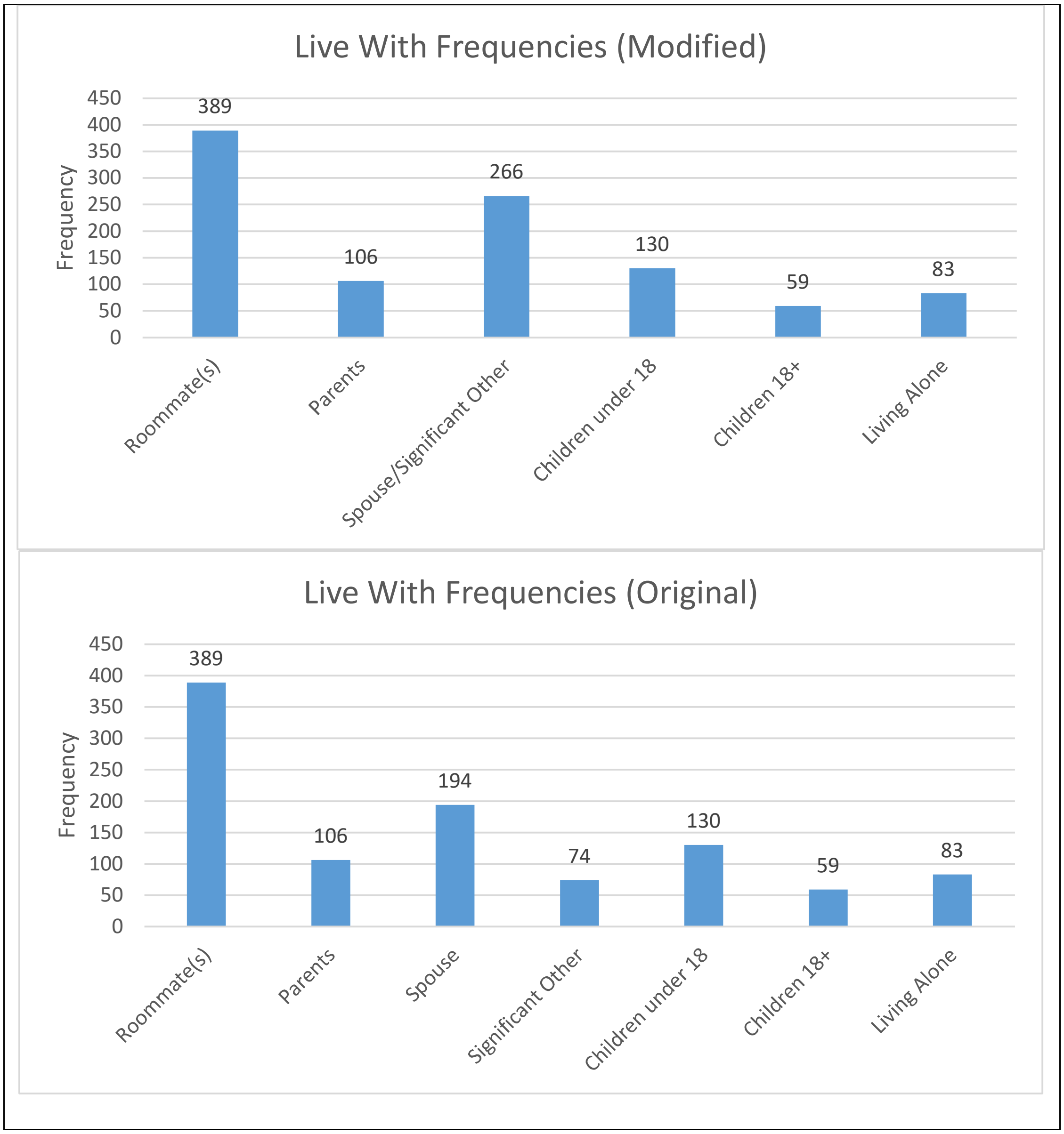
Figure 2 - Modified and Original Occupation Data



Live With

In the "Live With" category, the variables for living with a spouse and living with a significant other were originally separate in the dataset. However, they were combined as there was no definite difference between spouse and significant other. Figure 3 below shows the modified "Live With" data frequencies used in the models as well as the original data.

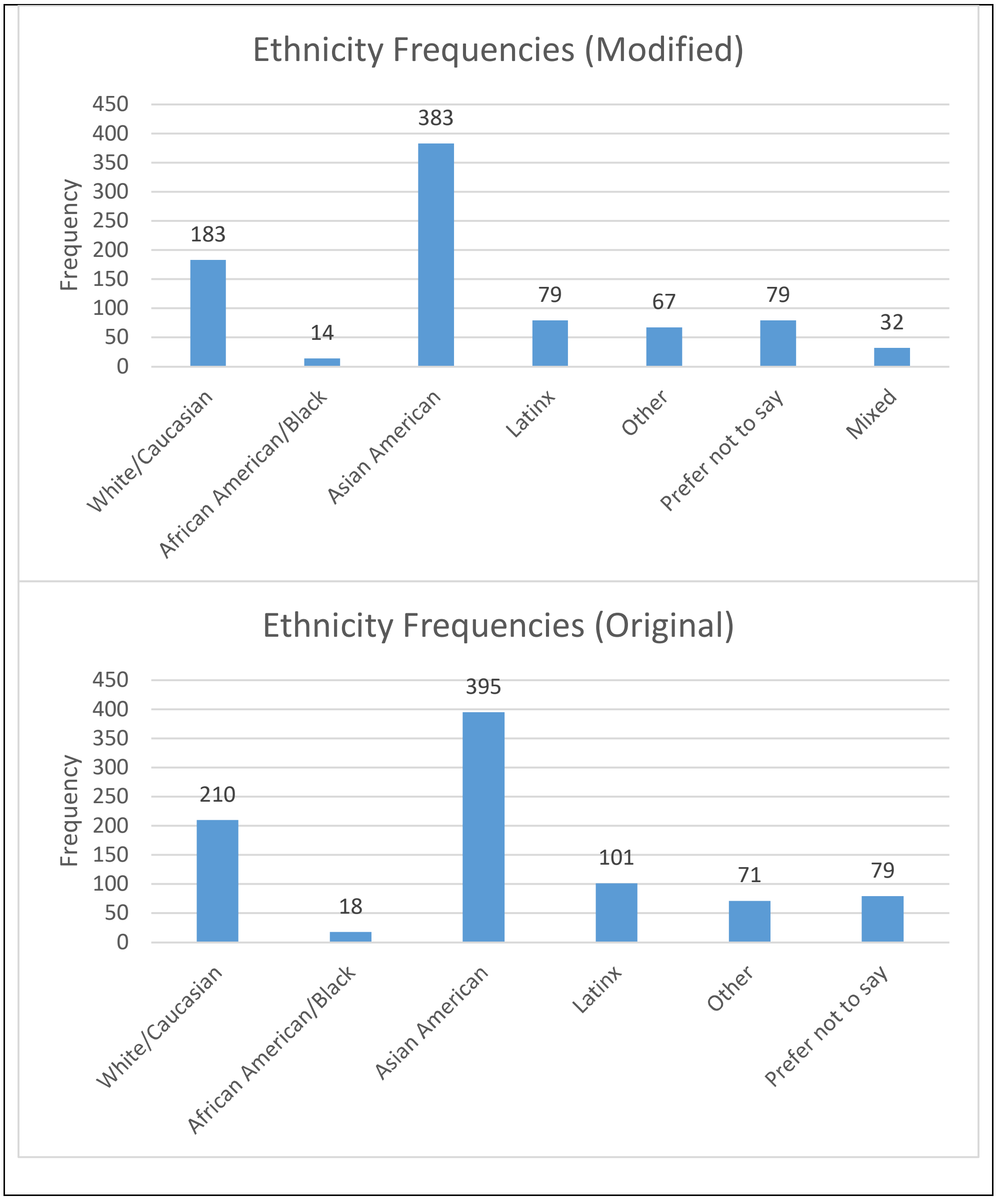
Figure 3 - Modified and Original "Live With" Data



<u>Ethnicity</u>

Out of the 837 observations, 32 users identified with more than one ethnicity and were thus placed into a new ethnicity variable for "Mixed" to represent users with a mixed ethnicity background. Figure 4 below shows the modified ethnicity data frequencies used in the models as well as the original data.

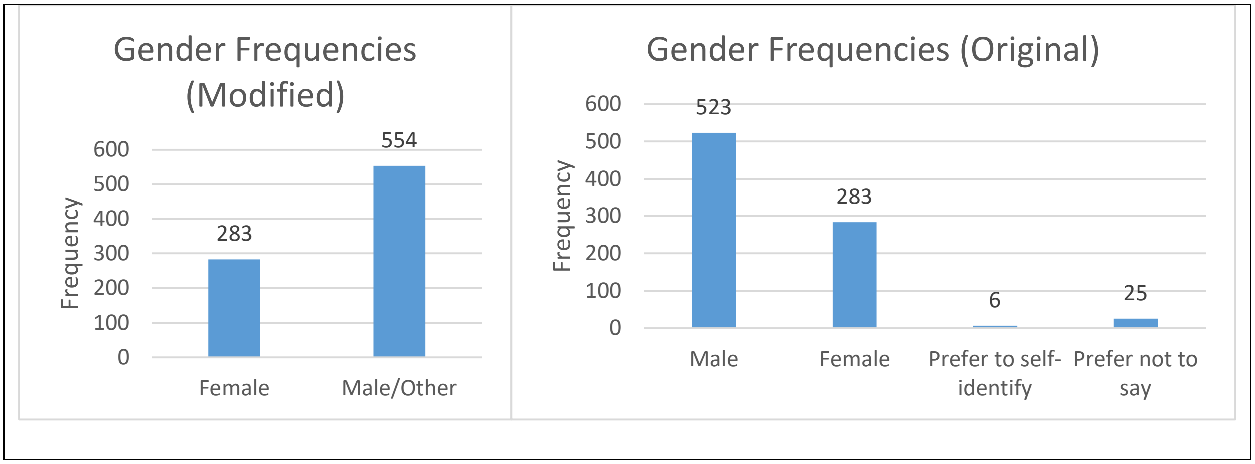
Figure 4 - Modified and Original Ethnicity Data



<u>Gender</u>

In the "Gender" category, the data for those that reported as male were combined with "prefer to self-identify" and "prefer not to say" since the latter two had a small number of responses. Figure 5 below shows the modified gender data frequencies used in the models as well as the original data.

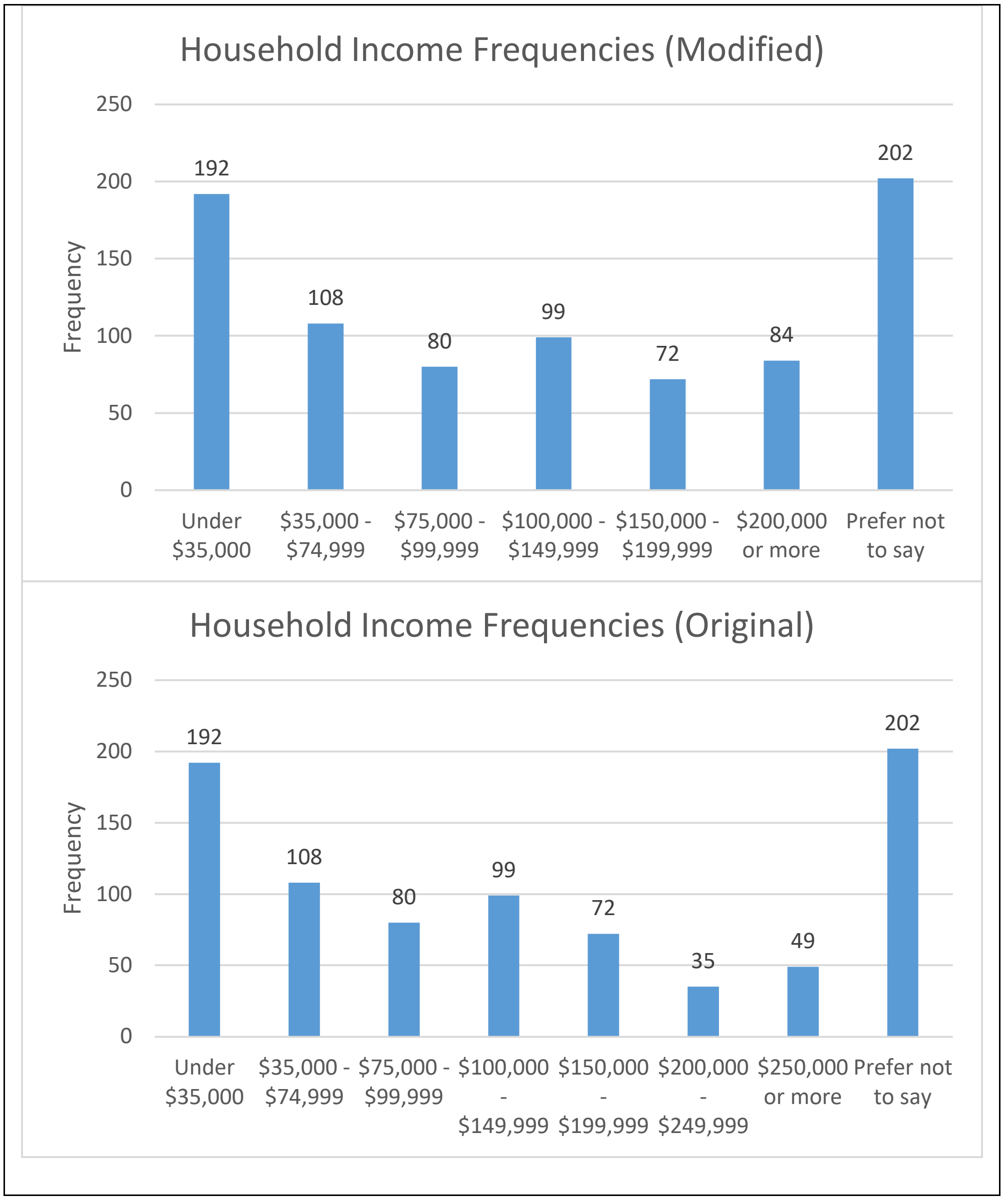
Figure 5 - Modified and Original Gender Data



<u>Household Income (HHI)</u>

The number of responses for those who made "$200,000 to $249,999" and "$250,000 or more" was small so it was combined to create a variable for those who earned "$200,000 or more" per year. Figure 6 below shows the modified HHI data frequencies used in the models as well as the original data.

Figure 6 - Modified and Original HHI Data



<u>Personal Vehicle</u>

Having access to a personal vehicle may have influenced the user's decision to use BotRide and was therefore included per the literature review.

Attitude Towards AVs

Emotion and attitude towards technology is likely to influence someone's decision to use that technology. Since there was a question in the BotRide data regarding user comfort with self-driving vehicles and AVs, it was included in this study.

Public Transit Experience

A user's experience with public transit may influence their comfort with riding new public transportation and new transportation technology. Therefore, variables for public transit experience were included.

Familiarity with AVs Based on Experience

Familiarity and experience with a service may allow the user to gain comfort, which could influence the user to use the service more. In addition, having more experience may lead the user to feel safer.

Convenience vs. Free

This variable looks at whether the user sought BotRide as a convenient solution or if they viewed BotRide in terms of saving money since the service was offered for free. Intuitively, I believe that either view may influence the usage of BotRide.

Function vs. Fun

This variable looks at whether the user sees BotRide as a practical solution for getting from point A to B or if BotRide is a fun option to test new technology. Intuitively, I believe that a functional opinion of BotRide may be associated with higher usage since things that are typically seen as fun are also seen as optional.

<u>Frequency Statistics</u>

Frequency statistics of my binary explanatory variables are shown in Table 3.

Table 3 - Explanatory Variable Frequency Statistics (Sample Size N=837)

Variable Category	Variable Name/Description	Yes (Count)	Yes (%)
Personality	Risktaker	313	37.4
	Spontaneous	259	30.9
	Saver	489	58.4
	Pessimist	166	19.8
Occupation	UCI Student/Faculty	430	51.4
	Irvine Resident	212	25.3
	Works fulltime in Irvine	124	14.8
	Other	71	8.5
Live With	Roommate(s)	389	46.5
	Parents	106	12.7
	Spouse/Significant Other	266	31.8
	Children under 18	130	15.5
	Children 18+	59	7.0
	Living Alone	83	9.9
Ethnicity	White/Caucasian	183	21.9
	African American/Black	14	1.7
	Asian American	383	45.8
	Latinx	79	9.4
	Other	67	8.0
	Prefer not to say	79	9.4
	Mixed	32	3.8
Gender	Female	283	33.8
HHI	Under $35,000	192	22.9
	$35,000 - $74,999	108	12.9
	$75,000 - $99,999	80	9.6
	$100,000 - $149,999	99	11.8
	$150,000 - $199,999	72	8.6
	$200,000 or more	84	10.0
	Prefer not the say	202	24.1
	Does not have regular access to a personal vehicle	174	20.8

Variable Category	Variable Name/Description	Yes (Count)	Yes (%)
Personal Vehicle	Has a personal vehicle	510	60.9
	Does not have a personal vehicle, but has regular access to someone else's	153	18.3
Attitude Towards AVs	Very comfortable with the idea of riding in a self-driving vehicle/AV	795	95.0
Public Transit Experience	Has a lot of experience	370	44.2
	Has some experience	388	46.4
	Does not have much experience	79	9.4
Familiarity with AVs Based on Experience	Has extensive knowledge	89	10.6
	Has solid knowledge	299	35.7
	Has a little bit of knowledge	394	47.1
	Does not really know anything	782	93.4
Convenience vs. Free	BotRide is more about convenience than saving	240	28.7
Function vs. Fun	Botride is more about function than fun	397	47.4

3.1.3 DATA LIMITATIONS

In Hyundai's dataset, age was not explicitly asked, but it is partially covered by occupation. For example, if someone classified their occupation as a student, one could assume that they are part of the younger generation. However, this may not always be the case, so not having specific data about age is a limitation. Occupation is also used in place of a work status variable, since direct information of work status is not provided in the dataset. For instance, although a user may describe their main occupation as a student, they may also work part time.

Other limitations include the absence of information related to whether elderly persons live in the user's household and the user's household location. In the "Live With" category, there is a variable for whether the users live with their parents, who may be elderly, but there is no other information about whether other elderly persons are part of the household. The "Occupation" category also partially covers household location since work location is an indicator of where one lives, but it does not fully define where the user may live.

As mentioned earlier, another limitation of the data is that I had no involvement in creating the survey that was distributed to BotRide users to take. Hence, I could not obtain information such as age and specific work status of the users.

3.2 MODELS

The models presented here draw relationships between the explanatory variables shown in Table 2, which contain sociodemographic, personality, and attitude characteristics of BotRide users, with the number of times users rode BotRide. To contrast Botride users with non-users, I first developed a logit model. To further understand how heavy and light users differ from non-users, I also estimated a multinomial logit model. They are presented below.

3.2.1 LOGIT MODEL

The logit model was developed to contrast Botride users (y=1) with non-users (y=0). It can be written as follows:

Logit Model:

$$y_i^* = X_i\beta + \varepsilon_i, \quad (3.1)$$
$$y_i = 1, if\ y_i^* \geq 0,$$
$$y_i = 0, if\ y_i^* < 0,$$

where:

y_i equals 1 if respondent i ever used Botride and 0 otherwise;

$y_i{*}$ is a latent variable that defines y_i;

X_i is a line vector of explanatory variables;

β is a column vector of unknown coefficients; and

ε_i is an error term assumed to be distributed according to the standard logistic.

To present my results, I relied on odds ratios. For explanatory variable k, the odds ratio is the odds (the probability that Y=1 divided by the probability that Y=0) after increasing variable k by one unit, divided by the odds for initial values of all explanatory variables. It can be shown that the odds ratio for explanatory variable k is exp(β_k) [16].

The baseline for this model were users with the following characteristics:

- Considers themselves cautious, a planner, a spender, and an optimist
- UCI student/faculty
- Lives alone
- Male, prefer to self-identity when it comes to gender, or preferred not to say gender
- Asian American ethnicity
- Has a household income of less than $35,000
- Does not have much public transit experience
- Does not really know anything about autonomous vehicles
- Views BotRide service as more about saving
- Views BotRide service as more about fun

3.2.2 MULTINOMIAL LOGIT MODEL

The multinomial logit model was developed to contrast light and heavy BotRide users with non-users.

Multinomial Logit Model:

To introduce this model, I adopted a random utility approach. In that context, for user "*i*" and alternative j = 0 (not a user), 1 (a light user), 2 (a heavy user), the probability that "*i*" belongs to category "*j*" is:

$$\hat{P}_{i,j} = \frac{\exp\,(V_{i,j|0})}{1 + \exp\,(V_{i,1|0}) + \exp\,(V_{i,2|0})} \tag{3.2}$$

where:

$$\exp\,(V_{i,j|0}) = \exp\,(\beta_{i,0} + \beta_{i,1}x_{i,1} + \beta_{i,2}x_{i,2} + \cdots + \beta_{i,35}x_{i,35}) \tag{3.3}$$

$V_{i,j|0}$ is the indirect utility for respondent "*i*" of each of the three options (non-user, light user, heavy user). The odds ratio for this model is the same as the one shown for the logit model.

The baseline for the explanatory variables in this model is the same as for the logit model.

<u>Multicollinearity</u>

To check for multicollinearity, I calculated VIF values for each explanatory variable using Stata. All of the VIF values were less than five with a mean VIF of 1.73. Since none of the VIFs were close to a value of ten, multicollinearity is not a problem here. It is also worth noting that explanatory variables are the same for both models. Calculated VIF values are provided in APPENDIX A, Table 8.

CHAPTER 4: RESULTS

4.1 LOGIT RESULTS

Table 4 below shows the coefficients and odds ratios to three decimal places for the explanatory variables, ten of which were found to be statistically significant at a p-value of 0.10. The count R^2 for the logit model was 68.1% and the adjusted count R^2 was 20.5%.

Table 4 – Logit Model Results Summary

No.	Variable	Coefficient	Odds Ratio
0	Constant	-0.780	N/A
Personality			
1	Risktaker	-0.047	0.954
2	Spontaneous	-0.104	0.901
3	Saver	0.189	1.208
4	Pessimist	-0.220	0.803
Occupation			
5	Irvine Resident	-0.955***	0.385
6	Works fulltime in Irvine	-0.789**	0.454
7	Other	-0.318	0.727
Live With			
8	Roommate(s)	-0.053	0.949
9	Parents	-0.415	0.660
10	Spouse/Significant Other	-0.138	0.871
11	Child under 18	0.007	1.007
12	Child 18+	0.318	1.375
Gender			
13	Female	-0.249	0.779
Ethnicity			
14	White/Caucasian	-0.247	0.781
15	African American/Black	-0.127	0.880
16	Latinx	-0.327	0.721
17	Other	0.217	1.242

No.	Variable	Coefficient	Odds Ratio
18	Prefer not to say	-0.271	0.763
19	Mixed	0.005	1.005
HHI			
20	$35,000 - $74,999	-0.556*	0.574
21	$75,000 - $99,999	0.193	1.213
22	$100,000 - $149,999	0.081	1.084
23	$150,000 - $199,999	-0.130	0.878
24	$200,000 or more	-0.296	0.744
25	Prefer not to say	-0.443*	0.648
Personal Vehicle			
26	Has a personal vehicle	-0.003	0.997
27	Does not have a personal vehicle, but has regular access to someone else's	0.208	1.231
Attitude Towards AVs			
28	Very comfortable with the idea of riding in a self-driving vehicle/AV	0.616*	1.851
Public Transit Experience			
29	Has a lot of experience	0.239	1.270
30	Has some experience	0.255	1.291
Familiarity with AVs Based on Experience			
31	Has extensive knowledge	1.343**	3.829
32	Has solid knowledge	1.896***	6.661
33	Has a little bit of knowledge	1.452***	4.272
Convenience vs. Free			
34	BotRide is more about convenience than saving	-0.392*	0.676
Function vs. Fun			
35	BotRide is more about function than fun	-0.409*	0.664

* significant at p = 0.10
** significant at p = 0.05
*** significant at p = 0.01

Variables "$35,000-$74,999", "Prefer not to say (HHI)", "Very comfortable with the idea of riding in a self-driving vehicle/AV", "BotRide is more about convenience than saving", and "BotRide is more about function than fun" were significant at the p = 0.10 level.

Both household income variables had negative coefficients, implying users with an income between \$35,000 and \$74,999 or a preference not to mention their income had a lower tendency to ride BotRide compared to the baseline conditions. People who make between \$35,000 and \$74,999 typically have entry level positions in their industry and may need a reliable form of transportation like a personal vehicle. Those who preferred not to say their income may have been embarrassed to report a low income and may be secretive of a very high income. In either case, these users may have been discouraged to use BotRide due to financial reasons or impracticality of using the service because they had better options.

The positive coefficient of “Very comfortable with the idea of riding in a self-driving vehicle/AV” suggests that users who are comfortable with riding AVs may be more likely to use BotRide. This made sense since people are generally more open to using services if they are perceived as low risk.

The variables related to convenience and function had negative coefficients. This implied that users who saw BotRide as more about convenience and function were less likely to use BotRide. Those who viewed BotRide in terms of convenience may not have personally believed BotRide was a feasible service due to schedule conflicts and other factors. For example, destinations were restricted to a set chosen by Hyundai. Therefore, these destination choices may not have been convenient. Those who viewed BotRide in terms of function may have determined that BotRide was not a practical use of their time or perhaps had access to a better option.

The variables “Works fulltime in Irvine” and “Has extensive knowledge on AVs” were significant at the p = 0.05 level. “Work fulltime in Irvine” had a negative coefficient, which implied that BotRide did not have high appeal for those who work fulltime in Irvine. Since fulltime employees typically need to be personally accountable for their own commute to and from their workplace, these persons may have a personal vehicle and therefore have little need for BotRide. Additionally, “Has extensive knowledge on AVs” had a positive coefficient. This made sense since having extensive knowledge on a service usually allows a user to become more familiar with the service, allowing the user to feel safer.

Variables “Irvine Resident”, “Has solid knowledge on AVs”, and “Has a little bit of knowledge on AVs” were highly significant at the p = 0.01 level. “Irvine Resident” had a negative coefficient. This may be due to the idea that Irvine residents may already have access to a personal vehicle or see little use in riding an AV to a location that is near their home. “Has solid knowledge on AVs” and “Has a little bit of knowledge on AVs” both had positive coefficients. This may be due to the idea that having more education on a topic typically allows users to gain familiarity, which in turn allows them to perceive something as safer.

It is also worth noting that most of the explanatory variables were not significant, notably the variable categories for “Personality”, “Live With”, “Gender”, “Ethnicity”, “Personal Vehicle” and “Public Transit Experience”. This implied that for this dataset, a user’s personality, living situation, gender, ethnicity, personal vehicle access, and experience with public transit as defined in this study are not as likely to influence BotRide usage. One reason these variables may have low influence could be due to the public curiosity surrounding AVs regardless of personal characteristics and situation. For example, a user who is not a risktaker and owns a personal vehicle may still want to experience BotRide for the novelty of the service. In addition, Irvine is generally considered a safe region, ranking within the top 20 safest cities in California, which may partially explain why gender and ethnicity have low influence [17].

4.1.1 LOGIT SPECIFICATION TEST

A linktest in Stata was used to find specification errors in the logit model. This test indicates whether relevant variables have been omitted or if the function is incorrectly specified [18]. To pass the linktest, additional statistically significant predictors should not be found except by chance. Linktest rebuilds the model by using the linear predicted value (_hat) and its square ($\_hat^2$) as predictors. The results of the linktest are shown in Table 5. The logit model passed the linktest since "_hat" was statistically significant and "$\_hat^2$" was not, indicating the linktest was not significant [18].

Table 5 – Logit Model Linktest Results

Logit Model	Coefficient
_hat	0.998***
$\_hat^2$	0.005
constant	-0.002

*** significant at p = 0.01

4.2 MULTINOMIAL LOGIT RESULTS

Table 6 below shows the coefficients and odds ratios to three decimal places for the explanatory variables. Variables were determined to be statistically significant at a p-value of 0.10. The count R^2 for the multinomial logit model was 50.7% and the adjusted count R^2 was 17.6%.

Table 6 – Multinomial Logit Model Results Summary

No.	Variable	No Rides vs. Some Rides		No Rides vs. A Lot of Rides	
		Coefficient	Odds Ratio	Coefficient	Odds Ratio
0	Constant	-1.068	N/A	-2.094*	N/A
Personality					
1	Risktaker	0.081	1.084	-0.213	0.808
2	Spontaneous	-0.142	0.867	-0.056	0.946
3	Saver	0.148	1.160	0.249	1.283
4	Pessimist	-0.287	0.751	-0.151	0.860
Occupation					
5	Irvine Resident	-0.831*	0.435	-1.095***	0.335
6	Works fulltime in Irvine	-0.490	0.612	-1.269***	0.281
7	Other	-0.248	0.781	-0.425	0.654
Live With					
8	Roommate(s)	-0.115	0.891	0.034	1.035
9	Parents	-0.340	0.712	-0.483	0.617
10	Spouse/Significant Other	-0.140	0.869	-0.140	0.869
11	Child under 18	0.006	1.006	0.046	1.047
12	Child 18+	0.370	1.447	0.269	1.309
Gender					
13	Female	-0.196	0.822	-0.336	0.715
Ethnicity					
14	White/Caucasian	-0.180	0.835	-0.365	0.694
15	African American/Black	0.124	1.132	-0.581	0.559
16	Latinx	-0.441	0.644	-0.189	0.828
17	Other	-0.081	0.922	0.494	1.639
18	Prefer not to say	-0.551	0.576	-0.028	0.973
19	Mixed	-0.185	0.831	0.214	1.238
HHI					
20	$35,000 - $74,999	-0.388	0.679	-0.752*	0.472
21	$75,000 - $99,999	0.607	1.359	0.084	1.088
22	$100,000 - $149,999	0.377	1.458	-0.378	0.685
23	$150,000 - $199,999	0.056	1.058	-0.313	0.732
24	$200,000 or more	-0.350	0.705	-0.194	0.824
25	Prefer not to say	-0.313	0.731	-0.515*	0.598
Personal Vehicle					
26	Has a personal vehicle	0.049	1.050	-0.022	0.978

No.	Variable	No Rides vs. Some Rides		No Rides vs. A Lot of Rides	
		Coefficient	Odds Ratio	Coefficient	Odds Ratio
27	Does not have a personal vehicle, but has regular access to someone else's	0.408	1.504	0.007	1.007
Attitude Towards AVs					
28	Very comfortable with the idea of riding in a self-driving vehicle/AV	0.509	1.663	0.727	2.069
Public Transit Experience					
29	Has a lot of experience	0.037	1.037	0.543	1.722
30	Has some experience	0.117	1.125	0.479	1.614
Familiarity with AVs Based on Experience					
31	Has extensive knowledge	0.895*	2.447	1.778**	5.917
32	Has solid knowledge	1.555***	4.737	2.245***	9.440
33	Has a little bit of knowledge	1.479***	4.388	1.330*	3.782
Convenience vs. Free					
34	BotRide is more about convenience than saving	-0.418*	0.658	-0.361	0.697
Function vs. Fun					
35	BotRide is more about function than fun	-0.722***	0.486	-0.134	0.986

* significant at p = 0.1
** significant at p = 0.05
*** significant at p = 0.01

Like the logit model, in the "No Rides vs Some Rides" results, the "Personality" variable category did not show significance, indicating that a user's personality had little influence on sometimes using BotRide. In contrast to the logit model, only one occupation "Irvine Resident" was significant and had a negative coefficient. This indicated that users who reside in Irvine are less likely to sometimes use BotRide. This may be due to Irvine residents already having access to a personal vehicle or the impractically of using BotRide in a city in which they already reside. However, the "Live With", "Gender", "Ethnicity", "HHI", "Personal Vehicle", "Attitude Towards AVs", and "Public Transit Experience" categories had no significant variables. This implied that a user's living situation, gender, ethnicity, income, access to a personal vehicle, comfort with AVs, and experience with transit are not as likely to influence a user to occasionally use BotRide. This is a little different to the logit model which had significant variables in "HHI" and "Attitude Towards AVs".

The variables for having extensive, solid, or a little knowledge in AVs had positive coefficient values and were also statistically significant. This suggested that having some amount of knowledge regarding AVs may be related to some BotRide usage. Since having sufficient knowledge on a service can lead to less anxiety with using the service, this result made sense. In addition, the significant variables for using BotRide due to convenience and function had negative values, implying that users wanting to use Botride for convenient or functional means had a lower likelihood of using the service. This may be due to BotRide neither fitting into their schedule nor being a practical mode of travel.

In the "No Rides vs. A Lot of Rides" results, the personality variables did not show significance similar to "No Rides vs. Some Rides", but the variables for users that live in Irvine and users that work fulltime in Irvine were highly significant with negative coefficients. This indicated that users who reside or work fulltime in Irvine are much less likely to use BotRide a lot. This may be due to Irvine residents and employees already having access to a personal vehicle for commute reliability and convenience. The next categories for "Live With", "Gender", and "Ethnicity" had no significant variables, which indicated that a user's living situation, gender, and ethnicity are less likely to influence a user to use BotRide a lot.

The variable for users who have a household income between \$35,000 and \$74,999 was significant with a negative coefficient as well, indicating that these users were less inclined to use BotRide frequently. Those with this income are typically entry level employees who may have less flexibility when it comes to travel. Hence, using BotRide may not have been a reliable form of travel due to factors such as wait time and travel time.

The variable for those who did not want to report their income was also significant with a negative coefficient. Users from this subcategory may either have a very low or high household income. As such, these users may have avoided using BotRide a lot due to it being unnecessary for their lifestyle due to financial burden or the availability of more convenient options. There was no significance in the "Personal Vehicle", "Attitude Towards AVs", and "Public Transit Experience" categories, which implied that a user's access to a personal vehicle, comfort with AVs, and experience with transit are not as likely to influence a user to frequently use BotRide. This is different to the logit model, which had significance in "Attitude Towards AVs".

Like "No Rides vs. Some Rides", the variables for having extensive, solid, or a little knowledge in AVs had positive coefficient values and were also statistically significant. Similarly, this implied that having some amount of knowledge regarding AVs may be related to higher BotRide usage. Furthermore, the "Convenience vs. Free" and "Function vs. Fun" categories showed no significance unlike the logit model that had significance in both. This suggested convenience or savings associated with BotRide, and the functionality or fun associated with BotRide do not have much influence on a user's choice to frequently use BotRide.

Overall, “No Rides vs. Some Rides” and the logit model shared six significant variables (“Irvine Resident”, “Has extensive knowledge”, “Has solid knowledge”, “Has a little bit of knowledge”, “BotRide is more about convenience than saving”, and “BotRide is more about function than fun”). Moreover, “No Rides vs. A Lot of Rides” and the logit model shared seven different significant variables (“Irvine Resident”, “Works fulltime in Irvine”, “$35,000 - $74,999”, “Prefer not to say (HHI)”, “Has extensive knowledge”, “Has solid knowledge”, and “Has a little bit of knowledge”). Although this comparison showed that the multinomial logit model produces results similar to the logit model, it also showed the multinomial logit model can further characterize the BotRide users.

4.2.1 MULTINOMIAL LOGIT SPECIFICATION TESTS

For a multinomial logit model, it is important to test for independence of irrelevant alternatives (IIA). Two common tests are the Hausman test and the Suest-based Hausman test. Their null hypothesis states that “odds are independent of other alternatives” [18]. These tests did not produce statistically significant results, which indicates that there was not enough evidence to reject the null hypothesis. In the Hausman test results, two of the multinomial logit outcomes “User took 1-2 rides” and “User took 0 rides” had a negative test statistic, which is also an indication that the IIA assumption is not violated [19]. The results of these test are shown in Table 7.

Table 7 - IIA Test Results

Outcome	Chi^2 (test statistic)	Degrees of Freedom	P > Chi^2
Hausman Test			
User took 1-2 rides	-19.063	36	.
User took 3+ rides	1.832	36	1.000
User took 0 rides	-0.382	36	.
Suest-Based Hausman Test			
User took 1-2 rides	15.477	36	0.999
User took 3+ rides	19.155	36	0.990
User took 0 rides	18.569	36	0.993

CHAPTER 5: CONCLUSIONS

This thesis aimed to analyze BotRide user characteristics collected via a survey and analyzed using discrete choice models. I relied on Stata to create my model variables and to estimate a logit model to contrast BotRide users with non-users, and a multinomial logit model to contrast light users and heavy users with non-users.

For my logit model, the statistically significant explanatory variables were variables “Irvine Resident”, “Works fulltime in Irvine”, “$35,000 - $74,999”, “Prefer not to say (HHI)”, “Very comfortable with the idea of riding in a self-driving vehicle/AV”, “Has extensive knowledge”, “Has solid knowledge”, “Has a little bit of knowledge”, “BotRide is more about convenience than saving”, and “BotRide is more about function than fun” as shown in Table 4. These results implied that users who have an annual household income between $35,000 and $74,999, prefer not to mention their income, view BotRide as more about convenience and function, work fulltime in Irvine, or are Irvine residents are less likely to use BotRide. Conversely, being comfortable with riding in an AV or having some level of familiarity with and knowledge about AVs results in an increased likelihood of using BotRide.

The multinomial logit model contrasted users who took 1 to 2 rides compared to those who took no rides, and users who took 3 or more rides again with users who took no rides. For the former, the statistically significant explanatory variables were variables “Irvine Resident”, “Has extensive knowledge”, “Has solid knowledge”, “Has a little bit of knowledge”, “BotRide is more about convenience than saving”, and “BotRide is more about function than fun" as shown in Table 6. Their signs imply that users who are Irvine residents or view BotRide as more about convenience and function may have a lower tendency to use BotRide. As for the logit model, having some level of familiarity with and knowledge about AVs implies an increased likelihood of sometimes using BotRide.

In the second category, the statistically significant explanatory variables were variables "Irvine Resident", "Works fulltime in Irvine", "$35,000 - $74,999", "Prefer not to say (HHI)", "Has extensive knowledge", "Has solid knowledge", and "Has a little bit of knowledge" as shown in Table 6. These results imply that users who are Irvine residents, work fulltime in Irvine, have an annual household income between $35,000 and $74,999, or preferred not to mention their income have a lower tendency to use BotRide a lot. As for the logit model and the first category, having some level of familiarity and knowledge about AVs implies an increased likelihood of using BotRide often.

Based on these results, to increase BotRide usage, Hyundai should consider targeting users who have a household income above $75,000 and do not live or work fulltime in Irvine. In addition, providing the public with more information about BotRide and how it operates could also encourage ridership since doing so will make users more familiar and comfortable with the technology. For example, a public information campaign to advertise BotRide through commercials, ads, or brochures may allow BotRide to become a more attractive option for travel. Ways to increase the convenience and function of BotRide should be explored as well. For instance, Hyundai could consider adding more BotRide trip destinations to appeal to a wider range of users.

Limitations of this study include the fact that I was not involved in the development of the survey that was distributed to the BotRide users in February 2020. Had I been part of the process, I could have asked to collect data about age, employment status, and household location. Even though BotRide was conducted in Irvine, some users likely live outside of the city or further away from BotRide destinations. If I had had access to household location data, I might have been able to draw relationships between household distance and BotRide usage to see for what types of land uses is BotRide most attractive.

For potential future studies, since there were questions in the survey related to pricing to BotRide, it would be useful to explore the influence of pricing on BotRide usage. Trip type data can also be found in the February 2020 survey data set, which may be used to analyze the types of trips BotRide was used for in relation to socioeconomic traits or even usage. Another suggestion for future studies is to look into the use of other types of models, like a zero-inflated negative binomial regression model to analyze BotRide usage.

REFERENCES

[1] G. Circella, F. Alemi, K. Tiedeman, S. Handy, and P. Mokhtarian, "The Adoption of Shared Mobility in California and Its Relationship with Other Components of Travel Behavior," p. 79p, 2018, [Online]. Available: https://merritt.cdlib.org/d/ark%253A%252F13030%252Fm5zm0cxx/1/producer%252FNCST-RR-201802.pdf%0Ahttps://ncst.ucdavis.edu/wp-content/uploads/2016/10/NCST-TO-033.1-Circella_Shared-Mobility_Final-Report_FEB-2018.pdf%0Ahttps://rosap.ntl.bts.gov/view/dot/35032.

[2] F. Alemi, G. Circella, P. Mokhtarian, and S. Handy, "What drives the use of ridehailing in California? Ordered probit models of the usage frequency of Uber and Lyft," *Transp. Res. Part C Emerg. Technol.*, 2019, doi: 10.1016/j.trc.2018.12.016.

[3] R. P. Payyanadan and J. D. Lee, "Understanding the ridesharing needs of older adults," *Travel Behav. Soc.*, 2018, doi: 10.1016/j.tbs.2018.08.002.

[4] F. F. Dias, P. S. Lavieri, T. Kim, C. R. Bhat, and R. M. Pendyala, "Fusing Multiple Sources of Data to Understand Ride-Hailing Use," *Transp. Res. Rec.*, 2019, doi: 10.1177/0361198119841031.

[5] S. Sikder, "Who Uses Ride-Hailing Services in the United States?," *Transp. Res. Rec.*, 2019, doi: 10.1177/0361198119859302.

[6] B. J. Tang, X. Y. Li, B. Yu, and Y. M. Wei, "How app-based ride-hailing services influence travel behavior: An empirical study from China," *Int. J. Sustain. Transp.*, 2020, doi: 10.1080/15568318.2019.1584932.

[7] N. Sadowsky and E. Nelson, "The Impact of Ride-Hailing Services on Public Transportation Use: A Discontinuity Regression Analysis," *Econ. Dep. Work. Pap. Ser.*, 2017.

[8] N. Merat, R. Madigan, and S. Nordhoff, "Human Factors, User Requirements, and User Acceptance of Ride-Sharing in Automated Vehicles," *Int. Transp. Forum*, 2017.

[9] J. Morales Sarriera, G. Escovar Álvarez, K. Blynn, A. Alesbury, T. Scully, and J. Zhao, "To share or not to share: Investigating the social aspects of dynamic ridesharing," *Transp. Res. Rec.*, 2017, doi: 10.3141/2605-11.

[10] Y. Zhang and Y. Zhang, "Examining the relationship between household vehicle ownership and ridesharing behaviors in the United States," *Sustain.*, 2018, doi: 10.3390/su10082720.

[11] S. Lee, S. Yoo, S. Kim, E. Kim, and N. Kang, "The Effect of Robo-Taxi User Experience On User Acceptance: Field Test Data Analysis," *arXiv*, 2020.

[12] J. Meurer, C. Pakusch, G. Stevens, D. Randall, and V. Wulf, "A wizard of oz study on passengers' experiences of a robo-taxi service in real-life settings," 2020, doi:

10.1145/3357236.3395465.

[13] A. Sanguinetti, B. Ferguson, J. Oka, E. Alston-Stepnitz, and K. Kurani, “Designing robo-taxis to promote ride-pooling,” 2021, doi: 10.1007/978-3-030-55307-4_2.

[14] R. Vosooghi, J. Kamel, J. Puchinger, V. Leblond, and M. Jankovic, “Robo-Taxi service fleet sizing: assessing the impact of user trust and willingness-to-use,” *Transportation (Amst).*, 2019, doi: 10.1007/s11116-019-10013-x.

[15] S. Yoo, S. Lee, S. Kim, E. Kim, H. Hwangbo, and N. Kang, “A Study on Anxiety about Using Robo-taxis: HMI Design for Anxiety Factor Analysis and Anxiety Relief Based on Field Tests,” *arXiv*, 2020.

[16] S. P. Washington, M. G. Karlaftis, and F. L. Mannering, *Statistical and Econometric Methods for Transportation Data Analysis*, 2nd ed. Taylor and Francis Group, LLC, 2011.

[17] K. McEntire, “California’s 50 Safest Cities of 2021,” *Safewise*, 2021. https://www.safewise.com/blog/safest-cities-california/.

[18] X. Chen, P. B. Ender, M. Mitchell, and C. Wells, “Stata Web Books Regression with Stata,” *UCLA Institute for Digital Research & Education Statistical Consulting*, 2006. https://stats.idre.ucla.edu/stata/webbooks/logistic/.

[19] D. Hausman, Jerry; McFadden, “Specification Tests for the Multinomial Logit Model Author (s): Jerry Hausman and Daniel McFadden Published by : The Econometric Society Stable URL : http://www.jstor.org/stable/1910997 Accessed : 22-06-2016 14 : 58 UTC Your use of the JSTOR archive ind,” *Econometrica*, vol. 52, no. 5, pp. 1219–1240, 1984.

APPENDICES

APPENDIX A

Table 8 - Model Variables and Corresponding VIF Values

Variable	VIF
Risktaker	1.15
Spontaneous	1.16
Saver	1.08
Pessimist	1.08
Irvine Resident	1.81
Works fulltime in Irvine	1.71
Other (Occupation)	1.34
Roommate(s)	2.62
Parents	1.52
Spouse/Significant Other	2.16
Children under 18	1.35
Children 18+	1.08
Female	1.25
White/Caucasian	1.31
African American/Black	1.08
Latinx	1.18
Other	1.13
Prefer not to say (Ethnicity)	1.30
Mixed	1.09
$35,000 - $74,999	1.52
$75,000 - $99,999	1.54
$100,000 - $149,999	1.68
$150,000 - $199,999	1.59
$200,000 or more	1.70
Prefer not the say (HHI)	1.88
Has a personal vehicle	2.27
Does not have a personal vehicle, but has regular access to someone else's	1.64
Very comfortable with the idea of riding in a self-driving vehicle/AV	1.07
Has a lot of experience	3.35
Has some experience	3.36
Has extensive knowledge	2.53
Has solid knowledge	4.44
Has a little bit of knowledge	4.50
BotRide is more about convenience than saving	1.19
BotRide is more about function than fun	1.11
Mean VIF	1.76