

## Project Report

# Attitudes and Behaviors Causal Relationships: Uncovering Latent Segments Within A Heterogeneous Population

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



By,

**Ram M. Pendyala**

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

**Shivam Sharda**

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

**Irfan Batur**

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

**Sara Khoeini**

Email: [sara.khoeini@asu.edu](mailto:sara.khoeini@asu.edu)

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

May 2020

## 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> Attitudes And Behaviors Causal Relationships: Uncovering Latent Segments Within A Heterogeneous Population				<b>5. Report Date</b> May 2020	
				<b>6. Performing Organization Code</b> N/A	
<b>7. Author(s)</b> Ram M. Pendyala, <a href="https://orcid.org/0000-0002-1552-9447">https://orcid.org/0000-0002-1552-9447</a> Shivam Sharda, Irfan Batur, <a href="https://orcid.org/0000-0002-8058-2578">https://orcid.org/0000-0002-8058-2578</a> Sara Khoeini, <a href="https://orcid.org/0000-0001-5394-6287">https://orcid.org/0000-0001-5394-6287</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 Tempe, AZ, 85281				<b>10. Work Unit No. (TRAIS)</b> N/A	
				<b>11. Contract or Grant No.</b> 69A3551747116	
<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 (February 2019 – February 2020)	
				<b>14. Sponsoring Agency Code</b> USDOT OST-R	
<b>15. Supplementary Notes</b> N/A					
<b>16. Abstract</b> This project aimed at unraveling the contemporaneous relationship that exists between attitudes and choice behaviors. Attitudes, perceptions, and preferences may shape behaviors; likewise, behavioral choices exercised by individuals may offer experiences that shape attitudes. While it is likely that these relationships play out over time, the question of whether attitudes affect behaviors or behaviors affect attitudes at a specific cross-section in time remains unanswered and a fruitful area of inquiry. Various studies in the literature have explored this question, but have done so without explicitly recognizing the heterogeneity that may exist in the population. In other words, the causal structure at play at any point in time may differ across individuals, thus motivating the development of an approach that can account for the presence of multiple segments in the population, each following a different causal structure. Results suggest that there is considerable heterogeneity in the population with the contemporaneous causal structures in which behaviors shape attitudes more prevalent than those in which attitudes affect choice behaviors. These findings have important implications for transport modeling and policy development.					
<b>17. Key Words</b> Causal analysis, Behavioral choices, Attitudes, Latent segments, Joint models				<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> 28	<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 study 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, USDOT, D-STOP, Complex Engineering Systems Institute 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 Chandra Bhat, and Sebastian Astroza for their contributions to the work presented in this report.*

**TABLE OF CONTENTS**

TECHNICAL REPORT DOCUMENTATION PAGE .....**Error! Bookmark not defined.**

DISCLAIMER ..... 3

ACKNOWLEDGMENTS ..... 3

EXECUTIVE SUMMARY ..... 5

INTRODUCTION ..... 6

DATA DESCRIPTION ..... 7

MODELING METHODOLOGY ..... 11

    The Joint Model of Behavioral Choices and Attitudinal Factor ..... 12

    Segmentation Model ..... 14

RESULTS ..... 15

SIZE AND CHARACTERISTICS OF LATENT SEGMENTS ..... 21

CONCLUSIONS AND POLICY IMPLICATIONS ..... 24

REFERENCES ..... 26

## **EXECUTIVE SUMMARY**

This project goal is unraveling the contemporaneous relationship that exists between attitudes and choice behaviors. Attitudes, perceptions, and preferences may shape behaviors; likewise, behavioral choices exercised by individuals may offer experiences that shape attitudes. While it is likely that these relationships play out over time, the question whether attitudes affect behaviors or behaviors affect attitudes at a specific cross-section in time remains unanswered and a fruitful area of inquiry. Various studies in the literature have explored this question, but have done so without explicitly recognizing the heterogeneity that may exist in the population. In other words, the causal structure at play at any point in time may differ across individuals, thus motivating the development of an approach that can account for the presence of multiple segments in the population, each following a different causal structure. However, the segments are unobserved to the analyst, necessitating the adoption of a latent segmentation approach to identify the extent to which alternative causal structures are prevalent in the population.

This study utilizes a data set that includes attitudinal variables to examine relationships among attitudes towards transit, residential location choice, and frequency of transit use (the latter two variables constituting choice behaviors). Four different latent segments are considered; two latent segments in which attitudes affect choice behaviors and two segments in which choice behaviors affect latent segments.

The overall finding is that the majority of the sample in the data set used in this study are assigned to the latent segments in which behavioral choices affect attitudes. Nearly two-thirds of the sample falls into these two segments, while only about one-third falls into the two segments where attitudes affect behaviors. In other words, the findings of this study appear to corroborate some recent evidence that people appear to modify their attitudes in response to their behaviors to reduce the cognitive dissonance that may exist. It appears that attitudes at any cross-section in time are shaped by the behavioral choices and experiences of the individual at that point in time. As time progresses, it is entirely possible that attitudes and behaviors will evolve; but within the context of a snapshot, the study results here clearly indicate that attitudes are shaped by behaviors more so than the other way around.

Travel demand forecasting models historically assume that the same causal structure applies to all agents in the population. Behavioral choice models are often specified with attitudinal variables and constructs as explanatory factors, when in fact it appears from this study that attitudinal factors are shaped by behavioral choices. It would be beneficial to probabilistically assign individuals in a population to different causal segments, and then forecast travel demand for different segments according to the causal structure that drives their decision-making process.

From a transportation policy perspective, it would appear that information campaigns and advertisements may not be all that effective in a world where the majority of the population has their attitudes shaped by behaviors. This implies that it is necessary to run campaigns where individuals actually get to experience modal options and different products first-hand; people need to be able to exercise alternative behavioral choices, learn through experience, and re-shape their attitudes in response to the experiences. It should, however, be recognized that a sizable portion of the sample was also allocated to segments where attitudes affect behaviors; hence programs that aim to change attitudes should not be discontinued, particularly for more mature segments of the population who may be rather set in their ways and formed habits that are difficult to break. To make different campaigns work most effectively, they need to be targeted to the appropriate segments depending on the causal structures that they follow.

## INTRODUCTION

This project is concerned with exploring the relationship between attitudes, perceptions, and values on the one hand and behavioral choices on the other hand. There is a vast body of literature in a number of disciplines that has clearly demonstrated a strong inter-dependent relationship between attitudes and behaviors (Wicker, 1969; Norman, 1975; Fishbein and Ajzen, 2010; Ahn and Back, 2018). In the transportation context, attitudes about various transportation options as well as personality traits that describe the innate proclivities and preferences of the individual are likely to be strongly associated with residential and work place location choices (Cao et al., 2010; Bhat, 2015a, Ettema and Nieuwenhuis, 2017), mode choice (Heinen et al., 2011; He and Thøgersen, 2017), parking choice (Soto et al., 2018), vehicle ownership and type choice (Acker et al., 2014; Choo and Mokhtarian, 2004), activity engagement and time use patterns (Archer et al., 2013; Frei et al., 2015), and willingness to participate in the sharing economy and adopt new technologies (Astroza et al., 2017; Lavieri et al., 2018; Egbue and Long, 2012; Alemi et al., 2018).

The question that motivates this research is: Do “attitudes affect behavioral choices” or “do experiences obtained through the exercise of behavioral choices shape attitudes”? A number of studies have utilized attitudinal variables and factors as explanatory variables to explain travel choices and behaviors (Ory and Mokhtarian, 2005; Seraj et al., 2012; Heinen et al., 2013; Bhat et al., 2016; Chen et al., 2017). These variables are combined with the usual socio-economic and demographic variables, built environment variables, and variables that describe the options in the choice set to predict behaviors. In most, if not all instances, these studies have reported that attitudinal variables contribute significantly to explaining the choice behaviors of interest.

More recently, however, a growing body of literature reports that the directionality of the relationship between attitudes and behaviors is actually one in which behaviors shape attitudes (Kroesen et al., 2017; Kroesen and Chorus, 2018). According to these studies, contrary to assumptions embedded in most models, behavior influences attitudes. These studies suggest that, when there is dissonance (inconsistency) between attitudes and behaviors, people are more prone to adjust their attitudes to align with behaviors as opposed to adjusting their behaviors to align with attitudes.

While attitudes and behaviors mutually influence each other over time (Kroesen et al., 2017), and attitudes themselves may change as more information becomes available (Sheela and Mannering, 2019), the question as to whether *attitudes affect behaviors* or *behaviors affect attitudes* at any point in time remains an intriguing one with very important implications for transportation demand forecasting and the design and implementation of policy interventions aimed at shaping behaviors. If it is true that behaviors affect attitudes (rather than the reverse), then information campaigns and strategies aimed at reshaping attitudes may not have the desired and intended effects. Policy interventions would need to directly target behavioral choices by providing individuals the opportunities to obtain alternative experiences first-hand by actually trying new and different mobility options; alternative behavioral experiences would then bring about changes in attitudes that would further reinforce desirable behaviors as individuals adjust their attitudes to reduce dissonance (Kroesen et al., 2017).

While previous literature has often characterized a uni-directional relationship between attitudes and behaviors, there is significant evidence of the existence of a bi-directional relationship as well (Dobson et al., 1978, Kroesen et al., 2017; Kroesen and Chorus, 2018). This study treats both attitudinal variables and behavioral choice variables as endogenous in nature, thus recognizing endogeneity associated with estimating relationships between these dimensions of interest. Treating both attitudes and behaviors as endogenous variables requires the specification

and estimation of joint equations model systems that accommodate error correlations, making it possible to account for the presence of correlated unobserved attributes that simultaneously affect both attitudes and behaviors.

This study aims to develop a joint equations model of attitudes and behaviors that explicitly recognizes the package nature of the relationship among them. However, unlike previous studies, this research effort explicitly recognizes that there may be population heterogeneity with respect to the nature of the relationship between attitudes and behaviors. While undoubtedly mutually reinforcing, attitudes may influence behaviors for some folks and behavioral choices may affect attitudes for others *at a specific cross-section in time*. A multitude of directional relationships between attitudes and behaviors may exist in the population and it would be of interest to determine the extent or degree to which each of the directional relationships is prevalent in the population *at a specific cross-section in time*. By determining the degree to which each relationship exists in the population, and the characteristics of each market segment (in terms of socio-economic and demographic characteristics, for example), it would be possible to design policy interventions, behavioral experiences, and information campaigns that are appropriately targeted and implemented to achieve desired outcomes.

Because the segments in the population are not known a priori, they are considered latent and determined endogenously within a joint modeling framework. Thus, the model estimated in this study takes the form of a joint equations model system with latent segmentation, similar to that presented in Astroza et al. (2018). The model system includes a model component that endogenously assigns individuals to different causal segments, and this component is coupled with a simultaneous equations model component that relates attitudes and behaviors to one another in a manner consistent with the latent segment to which the behavioral unit has been probabilistically assigned. This methodology makes it possible to identify the characteristics of the subgroups that predominantly depict alternative causal structures.

The model system in this study is estimated on a data set derived from the 2014 Who's On Board Mobility Attitudes Survey conducted in the United States. In addition to an extensive battery of attitudinal variables, the survey includes information about people's behavioral choices including use of various modes of transportation, residential location type choice, and car ownership. This particular study examines the nature of the relationships between *attitudes toward transit* and two behavioral choice variables, namely, *residential location choice* and *frequency of use of transit*. By considering multiple behavioral dimensions, this study sheds light on the extent to which attitudes affect behavior (or vice versa) in the context of different behavioral choices, and identifies the relative presence of different latent segments (following different decision structures) in the population.

The remainder of this report is organized as follows. The next section offers a description of the data. The methodology is presented in the third section, model estimation results are presented in the fourth section, and the description of the latent segments is presented in the fifth section. Concluding thoughts are offered in the sixth and final section.

## **DATA DESCRIPTION**

The data set used in this study is derived from the 2014 Who's On Board Mobility Attitudes Survey (Transit Center, 2014), an online survey administered to a sample residing in 46 diverse metropolitan areas in the United States. The data set includes information for 11,842 respondents who responded to the survey. After filtering records for missing data, 9,600 observations were retained for analysis and model estimation. Table 1 presents a socio-economic and demographic

profile of the sample.

Overall, the sample provides the richness of variation and diversity of information necessary to undertake a study of this nature. Among individual characteristics, the sample has a slightly higher proportion of women. About one-fifth of the respondents in this sample are 65 years and above and more than one-half of the sample has an educational attainment of college graduate or higher. About 40 percent of the sample is employed full-time, while another 12.5 percent are employed part-time. The sample spends a fair amount of time online, with 34 percent indicating that they spend 4-8 hours online per day while five percent of the sample indicated an hour or few hours per week.

Among household attributes (the right column of Table 1), just about 20 percent of the sample has household income less than \$35,000, while 23.7 percent of the sample has household income greater than or equal to \$100,000. Just about 38 percent of the sample reports household sizes of three or more, and nearly 70 percent of the sample resides in detached housing units – which is consistent with the statistic that 61 percent of the sample resides in housing units owned by the household. With respect to transit richness, 61 percent of the sample reports residing in cities that may be characterized as transit progressive (Transit Center, 2014), i.e., cities where there is a substantial presence of transit modes. Only four percent of the sample resides in households with zero vehicles, and 25 percent of the sample reported residing in households with no workers (consistent with the age distribution noted earlier). About 40 percent of the sample indicated that the distance to the nearest transit station is less than 0.5 mile, while 38.6 percent reported that the nearest transit station is more than one mile from the residence. The sample is well distributed across the country, with the largest proportion (23.9 percent) drawn from the West Coast.

Among endogenous variables (left column bottom of Table 1), urban dwellers account for 27.8 percent of the sample. Another 32.4 percent of the sample resides in suburban and small town locations that have mixed land use; the remaining 39.8 percent reside in suburban and small town/rural locations that would not be characterized as having mixed land use. Just about one-half of the sample reports that they never use transit at all even though it is available. Seventeen percent report using transit at least once per week.

The third endogenous variable of interest in this study is the *attitudes towards transit (transit proclivity)*. This endogenous variable constitutes a factor derived by conducting a factor analysis on 10 attitudinal statements in the survey data set. These attitudinal statements pertain to feelings about transit and are therefore used to derive a transit proclivity or propensity factor. Table 2 presents the attitudinal statements, the percent of the sample agreeing or disagreeing with each statement, and the factor loadings. After a number of trials, it was found that three of the statements had insignificant factor loadings, and hence the final factor was based on seven of the ten attitudinal statements. The loadings are intuitive and suggest that the factor represents a propensity or proclivity towards using transit as a mode of transportation. The results of the factor analysis were used to compute factor scores for each individual in the sample. This continuous factor score was used in the model estimation effort to retain the variation in transit proclivity represented by the factor. This continuous factor score does not have a specific underlying scale, but simply represents the range of lower and higher positive attitudes towards transit. Thus, we do not show any specific descriptive statistics for this variable in Table 1.



**TABLE 1 Socio-economic and demographic characteristic of the sample (N=9600)**

<b>Individual Characteristics</b>		<b>Household Characteristics</b>	
<i>Exogenous Variables</i>	Value (%)	<i>Exogenous Variables</i>	Value (%)
<i>Gender</i>		<i>Household income</i>	
Female	53.5	< \$25,000	11.1
Male	46.5	\$25,000 to \$34,999	9.8
<i>Age category</i>		\$35,000 to \$49,999	14.2
16-17 years	0.2	\$50,000 to \$74,999	22.9
18-24 years	17.2	\$75,000 to \$99,999	18.3
25-34 years	22.8	≥ \$100,000	23.7
35-54 years	19.2	<i>Household size</i>	
55-64 years	19.2	One	17.9
65 years and above	21.4	Two	44.2
<i>Education attainment</i>		Three and more	37.9
High school or less	17.0	<i>Housing unit type</i>	
Technical/training beyond high school	5.1	Detached Housing	69.4
Some college	26.2	Apartment housing	28.2
College graduate or higher	51.7	Others	2.4
<i>Employment status</i>		<i>Home ownership</i>	
Employed full-time	40.4	Rent	28.1
Employed Part-Time	12.5	Own	61.0
Not Employed	6.1	Living family rent-free or other	10.9
Other (student, retired, homemaker)	41.0	<i>Presence of kids</i>	
<i>Time spent online</i>		Presence of kids 0-4 years	8.0
More than 8 hours per day	18.0	Presence of kids 5-15 years	10.0
4 to 8 hours per day	34.0	Presence of kids 16-18 years	4.0
1 to 4 hours per day	42.0	<i>Transit Richness</i>	
A few hours per week or an hour per week	5.0	Deficient	38.6
<b><i>Endogenous Variables</i></b>		Progressive	61.4
<i>Residential location choice (RLC)</i>		<i>Vehicle ownership</i>	
Urban	27.8	Zero	4.0
Suburban and small town – mixed land use	32.4	One	30.7
Other suburban and small town + rural	39.8	Two	42.6
<i>Frequency of transit use (FTU)</i>		Three or more	22.8
Frequent: once per week or more	17.0	<i>Number of employed persons</i>	
Infrequent: less than once per week	32.6	Zero	25.0
Never (but has available)	50.4	One	35.3
<i>Attitudes Toward Transit (ATT) – Factor Score</i>		Two or more	39.7
Scale-less underlying continuous variable		<i>Distance from home to nearest transit station</i>	
		Less than 0.5 mile	40.0
		0.5-1 mile	21.4
		More than 1 mile	38.6
		<i>Geographic Region</i>	
		Northeast	16.3
		South	18.5
		West/Southwest	19.0
		West Coast	23.9
		Midwest	22.3

**TABLE 2 Transit Attitudes and Factor Loadings**

Attitudinal Statements	Agree (%)	Neutral (%)	Disagree (%)	Factor Loading (Std Error)
I like the idea of doing something good for the environment when I ride transit	50.9	39.8	9.3	1.00 (base)
I am not sure I know how to do all the things to make the bus or train trip work	39.3	26.2	34.5	—
I worry about crime or other disturbing behavior on public forms of transportation	51.8	28.0	20.2	—
I feel safe when riding public transportation	39.4	41.2	19.4	0.972 (0.037)
Public transit does not go where I need to go	52.2	27.1	20.7	-0.298 (0.023)
Riding transit is less stressful than driving on congested highways	50.8	29.2	20.0	1.275 (0.049)
It would be easier for me to use transit more if I were not so concerned about traveling with people I do not know	24.3	25.4	50.3	0.378 (0.025)
My family and friends typically use public transportation	17.2	20.3	62.5	1.483 (0.062)
I like to make productive use of my time when I travel	62.8	29.5	7.7	—
I sometimes take public transit to avoid traffic congestion	31.0	20.0	49.0	2.553 (0.135)

The three endogenous variables considered in this study are as follows:

- Residential Location Choice (RLC): Three categories – Urban; Suburban + Small Town with Mixed Land Use; and Suburban + Small Town/Rural without Mixed Land Use
- Frequency of Transit Use (FTU): Three categories – Frequent (once or more per week); Infrequent (less than once per week); and Never
- Attitude Towards Transit (ATT): Continuous factor score

The three endogenous variables may be related in six possible different causal structures. It is entirely possible that all six causal structures are prevalent in the population, i.e., there is at least some fraction of the population following each of the causal structures. However, the estimation of a joint simultaneous equations model system that involves three mixed endogenous variables and six different latent segments is computationally challenging, and the interpretation of results obtained from such a large-scale model estimation effort may prove difficult. In order to reduce the size of the problem, four plausible causal structures (and hence, four possible latent segments) are considered and included within the scope of this study. Because the intent of this study is to unravel relationships between attitudes and behaviors, the four causal structures where attitudes (ATT) come first in the causal hierarchy and attitudes (ATT) come last in the causal hierarchy are considered. The two causal structures where attitudes (ATT) act as a mediator between residential location choice (RLC) and frequency of transit use (FTU) are omitted from the scope of this modeling exercise. The four causal structures may be depicted as follows:

<b>Structure 1</b>	RLC (R) RLC $\rightarrow$ FTU FTU + RLC $\rightarrow$ ATT	<b>Structure 3</b>	ATT (A) ATT $\rightarrow$ RLC RLC + ATT $\rightarrow$ FTU
<b>Structure 2</b>	FTU (F) FTU $\rightarrow$ RLC RLC + FTU $\rightarrow$ ATT	<b>Structure 4</b>	ATT (A) ATT $\rightarrow$ FTU FTU + ATT $\rightarrow$ RLC

*Note:* RLC = Residential Location Choice  
FTU = Frequency of Transit Use  
ATT = Attitude towards transit (Transit propensity)

The first two structures are those where behaviors affect attitudes towards transit (ATT), while the latter two structures are those where attitudes towards transit (ATT) influence behaviors. The relationship between residential location choice (RLC) and frequency of transit use (FTU) may go either way. On the one hand, residential location may engender transit use; on the other hand, the frequency of transit use may motivate an individual to seek a residential location that supports the level of transit use undertaken and desired by an individual.

## MODELING METHODOLOGY

In the case where both attitudinal and behavioral choice variables are represented as continuous variables, it is econometrically feasible to identify and estimate bidirectional causal models – thus enabling an explicit portrayal of the mutually reinforcing relationship that exists between attitudes and behaviors. However, when the behavioral choice variables of interest are not continuous (and, are often discrete in the context of travel behavior), then a bidirectional causal model is not identified, and identification restrictions must be imposed for logical consistency purposes (Pendyala and Bhat, 2004). This necessitates the estimation of recursive joint equations model systems when considering multiple endogenous variables of different types. In other words, when dealing with discrete choice variables (or, more generally, limited dependent variables), the joint equations model system can reflect the influence of attitudes on behaviors *or* the influence of behaviors on attitudes, *but not both* (after accommodating for unobserved covariance effects) It should be noted, however, that the recursive joint equations model system that depicts unidirectional relationships does not necessarily imply a sequential ordering in the decision mechanism. By estimating both attitudes and behaviors in a joint equations framework, while recognizing the presence of unobserved correlated attributes that affect multiple dimensions, the system of equations portrays *jointness* in the determination of attitudes and behaviors while recognizing that one dimension influences the other. A more detailed discussion about the important distinction between sequentiality and simultaneity in the choice processes at play may be found in Astroza et al. (2018).

Another important note here is that inference about causality is inextricably tied to observations of individuals and their choices over time. In other words, longitudinal data is very desirable for any effort aimed at unraveling and identifying causal relationships and structures. Generally, cause-and-effect patterns play out over time, involve leads and lags, and are inherently dynamic in nature. Although the profession has seen the collection of longitudinal panel survey data on occasion, the prevailing norm continues to be the collection of (repeated) cross-sectional

data from a sample of the population. In the absence of true longitudinal panel data, it is virtually impossible to unravel cause-and-effect relationships that transpire over time. Therefore, the analysis in this study should be construed as depicting contemporaneous causation, i.e., the causal relationships that exist at a single snapshot in time. Individuals are making a bundle of choices jointly (involving attitudes, residential location choice, and frequency of transit use), and the causal relationships depict the nature and direction of influence among the endogenous variables and capture the reasoning or logical flow of thought that an individual may exercise. For example, an individual may reason *at any cross-section in time* that he or she likes the idea of riding transit (positive attitude) and therefore resides in a residential location that facilitates a high level of transit use. The logical flow of relationships among the dimensions represents a contemporaneous causation, the notion that “behavior is caused at the moment of its occurrence by all the influences that are present in the individual at that moment” (Lewin, 1936).

### **The Joint Model of Behavioral Choices and Attitudinal Factor**

The remainder of this section describes in detail the model formulation adopted in this study. Consider an individual  $q$  ( $q=1, 2, 3, \dots, Q$ ) facing a multi-dimensional choice system comprised by one continuous variable (attitudes towards transit), one ordinal variable (frequency of transit use), and one nominal variable (residential location). The discussion starts with the formulation for each type of variable, and then presents the structure and estimation procedure for the multi-dimensional system. For this section, assume that the individual belongs to a specific segment  $h$ .

Let  $y_{qh}$  be the continuous variable (corresponding to the attitudes towards transit score) for individual  $q$  given that he/she belongs to segment  $h$ . Let  $y_{qh} = \gamma'_h s_{qh} + \eta_{qh}$  in the usual linear regression fashion, where  $s_{qh}$  is a column vector of exogenous attributes as well as possibly the observed values of other endogenous variables,  $\gamma_h$  is a column vector of corresponding coefficients, and  $\eta_{qh}$  is a normal standard scalar error term (the variance of  $\eta_{qh}$  is normalized to one for all segments  $h$ , because, though  $y_{qh}$  is a continuous variable, it represents a scale-less latent factor score in our empirical analysis that is constructed from other observed indicators). Note that some elements of  $\gamma_h$  can be zero for some of the exogenous variables, indicating that the corresponding exogenous variables do not impact choice-making in segment  $h$ . Further, because latent segmentation is used as a way to introduce, across the segments, heterogeneity in the recursive effects among the endogenous variables,  $\gamma_h$  will necessarily be zero on some of the endogenous variables within each segment (see Astroza et al., 2018 for a detailed explanation).

Let there be one ordinal variable for the individuals. In the empirical context of the current study, the ordinal variable corresponds to the frequency of transit use and has three different levels: never, infrequent (less than once per week), and frequent (once per week or more). Let the ordinal index for the individual given that he/she belongs to segment  $h$  be  $j_{qh}$  ( $j_l = 1, 2, 3$ ) and let  $n_q$  be the actual observed value. Then, assume an ordered-response probit (ORP) formulation as:  $y_{qh}^* = \phi'_h z_{qh} + \xi_{qh}$ ,  $j_{qh} = n_q$  if  $\psi_{h,n_q-1} < y_{qh}^* < \psi_{h,n_q}$ ,  $j_{qh} \in \{1, 2, 3\}$ , where  $z_{qh}$  is a column vector of exogenous attributes as well as possibly the observed values of other endogenous variables,  $\phi_h$  is a column vector of corresponding coefficients, and  $\xi_{qh}$  is a standard normal scalar error term. Similar to the case of the continuous variable,  $\phi_h$  can be zero on some of the endogenous variables

within each segment (structural heterogeneity). For identification conditions, set  $\psi_{h,0} = -\infty$ ,  $\psi_{h,3} = +\infty$ , and  $\psi_{h,1} = 0$ . Only one threshold,  $\psi_{h,2}$ , is then estimated.

Let there be one nominal (unordered-response) variable for the individuals. In the empirical context of the current study, the nominal variable is residential location, which has  $I=3$  alternatives (shown in Table 1). Using the typical utility maximizing framework, it is possible to write the utility for alternative  $i$  for individual  $q$  given that he/she belongs to segment  $h$  as:  $U_{qih} = \beta'_h \mathbf{x}_{qih} + \varepsilon_{qih}$ , where  $\mathbf{x}_{qih}$  is a column vector of exogenous attributes as well as possibly the observed values of other endogenous variables,  $\beta_h$  is a column vector of corresponding coefficients, and  $\varepsilon_{qih}$  is a normal scalar error term. Let the variance-covariance matrix of the vertically stacked vector of errors  $\boldsymbol{\varepsilon}_{qh} = [(\varepsilon_{q1h}, \varepsilon_{q2h}, \varepsilon_{q3h})']$  be  $\Lambda_h$ . Again,  $\beta_h$  can be zero on some of the endogenous variables within each segment. Define  $\mathbf{U}_{qh} = (U_{qh1}, U_{qh2}, U_{qh3})'$ . Several important identification issues need to be addressed for the nominal variable. First, one of the alternatives has to be used as the base when introducing alternative-specific constants and variables that do not vary across the alternatives. This is because only utility differences matter in terms of the nominal variable choice. For future reference, let  $\mathbf{u}_{qh}$  be the vector of utility differences with respect to the chosen alternative for the nominal variable and let  $\tilde{\Lambda}_{qh}$  be the corresponding covariance matrix. Also, because only utility differences matter, only the covariance matrix of the error differences is estimable. Taking the difference with respect to the first alternative, only the elements of the covariance matrix  $\tilde{\Lambda}_h$  of  $\tilde{\mathbf{u}}_{qh} = (U_{qh2} - U_{qh1}, U_{qh3} - U_{qh1})$  is estimable.

The jointness across the different types of dependent variables may be specified by writing the covariance matrix of the  $[4 \times 1]$  vector  $\tilde{\mathbf{y}}_{qh} = (\mathbf{u}_{qh}, y_{qh}^*, y_{qh})$  as:

$$\text{Var}(\tilde{\mathbf{y}}_{qh}) = \tilde{\Omega}_{qh} = \begin{bmatrix} \tilde{\Lambda}_{qh} & \boldsymbol{\Sigma}_{uy^*h} & \boldsymbol{\Sigma}_{uyh} \\ \boldsymbol{\Sigma}'_{uy^*h} & 1 & \sigma_{y^*yh} \\ \boldsymbol{\Sigma}'_{uyh} & \sigma_{y^*yh} & 1 \end{bmatrix}, \quad (1)$$

where  $\boldsymbol{\Sigma}_{uy^*h}$  is a  $2 \times 1$  vector capturing covariance effects between the  $\mathbf{u}_{qh}$  vector and the scalar  $y_{qh}^*$ ,  $\boldsymbol{\Sigma}_{uyh}$  is a  $2 \times 1$  vector capturing covariance effects between the  $\mathbf{u}_{qh}$  vector and the scalar  $y_{qh}$ , and  $\boldsymbol{\Sigma}_{y^*yh}$  is the covariance between  $y_{qh}^*$  and  $y_{qh}$ . The covariance matrix in Equation (1) needs to be mapped appropriately in terms of a corresponding covariance matrix (say  $\Omega_h$ ) for the vector  $(\mathbf{U}_{qh}, y_{qh}^*, y_{qh})$ , with appropriate identification conditions imposed on  $\Omega_h$  to recognize that only utility differences matter for the nominal variable. The approach to do so is discussed in detail in Bhat (2015b). This needs some additional notations and discussion, which are omitted in the interest of brevity.

Next, let  $\boldsymbol{\theta}_h$  be the collection of parameters to be estimated:  $\boldsymbol{\theta}_h = [\boldsymbol{\gamma}'_h, \boldsymbol{\phi}'_h, \psi_{h,2}, \boldsymbol{\beta}'_h; \text{Vech}(\Omega_h)]$ , where  $\text{Vech}(\Omega_h)$  represents the vector of estimable parameters of  $\Omega_h$ . Then the likelihood function for the individual  $q$  given that he/she belongs to segment  $h$  may be written as:

$$\begin{aligned}
L_q(\boldsymbol{\theta}_h) &= \phi_1(y_{qh} - \boldsymbol{\gamma}'_h \mathbf{s}_{qh}) \times \Pr \left[ \tilde{\boldsymbol{\psi}}_{low,qh} \leq \tilde{\mathbf{u}}_{qh} \leq \tilde{\boldsymbol{\psi}}_{up,qh} \right], \\
&= \phi_1(y_{qh} - \boldsymbol{\gamma}'_h \mathbf{s}_{qh}) \times \int_{D_{\tilde{\mathbf{u}}_{qh}}} \phi_3(\tilde{\mathbf{u}}_{qh} | \tilde{\boldsymbol{\Omega}}_{qh}) d\tilde{\mathbf{u}}_{qh},
\end{aligned} \tag{2}$$

where  $\tilde{\mathbf{u}}_{qh} = (\mathbf{u}_{qh}, y_{qh}^*)'$ , the integration domain for the probability  $D_{\tilde{\mathbf{u}}_{qh}} = \{\tilde{\mathbf{u}}_{qh} : \tilde{\boldsymbol{\psi}}_{low,qh} \leq \tilde{\mathbf{u}}_{qh} \leq \tilde{\boldsymbol{\psi}}_{up,qh}\}$  is simply the multivariate region of the elements of the  $\tilde{\mathbf{u}}_{qh}$  vector determined by the range  $(-\infty, 0)$  for the nominal variable and by the observed outcome of the ordinal variable. That is,  $\tilde{\boldsymbol{\psi}}_{low,qh} = (-\infty, -\infty, \psi_{h,n_q-1})$  and  $\tilde{\boldsymbol{\psi}}_{up,qh} = (0, 0, \psi_{h,n_q})$ , and  $\phi_R(\cdot)$  is the multi-variate normal density function of dimension  $R$ .

### Segmentation Model

The derivation thus far is based on the notion that individual  $q$  belongs to a single segment  $h$ . However, the actual assignment of individual  $q$  to a specific segment is not observed; but 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:

$$\pi_{qh} = \frac{\exp(\boldsymbol{\mu}'_h \mathbf{w}_q)}{\sum_{j=1}^H \exp(\boldsymbol{\mu}'_j \mathbf{w}_q)}, \tag{3}$$

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

$$L_q(\boldsymbol{\theta}) = \sum_{h=1}^H \pi_{qh} [L_q(\boldsymbol{\theta}_h) | q \in \text{segment } h], \tag{4}$$

and the overall likelihood function is then given as:

$$L(\boldsymbol{\theta}) = \prod_q L_q(\boldsymbol{\theta}). \tag{5}$$

Typical simulation-based methods to approximate the multivariate normal cumulative distribution function in Equation 1 can prove inaccurate and time-consuming. As an alternative, the Maximum Approximate Composite Marginal Likelihood (MACML) approach (Bhat, 2011), which is a fast analytic approximation method, is used. The MACML estimator is based solely on univariate and bivariate cumulative normal distribution evaluations, regardless of the dimensionality of integration, which considerably reduces computation time compared to other simulation techniques used to evaluate multidimensional integrals. For a detailed description of the MACML approach in the specific case of a joint system of continuous, ordinal, and nominal variables, the reader is referred to Bhat (2015b).

## RESULTS

Model specifications that incorporate latent segments can prove to be computationally challenging to estimate (Astroza et al., 2018). To help facilitate the identification of good starting values for model parameters, the study employed a strategy of first estimating four different causal structures separately and independently, assuming that the entire sample constituted a single segment. The parameter estimates from these independent models were used as starting values for the full-fledged model with latent segmentation. Also, to help inform the specification of the joint model, we started with obtaining best specifications for the individual models corresponding to Residential Location Choice (RLC), Frequency of Transit Use (FTU), and Attitude Towards Transit (ATT), and using these to inform the joint modeling with multiple segments.

Models with different numbers of latent segments were estimated and compared. It was found that the model with four latent segments (i.e., all four causal structures considered in this study) offered the best fit compared to models with one, two, or three latent segments. For the four-segment model, the log-likelihood value at convergence is  $-164,377.29$  and, with 242 parameters, the Bayesian Information Criterion (BIC) is 165,486.8; the corresponding BIC values for the one, two, and three segment models are larger at 166,257.3, 165,932.5, and 165,599.2 respectively. Just to explore further, a five-segment model was also estimated and evaluated (by adding one of the causal structures in which attitudes act as a mediator between residential location choice and frequency of transit use), and the fit was found to be inferior to the four segment model (BIC for the five segment model was 165,525.1). As such, the remainder of this section is dedicated to discussing results for the four-segment model.

In the interest of brevity, the joint equations model estimation results for each of the four causal structures are not presented in full. Rather, complete estimation results are presented for one causal structure for illustrative purposes (Table 3). In general, the effects of exogenous variables on endogenous variables do not vary by causal structure, and there is no reason that they should. The exogenous variable influences are largely based on patterns of relationships within the data set and there is no reason for these relationships to vary across the causal structures considered. Indeed, an examination of the detailed model estimation results for the four causal structures shows that the exogenous variables depict similar coefficient values and signs. A brief description of the influence of various exogenous variables on the endogenous variables of interest is provided here. These relationships can be seen in Table 3.

An examination of exogenous variable influences shows that women respondents show a lower inclination to reside in urban areas relative to non-urban areas. Admittedly, this result needs to be interpreted with care, because residential locations are likely to be based on all individuals in a household. However, since the survey used here was an individual-based survey (only one individual responded per household), and this result came out to be statistically significant, we left it in to potentially reflect the notion that, at least within the group of single adult households, women prefer non-urban settings. Women respondents also are more likely than men to use transit and have a more positive attitude towards transit. Younger individuals (particularly below 35 years of age) are more likely to be urban dwellers when compared with older individuals. Older individuals (35 years or above) use transit less frequently than their younger counterparts; consistent with this finding, younger individuals below the age of 35 years are found to have a more positive attitude towards transit. College graduates are found to favor urban residential location type, as do those employed full time. Time spent online is significantly related to the endogenous variables considered in this study; those who spend more than eight hours per day online are more likely to reside in urban and suburban mix areas, show a propensity towards higher

frequency of transit use, and demonstrate a more positive attitude towards transit. It is likely that those who are technology oriented prefer transit-oriented urban lifestyles (Hong and Thakuriah, 2018).

Among household attributes, home ownership is negatively associated with urban and suburban mix residential choice and negatively associated with transit use, but positively associated with attitudes towards transit. It appears that home owners are positively disposed towards transit, and their infrequent (or non-existent) use of the service does not provide a sufficient basis to change that perspective. Lower incomes are associated with urban living and higher propensity to use transit. Individuals in larger households are less likely to favor urban residential locations and are less inclined to use transit, presumably because of the lifecycle stage and need to fulfill household obligations. This is further reinforced by the finding that the presence of children negatively impacts urban residential location choice and propensity to use transit. As expected, households with high levels of vehicle ownership (three or more vehicles) are less likely to reside in urban and suburban mix areas, depict a lower propensity to use transit, and have more negative attitudes towards transit. The exact nature of the causal relationships involving vehicle ownership is unknown and merits further investigation. Vehicle ownership is actually an endogenous mobility choice variable, but has been treated in this study as an exogenous variable for simplicity. It is entirely possible that vehicle ownership is affected by residential location choice, propensity to use transit, and transit attitudes; exploring the causal influences that shape vehicle ownership remains a task for future research efforts. Those who reside in transit progressive cities are more prone to using transit and have a more positive attitude towards transit, while those in the South region of the United States (which is generally more sprawled and auto-oriented) have a lower propensity to use transit and a have a more negative attitude towards transit. In general, all of the exogenous variable impacts are consistent with expectations and demonstrate that socio-economic and demographic variables play a significant and important role in shaping attitudes and mobility/location choices.

For each of the segments, we could not reject the hypothesis that the diagonal terms in the 2×2 covariance sub-matrix of the differenced error terms corresponding to the residential location choice alternative utilities were 1.0 and that all the off-diagonal elements in the sub-matrix were 0.5. This implies that the error terms of the residential location choice alternatives are independently and identically distributed. Assuming that the error term in the base alternative in each dimension is independent of the error terms in other dimensions, and scaling the variance of the utilities of each alternative error term in the residential location choice model to one, the implied covariance (correlation) matrix among (1) the urban residential location utility (UL), (2) the suburban/small town mix residential location utility (SUBT), (3) the propensity underlying frequency of transit use (FTU), and (4) ATT factor score is presented toward the bottom of Table 3 for causal structure 1 (we present only the lower diagonal elements because of the symmetric nature of this matrix). There are statistically significant error correlations, and we found this to be the case for every causal structure considered in this study. In general, the error correlations for the other causal structures had the same signs as those for the first causal structure in Table 3, clearly indicating that, in each segment, there is a residual association between the dependent variables not captured by the explanatory variables included in the model specification. This result justifies the use of a joint package approach to model relationships among the endogenous variables considered in this study. Not surprisingly, the positive correlation in the second column and last row of the covariance (correlation) matrix suggests that unobserved factors that increase the utility of residing in an urban area increases positive views of transit, even if these factors do



not necessarily increase the actual use of transit. A possible explanation is that a variety seeking individual (who likes to try different experiences) may like to reside in an urban location (where there is a variety of amenities in close proximity) and have a positive attitude towards alternative (a variety of) modes of transportation. The correlations in the third column suggest that unobserved factors that lead to residing in suburban and small towns reduce transit use propensity as well as positive attitudes toward transit. These results are clear evidence of unobserved residential self-selection effects (see Bhat and Guo, 2007 for a detailed discussion). Those who intrinsically (due to unobserved individual factors) do not have positive views about transit and are not very likely to use transit and are likely to locate themselves in suburbia.

Table 4 presents a summary of the endogenous variable effects, which are of interest in the context of understanding relationships among dependent variables under different causal structures (and so the estimates of all the causal structures are shown in Table 4). Note that these are “true” causal effects after “cleansing” any relationships among the endogenous variables caused by “spurious” unobserved correlation effects. In general, it can be seen that the relationships are significant and consistent with expectations, indicating that these three endogenous variables affect one another in behaviorally intuitive ways after accommodating unobserved covariances. In causal structure 1 ( $RLC \rightarrow FTU$ ;  $RLC + FTU \rightarrow ATT$ ), it is found that those in suburban and small town locations show a lower propensity to use transit. Compared to those in suburban and small town/rural areas with no mixed land use, the residents of urban and suburban mix areas have a more positive attitude towards transit (again, this is after accommodating unobserved factors that may influence these endogenous variables). Likewise, frequent and infrequent transit users have a more positive attitude towards transit than those who never use transit; between these two groups, frequent users have a more positive attitude than infrequent users. In causal structure 2 ( $FTU \rightarrow RLC$ ;  $FTU + RLC \rightarrow ATT$ ), it is found that frequent users of transit are more likely to reside in urban areas and suburban and small town areas with mixed land use areas rather than suburban and small town areas without mixed land use. Transit users also have a more positive attitude towards transit. Similarly, urban dwellers are likely to have a more positive attitude towards transit. In causal structure 3 ( $ATT \rightarrow RLC$ ;  $ATT + RLC \rightarrow FTU$ ), those with a positive attitude towards transit are more likely to favor urban and suburban mix residential locations and exhibit a greater propensity to use transit. Those residing in suburban mix locations depict a lower propensity to use transit than their counterparts in other urban and suburban/rural areas. In causal structure 4 ( $ATT \rightarrow FTU$ ;  $ATT + FTU \rightarrow RLC$ ), positive attitudes towards transit lead to a more urban and suburban mix residential location choice (relative to those residing in suburban/rural locations) and a higher propensity to use transit. Similar to indications in other causal structures, those who use transit more frequently are more likely to choose urban and suburban mix locations for residence (relative to suburban/rural locations), with this tendency being particularly high for urban locations.

**TABLE 3 Illustrative Model Estimation Results: Causal Structure 1 (RLC → FTU; RLC + FTU → ATT)**

Explanatory Variables	Residential Location Choice RLC (base: other suburban & small town + rural)				Frequency of Transit Use FTU (never, infrequent, and frequent)		Attitude Towards Transit ATT (continuous factor scores)	
	Urban dwellers		Suburban and small town mix		Coef	t-stat	Coef	t-stat
	Coef	t-stat	Coef	t-stat				
Constant	-0.779	-6.12	-0.561	-7.21	0.206	18.86	-0.653	-17.42
<b>Individual Characteristics</b>								
<i>Gender</i>								
Female	-0.193	-3.71	—	—	0.112	3.23	0.099	4.12
<i>Age category</i>								
18-24 years	0.510	4.19	—	—	—	—	0.157	5.88
25-34 years	0.294	3.94	—	—	—	—	0.111	3.21
18-34 years	—	—	0.163	2.11	—	—	—	—
35-54 years	—	—	—	—	-0.300	-5.95	—	—
55-64 years	—	—	—	—	-0.412	-6.32	—	—
65 years and above	—	—	—	—	-0.587	-7.35	—	—
<i>Education attainment</i>								
College graduate or higher	0.189	2.63	—	—	—	—	—	—
<i>Employment Status</i>								
Employed full-time	0.265	4.71	-0.105	-4.19	—	—	—	—
<i>Time spent online</i>								
More than 8 hours per day	0.322	3.28	0.224	2.96	0.702	3.29	0.061	2.11
<b>Household Characteristics</b>								
<i>Home ownership</i>								
Own	-0.642	-5.39	-0.206	-3.12	-0.131	-4.12	0.075	2.42
<i>Household income</i>								
Less than \$35,000	0.203	3.14	-0.241	-4.51	0.073	2.12	—	—
More than \$75,000	—	—	—	—	—	—	—	—
<i>Household size</i>								
Two or more	-0.245	-3.21	—	—	-0.131	-3.78	—	—
<i>Presence of children</i>								
Presence of children 0-4 years	-0.110	-2.11	-0.125	-2.02	-0.102	-4.12	—	—
Presence of children 0-15 years	—	—	—	—	—	—	0.124	5.63

**TABLE 3 Illustrative Model Estimation Results: Causal Structure 1 (RLC → FTU; RLC + FTU → ATT) (continued)**

Explanatory Variables	Residential Location Choice RLC (base: other suburban & small town + rural)				Frequency of Transit Use FTU (never, infrequent, and frequent)		Attitude Towards Transit ATT (continuous factor scores)	
	Urban dwellers		Suburban and small town mix		Coef	t-stat	Coef	t-stat
	Coef	t-stat	Coef	t-stat				
<b>Household Characteristics</b>								
<i>Vehicle ownership</i> Three or more	-0.710	-8.22	-0.321	-6.10	-0.239	-3.29	-0.104	-4.62
<b>Location Characteristics</b>								
<i>Lives in Transit Rich City</i> Progressive	—	—	—	—	0.412	9.55	0.086	3.06
<i>Region</i> South	—	—	—	—	-0.183	-4.90	-0.098	-3.10
<b>Threshold Parameter</b>	—	—	—	—	1.217	19.96	—	—
<b>Correlation Between Error Terms</b>								
	URB	SUB	FTU	ATT				
URB	1.000				URB: Urban residence utility SUB: Suburban and small town mix utility			
SUB	0.000	1.000						
FTU	0.000	-0.167	1.000					
ATT	0.121	-0.098	0.221	1.000				

**TABLE 4 Relationships Among Endogenous Variables for the Four Causal Structures/Segments**

Variables	Residential location choice (base: suburban and small town+rural)				Frequency of Transit Use (never, infrequent, and frequent)		Attitude Towards Transit (continuous factor score)	
	Urban Dwellers		Suburban & Small Town Mix		Coef	t-stat	Coef	t-stat
	Coef	t-stat	Coef	t-stat				
<i>Segment 1 (RLC → FTU; RLC+FTU → ATT)</i>								
<i>Residential Location Choice (base: suburban, small town +rural area)</i>								
Urban dwellers	—	—	—	—	—	—	0.146	4.98
Suburban and small town mix	—	—	—	—	-0.089	-3.34	0.078	2.87
<i>Frequency of Transit Use</i>								
Frequent (≥ once per week)	—	—	—	—	—	—	1.298	21.43
Infrequent (< once per week)	—	—	—	—	—	—	0.653	25.31
<i>Segment 2 (FTU → RLC; FTU+RLC → ATT)</i>								
<i>Frequency of Transit Use</i>								
Frequent (≥ once per week)	1.122	4.10	0.308	3.55	—	—	1.311	18.32
Infrequent (< once per week)	0.462	3.92	0.237	4.21	—	—	0.703	22.01
<i>Residential Location Choice (base: suburban/small town + rural area)</i>								
Urban dwellers	—	—	—	—	—	—	0.127	3.22
Suburban and small town mix	—	—	—	—	—	—	0.083	2.04
<i>Segment 3 (ATT → RLC; ATT+RLC → FTU)</i>								
Attitude Towards Transit	0.312	4.62	0.119	5.32	0.624	19.05	—	—
<i>Residential Location Choice (base: suburban/small town + rural area)</i>								
Urban dwellers	—	—	—	—	—	—	—	—
Suburban and small town mix	—	—	—	—	-0.110	-3.46	—	—
<i>Segment 4 (ATT → FTU; ATT+FTU → RLC)</i>								
Attitude Towards Transit	0.156	4.63	0.0799	2.63	0.631	24.12	—	—
<i>Frequency of Transit Use</i>								
Frequent (≥ once per week)	—	—	—	—	—	—	—	—
Infrequent (< once per week)	0.347	4.32	0.180	2.98	—	—	—	—

*Goodness of Fit Statistics (Four-Segment Model System)*

Log likelihood at convergence, L(β) = -164,377.29 (242 parameters); Log likelihood with constants, L(c) = -217,269.31

Log likelihood with no constants, L(0) = -278,366.45; Adjusted ρ<sup>2</sup>(c) = 0.2424; Adjusted ρ<sup>2</sup>(0) = 0.4086

## SIZE AND CHARACTERISTICS OF LATENT SEGMENTS

This section presents information about the latent segments in the population. As posited earlier in this study, it is hypothesized that different segments in the population follow different causal structures in their contemporaneous decision-making processes. This section offers information about the size and characteristics of the latent segments to determine the extent to which behaviors affect attitudes or attitudes affect behaviors in the survey sample of this study. Table 5 presents the results of the latent segmentation membership model.

**TABLE 5 Latent Segmentation Model**

Segmentation Variables	Segment 1 (base)	Segment 2 Coef (t-stat)	Segment 3 Coef (t-stat)	Segment 4 Coef (t-stat)
Constant	—	-0.289 (-6.23)	-0.302 (-8.11)	-0.561 (-9.32)
Age 18-34 years	—	-0.134 (-2.88)	-0.309 (-3.01)	-0.481 (-3.76)
Age 35-64 years	—	-0.163 (-2.11)	-0.235 (-2.34)	-0.432 (-3.08)
Gender: Female	—	-0.193 (-3.53)	0.187 (5.03)	0.059 (3.32)
Lives in transit rich city	—	—	-0.205 (-3.31)	-0.223 (-4.10)
College graduate or higher	—	—	-0.211 (-2.22)	-0.231 (-2.57)
Distance to nearest transit station < 0.5 mile	—	0.103 (4.12)	-0.254 (-3.21)	-0.102 (-2.18)
Hhld Income > \$75,000	—	-0.131 (-3.25)	0.138 (5.19)	0.064 (2.74)
Segment Size	41% 3,936	25% 2,400	21% 2,016	13% 1,248

Segment 1 Causal Structure: RLC (R) → FTU (F); RLC (R) + FTU (F) → ATT (A)  
 Segment 2 Causal Structure: FTU (F) → RLC (R); FTU (F) + RLC (R) → ATT (A)  
 Segment 3 Casual Structure: ATT (A) → RLC (R); ATT (A) + RLC (R) → FTU (F)  
 Segment 4 Causal Structure: ATT (A) → FTU (F); ATT (A) + FTU (F) → RLC (R)

The model offers a first glimpse into the profile of the segments. In general, it appears that individuals are more likely to belong to the first segment in which residential location choice affects frequency of transit use, and these two behavioral choices together impact attitudes (see the last row of Table 5 for the segment size information). It is found that 41 percent of the sample is assigned to this first segment, with all other segments substantially smaller in size (the size of each segment may be determined based on the procedure discussed in Bhat (1997)). The second largest segment is the second segment in which frequency of transit use affects residential location choice, and these two choice behaviors together shape attitudes. In other words, the two causal structures (the first and second) in which behaviors shape attitudes account for two-thirds of the sample. The other one-third of the sample is collectively assigned to the other causal segments (the third and fourth segments) in which attitudes affect behaviors. It appears that, in the context of this sample (which is a rather large sample drawn from diverse areas in the United States), behaviors influence attitudes for a majority of the respondents, consistent with recent evidence in the literature (Kroesen et al., 2017) which suggests that people adjust their attitudes according to behavioral choices and experiences in an effort to reduce cognitive dissonance.

The results of the effects of exogenous variables in Table 5 indicate that individuals younger than 65 years of age are increasingly less likely to belong to the second, third, or fourth segments (see the progression of coefficients from left to right for the two age groups). Women, however, are more likely to belong to the third and fourth causal segments than the first two causal segments. Compared to men, women appear to be more set with respect to their attitudes and likely

to exhibit behavioral choices according to their attitudes. On the other hand, those who live in transit-rich cities and those who are college graduates are more likely to belong to the first two segments in which behaviors shape attitudes (notice the negative signs on these variables associated with the third and fourth segments). Those who live close to a transit station are also more likely to belong to the first two segments; perhaps their attitudes are shaped by the proximity to transit that engenders greater level of transit use. On the other hand, higher income individuals are more likely to belong to the third and fourth segments where attitudes shape behaviors. It is possible that individuals who have reached this level of income have opinions and attitudes that have matured, and also have the wealth firepower to actually take their attitudes/opinions to fruition. That is, there is perhaps less presence of cognitive dissonance for such individuals than their lower income counterparts (lower income individuals may be less able to get out of a less-than-desirable situation, and may change their attitudes as a coping mechanism).

It is interesting to note that, within the two distinct sets of causal structures (one where behavior shapes attitudes, and the other set where attitudes influence behavior), the causal structure that is more dominant is the one where residential location choice affects frequency of transit use. In other words, the longer term choice (residential location) influences the shorter term mode use decision (frequency of transit use). This type of relationship is quite consistent with that often invoked in integrated models of transport and land use where land use choices are often considered higher in the hierarchy and assumed to influence shorter term activity-travel choices. However, it is also found that the sizes of the segments in the causal structures where frequency of transit use influences residential location choice are not trivial. These segments (Segments 2 and 4) are quite sizable in their own right. Individuals in these latent segments appear to be choosing a residential location choice that is conducive to the level of transit use that they undertake. Overall, it can be concluded that there is considerable *structural heterogeneity* in the sample, and any travel forecast that assumes the same causal structure for the entire sample is likely to yield erroneous estimates of impacts of alternative transport policies and investments.

Table 6 presents a detailed overview of the profile of the various latent segments in the sample. The left half of the table shows the percent of individuals in each latent segment that belong to a socio-economic group; the right half of the table shows the percent of individuals in each socio-economic group that is assigned to each of the latent segments. The percent of individuals in each socio-economic group that belongs to a specific segment does not vary greatly. This is a reflection of the strong effect of the constants in Table 5 in determining segment membership, relative to other observed exogenous variables. This suggests that there is still room for improvement in determining the factors that influence segment membership, which may be explored in future studies with a more exhaustive set of demographic variables as well as built environment contextual variables. However, while the latent segments may appear rather similar in profile, distinct patterns can be gleaned as one transitions across segments. For example, consider the age profile of the segments. In the first segment, 58.4 percent of individuals belong to the 35+ age group ( $R \rightarrow F \rightarrow A$ ); this percentage gradually increases from left to right, ending with 63 percent of those in the last segment ( $A \rightarrow F \rightarrow R$ ) belonging to the 35+ year age group. In other words, the segments in which attitudes affect behaviors have a slightly older age profile than the first two segments where behaviors affect attitudes. It is entirely plausible that there are more people in the older age groups whose attitudes have matured and hardened, and their choice behaviors are influenced by their attitudes.

**TABLE 6 Profile of the Four Latent Segments**

Person Characteristics		Percent (%) within segment				Percent (%) within attribute				Overall Sample
Attribute	Categories	R→F→A	F→R→A	A→R→F	A→F→R	R→F→A	F→R→A	A→R→F	A→F→R	
Age Categories (years)	16-24	17.8	17.8	16.6	16.4	42.3	25.3	20.0	12.4	17.4
	25-34	23.8	23.6	21.1	20.6	43.2	25.6	19.4	11.8	22.8
	35 or more	58.4	58.6	62.3	63.0	40.3	24.1	21.8	13.8	59.9
Gender	Female	53.9	49.0	57.6	54.1	41.6	22.6	22.6	13.2	53.5
	Male	46.1	51.0	42.4	45.9	40.9	27.1	19.1	12.9	46.5
Marital status	Single	30.8	31.3	27.9	27.9	42.5	25.8	19.5	12.2	29.9
	Married	57.9	57.1	60.0	59.6	40.9	24.1	21.6	13.4	58.4
	Divorced	11.3	11.5	12.1	12.5	40.0	24.3	21.7	14.0	11.7
Frequency of transit use	≥ Once per week	17.8	17.8	16.6	16.4	42.3	25.3	20.0	12.4	17.0
	< Once per week	23.8	23.6	21.1	20.6	43.2	25.6	19.4	11.8	32.6
	Never	58.4	58.6	62.3	63.0	40.3	24.1	21.8	13.8	50.4
Distance from Home to Transit Station	< 0.5 mile	41.3	44.1	34.3	37.8	42.6	27.1	17.9	12.4	40.1
	≥ 0.5, <1 mile	20.9	20.2	23.2	22.2	40.4	23.2	22.8	13.6	21.4
	≥ 1 mile	37.7	35.7	42.4	40.0	40.4	22.9	23.1	13.6	38.5
Vehicle ownership	Zero	4.0	4.2	3.5	3.7	42.2	26.6	18.8	12.4	3.9
	1 vehicle	30.8	31.7	29.6	30.4	41.4	25.5	20.2	13.0	30.8
	2+ vehicle	65.2	64.1	66.9	65.9	41.2	24.2	21.4	13.2	65.4
Household size	1 person	18.0	18.8	17.1	17.8	39.6	25.1	21.2	14.1	18.0
	2 person	43.3	43.8	44.9	45.7	39.5	23.9	22.3	14.3	44.1
	3+ person	38.7	37.4	37.9	36.5	42.1	24.3	21.0	12.6	37.9
Annual household income	< \$35K	21.0	21.9	19.9	20.6	41.4	25.8	19.9	12.9	21.0
	≥ \$35K, < \$50K	14.2	14.8	13.8	14.2	41.1	25.6	20.3	13.1	14.2
	≥ \$50K, < \$75K	22.8	24.2	21.7	22.4	41.2	26.1	19.9	12.8	22.9
	≥ \$75K	42.0	39.1	44.6	42.8	41.3	23.0	22.3	13.4	41.9
Residential location choice	Urban dweller	28.4	29.3	25.7	26.5	42.2	25.9	19.4	12.5	27.8
	Suburban & small town mix	32.3	32.4	32.3	32.5	41.2	24.7	20.9	13.2	32.4
	Other suburban & small town + rural	39.3	38.3	41.9	40.9	40.7	23.7	22.1	13.5	39.8
Segment Size						41%	25%	21%	13%	100%
						3,936	2,400	2,016	1,248	9,600

Similar differential patterns across segments can be seen throughout the table. When compared with males, females are more likely to belong to causal structures in which attitudes shape behaviors. Single individuals who have never been married are more likely to belong to segments in which behaviors shape attitudes when compared with individuals who have been married or divorced. It generally appears that those in younger stages of life (from an age and lifecycle perspective) are less likely to have attitudes that have matured and hardened in comparison to those in later stages of life. Attitudes for these demographic groups may still be evolving to a slightly greater extent than others in the population.

Those who use transit more frequently are more likely to fall into the first two segments than those who never use transit. Individuals in households with no vehicles are similarly likely to fall into segments where behaviors shape attitudes, in comparison to those in households with more vehicles. Urban dwellers are more likely to be in the categories where behaviors shape attitudes in comparison to those in suburban and small town or rural settings. Again, all of these comparisons should be viewed carefully in *relative* terms because the differences are quite small. Although this analysis is not based on longitudinal data, the patterns in the table may be indicative of a transition process that may be at play. Broadly speaking, a majority of individuals fall into the segments where behaviors affect attitudes, but it appears that some individuals transition into other segments (where attitudes influence behaviors) as they age through lifecycle stages.

## **CONCLUSIONS AND POLICY IMPLICATIONS**

Transportation analysts are increasingly concerned with the relationships between human attitudes and perceptions on the one hand and behavioral choices on the other. From a travel demand forecasting perspective, there is interest in exploring the possibility of using attitudinal variables and constructs to better explain and more accurately predict travel demand under a variety of scenarios, particularly in the context of the emergence of new and disruptive technologies. From a sustainable transportation policy development perspective, there is interest in influencing attitudes of people (say, through information campaigns) to bring about more sustainable activity-travel behaviors.

Across a number of disciplines, the relationships between attitudes and behaviors have been well documented. Various studies, however, assume different causal relationships between attitudes and behaviors. Most studies appear to treat attitudes as affecting behavioral choices, but there are a number of studies (as noted in the introductory section) where behavioral choices are assumed to affect attitudes. A few studies have attempted to treat the attitude – behavior relationship as a bi-directional one, but econometric identification issues render the estimation of such models challenging when the endogenous variables are not continuous in nature. There is considerable uncertainty as to the direction of causality between attitudes and behaviors, and this study constitutes an attempt at shedding deep insights into the nature of the relationship. More specifically, this study recognizes that different causal structures may be prevalent in a population, leading to the presence of multiple population segments. Population heterogeneity may arise not only in terms of varied sensitivity to different attributes of alternatives, but also in terms of differing causal structures driving decision-making processes.

In an effort to unravel the extent to which different causal structures relating attitudes and behaviors are prevalent in the population, this study adopts a latent segmentation approach to reflect the notion that the analyst does not observe and is not aware of the causal structure adopted by each individual in the population. The latent segmentation approach endogenously assigns individuals to different causal structures, thus enabling the identification of segments in the



population and the degree of heterogeneity that may be prevalent. In this study, a joint equations model that relates residential location choice, frequency of transit use, and attitudes towards transit is estimated. The former two variables constitute behaviors, while the third variable is an attitudinal factor score. The model system is estimated on a large sample data set that includes both attitudinal and behavioral choice variables. Four different latent segments are considered; two latent segments in which attitudes affect choice behaviors and two segments in which choice behaviors affect latent segments. The two causal structures in which attitudes appear as a mediating factor between the two behavioral choice variables are ignored in this study.

The overall finding is that the majority of the sample in the data set used in this study are assigned to the latent segments in which behavioral choices affect attitudes. Nearly two-thirds of the sample falls into these two segments, while only about one-third falls into the two segments where attitudes affect behaviors. In other words, the findings of this study appear to corroborate some recent evidence that people appear to modify their attitudes in response to their behaviors to reduce the cognitive dissonance that may exist. It appears that attitudes at any cross-section in time are shaped by the behavioral choices and experiences of the individual at that point in time. As time progresses, it is entirely possible that attitudes and behaviors will evolve; but within the context of a snapshot, the study results here clearly indicate that attitudes are shaped by behaviors more so than the other way around.

It is clear, however, that there is considerable population heterogeneity. Four causal structures were considered in this study; the largest segment accounted for just over 40 percent of the sample and the smallest segment accounted for 13 percent of the sample. In other words, no latent causal segment is too small to be ignored. All segments are sizeable and hence it is important to recognize this population heterogeneity both in travel demand forecasting and in transport policy formulation. Travel demand forecasting models historically assume that the same causal structure applies to all agents in the population. Behavioral choice models are often specified with attitudinal variables and constructs as explanatory factors, when in fact it appears from this study that attitudinal factors are shaped by behavioral choices. Travel demand forecasting models that assume the same causal structure across the entire population are likely to return erroneous predictions of travel demand in response to policy and investment scenarios. It would be beneficial to probabilistically assign individuals in a population to different causal segments, and then forecast travel demand for different segments according to the causal structure that drives their decision-making process.

From a transportation policy perspective, it would appear that information campaigns and advertisements may not be all that effective in a world where the majority of the population has their attitudes shaped by behaviors. This implies that it is necessary to run campaigns where individuals actually get to experience modal options and different products first-hand; people need to be able to exercise alternative behavioral choices, learn through experience, and re-shape their attitudes in response to the experiences. Programs in which individuals are able to actually try out new and different alternatives (modes and services, for example) may yield greater benefit than messaging aimed at directly influencing attitudes. It should, however, be recognized that a sizable portion of the sample was also allocated to segments where attitudes affect behaviors; hence programs that aim to change attitudes should not be discontinued, particularly for more mature segments of the population who may be rather set in their ways and formed habits that are difficult to break. To make different campaigns work most effectively, they need to be targeted to the appropriate segments depending on the causal structures that they follow.

## REFERENCES

- Acker, V., Mokhtarian, P.L., & Witlox, F. (2014). Car availability explained by the structural relationships between lifestyles, residential location, and underlying residential and travel attitudes. *Transport Policy*, 35, 88–99.
- Ahn, J., & Back, K.-J. (2018). Influence of brand relationship on customer attitude toward integrated resort brands: a cognitive, affective, and conative perspective. *Journal of Travel & Tourism Marketing*, 35(4), 449–460.
- Alemi, F., Circella, G., Handy, S.L., & Mokhtarian, P.L. (2018). What influences travelers to use Uber? Exploring the factors affecting the adoption of on-demand ride services in California. *Travel Behaviour and Society* 13, 88–104.
- Archer, M., Paleti, R., Konduri, K.C., Pendyala, R.M., & Bhat, C.R. (2013). Modeling the connection between activity-travel patterns and subjective well-being. *Transportation Research Record*, 2382, 102–111.
- Astroza, S., Garikapati, V.M., Bhat, C.R., Pendyala, R.M., Lavieri, P.S., & Dias, F.F. (2017). Analysis of the impact of technology use on multimodality and activity travel characteristics, *Transportation Research Record*, 2666, 19–28
- Astroza, S., Garikapati, V.M., Pendyala, R.M., Bhat, C.R., Mokhtarian, P.L. 2018. Representing heterogeneity in structural relationships among multiple choice variables using a latent segmentation approach. *Transportation*, forthcoming. <https://doi.org/10.1007/s11116-018-9882-7>.
- Bhat, C.R. (1997). An endogenous segmentation mode choice model with an application to intercity travel, *Transportation Science*, 31(1), 34–48.
- Bhat, C.R. (2011). The maximum approximate composite marginal likelihood (MACML) estimation of multinomial probit-based unordered response choice models. *Transportation Research Part B*, 45(7), 923–939.
- Bhat, C.R. (2015a). A new generalized heterogeneous data model (GHDM) to jointly model mixed types of dependent variables. *Transportation Research Part B*, 79, 50–77.
- Bhat, C.R. (2015b). A comprehensive dwelling unit choice model accommodating psychological constructs within a search strategy for consideration set formation, *Transportation Research Part B*, 79, 161–188
- Bhat, C.R., & Guo, J.Y. (2007). A comprehensive analysis of built environment characteristics on household residential choice and auto ownership levels. *Transportation Research Part B*, 41(5), 506–526.
- Bhat, C.R., Astroza, S., Bhat, A.C., & Nagel, K. (2016). Incorporating a multiple discrete-continuous outcome in the generalized heterogeneous data model: Application to residential self-selection effects analysis in an activity time-use behavior model, *Transportation Research Part B*, 91, 52–76
- Cao, X.J., Xu, Z., & Fan, Y. (2010). Exploring the connections among residential location, self-selection, and driving: Propensity score matching with multiple treatments. *Transportation Research Part A*, 44(10), 797–805.
- Chen, F., Wu, J., Chen, X., & Wang, J. (2017). Vehicle kilometers traveled reduction impacts of Transit-Oriented Development: Evidence from Shanghai City. *Transportation Research Part D*, 55, 227–245.
- Choo, S., & Mokhtarian, P.L. (2004). What type of vehicle do people drive? The role of attitude and lifestyle in influencing vehicle type choice. *Transportation Research Part A*, 38(3), 201–

- Dobson, R., Dunbar, F., Smith, C.J., Reibstein, D., & Lovelock, C. (1978). Structural models for the analysis of traveler attitude-behavior relationships. *Transportation*, 7(4), 351–363.
- Egbue, O., & Long, S. (2012). Barriers to widespread adoption of electric vehicles: An analysis of consumer attitudes and perceptions. *Energy Policy*, 48, 717–729.
- Ettema, D., & Nieuwenhuis, R. (2017). Residential self-selection and travel behaviour: what are the effects of attitudes, reasons for location choice and the built environment? *Journal of Transport Geography*, 59, 146–155
- Fishbein, M., & Ajzen, I. (2010). *Predicting and changing behavior: The reasoned action approach*. Psychology Press, Taylor & Francis Group, New York, NY.
- Frei, C., Mahmassani, H.S., & Frei, A. (2015). Making time count: Traveler activity engagement on urban transit. *Transportation Research Part A*, 76, 58–70.
- He, S.Y., & Thøgersen, J. (2017). The impact of attitudes and perceptions on travel mode choice and car ownership in a Chinese megacity: The case of Guangzhou. *Research in Transportation Economics*, 62, 57–67.
- Heinen, E., Maat, K., & van Wee, B. (2011). The role of attitudes toward characteristics of bicycle commuting on the choice to cycle to work over various distances. *Transportation Research Part D*, 16(2), 102–109.
- Heinen, E., Maat, K., & van Wee, B. (2013). The effect of work-related factors on the bicycle commute mode choice in the Netherlands. *Transportation*, 40(1), 23–43.
- Hong, J., & Thakuriah, P.V. (2018). Examining the relationship between different urbanization settings, smartphone use to access the Internet and trip frequencies. *Journal of Transport Geography*, 69, 11–18.
- Kroesen, M., Handy, S.L., & Chorus, C. (2017). Do attitudes cause behavior or vice versa? An alternative conceptualization of the attitude-behavior relationship in travel behavior modeling. *Transportation Research Part A*, 101, 190–202.
- Kroesen, M., & Chorus, C. (2018). The role of general and specific attitudes in predicting travel behavior – A fatal dilemma? *Travel Behaviour and Society*, 10, 33–41.
- Lavieri, P.S., Dai, Q., & Bhat, C.R. (2018). Using virtual accessibility and physical accessibility as joint predictors of activity-travel behavior. *Transportation Research Part A*, 118, 527–544.
- Lewin, K. (1936). Principles of topological psychology. *McGraw-Hill Book Company*, New York, NY.
- Norman, R. (1975). Affective-cognitive consistency, attitudes, conformity, and behavior. *Journal of Personality and Social Psychology*, 32(1), 83–91.
- Ory, D.T., & Mokhtarian, P.L. (2005). When is getting there half the fun? Modeling the liking for travel. *Transportation Research Part A*, 39(2-3), 97–123.
- Pendyala, R.M., & Bhat, C.R. (2004). An exploration of the relationship between timing and duration of maintenance activities. *Transportation*, 31(4), 429–456.
- Seraj, S., Sidharthan, R., Bhat, C.R., Pendyala, R.M., & Goulias, K.G. (2012). Parental attitudes toward children walking and bicycling to school: multivariate ordered response analysis. *Transportation Research Record*, 2323, 46–55.
- Sheela, P., & Mannering, F. (2019). The effect of information on changing opinions toward autonomous vehicle adoption: An exploratory analysis. *International Journal of Sustainable Transportation*, forthcoming.
- Soto, J.J., Márquez, L., & Macea, L.F. (2018). Accounting for attitudes on parking choice: An integrated choice and latent variable approach. *Transportation Research Part A*, 111, 65-77.

- Transit Center. (2014). *Who's On Board: 2014 Mobility Attitudes Survey*. Transit Center, New York, NY.
- Wicker, A.W. (1969). Attitudes versus actions: The relationship of verbal and overt behavioral responses to attitude objects. *Journal of Social Issues*, 25(4), 41–78.