

## Final Project Report

# Does Ridehailing Use Affect Vehicle Ownership or Vice Versa? An Exploratory Investigation of the Relationship Using a Latent Market Segmentation Approach

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



By

**Irfan Batur**

Email: [ibaturl@asu.edu](mailto:ibaturl@asu.edu)

**Abbie C. Dirks**

Email: [acdirks@asu.edu](mailto:acdirks@asu.edu)

**Ram M. Pendyala**

Email: [ram.pendyala.@asu.edu](mailto:ram.pendyala.@asu.edu)

School of Sustainable Engineering and the Built Environment  
Arizona State University  
660 S. College Avenue, Tempe, AZ 85287-3005

September 2023

## TECHNICAL REPORT DOCUMENTATION PAGE

<b>1. Report No.</b> N/A	<b>2. Government Accession No.</b> N/A	<b>3. Recipient's Catalog No.</b> N/A	
<b>4. Title and Subtitle</b> Does Ridehailing Use Affect Vehicle Ownership or Vice Versa? An Exploratory Investigation of the Relationship Using a Latent Market Segmentation Approach		<b>5. Report Date</b> September 2023	
		<b>6. Performing Organization Code</b> N/A	
<b>7. Author(s)</b> Irfan Batur, <a href="https://orcid.org/0000-0002-8058-2578">https://orcid.org/0000-0002-8058-2578</a> Abbie C. Dirks, <a href="https://orcid.org/0000-0003-4590-468X">https://orcid.org/0000-0003-4590-468X</a> Ram M. Pendyala, <a href="https://orcid.org/0000-0002-1552-9447">https://orcid.org/0000-0002-1552-9447</a>		<b>8. Performing Organization Report No.</b> N/A	
		<b>9. Performing Organization Name and Address</b> School of Sustainable Engineering and the Built Environment Arizona State University 660 S. College Avenue, Tempe, AZ 85287-3005	
<b>11. Contract or Grant No.</b>			
<b>12. Sponsoring Agency Name and Address</b> U.S. Department of Transportation, University Transportation Centers Program, 1200 New Jersey Ave, SE, Washington, DC 20590		<b>13. Type of Report and Period Covered</b> Research Report (2022-2023)	
		<b>14. Sponsoring Agency Code</b>	
<b>15. Supplementary Notes</b> N/A			
<b>16. Abstract</b> This report presents an examination of the inter-relationship between household vehicle ownership and ridehailing use frequency. Both variables constitute important mobility choices with significant implications for the future of transport. While it is generally known that these two behavioral phenomena are inversely related to one another, the direction of causality is rather ambiguous. Do vehicle ownership levels affect ridehailing use frequency, or does the adoption and use of ridehailing services affect vehicle ownership? If ridehailing services affect vehicle ownership, then it is plausible that a future of mobility-as-a-service would be characterized by lower levels of vehicle ownership. To explore the degree to which these causal relationships are prevalent in the population, a joint latent segmentation model system is formulated and estimated on a survey data set collected in four automobile-oriented metropolitan areas of the United States. The latent segmentation model system recognizes that the causal structures driving mobility choices of individuals are not directly observed. Model estimation results show that 58 percent of the survey sample follow the causal structure in which ridehailing use frequency affects vehicle ownership. This finding suggests that there is considerable structural heterogeneity in the population with respect to causal structures, and that ridehailing use does indeed hold considerable promise to effect changes in private vehicle ownership in the future.			
<b>17. Key Words</b> ridehailing use, vehicle ownership, vehicle availability, causal relationship, latent segments, attitudinal factors, joint models, traveler behavior		<b>18. Distribution Statement</b> No restrictions.	
<b>19. Security Classif.(of this report)</b> Unclassified	<b>20. Security Classif.(of this page)</b> Unclassified	<b>21. No. of Pages</b> 23	<b>22. Price</b> N/A

Form DOT F 1700.7 (8-72)

Reproduction of completed page authorized

## **DISCLAIMER**

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

## **ACKNOWLEDGMENTS**

*This project was funded by a grant from A USDOT Tier 1 University Transportation Center, supported by USDOT through the University Transportation Centers program. The authors would like to thank the TOMNET and USDOT for their support of university-based research in transportation, and especially for the funding provided in support of this project. The authors would like to thank Aupal Mondal and Chandra R. Bhat for their contributions to the work presented in this report.*

## TABLE OF CONTENTS

LIST OF TABLES .....	5
LIST OF FIGURES .....	5
EXECUTIVE SUMMARY .....	6
1. INTRODUCTION .....	7
2. DATA DESCRIPTION .....	8
2.1 Endogenous Variables and Attitudinal Indicators .....	9
3. MODELING FRAMEWORK .....	13
3.1. Model Structure .....	13
3.2. The Joint Model of Behavioral Choices .....	14
3.3. Latent Segmentation Model.....	15
4. MODEL ESTIMATION RESULTS.....	16
4.1. Bivariate Model of Behavioral Outcomes .....	16
4.2. Characteristics and Size of the Latent Segments .....	19
5. DISCUSSION AND CONCLUSIONS .....	21
REFERENCES .....	22

**LIST OF TABLES**

Table 1 Socio-Economic and Demographic Characteristics of the Sample ..... 10  
Table 2 Estimation Results for Endogenous Variables Within Each Segment ..... 17  
Table 3 Profiles of the Two Latent Market Segments ..... 20

**LIST OF FIGURES**

Figure 1 Household Vehicle Availability by Ridehailing Use Frequency (N=3,146)..... 11  
Figure 2 Agreement with Attitudinal Indicators Defining Latent Constructs (N=3,146)..... 12  
Figure 3 Latent Segmentation Model Framework ..... 14

## **EXECUTIVE SUMMARY**

This report presents an examination of the inter-relationship between household vehicle ownership and ridehailing use frequency. Both variables constitute important mobility choices with significant implications for the future of transport. While it is generally known that these two behavioral phenomena are inversely related to one another, the direction of causality is rather ambiguous. Do vehicle ownership levels affect ridehailing use frequency, or does the adoption and use of ridehailing services affect vehicle ownership? If ridehailing services affect vehicle ownership, then it is plausible that a future of mobility-as-a-service would be characterized by lower levels of vehicle ownership. To explore the degree to which these causal relationships are prevalent in the population, a joint latent segmentation model system is formulated and estimated on a survey data set collected in four automobile-oriented metropolitan areas of the United States. The latent segmentation model system recognizes that the causal structures driving mobility choices of individuals are not directly observed. Model estimation results show that 58 percent of the survey sample follow the causal structure in which ridehailing use frequency affects vehicle ownership. This finding suggests that there is considerable structural heterogeneity in the population with respect to causal structures, and that ridehailing use does indeed hold considerable promise to effect changes in private vehicle ownership in the future.

## 1. INTRODUCTION

Arguably, the most notable and impactful mobility innovation of the past decade is *ridehailing services*, which allow individuals to summon and pay for a ride in real-time using the convenience of a mobile app. Vehicles are generally owned, operated, and maintained by individual drivers, who are similar to freelance workers setting their own working hours and operating as independent contractors. Individuals, acting as drivers, can then provide rides to any individual who signs up to use the ridehailing service platform. Examples of ridehailing services include Lyft in the US, Uber in many countries, Didi in China, and Ola in India. These services are sometimes called mobility-on-demand (MOD) services or Mobility-as-a-Service (MaaS). Despite subtle distinctions between these terms, they will be used interchangeably within the context of this report.

Ridehailing services have become very popular in many cities around the world, with Uber arguably the world's largest ride-hailing service provider. Although official statistics are hard to come by, informal sources (Dean, 2021) report that Uber has nearly 100 million active users and completes 1.44 billion rides every quarter. Major ridesharing platforms in other countries report similarly impressive numbers. Even though the mode share of ridehailing services remains modest, especially in the US, it is fair to say that ridehailing is a well-established and entrenched mode of transportation in most metropolitan areas. Ridehailing services have grown to such an extent that decreases in transit ridership and increases in urban congestion are being attributed, at least in part, to the rise of ridehailing services (Diao et al., 2021; Erhardt et al., 2022).

Because of the widespread adoption of ridehailing services as a mode of transportation, transportation demand forecasting models need to be enhanced to reflect ridehailing service usage patterns and their impacts on other modes of transportation. In addition, metropolitan areas and planning agencies have been grappling with implementing policies and strategies to ameliorate any adverse impacts of ridehailing services in their jurisdictions. Due to these myriad and complex planning considerations and modeling needs, a vast body of literature exploring and documenting the adoption and impacts of ridehailing services has emerged (Tirachini, 2019). There have been many studies examining various facets of ridehailing adoption, including an exploration of market segments more and less likely to use such services, the travel characteristics of trips undertaken by ridehailing services, and the extent to which ridehailing services may be contributing to vehicle miles of travel (VMT) due to deadhead miles and zero-occupant travel (Lavieri and Bhat, 2019b; Alemi et al., 2019; Henao and Marshall, 2019; Wu and MacKenzie, 2021).

Among the many aspects of interest is the inter-relationship between ridehailing service usage and vehicle ownership. Multiple strands of research have explored this intricate relationship. With a focus on modeling ridehailing usage as a function of socio-economic and demographic attributes, built environment attributes, and vehicle availability, many studies have documented an inverse relationship between ridehailing usage and vehicle ownership, with individuals in households of higher vehicle ownership exhibiting a lower level of ridehailing usage (Dias et al., 2017; Alemi et al., 2019; Sikder, 2019). This is behaviorally intuitive; increased access to a personal vehicle would decrease the need to use ridehailing services. Individuals in such households likely use such services only under special circumstances (e.g., trips to/from the airport, when a personal vehicle breaks down).

Another strand of research has focused on the vehicle ownership implications of ridehailing services. With the widespread availability of ridehailing services, it is potentially feasible for households to downsize the number of vehicles they own. In other words, with time, households may shed vehicles and not replace them due to their use of ridehailing services. A number of studies have focused on the potential for ridehailing services to contribute to lower levels of private vehicle

ownership in the future (Wang et al., 2021; Wu and Mackenzie, 2021). Indeed, some studies reported that individuals who embraced ridehailing services as a mobility option have reduced the number of vehicles they own or are contemplating such a reduction in the future (Sabouri et al., 2020; Tang et al., 2020).

Past research suggests a probable two-way interaction between ridehailing usage and vehicle ownership. On the one hand, vehicle ownership levels may dictate the extent of ridehailing service usage, and on the other hand, the extent of ridehailing service adoption may impact vehicle ownership. These two causal directions are likely to co-exist in the population, and it would be useful to determine the extent to which each causal structure is prevalent in a population. Virtually all travel demand forecasting models in practice assume that ridehailing usage is influenced by vehicle ownership, without considering the possibility that the other direction may also hold true. If both causal directions are prevalent to a significant degree, then transportation demand forecasting models should reflect this duality accurately.

This project adopts a novel latent segmentation modeling approach to decipher the extent to which the two different causal directions are prevalent in the population. The project utilizes survey data collected in 2019 in four automobile-oriented metropolitan areas of the United States. The survey has detailed information about individual ridehailing usage and vehicle ownership patterns besides a host of socio-economic and demographic attributes. A latent segmentation approach is adopted because the causal relationship between ridehailing use and vehicle ownership is not observed for each observation in the data set. The causal relationship is unobserved and hence treated as *latent*. Each observation may belong to one or the other of the causal structures, but which one is not observed. So, based on a mixing approach, we estimate a probability for each observation belonging to each segment, thus providing the ability to calculate the size of each causal market segment. In addition, the profiles of each latent market segment may be derived, thus providing valuable insights on their characteristics. Armed with such knowledge, it will be possible to enhance transportation demand forecasting models so that they reflect the appropriate causal structure for different subgroups in the population. It is recognized that vehicle transactions (turnover) occur slowly over long periods of time. However, as ridehailing services have been in vogue for more than a dozen years now, enough time has elapsed for ridehailing services to potentially affect household vehicle ownership levels. The latent segmentation approach will be able to uncover the two causal structures in the population using the 2019 survey data set employed in this project.

The remainder of this report is organized as follows. The next section presents a description of the data set used in this project. The third section presents the modeling framework and methodology while the fourth section presents model estimation results. The fifth section offers discussion and concluding remarks.

## **2. DATA DESCRIPTION**

The data used in this project is derived from a survey conducted in 2019 in four automobile-oriented metropolitan areas in the United States, namely, Phoenix (Arizona), Austin (Texas), Atlanta (Georgia), and Tampa (Florida). The primary objective of the survey was to gather detailed information about attitudes and perceptions towards emerging mobility services and transportation technologies, lifestyle preferences and mobility choices, and socio-economic and demographic attributes. The survey was comprehensive in nature and provided an in-depth perspective on how individuals felt about ridehailing services and the extent to which they currently use ridehailing services. The survey was administered using various survey administration methods, with



recruitment of survey respondents done through e-mail and postal mail communications, Facebook advertisements, and news and media releases. To maximize response rates, rigorous reminder protocols were implemented, and respondents were given incentives. These efforts resulted in a total respondent sample of 3,465 individuals, with each respondent belonging to a unique household. The same survey instrument was deployed in all regions, thus enabling consistent pooling of data sets across areas. More details about the survey instrument, sampling strategies, response rates, and respondent profiles may be found elsewhere (Khoeni, et. al., 2021).

In accordance with the project objectives, respondents retained in the analysis sample were limited to those familiar with ridehailing services (regardless of whether they use the services or not). It is unlikely that there is any relationship between ridehailing use and vehicle ownership for those unfamiliar with ridehailing services. After removing the individuals unfamiliar with such services, and cleaning missing and obviously erroneous records, the final sample included 3,146 individuals.

Table 1 presents an overview of sample characteristics. In general, the sample exhibits the desired level of variability for conducting a modeling exercise of the nature undertaken in this project. It is found that females comprise 57 percent of the sample. More than one-quarter of the sample falls into the lowest age bracket of 18-30 years, with individuals well distributed across all other age groups. Nearly 94 percent of respondents have a driver's license. Over 64 percent of the sample are either full-time or part-time workers and roughly one-quarter are neither workers nor students. As is commonly the case with surveys of this nature, the sample is skewed towards individuals with a higher level of education. Only 8.5 percent have a high school diploma or less, while one-quarter possess a graduate degree. With respect to race, 71 percent of the sample is White, nine percent are Asian or Pacific Islanders, and 7.8 percent are Black. A little more than one-third of the sample reside in households with an annual income between \$50,000 and \$99,999. The rest of the sample is well distributed across other income groups, offering a good representation of all income levels. About 40 percent of individuals in the sample reside in larger households with three or more members, while 21 percent report being in single-person households. There is a strong relationship between housing unit type and home ownership. It is found that 70 percent of respondents reside in stand-alone homes and two-thirds own their homes. With respect to vehicle ownership, only four percent of the respondents live in households with no vehicles. About 40 percent reside in households with two vehicles, and another 32.6 percent reside in households with three or more vehicles. Given the automobile-oriented nature of the survey locations, the high auto ownership level is consistent with expectations. The sample is rather evenly split between Phoenix, Austin, and Atlanta, with a smaller percentage residing in Tampa.

## **2.1 Endogenous Variables and Attitudinal Indicators**

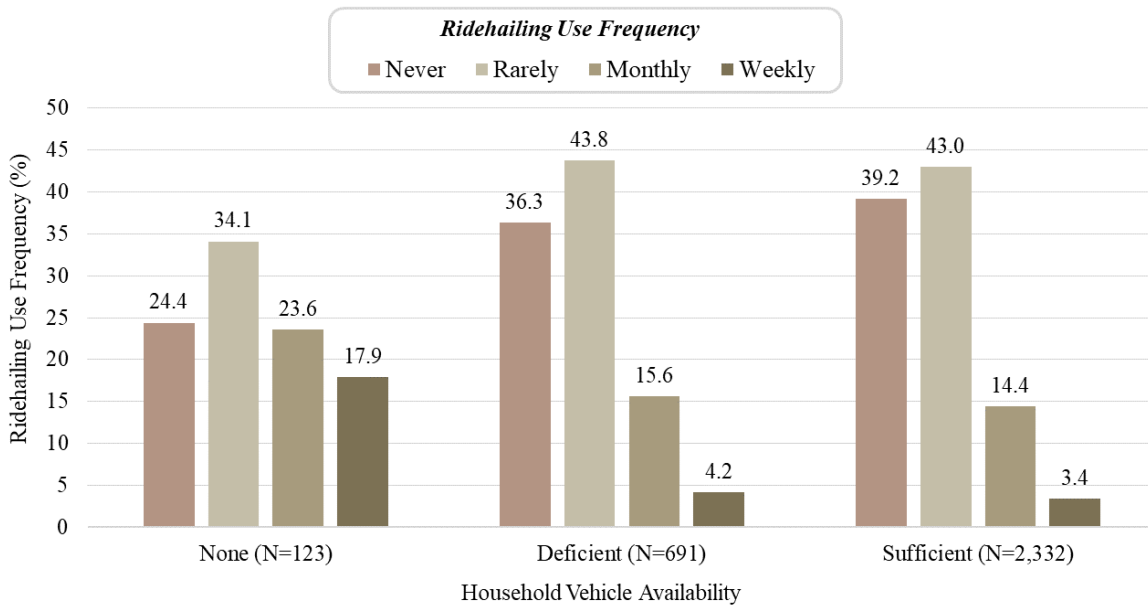
The two endogenous variables of interest in this project constitute ridehailing use frequency and vehicle availability (which represents vehicle ownership, but in the form of a per-adult vehicle ownership level). Thus, access to household vehicles is computed as the number of vehicles per adult (18+) for each record. This is used to define three categories of vehicle availability: none (zero cars), deficient (fewer vehicles than adults), and sufficient (at least as many vehicles as adults). As shown at the bottom of Table 1, just about 4 percent report no vehicles available, 22 percent reside in vehicle-deficient households, and 74 percent reside in vehicle-sufficient households. With respect to ridehailing use, 38 percent indicate that they never use such services. About 43 percent use ridehailing services rarely (less than once a month), while 15 percent report using ridehailing services about once a month. About 4 percent use the services frequently (at least

once a week). These statistics are all consistent with expectations and aligned with low transit ridership levels in these markets (Tirachini, 2019).

**Table 1 Socio-Economic and Demographic Characteristics of the Sample**

<b>Individual Demographics (N=3,146)</b>		<b>Household Characteristics (N=3,146)</b>	
<b>Variable</b>	<b>%</b>	<b>Variable</b>	<b>%</b>
<b>Gender</b>		<b>Household annual income</b>	
Female	57.1	Less than \$25,000	10.2
Male	42.9	\$25,000 to \$49,999	14.9
<b>Age category</b>		\$50,000 to \$99,999	34.0
18-30 years	26.1	\$100,000 to \$149,999	21.8
31-40 years	11.9	\$150,000 to \$249,999	12.8
41-50 years	15.5	\$250,000 or more	6.3
51-60 years	16.4	<b>Household size</b>	
61-70 years	16.1	One	21.2
71+ years	14.0	Two	38.5
<b>Driver's license possession</b>		Three or more	40.3
Yes	93.6	<b>Housing unit type</b>	
No	6.4	Stand-alone home	69.9
<b>Employment status</b>		Condo/apartment	21.1
A student (part-time or full-time)	10.0	Other	9.0
A worker (part-time or full-time)	53.6	<b>Home ownership</b>	
Both a worker and a student	10.7	Own	67.9
Neither a worker nor a student	25.7	Rent	26.3
<b>Education attainment</b>		Other	5.8
Completed high school or less	8.5	<b>Vehicle ownership</b>	
Some college or technical school	28.5	Zero	3.9
Bachelor's degree(s)	37.7	One	23.7
Graduate degree(s)	25.3	Two	39.8
<b>Race</b>		Three or more	32.6
Asian or Pacific Islander	9.0	<b>Location</b>	
Black or African American	7.8	Atlanta, GA	30.1
Native American	0.4	Austin, TX	32.3
White or Caucasian	71.0	Phoenix, AZ	30.4
Other	11.8	Tampa, FL	7.2
<b>Main Outcome Variables</b>			
<b>Ridehailing use frequency</b>		<b>Household vehicle availability</b>	
Never	38.0	None	3.9
Rarely (< once per month)	42.8	Deficient (less than one per adult)	22.0
Monthly (about once per month)	15.1	Sufficient (at least one per adult)	74.1
Weekly (at least once per week)	4.1	---	---

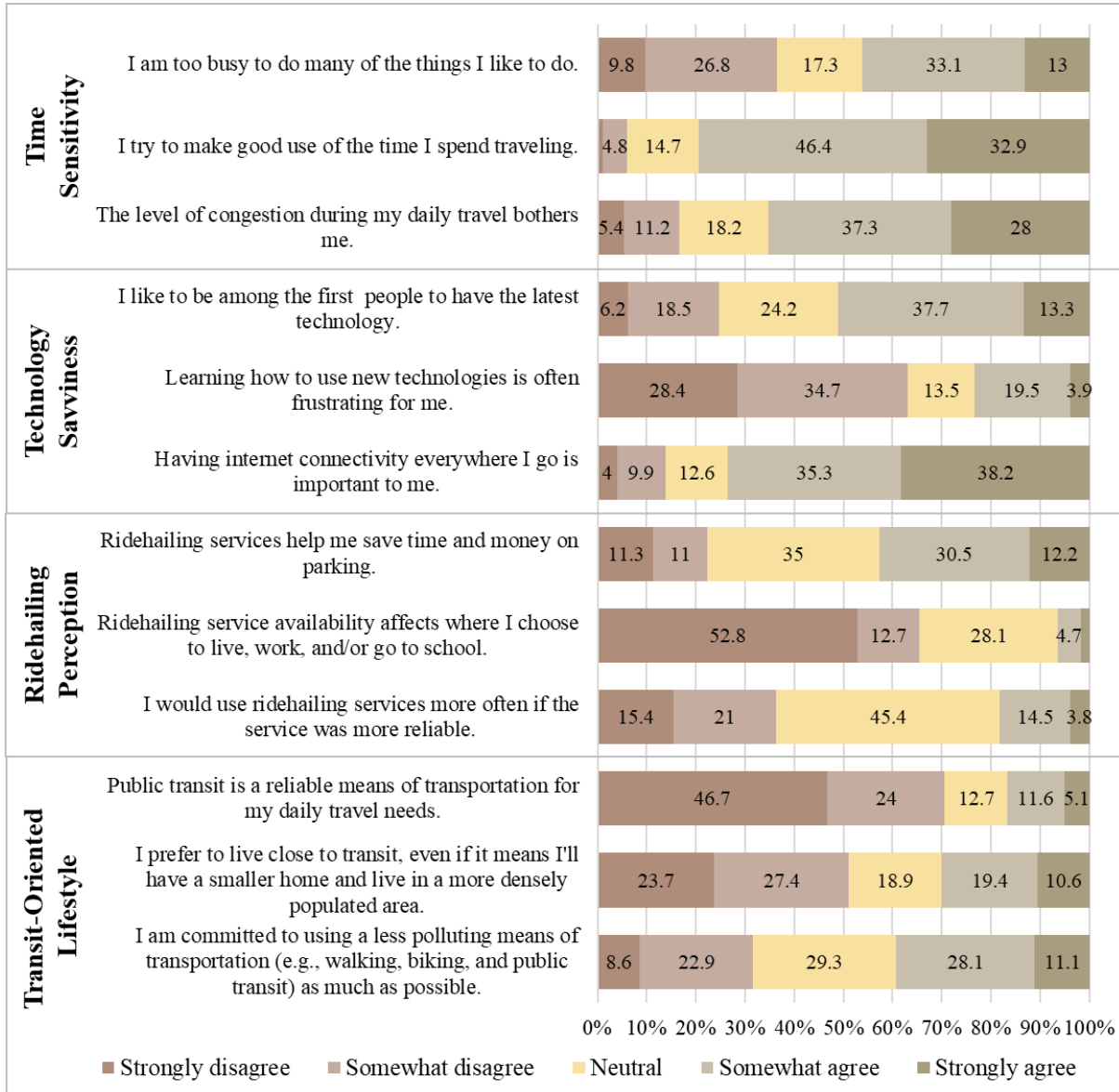
Figure 1 shows the bivariate relationship between the two endogenous variables. A pattern is discernible. Among individuals residing in zero-vehicle households, 18 percent use ridehailing services weekly, and another 23.6 percent use the services monthly. These percentages stand in stark contrast to those who reside in households with vehicles available. The two categories in which at least one vehicle is available in the household (deficient and sufficient) exhibit a different pattern of ridehailing usage compared with individuals in zero-vehicle households. In vehicle-deficient households, only 4.2 percent of individuals use ridehailing services frequently, suggesting that households are effective at sharing limited vehicles amongst one another. In vehicle-sufficient households, only 3.4 percent use ridehailing services on a weekly basis. Likewise, while 15.6 percent in vehicle-deficient households use ridehailing services monthly, a slightly smaller 14.4 percent of individuals in vehicle-sufficient households use ridehailing services with such a frequency. Also, 36.3 percent of individuals in vehicle deficient households never use ridehailing services; the corresponding percent for individuals in vehicle-sufficient households is higher at 39.2 percent. Overall, the bivariate distribution shows a strong relationship between vehicle availability and ridehailing frequency. This project aims to shed light on the nature of the causal relationship between these two endogenous choice variables.



**Figure 1 Household Vehicle Availability by Ridehailing Use Frequency (N=3,146)**

A key aspect of this project’s methodological approach involves incorporating and explicitly accounting for attitudinal variables that capture perceptions, values, and preferences. The survey included a large battery of attitudinal statements to elicit perceptions, values, and preferences with respect to mobility options, lifestyle proclivities, and outlook towards ridehailing services. As such, four latent attitudinal constructs were adopted in this project: *time sensitivity*, *technology savviness*, *(positive) ridehailing service perception*, and *transit-oriented lifestyle (proclivity)*. These latent attitudinal constructs were developed based on evidence in the literature (Alemi et al., 2019; Lavieri and Bhat, 2019a), behavioral intuitiveness, and consideration of the types of variables that are most likely to play a role in shaping vehicle availability and ridehailing frequency choices. Each latent construct is represented by three highly correlated attitudinal

indicators, thus calling for the estimation of latent factors that serve as a composite representation of disparate attitudinal dimensions. Figure 2 depicts the latent constructs and the attitudinal statements defining them. A detailed discussion is not furnished here in the interest of brevity. The figure shows that the attitudinal statements are intuitively related to the construct of interest and are distributed in the sample in a manner consistent with expectations.



**Figure 2 Agreement with Attitudinal Indicators Defining Latent Constructs (N=3,146)**

A standard factor analysis (principal components with varimax rotation) was conducted to develop the latent factors and compute latent factor scores for each observation in the data set. The latent factors were then used in the model estimation exercise. Although latent attitudinal constructs are endogenous variables themselves, they are treated as exogenous explanatory variables in this project. Treating them as endogenous variables within the context of a latent segmentation modeling framework that aims to simultaneously capture multiple causal relationships between endogenous choice variables presents an analytical and computational

challenge. The number of possible permutations of causal relationship structures becomes very large, thus presenting computational complexity. As such, only ridehailing use and vehicle availability are treated as endogenous variables and the model structure focuses on the nature of the causal relationship between them.

### **3. MODELING FRAMEWORK**

This section presents the modeling framework adopted in this project. This project is concerned with unraveling the direction of the causal relationship between ridehailing frequency and vehicle availability (both are frequencies or counts with a natural ordered representation). Estimation of bidirectional causal models is only feasible when both behavioral choice variables are continuous; under that restrictive scenario, a mutually reinforcing relationship between two dependent variables may be explicitly estimated. However, when the choice variables are discrete or limited dependent in nature (not continuous), which is often the case in travel behavior research, then a bidirectional relationship is not identified, and identification restrictions must be imposed for logical consistency and estimability (Pendyala and Bhat, 2004). This necessitates the estimation of recursive simultaneous equation models, where a specific direction of causality is assumed for all observations. However, all individuals in a population are unlikely to follow the same single causal structure; this project is motivated by the desire to identify the extent to which multiple causal relationships co-exist in the project sample and to understand the differences in socio-economic and demographic characteristics between market segments defined by two different causal structures. Such insights will help inform demand forecasting models, enabling them to better represent structural heterogeneity (in causal relationships between mobility choice variables) in the population. The insights will also guide policies and the design of behavioral change interventions.

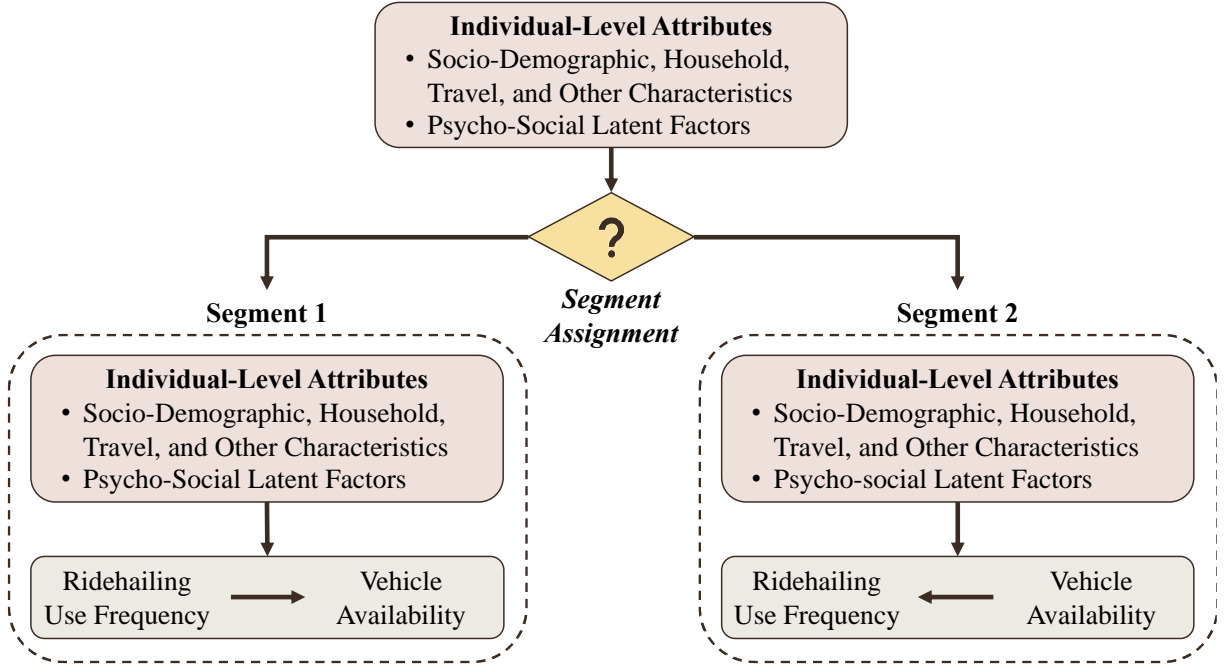
It must be recognized that cause-and-effect patterns, in general, unfold over time, involve leads and lags, and are inherently dynamic in nature. Hence longitudinal panel data is needed to elucidate and identify causal relationships. Although such data have been collected occasionally in the profession, the prevailing norm continues to be the collection of (repeated) cross-sectional data from a sample of the population. It is nearly impossible to unravel cause-and-effect relationships that occur over time in the absence of true longitudinal panel data. Hence the travel behavior field has had to infer causal relationships based on cross-sectional survey data (this is the norm in the vast body of transportation modeling literature). Because cross-sectional data are used in this project, the analysis should be interpreted as invoking the notion of *contemporaneous* causation (Hicks, 1980), which is generally defined as the concept that behavior is caused at the moment of its occurrence by all influences that are present in the individual at that moment (Lewin, 1947). The authors fully recognize that vehicle transaction decisions play out over time. Hence, a strong underlying assumption of this project is that causal relationships involving vehicle availability can be modeled within the psychological construct of contemporaneous causation. Future research needs to relax this assumption and employ longitudinal survey data to address the limitations of cross-sectional treatment of this causal relationship.

#### **3.1. Model Structure**

The model structure adopted in this project posits that each individual in the project sample follows one of two causal structures representing the relationship between ridehailing frequency and vehicle availability. The actual causal structure that drives behaviors of each individual, however, is not observed explicitly. Hence, a latent segmentation modeling approach is adopted; the

approach facilitates the identification of two latent classes and the probabilistic allocation of each individual to one of the two latent market segments. This approach recognizes the inter-relationship between the two choice variables. Vehicle availability could impact the need for and usage frequency of ridehailing services, and conversely, the frequency of ridehailing use could impact the need to own (or dispose of) household vehicles.

The model system comprises multiple elements. A latent segmentation model component is used to assign individuals to one of the two causal segments based on their individual and household-level attributes. Then, within these segments, both variables of interest are jointly modeled as a function of socio-economic, demographic, travel, built environment, and attitudinal variables. Figure 3 offers a simplified representation of the model framework.



**Figure 3 Latent Segmentation Model Framework**

### 3.2. The Joint Model of Behavioral Choices

Consider an individual  $q$  ( $q=1, 2, 3, \dots, Q$ ) facing a multi-dimensional ordered choice system. Let  $c$  be the index for the ordinal outcome ( $c = 1, 2, \dots, C$ ;  $C=2$  in our case). For presentation ease, subscript  $q$  is dropped for the individual. Assume that the individual belongs to a specific segment  $h$ . Define a latent propensity  $y_{ch}^*$  underlying the count variable  $y_c$  for the outcome  $c$  and for segment  $h$ . Then,

$$y_{ch}^* = \beta_{ch}' \mathbf{x} + \varepsilon_{ch}, \quad y_c = k_c \text{ if } \psi_{ch, k_c - 1} < y_{ch}^* < \psi_{ch, k_c}, \quad (1)$$

where  $\mathbf{x}$  is a  $(L \times 1)$  vector of exogenous attributes (not including a constant) as well as possibly the observed values of other endogenous variables,  $\beta_{ch}$  is a corresponding  $(L \times 1)$  vector of channel-specific coefficients to be estimated (note that by restricting specific elements of  $\beta_{ch}$  to be zeros, it is possible to control which variables to estimate specific to the segment  $h$ ; also,  $\beta_{ch}$  can be zero on the endogenous variables within each segment) and  $\varepsilon_{ch}$  is a random error term

assumed to be standard normally distributed.  $k_c$  represents a specific value of  $y_c$ , which can range from the value of 0 to a maximum of  $K_c$  in the sample ( $y_c \in \{0, 1, 2, \dots, K_c\}$ ). The latent count propensity  $y_{ch}^*$  is mapped to the observed count variable  $y_c$  by the thresholds  $\psi_{ch,k_c}$ , which should satisfy the ordering conditions ( $\psi_{ch,-1} = -\infty$ ;  $-\infty < \psi_{ch,0} < \psi_{ch,1} < \dots < \psi_{ch,K_c-1} < \infty$ ) in the usual ordered-response fashion.

Next, vertically stack the  $C$  latent variables  $y_{ch}^*$  into a  $(C \times 1)$  vector  $\mathbf{y}_h^*$ , and the  $C$  error terms  $\varepsilon_{ch}$  into another  $(C \times 1)$  vector  $\boldsymbol{\varepsilon}_h$ . Let  $\boldsymbol{\varepsilon}_h \sim MVN_C(\mathbf{0}_C, \boldsymbol{\Xi}_h)$ , where  $MVN_C(\mathbf{0}_C, \boldsymbol{\Xi}_h)$  represents the  $C$  dimensional multivariate normal distribution with mean vector  $\mathbf{0}_C$  (a  $(C \times 1)$  vector of zeros) and a correlation matrix of  $\boldsymbol{\Xi}_h$  specific to segment  $h$ . The off-diagonal terms of  $\boldsymbol{\Xi}_h$  capture the error covariance across the underlying latent continuous propensities of the different ordered outcomes. For future use, also define the vector of thresholds for each outcome  $c$  as:  $\tilde{\boldsymbol{\Psi}}_{ch} = (\psi_{ch,0}, \psi_{ch,1}, \dots, \psi_{ch,K_c-1})'$ , and further vertically stack all the  $\tilde{\boldsymbol{\Psi}}_{ch}$  vectors into a single  $\tilde{\boldsymbol{\Psi}}_h$  vector.

Let an individual under consideration be observed to have the count values of  $k_c$  ( $c=1, 2, \dots, C$ ). Accordingly, stack the lower thresholds  $\psi_{ch,k_c-1}$  ( $c=1, 2, \dots, C$ ) corresponding to the observed ordered values of the individual into a  $(C \times 1)$  vector  $\boldsymbol{\Psi}_{low,h}$ , and the upper thresholds  $\psi_{ch,k_c}$  ( $c=1, 2, \dots, C$ ) into another  $(C \times 1)$  vector  $\boldsymbol{\Psi}_{high,h}$ . Also, define  $\boldsymbol{\beta}_h = (\boldsymbol{\beta}_{1h}, \boldsymbol{\beta}_{2h}, \dots, \boldsymbol{\beta}_{Ch})'$  [ $(C \times L)$  matrix]. With these notational preliminaries, the latent propensities underlying the multivariate ordered outcomes may be written in matrix form as:

$$\mathbf{y}_h^* = \boldsymbol{\beta}'_h \mathbf{x} + \boldsymbol{\varepsilon}_h, \quad \boldsymbol{\Psi}_{low,h} < \mathbf{y}_h^* < \boldsymbol{\Psi}_{high,h}, \quad \text{where } \mathbf{y}_h^* \sim MVN_C(\boldsymbol{\beta}'_h \mathbf{x}, \boldsymbol{\Xi}_h). \quad (2)$$

Let  $\boldsymbol{\delta}_h$  be the collection of parameters to be estimated for segment  $h$ :

$$\boldsymbol{\delta}_h = \left( \left[ \text{Vech}(\boldsymbol{\beta}_h) \right]', \tilde{\boldsymbol{\Psi}}_h', \left[ \text{Vechup}(\boldsymbol{\Xi}_h) \right]' \right)', \quad \text{where the operator "Vech(.)" row-vectorizes all the}$$

non-zero elements of the matrix/vector on which it operates, and the operator Vechup(.) row-vectorizes the upper diagonal elements of a matrix. Then the likelihood function of a single individual  $q$  may be written as:

$$L(\boldsymbol{\delta}_h) = \Pr \left[ \boldsymbol{\Psi}_{low,h} < \mathbf{y}_h^* < \boldsymbol{\Psi}_{high,h} \right], \quad (3)$$

$$= \int_{D_r} f_C(\mathbf{r} \mid \boldsymbol{\beta}'_h \mathbf{x}, \boldsymbol{\Xi}_h) d\mathbf{r}, \quad (4)$$

where the integration domain  $D_r = \{\mathbf{r} : \boldsymbol{\Psi}_{low,h} < \mathbf{r} < \boldsymbol{\Psi}_{high,h}\}$  is simply the multivariate region of the  $\mathbf{y}_h^*$  vector determined by the upper and lower thresholds.  $f_C(\mathbf{r} \mid \boldsymbol{\beta}'_h \mathbf{x}, \boldsymbol{\Xi}_h)$  is the MVN density function of dimension  $C$  with a mean of  $\boldsymbol{\beta}'_h \mathbf{x}$  and a correlation matrix  $\boldsymbol{\Xi}_h$ . Bhat's (2018) matrix-based approximation method for evaluating the multivariate normal cumulative distribution (MVNCD) function was employed to evaluate this integral, which provides an efficient and tractable formulation to approximate the integral.

### 3.3. Latent Segmentation Model

The derivation thus far is based on the notion that individual  $q$  belongs to a single segment  $h$ .

Although the actual assignment of individual  $q$  to a specific segment is not observed, it is possible to attribute a probability  $\pi_{qh}$  ( $h=1,2,\dots,H$ ) to individual  $q$  belonging to segment  $h$ . The conditions that  $0 \leq \pi_{qh} \leq 1$  and  $\sum_{h=1}^H \pi_{qh} = 1$  must be met. To enforce these restrictions, following

Bhat (1997), the following logit link function is used:

$$\pi_{qh} = \frac{\exp(\boldsymbol{\mu}'_h \mathbf{w}_q)}{\sum_{j=1}^H \exp(\boldsymbol{\mu}'_j \mathbf{w}_q)}, \quad (5)$$

where  $\mathbf{w}_q$  is a  $(J \times 1)$  vector of individual exogenous variables,  $\boldsymbol{\mu}_h$  is the corresponding  $(J \times 1)$  vector of parameters, and  $\boldsymbol{\mu}_1 = \mathbf{0}$  serves as a vector identification condition. Defining  $\boldsymbol{\delta} = [\boldsymbol{\delta}'_1, \dots, \boldsymbol{\delta}'_h; \boldsymbol{\mu}'_1, \dots, \boldsymbol{\mu}'_h]'$ , then the likelihood function for individual  $q$  is:

$$L_q(\boldsymbol{\delta}) = \sum_{h=1}^H \pi_{qh} [L(\boldsymbol{\delta}_h) | q \in \text{segment } h], \quad (6)$$

and the overall likelihood function is then given as:

$$L(\boldsymbol{\delta}) = \prod_q L_q(\boldsymbol{\delta}). \quad (7)$$

#### 4. MODEL ESTIMATION RESULTS

This section summarizes model estimation results. Before estimating the joint model system, separate confirmatory factor analysis (CFA) was conducted to construct the latent attitudinal factors that serve as exogenous variables in the model specification. These constructs were explained in detail earlier. All the indicators used to define the latent constructs were significant and loaded heavily on their designated latent constructs following a varimax rotation. The CFA results are suppressed and not presented in detail in the interest of brevity. It should be noted that factor loadings are all intuitive, and the latent constructs capture a range of proclivities that are likely to influence an individual's propensity for vehicle ownership and using ridehailing services.

##### 4.1. Bivariate Model of Behavioral Outcomes

Table 2 presents model estimation results for the bivariate model of ordered behavioral outcomes. Ridehailing frequency is represented by outcomes of *never*, *rarely*, *monthly*, and *weekly*. Vehicle availability is represented by the outcomes of *none*, *vehicle deficient*, and *vehicle sufficient*. The table shows coefficient estimates for each of the two latent causal segments. In both segments, the endogenous variables depict a significant inverse (negative) relationship, suggesting that higher vehicle availability is associated with a lower level of ridehailing frequency, and vice versa. These relationships are significant (in either causal direction), behaviorally intuitive, and consistent with previous findings (Dias et al., 2017; Sikder, 2019; Sabouri et al., 2020; Wu and Mackenzie, 2021).



**Table 2 Estimation Results for Endogenous Variables Within Each Segment**

Explanatory Variables (base category)		Segment 1 (Ridehailing → Vehicle Avail)				Segment 2 (Vehicle Avail → Ridehailing)			
		Ridehailing Frequency (4-level: never to weekly)		Vehicle Availability (3-level: none, deficient, sufficient)		Ridehailing Frequency (4-level: never to weekly)		Vehicle Availability (3-level: none, deficient, sufficient)	
		Coef.	t-stat	Coef.	t-stat	Coef.	t-stat	Coef.	t-stat
<b>Endogenous variables</b>	Ridehailing frequency	na	na	-0.49	-2.88	na	na	na	na
	Vehicle availability	na	na	na	na	-1.04	-5.58	na	na
<b>Latent constructs</b>	Time sensitivity	na	na	0.30	2.91	na	na	na	na
	Technology savviness	0.09	1.89	na	na	0.09	1.70	na	na
	Ridehailing perception	0.25	5.05	na	na	0.14	1.84	na	na
	Transit-oriented lifestyle	na	na	-0.43	-5.05	0.09	1.62	-0.22	-3.56
<b>Individual characteristics</b>									
Age (*)	18-40 years	na	na	na	na	na	na	-0.86	-4.14
	65 years or older	na	na	na	na	na	na	0.57	4.31
	71 years or older	-0.40	-3.82	na	na	-0.56	-2.53	na	na
Education (< Bachelor's degree)	Higher education	0.19	2.88	0.16	1.54	0.24	3.00	na	na
Race (not White or Caucasian)	White or Caucasian	na	na	0.32	2.92	na	na	0.14	1.46
Ethnicity (not Hispanic)	Hispanic	na	na	-0.20	-1.65	na	na	-0.20	-1.49
Employment (*)	Worker	na	na	0.28	2.34	na	na	0.39	3.23
	Non-worker	-0.22	-3.10	na	na	-0.29	-3.08	na	na
<b>Household and other characteristics</b>									
Household income (*)	Less than \$25,000	na	na	-0.61	-4.67	na	na	na	na
	Less than \$50,000	na	na	na	na	-0.19	-2.14	na	na
	\$100,000 or more	na	na	0.40	2.63	na	na	na	na
	\$150,000 or more	0.64	8.00	na	na	0.50	4.42	na	na
Household size (>1)	One	0.11	1.62	na	na	na	na	na	na
Housing unit type (other)	Stand-alone home	na	na	0.41	2.71	na	na	0.25	1.86
	Apartment	0.49	6.28	na	na	0.28	2.46	na	na
Population density (high)	Low (<3000 person/mi <sup>2</sup> )	-0.19	-3.15	na	na	na	na	na	na
City (Austin, Phoenix, Tampa)	Atlanta	na	na	na	na	na	na	0.42	4.28
Commute distance (0 or 5+)	>0 to 5 mi	0.45	5.60	na	na	na	na	na	na
<b>Thresholds</b>	1 2	-0.30	-3.46	-1.95	-6.09	-2.62	-5.67	-4.22	-0.37
	2 3	1.13	13.30	-1.82	-5.36	-1.37	-2.76	0.48	2.32
	3 4	2.14	22.11	na	na	-0.48	-0.89	na	na
<b>Correlation</b>	Ridehailing frequency	na	na	0.32	1.43	na	na	0.50	2.89

\*Base category is not identical across the model equations and corresponds to all omitted categories.

Goodness-of-Fit Measures: Adjusted  $\rho^2 = 0.147$ ; BIC = 5135.04; Log-likelihood (Joint Model) = -4901.47; Log-likelihood (Constants-only Model) = -5,746.48

Average probability of correct prediction: Joint Model = 0.285; Constants-only Model = 0.215

All other results are behaviorally intuitive, with largely consistent indications between the two segments. In the segment where ridehailing affects vehicle availability, time sensitivity is found to have a positive influence on the propensity for higher vehicle availability. This finding echoes the notion that time sensitive individuals who feel rushed are likely to prefer a higher level of access to the automobile, which is generally the fastest mode in the metropolitan areas where the survey was conducted. Also, in both segments, technology savviness and a positive perception of ridehailing services enhance the proclivity towards using ridehailing services more frequently.

Similar findings were reported in the literature (Lavieri and Bhat, 2019b; Alonso-González et al., 2020). A transit-oriented lifestyle is associated with a lower level of vehicle availability, consistent with the findings reported by Cervero (2007), and positively influencing ridehailing frequency in the segment where vehicle availability affects ridehailing frequency.

Among individual characteristics, younger individuals (18-40 years) show a lower proclivity towards vehicle availability in the segment where vehicle availability affects ridehailing. The older age group (71+ years) exhibits a lower propensity towards ridehailing frequency, consistent with the notion that ridehailing users tend to be younger (Alemi et al., 2019). Also consistent with prior research is the finding that higher education levels are associated with a proclivity towards higher frequency of ridehailing use (Dias et al., 2017). In the segment where ridehailing affects vehicle availability, higher education levels are associated with a greater proclivity to higher vehicle availability. Whites have a greater proclivity for higher vehicle availability levels in both causal segments, whereas Hispanics have a lower proclivity; prior research has also documented these racial differences (Klein and Smart, 2017; Sabouri et al., 2020). Workers are likely to prefer higher vehicle availability (presumably for commute needs), while non-workers exhibit a lower propensity towards frequent ridehailing use – aligned with the findings reported previously (Blumenberg et al., 2021; Potoglou and Kanaroglou, 2008).

In the segment where ridehailing frequency affects vehicle availability, income effects echo prior research (Blumenberg et al., 2021). A lower income level (less than \$50,000 per year) is associated with lower levels of ridehailing frequency – a finding consistent with the literature (Dias et al., 2017). The highest income category exhibits a positive association with higher ridehailing usage, while the second highest income bracket is associated with higher levels of vehicle availability. Single persons are more likely to use ridehailing, consistent with earlier findings (Sikder, 2019). Stand-alone home residents are likely to have higher levels of vehicle availability, while individuals in apartments (presumably in higher-density locales) tend to embrace higher levels of ridehailing use. Low-density living is associated with lower levels of ridehailing use, while those with short commutes are likely to adopt higher levels of ridehailing use, confirming previous findings (Lavieri and Bhat 2019a). Residents of Atlanta appear to have a proclivity towards higher levels of vehicle availability, but this finding appears only in the segment where vehicle availability affects ridehailing frequency.

The correlations between the two outcomes are positive in both segments, possibly indicating underlying correlated unobserved factors that favor private vehicle usage (personal cars as well as ridehailing vehicles). In contrast to transit and other non-motorized modes, people generally prefer to travel in private vehicles due to the greater convenience, comfort, and efficiency (Magassy et al., 2022). The finding that this correlation is statistically significant and larger in Segment 2 (where vehicle availability influences ridehailing frequency) also highlights the underlying tendency towards the auto mode. This result speaks not only to the importance of the self-selection effect but also to the importance of joint modeling. If positive correlations are ignored, the unexplained error correlation between the two variables will be included in the direct

effect of one outcome on another (depending on which causal direction is considered). As a result, the magnitude of the negative impact of ridehailing frequency on vehicle availability (or vice versa) will be underestimated by independent model systems that ignore error correlation. The direct impacts estimated in this joint model system are thus the *true* (cleansed) effects of one outcome on the other, after controlling for the self-selection effect arising from unobserved factors that affect both outcome variables. The need for a joint model in examining inter-related mobility choices is further supported by goodness-of-fit measures.

#### 4.2. Characteristics and Size of the Latent Segments

To probabilistically assign individuals to a causal structure, a binary latent market segmentation model was estimated. In the interest of brevity, the estimation results for this binary logit model are not presented in tabular form. Table 3 offers a detailed description of the segment profiles, thus obviating the need to present the estimation results explicitly (they essentially mirror the profiles in Table 3). This section, therefore, focuses on presenting the latent segment profiles.

Each segment size is reported at the bottom of Table 3, and it is noteworthy that a majority (58 percent) of the observations are probabilistically assigned to the market segment where ridehailing frequency affects vehicle availability. The remainder is assigned to the segment where vehicle availability affects ridehailing frequency. This is *counter* to what is often represented in transportation demand forecasting models in practice, which generally tend to predict mode choice (including ridehailing use) as a function of vehicle ownership levels. While vehicle ownership is affected by composite modal accessibility measures (such as logsums that presumably reflect the presence of ridehailing services as well), these measures rarely (if ever) capture the *frequency* of ridehailing use. As a result, models do not reflect the influence of the *extent* of ridehailing use on household vehicle ownership. Both segments have a sizeable proportion of sample observations, reflecting the need to incorporate multiple causal structures (reflecting different market segments) in transportation demand forecasting models (as opposed to assuming a single causal structure for all agents in the population).

Table 3 also shows variations of two segments by demographic attributes. In general, attributes with substantial differences in the table appeared statistically significant in the binary segmentation model. The first broad numeric column “Percent within segment” provides the split of a variable within each segment; thus, within the first segment where ridehailing frequency affects vehicle availability, 56.2 percent are women and 43.8 percent are men. Within the second segment where vehicle availability affects ridehailing frequency, the corresponding split between women and men is 58.4 and 41.6 percent, respectively. This indicates that women populate segment 2 more than men. Another way to see this is the entries corresponding to the broad column entitled “Percent within attribute”. This shows that 57.1 percent of women belong to segment 1 (compared to 59.3 percent of men), while 49.2 percent of women belong to segment 2 (compared to only 40.7 percent of men). Other entries may be similarly interpreted.

According to our results, age also is a distinguishing characteristic between the two segments. While 40.6 percent of individuals following the pattern where ridehailing frequency affects vehicle availability fall in the younger age group of 18-40 years, the corresponding percentage for the other segment is lower at 34 percent. This is intuitive since younger individuals are more likely to embrace new mobility options. They were early adopters of ridehailing services, and likely to have used such services frequently enough and for a duration long enough to influence their decisions about vehicle ownership. Such differences are discernible throughout the table.

**Table 3 Profiles of the Two Latent Market Segments**

Attributes		Percent (%) within segment		Percent (%) within attribute		Overall Sample (%)
		Segment 1 RF→VA	Segment 2 VA→RF	Segment 1 RF→VA	Segment 2 VA→RF	
<b>Individual characteristics</b>						
Gender	Female	56.2	58.4	57.1	42.9	57.1
	Male	43.8	41.6	59.3	40.7	42.9
Age	18-40	40.6	34.0	62.3	37.7	37.8
	41-60	25.2	40.8	46.1	53.9	31.8
	61 or older	33.7	24.6	65.5	34.5	29.9
Education	High school or less	7.2	10.4	49.1	50.9	8.5
	Assoc. degree or some college	26.0	31.7	53.2	46.8	28.4
	Bachelor's degree or higher	66.6	57.7	61.5	38.5	62.9
Race	Asian	8.2	10.1	53.1	46.9	9.0
	Black	8.3	7.1	61.6	38.4	7.8
	White or Caucasian	72.3	69.2	59.1	40.9	71.0
	Other	11.2	13.6	53.3	46.7	12.2
Employment	Worker	63.2	66	57.0	43.0	64.4
	Non-worker	36.8	34	60.0	40.0	35.6
<b>Household characteristics</b>						
Household income	Up to \$50,000	25.4	24.4	59.0	41.0	25.0
	\$50,000 to \$99,999	30.7	37.8	53.0	47.0	33.7
	\$100,000 to \$149,999	21.7	21.7	58.0	42.0	21.7
	\$150,000 or more	21.7	15.1	66.5	33.5	18.9
Household size	One	27.7	12.1	76.0	24.0	21.1
	Two	51.6	20.4	77.8	22.2	38.5
	Three or more	20.7	67.5	29.9	70.1	40.3
Household children	Zero	85.2	57.2	67.4	32.6	73.5
	One or more	14.8	42.8	32.4	67.6	26.5
Housing unit type	Stand-alone home	65.6	76.0	54.5	45.5	69.9
	Apartment	24.5	16.3	67.5	32.5	21.1
	Other	8.5	6.5	64.5	35.5	7.7
Household vehicle	Zero	4.7	2.7	70.5	29.5	3.9
	One	28.1	17.6	68.9	31.1	23.7
	Two or more	67.1	79.6	53.9	46.1	72.4
<b>Other characteristics</b>						
Population density	Low (<3000 person/mi <sup>2</sup> )	48.7	51.1	56.9	43.1	49.7
	High (≥ 3000 person/mi <sup>2</sup> )	51.3	48.9	59.2	40.8	50.3
City	Atlanta	31.7	27.9	61.2	38.8	30.1
	Austin	32.4	32.3	58.1	41.9	32.3
	Phoenix	29.2	32.1	55.8	44.2	30.4
	Tampa	6.7	7.7	54.7	45.3	7.2
Segment size		Percent (%)		<b>58.1</b>	<b>41.9</b>	<b>100</b>
		N		<b>1,828</b>	<b>1,318</b>	<b>3,146</b>

Other variables depicting key differences between the two segments include education, household income, household size, presence of children, housing unit type, and household vehicle ownership itself. All of these constitute variables that may be used to define market segments so that the appropriate causal structure can be applied in demand forecasting models. Consistent with the historical evidence on who has tended to be early and more frequent adopters of ridehailing services, this project finds that the market segment where ridehailing frequency influences vehicle availability exhibits higher shares of individuals who are highly educated and affluent, live in small households, reside in apartments, and own fewer vehicles. If one were to examine the percentages within segments, 67 percent of the market segment where ridehailing frequency affects vehicle availability has a Bachelor's degree or higher, as opposed to a smaller 57.7 percent for the market segment where vehicle availability affects ridehailing frequency. Also, 21.7 percent of those in the market segment where ridehailing frequency affects vehicle availability fall in the income category of \$150,000 or higher; this percentage is only 15.1 percent for the other segment.

Nearly 80 percent of individuals in the segment where ridehailing frequency affects vehicle availability belong to one- or two-person households; the corresponding percentage for the other segment is merely 33 percent. Similar differences can be seen with respect to the presence of children (just 14.8 percent for the segment where ridehailing frequency affects vehicle availability, but 42.8 percent for the other segment). In general, this project finds that larger households with children in stand-alone housing units in suburban locales are more likely to embrace vehicle ownership-oriented lifestyle (because of their household mobility needs, patterns, and constraints), and this consequently impacts the use of ridehailing services. These findings are very consistent with expectations and demonstrate the importance of reflecting multiple causal structures in transportation demand forecasting models.

## **5. DISCUSSION AND CONCLUSIONS**

This project is concerned with the complex inter-relationship between ridehailing service usage and vehicle ownership. There are essentially two plausible (causal) relationships between these variables, and this project attempts to determine the degree to which these two causal relationships co-exist in a population. In addition, the project seeks to determine the profiles of the market segments following the two different causal structures. Because the causal relationship is not directly observed, a latent segmentation modeling approach is adopted. This approach allows individuals to be probabilistically assigned to different causal market segments based on their attributes. A joint bivariate ordered probit model of ridehailing frequency and household vehicle availability is estimated that incorporates the two plausible causal structures, one in which ridehailing frequency affects vehicle ownership and the other in which the opposite causal direction exists.

A majority of the sample (58 percent) is found to follow the causal structure in which ridehailing frequency affects vehicle availability. The two latent market segments are found to differ substantially with respect to age, income, household size, housing unit type, and presence of children. Two key conclusions may be drawn from these findings. First, the two causal structures are prevalent in this particular sample to a substantial degree. While it may be acceptable to ignore a specific causal structure if it is rare, either causal structure cannot be ignored in this empirical context. Second, certain demographics (young, highly educated, affluent, adults in small households with no children and residing in apartments) appear to have used ridehailing services frequently enough and for a duration long enough to have had an impact on their vehicle ownership.

This suggests that ridehailing services do exhibit the potential to (negatively) influence the

levels of vehicle ownership in the future (as the services continue to grow). This project lends credence to the notion that a future characterized by MaaS may indeed see lower levels of private auto ownership as households become increasingly comfortable with downsizing their private vehicle fleet. At the same time, however, there is a sizeable segment of the population for whom vehicle ownership levels affect the degree to which they use ridehailing services. Targeted marketing campaigns and interventions that enhance the ability to embrace ridehailing services may help accelerate a future of lower vehicle ownership; these campaigns should target older individuals and larger households (with children) residing in stand-alone housing units in suburban locales.

The report findings also indicate the need to reflect multiple causal structures in transportation demand forecasting models. Model systems that are based on a single causal structure (where vehicle availability affects mode choice and ridehailing usage) do not reflect the structural heterogeneity prevalent in the population. Model systems need to be enhanced to define specific market segments in the population based on a multitude of socio-economic dimensions. Furthermore, inter-related mobility choices should be modeled jointly with explicit accounting for error correlations to enable computation of the true effect between the choice variables. Through a market segmentation approach that employs joint model specifications, it will be possible to simultaneously reflect alternative causal structures driving mobility choices, more accurately reflect true behavioral phenomena at play, obtain more reliable estimates of policy impacts/effects, and target interventions more effectively.

## REFERENCES

- Alemi, F., 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. *Transportation Research C*, 2019.
- Alonso-González, M.J., S. Hoogendoorn-Lanser, N. van Oort, O. Cats, and S. Hoogendoorn. Drivers and Barriers in Adopting Mobility as a Service (MaaS)—a Latent Class Cluster Analysis of Attitudes. *Transportation Research A*, 2020.
- Bhat, C.R. An Endogenous Segmentation Mode Choice Model with an Application to Intercity Travel. *Transportation Science*, 1997.
- Bhat, C.R. New Matrix-Based Methods for the Analytic Evaluation of the Multivariate Cumulative Normal Distribution Function. *Transportation Research B*, 2018.
- Blumenberg, E., J. Paul, and G. Pierce. Travel in the Digital Age: Vehicle Ownership and Technology-Facilitated Accessibility. *Transport Policy*, 2021.
- Cervero, R. Transit-Oriented Development's Ridership Bonus: A Product of Self-Selection and Public Policies. *Environment and Planning A*, 2007.
- Dean, B. Uber Statistics 2022: How Many People Ride with Uber?. 2021. <https://backlinko.com/uber-users>. Accessed Aug 1, 2022.
- Diao, M., H. Kong, and J. Zhao. Impacts of Transportation Network Companies on Urban Mobility. *Nature Sustainability*, 2021.
- Dias, F., P. Lavieri, V. Garikapati, S. Astroza, R.M. Pendyala, and C.R. Bhat. A Behavioral Choice Model of the Use of Car-Sharing and Ride-Sourcing Services. *Transportation*, 2017.
- Erhardt, G.D., J.M. Hoque, V. Goyal, S. Berrebi, C. Brakewood, and K.E. Watkins. Why Has Public Transit Ridership Declined in the United States?. *Transportation Research A*, 2020.
- Henao, A. and W.E. Marshall. The Impact of Ride-hailing on Vehicle Miles Traveled.

- Transportation*, 2019.
- Hicks, J. *Causality in Economics*. Australian National University Press, 1980.
- Khoeini, S. and R.M. Pendyala. T4 Survey. <https://tomnet-utc.engineering.asu.edu/data/t4-survey/>. Accessed Aug 1, 2022.
- Klein, N.J. and M.J. Smart. Millennials and Car Ownership: Less Money, Fewer Cars. *Transport Policy*, 2017.
- Lavieri, P.S. and C.R. Bhat. Modeling Individuals' Willingness to Share Trips with Strangers in an Autonomous Vehicle Future. *Transportation Research A*, 2019a.
- Lavieri, P.S. and C.R. Bhat. Investigating Objective and Subjective Factors Influencing the Adoption, Frequency, and Characteristics of Ride-Hailing Trips. *Transportation Research C*, 2019b.
- Lewin, K. *Field Theory in Social Science: Selected Theoretical Papers (Edited by Dorwin Cartwright)*. 1951.
- Magassy, T.B., I. Batur, A. Mondal, K.E. Asmussen, S. Khoeini, R.M. Pendyala, and C.R. Bhat. Influence of Mode Use on Level of Satisfaction with Daily Travel Routine: A Focus on Automobile Driving in the United States. *Transportation Research Record*, 2022.
- Potoglou, D. and P.S. Kanaroglou. Modelling Car Ownership in Urban Areas: A Case Study of Hamilton, Canada. *Transport Geography*, 2008.
- Pendyala, R.M. and C.R. Bhat. An Exploration of the Relationship between Timing and Duration of Maintenance Activities. *Transportation*, 2004.
- Sabouri, S., S. Brewer, and R. Ewing. Exploring the Relationship between Ride-Sourcing Services and Vehicle Ownership, Using Both Inferential and Machine Learning Approaches. *Landscape and Urban Planning*, 2020.
- Sikder, S. Who Uses Ride-Hailing Services in the United States? *Transportation Research Record*, 2019.
- Tang, B.J., X.Y. Li, B. Yu, and Y.M. Wei. How App-Based Ride-Hailing Services Influence Travel Behavior: An Empirical Study from China. *International Journal of Sustainable Transportation*, 2020.
- Tirachini, A. Ride-hailing, Travel Behaviour and Sustainable Mobility: An International Review. *Transportation*, 2019.
- Wang, Y., W. Shi, and Z. Chen. Impact of Ride-Hailing Usage on Vehicle Ownership in the United States. *Transportation Research D*, 2021.
- Wu, X. and D. MacKenzie. Assessing the VMT Effect of Ridesourcing Services in the US. *Transportation Research D*, 2021.