The Thesis Committee for Pragun Vinayak
Certifies that this is the approved version of the following thesis:

Accounting for Multi-Dimensional Dependencies Among Decision-makers Within a Generalized Model Framework: An Application to Understanding Shared Mobility Service Usage Levels

APPROVED BY
SUPERVISING COMMITTEE:

Supervisor:

____
Chandra R. Bhat

____
Randy B. Machemehl

Accounting for Multi-Dimensional Dependencies Among Decision-makers Within a Generalized Model Framework: An Application to Understanding Shared Mobility Service Usage Levels

by

Pragun Vinayak

Thesis

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Science in Engineering

The University of Texas at Austin

December 2017

Dedicated to my parents and my brother, who inspired me to dream beyond the stars

Acknowledgements

Graduate school has been a challenging, yet rewarding expedition of transformation and self-discovery. I have crossed paths with many great people, who deserve credit for making the journey so enriching and fulfilling. Right at the top of the list is my advisor Dr. Chandra R. Bhat. His unflinching commitment to excellence, invigorating passion for transportation studies and sheer humility has rubbed off on me and sure to stay with me in the times to come. Our short, yet stimulating interactions, have helped me grow as an engineer, a thinker and as a person.

I would like to thank my reader, Dr. Randy Machemehl, for his constructive feedback and valuable insights on my thesis work. My amazing teachers at UT Austin – Professors Boyles, Machemehl, Walton, Lin, Zhang, Kockelman, Balakrishnan and Lingham have been instrumental in nurturing a genuine love for scientific investigation and critical thinking. I would also like to thank Dr. Jen Duthie and James Kuhr at the Center for Transportation Research, who's mentoring and enthusiasm for connected/autonomous vehicle technologies has opened up so many frontiers for me. My passion for transportation planning traces its roots to my interactions with Dr. A.K. Sarkar and Dr. Ashish Verma. I cannot thank you enough for spotting the curious child in me and patiently advised me through all these years.

My life at Austin wouldn't have been so exciting without my collegues and friends. Shout out to my fellow sufferers in the graduate school grind – Sebastian, Cesar, Stefanos, Natalia, Neil, Scott, Abhilash, Priyadarshan, Chang and Bharathwaj (aka the Godfather). Felipe, my go-to guy for everthing thesis and CrossFit, and Patricia, the

superstar project partner – thank you for being a part of this journey and I look forward to staying in touch. I couldn't possibly forget to mention my Traffic Bowl Team – Murthy, Qichun, Rydell and Venktesh. Preparing for the competition was merely an excuse to spend time with you folks - debating, laughing and travelling. Murthy, Kaavya and Jayashree, my constants in this roller-coaster ride; Tanvi and Mehta, my daily dose of humour and pep talk; Rahul, Ashish, Gaurav and Tapan, my comarades in the Coup gang; Dev, Gopindra and Srijith – my fantastic roommates who put up with all my follies.

My friends from Pilani played a pivotal role in supporting my eclectic endeavors – Prajakta, Manasi, Anupam, Kalyan, Shubhangi, Dosi, Anshuman, "Hippo", Jyoti, and Mansi. Thank you all for staying in touch despite my "impressive" track-record in the same. This section would be incomplete if I didn't mention my best friends from high school – Siddharth, Guragam, Piyush, Saransh, Vishwajeet and Shubham. They've got me through some tough times and I am really lucky to have grown up with these fine gentlemen.

Lastly, but most importantly I would like to my family – "daddyji", "biji", mom, dad and Rohin. The values I imbibed from y'all has made me the man I am today. You've been my fiercest critics, loudest supporters and life-long role-models. Your love, encouragement and empathy will continue to guide me in the future.

Abstract

Accounting for Multi-Dimensional Dependencies Among Decision-makers Within a Generalized Model Framework: An Application to Understanding Shared Mobility Service Usage Levels

Pragun Vinayak, M.S.E.

The University of Texas at Austin, 2017

Supervisor: Chandra R. Bhat

Activity-travel choices of decision makers are influenced by spatial dependency effects. As decision makers interact and exchange information with, or observe the behaviors of, those in close proximity of themselves, they are likely to shape their behavioral choices accordingly. For this reason, econometric choice models that account for spatial dependency effects have been developed and applied in a number of fields, including transportation. However, spatial dependence models to date have largely defined the strength of association across behavioral units based on spatial or geographic proximity. In the current context of social media platforms and ubiquitous internet and mobile connectivity, the strength of associations among decision makers is no longer solely dependent on spatial proximity. Rather, the strength of associations among decision makers may be based on shared attitudes and preferences as well. In other words, behavioral choice models may benefit from defining dependency effects based on attitudinal constructs in addition to geographical constructs. In this thesis, the frequency of usage of car-sharing and ride-sourcing services, collectively termed as shared mobility services, is modeled using a sequential generalized heterogeneous data model – spatial ordered response probit (GHDM - SORP) framework that incorporates multi-dimensional dependencies among decision-makers.

The model system is estimated on the 2014-2015 Puget Sound Regional Travel Study survey sample, with inter-dependence in attitudinal space defined using latent psychometric constructs reflecting inherent attitudes, lifestyle preferences and habits. These latent constructs are based on variables in the data set that represent observed travel and locational choice behavior, as well as responses to attitudinal questions. Model estimation results show that social dependency effects arising from similarities in attitudes and preferences are significant in explaining shared mobility service usage, over and above what is explained by spatial dependency. Ignoring such effects may lead to erroneous estimates of the adoption and usage of future transportation technologies and mobility services.

Keywords: *spatial dependence, social interactions, attitudinal proximity, values and behavior, shared mobility service usage, latent constructs*

Table of Contents

List of Tables x

List of Figures xi

Chapter 1: Introduction and Background..1

1.1. Overview 1

1.2. Shared Mobility Services: Redefining Urban Mobility Ecosystem.......4

1.2.1. An Overview 4

1.2.2. Literature and Gaps 6

1.3. Beyond Spatial Measures of Dependence..7

Chapter 2: Data Preparation and Descriptive Analysis...10

2.1. Imputation of variables 10

2.3. Data Description 13

Chapter 3: Behavioral and Methodological Frameworks 16

3.1 Linking Latent Constructs with Usage Patterns – Behavioral Framework16

3.2. Modelling Framework 22

3.2.1. Latent Structural Equation Model (SEM).........................22

3.2.2. Latent Measurement Equation Model (MEM) 23

3.2.3. Capturing Dependency Effects Using a Spatial Lag Structure 24

3.2.5. Standard Error Corrections 29

Chapter 4: Model Estimation Results 33

4.1 Structural Equation Model (SEM) Component of GHDM.....................33

4.2 Measurement Equation Model (MEM) Component of GHDM 35

4.3 Spatial Ordered Response Probit (SORP) Model with Dependency Effects 39

4.4 Model Fit...40

Chapter 5: Discussion and Conclusions...44

References...47

List of Tables

Table 2.1: Descriptive Statistics of Sample 15
Table 3.1: Descriptive Statistics of Indicator Variables 21
Table 4.1: Estimation Results for Structural Equation Model of GHDM 35
Table 4.2: Estimation Results for Nominal Measurement Equations 38
Table 4.3: Estimation Results for Non-Nominal Measurement Equations 38
Table 4.4: SORP Model with Spatial and Non-Spatial Dependencies 42
Table 4.5: Measures of fit 43

List of Figures

Figure 3.1: Overview of Behavioral Framework 17

Chapter 1: Introduction and Background

1.1. Overview

Incorporating notions of interdependency in explaining travel patterns and locational choice behavior of decision makers has garnered much interest in the recent past (Dugundji and Walker, 2005; Blume and Durlauf, 2003; Bhat et al, 2016). A key differentiating factor in these studies is that they account for the nature of proximity amongst decision makers, which results in varied forms of networks over which feedback or inter-dependency effects propagate. Proximity is defined as the degree of closeness between decision makers and can be measured along different dimensions - geographic space, social space, and attitudinal space (lifestyle preferences, attitudes and values). Proximity in geographic space has traditionally accrued importance in econometric models that account for dependency amongst decision makers (Dugundji and Walker, 2005; Bhat et al, 2016), largely due to the idea that decision makers' preferences and choice behavior are shaped by dyadic exchanges between decision makers in close spatial proximity of one another. However, several studies have pointed out that social influence is pervasive, and a decision maker's choices are not isolated from the influence of other decision makers in his or her social sphere (Brock and Durlauf, 2001; Arentze and Timmermans, 2008).

Recent advances in technology and the accompanying growth in social media platforms such as Facebook and Twitter have rendered spatial separation practically moot as much of social interaction occurs virtually (Hackney and Axhausen, 2006). Research in social interactions has considered associations within tight social networks such as among family members (Arentze and Timmermans, 2009) as well as wider networks extending to colleagues, friends, and virtual social media connections (McPherson et al, 2001; Axhausen, 2008; Carrasco et al, 2008; Bhat, 2015a). However, there is limited knowledge of (a) the topology of such networks and their influence on transportation decisions, (b) the feasibility of using global networks of decision makers in such a social space, and (c) methods to operationalize the strength of relationships in such networks (Hackney and Axhausen, 2006). Adding to this is the arduous and often intractable task of extracting information about social network connections from conventional travel and land-use survey data (Axhausen, 2008). As a result, research that accounts for the influence of social networks in shaping travel behavior is rather sparse. Even in the limited literature on this topic, studies have utilized associative, aggregate-level networks where decision makers are grouped by planning zone and observed socio-demographic or economic characteristics (Yang and Allenby, 2003; Dugundji and Walker, 2005) as opposed to innate lifestyle preferences, values, and attitudes). In such a socio-spatial network, the feedback effects within apparently homogenous spatial or socio-economic groups do not account for self-selection effects attributable to the decision maker's underlying perceptions, attitudes and preferences towards built environment and travel behavior (van Wee et al., 2002; Anable, 2005; Mokhtarian and Cao, 2008; Van Acker et at., 2010).

In pursuit of a framework that can accommodate social dependency effects in studying travel behavior, this thesis extends the concept of proximity-based dyadic interactions by introducing the idea of attitudes, habits and lifestyle preferences as a new dimension and measure of proximity. As opposed to the physical networks that are based on observable socio-spatial variables, *latent social networks* are introduced in this thesis. In this paradigm, the inter-dependency among decision makers originates from similarity in the attitudinal space. Unlike previous formulations, where extracting the topology of social networks and operationalizing strength of influence amongst decision makers in such complex networks maybe infeasible (if not impossible), social dependence can be parsimoniously expressed using latent psychometric constructs, which link decision makers with similar attitudes and lifestyle preferences.

The methodology applied in this study accounts for both interdependencies amongst decision makers in spatial-attitudinal space and dynamics of self-selection due to inherent attitudes, preferences, and habits affecting a decision maker's frequency of using car-sharing and ride-sourcing mobility services. This topic is of particular relevance as the urban transportation landscape has been significantly disrupted by the emergence of shared mobility services, inspired by the concept of a sharing economy (Hannon et al, 2016). Two such services that figure prominently in this era of smart- mobility are car-sharing and ride-sourcing services. While many studies on car-sharing and ride-sourcing services have explored the role of socio-economic and built environmental factors (Coll et al, 2014; Kim, 2015; Clewlow, 2016; Rayle et al, 2016) in shaping usage of such services, there is a paucity of literature that examines inter-dependencies in attitudinal space that impact usage patterns of these shared mobility services.

The effort reported in this thesis uses data from the 2015 Puget Sound Regional Travel Study (PSRC, 2015) to model the monthly usage of ride-sourcing and car-sharing services for adults, which constitutes the ordinal variable of interest. The study considers two latent constructs relevant to urban travel and locational behavior: pro-environment attitude and neo-urban (active) lifestyle propensity. It should be noted that the thesis's focus is only on short-term travel choices, and hence variables reflecting long-term household decisions, such as residence type and vehicle ownership, are included only as exogenous covariates to explain the ordinal variable of interest. The next section presents an overview of shared mobility services and gaps in literature that motivate this study. The third section in this chapter presents the foundation for accommodating social dependency in attitudinal space.

1.2. SHARED MOBILITY SERVICES: REDEFINING URBAN MOBILITY ECOSYSTEM

1.2.1. An Overview

The emergence of ride-sourcing and car-sharing services have ushered in a new era of shared mobility that leverages technology to connect service providers with the customers (Hannon et al, 2016). These services have changed how people move around and participate in different activities – commuters can forgo the need to own cars or pre-arrange car pools; social and recreational activities can be scheduled on-the-fly without worrying about multiple trips or parking; cabs can be hailed at any point in time with a mere tap on people's smartphones.

Ride-sourcing, offered by Transportation Network Companies (TNC), refers to a mobility-on-demand service that offers a lower cost alternative to taxis, provides door-to-door service and hailed, monitored and paid for using technology-based platforms (e.g. smart-phones) (Dias et al. 2017). Besides the more popular ride-sourcing service providers like Uber and Lyft, many other services across the globe such as Ola (India), Didi Chuxing (China), Grab (South-East Asia) have seen consistent increases in their riderships. Lyft completed 160 million trips in 2016 – more than a 300% increase over 2015 (Lyft, 2017). Similarly, in June 2017, Uber ferried its 5 billionth customer in just over a year after surpassing the two billion mark (Uber, 2017). Despite the obviously increasing demand for ride-sourcing services, such services have elicited mixed reactions from policy makers and planners. On one hand, it affords an alternative to driving and a prospective last-mile connector to public transportation systems, it can potentially reduce auto-ownership and, hence, ameliorate environmental concerns (Metcalfe and Warburg, 2012; Silver and Fischer-Baum, 2015). Critics, on the other hand, argue that ride-sourcing services increase the vehicle-miles travelled by inducing a latent demand, compete and erode the share of green modes (public transportation, walking and bicycling) and cater to mostly young, economically sound decision makers (Sabatini, 2014; Rayle et. al, 2016). Latent demand refers to the additional trips on the transportation network, previously suppressed due to behavioral reasons and limited service supply levels, that originate due to the availability of a new service. However, the true impacts of ride-sourcing on vehicle miles traveled and impacts on other transportation modes are still unclear, owing mostly to the lack of disaggregate data sources (Rayle et al. 2016) and limited understanding of the true nature of such services.

Car-sharing services, which in their most basic form are car-rentals by the hour or minute, afford consumers all the benefits of automobile ownership without incurring high fixed costs of purchase, insurance and maintenance (Shaheen et al., 2009). ZipCar and Car2Go are two of the main examples of commercial car-sharing services. While similar programs have existed since the 90's, ZipCar and Car2Go have heavily benefitted by riding the Internet wave and have rapidly expanded their customer base, with young and educated decision makers in metropolitan areas constituting the biggest chunk of the 1.5 million estimated members in 2015 (Shaheen, 2016). Car-sharing systems have potentially substantive benefits, such as efficient mobility with lower car-ownership levels, lower demand for parking, and lower acquisition and usage costs (Baptista, 2004). Firkorn and Muller (2011) found that more than 25% of German respondents would be willing to forgo personal cars if they had access to car-share services.

1.2.2. Literature and Gaps

A growing body of literature has explored the interplay of socio-demographic, socio-economic and built environmental factors on usage patterns of ride-sourcing and car-sharing services (Clewlow 2016, Coll et al. 2014, Rayle et al., 2016). However, the social dependence amongst decision makers stemming from attitudes and lifestyle preferences is yet to be fully explored. Studies by Costain et al. (2012), Efthymous et al. (2013) and Dias et al. (2017) have acknowledged the crucial role of underlying attitudes and lifestyle preferences in the adoption of and participation in such services. Anable (2005) found that attitudes are important predictors of an decision maker's mode switching potential.

Almost all early studies of shared mobility services (except Dias et al. 2017) have scrutinized the two mobility platforms independently. However, both services are technology enabled, involve vehicles not owned by decision makers, are off-shoots of shared economy, and are predominantly urban phenomena, which means there ought to be underlying unobserved factors that simultaneously affect the usage of both services. Due to these reasons, both services are considered concomitantly in this thesis, and the analysis does not evaluate whether such services are synergistic or competitive with each other. Inspired by the availability of a rich dataset and recent studies that have alluded to such unobserved factors, this thesis posits latent constructs to capture underlying attitudes and lifestyle preferences that may influence the usage patterns of these services. Additionally, the present work accommodates dependency effects based on interactions in a unique spatial-attitudinal space, and evaluate the complex interplay between different dimensions of proximity.

1.3. Beyond Spatial Measures of Dependence

The study of attitudes, perceptions, habits, and lifestyle preferences has been of interest to travel behavior researchers due to their role in shaping human activity-travel choices (Kitamura et al, 1997; Bagley and Mokhtarian, 2002). This notion is further reinforced by theories in social psychology which evaluate how such personality traits shape short-term and long-term behavior, and recognize that a decision maker's behavior often tends to conform to the social constraints and norms of the individual's cohort or reference group. Theory of Reasoned Action (Fishbein, 1980) and Theory of Planned Behavior (Ajzen, 1991) suggest that attitudes and lifestyle preferences play an important role in shaping behavior in different contexts. Subjective norms – the sum of normative beliefs due to social pressure to conform to one's reference group – also influence behavior. For example, people who perceive themselves to be pro-environmental may bicycle to work or buy a clean-fuel vehicle to align their actions with those of other pro-environmental decision makers. These three influences (attitudes, lifestyle preferences, and subjective norms), which contribute to consistent patterns of behavior, have been termed as *reasoned influences*. In addition to reasoned influences, Van Acker et al (2010) consider *unreasoned influences* as an additional determinant of travel behavior.

Unreasoned influences include habits and dependencies, and trace their origins to the Theory of Repeated Behavior (Ronis et al, 1989). This theory suggests that repeated behavior is motivated more by habit than attitudes.

Unlike some of the social networks mentioned previously, a decision maker may not necessarily interact with group members in the same attitudinal space either physically or virtually (refer to example of bicycling to work). Social inter-dependency engendered through *passive observation* of individuals in a similar attitudinal space is a simple and powerful construct that is yet to be fully explored. It is therefore hypothesized that a decision-makers' position in attitudinal space can suppress or promote different courses of action, a behavioral phenomenon that policy makers can leverage to achieve mobility goals. Within the context of accommodating dependencies, this study adopts a spatial lag structure for the outcome variable of interest. The latent constructs reflecting attitudes, habits, and preferences are based on observed psychometric indicators and/or other variables describing observed behavior (e.g., smartphone ownership) and scores for these latent constructs are estimated using Bhat's (2015b) Generalized Heterogeneous Data Model (GHDM). These latent constructs serve to introduce dependencies amongst decision makers in the attitudinal space.

In conventional spatial econometric models, the autocorrelation among decision makers is diffused via a weight matrix that is based on a spatial network measuring distances between decision makers (e.g., Paleti et al, 2013). Elements in each row of the matrix reflect the absolute spatial influence of all decision makers on a given decision maker. In this paper, the network topology is determined by both spatial (geographical) and attitudinal (non-spatial) proximities, the latter incorporating attitudes, preferences, and habits. The influence of attitudinal and spatial networks is disentangled by using coefficients for each proximity measure. This opens up the possibility for one measure counteracting the influence of another; for example, even when decision-makers are in close geographical proximity, differences in their attitudes, preferences, and habits may outweigh their spatial proximity.

Chapter 2: Data Preparation and Descriptive Analysis

The data for this study is derived from the Puget Sound Regional Travel Study that involved survey data collection in 2014 and 2015 covering a five-county area in the State of Washington. In addition to collecting information about socio-economic, demographic, and activity-travel characteristics, the survey asked respondents to provide information about attitudes, preferences, and technology (e.g., smartphone) ownership and usage. Data about residential location choice preferences, and membership and usage of shared mobility services such as ride-hailing, bike-share, and car-share services, was collected through the survey. All relevant variables used in this study were extracted from the 2015 edition of the survey data set, except for two variables that capture the usage patterns of technology platforms (frequency of use of smartphone apps and frequency of use of websites) for obtaining travel-related information. These two variables are available in the 2014 edition of the survey; these variables are imputed into the 2015 data set based on ordered response probit models of technology use estimated on the 2014 data. The imputation exercise is described below, followed by descriptive statistics for the final sample that is used for modelling purposes.

2.1. Imputation of Variables

For this exercise, only respondents above the age of 18 were chosen. Respondents in 2014 provided information frequency of deriving travel-related information in the past month via (1) smart-phone apps and (2) websites using a scale of one to seven, with an increasing level of usage. The ordinal frequency variables (frequency of using smart-phone apps and frequency of using websites for travel-related information) are recoded into a five-level scale to ensure there are enough sample data points under each level of usage:

1) Never
2) Less than once a week
3) One day per week
4) Two to four days per week
5) More than 4 days per week

Since the levels indicate an increasing degree of use, the two frequency variables from the 2014 sample are treated as ordinal outcomes and separate ordered response probit models are developed using exogenous variables reflecting individual socio-demographic characteristics (age, education level, employment status, income) and other relevant variables (smart-phone ownership and household density expressed in number of households per square mile). The results are consistent with expectations as well as findings reported in the literature (Pew Research Center, 2014). Assuming the effects of the above variables remain constant over the course of a year, the estimated models are applied to the 2015 sample, with coefficients fixed at their 2014 values. The model application provides a probabilistic assignment to the five usage levels for each ordinal frequency variable, thus, resulting in two sets of five probability values for every respondent. The naive approach to impute the 2015 usage levels for smart-phone apps and websites for travel information is adopting the levels corresponding to the highest probability values. However, the sample, though considered representative of the population of interest, cannot fully capture behavioral heterogeneity amongst decision makers. To account for this heterogeneity and introduce greater variability in the sample, pseudo-random draws from a uniform distribution (0,1) are performed. The interval (0,1) is split into five blocks (labeled as five levels of the ordinal variable) with thresholds based on the predicted probabilities from the model application step. These thresholds vary amongst respondents due to differences in socio-demographic and other characteristics. The respondent's imputed level of usage is decided by the block into which the pseudo-random number falls. This procedure results in the imputed levels of usage for smart-phone apps and websites for travel information in 2015. To generate a new composite variable, frequency of using technology-based platforms for travel information, we transform the two imputed ordinal variables into monthly counts using the follow rubric:

1) 0 times
2) 0.3 times
3) 4 times
4) 8 times
5) 16 times

The monthly counts for app and website usage are summed to obtain monthly instances of using technology-based platforms, and then reconverted to an ordinal scale variable "frequency of using technology-based platforms for travel info" with four levels (five levels are collapsed into four to ensure a more equitable distribution of sample points and improve computational tractability) namely,

1) Never
2) Less than once a week
3) One day per week
4) Two or more days per week

2.3. DATA DESCRIPTION

The analysis is limited to adults (age 18 years or above). Respondents below the age of 18 were excluded since perceptions about attitudes and lifestyle preferences amongst youngsters are not expected to be fully matured. Additionally, all records with proxy reporting were filtered out as it was deemed potentially challenging to report attitudes and preferences, and true usage patterns of shared mobility services on behalf of other household members. The dependent variable of interest is the frequency of using ride-sourcing services (e.g., Uber and Lyft) and/or car-sharing services (e.g., ZipCar and car2go) in the past 30 days. Information on this variable is derived from ordinal indicators measuring level of usage as reported by the respondents. The seven-level ordinal scale includes the following:

1) Never
2) I do this, but not in the past 30 days
3) 1-3 times in the past 30 days
4) 1 day per week
5) 2-4 days per week
6) 5 days per week
7) 6-7 days per week

The two disruptive mobility services are considered together in this study because both are technology-enabled, and involve the use of vehicles not owned by the traveler. To account for very small sample sizes in some categories, and for computational tractability, a more aggregate three-point ordinal scale was used to represent the level of usage:

1) Never

2) Occasionally, but not in the past 30 days
3) Used service in past 30 days with any frequency

The final cleaned and filtered sample used for analysis and model estimation included 2170 adults. Table 2.1 summarizes the characteristics of the final sample. A majority of the decision makers in the analysis sample are in the middle age groups. There are more females than males, and full-time employed decision makers constitute nearly one-half of the sample. About 36 percent of the sample is unemployed. Only about six percent of the sample reported being a student, a similar percent reported not having a driver's license, and about 70 percent of the sample reported owning a smartphone. About 12 percent of the sample resides in households with no vehicles, about 30 percent of the sample report living in high-density census blocks of 5000 or more households per square mile. Nearly 20 percent of the sample reside in single-person households, and an almost equal percent reside in nuclear family households with children. Most of the respondents (over 68%) have bachelor or graduate degrees, indicative of Puget Sound Region's prominence as one of the technology hubs in the US, that attracts highly skilled workers. The income distribution shows that 34 percent of decision makers reside in households that make over $100,000 per year. Only 10 percent of the sample has membership in car- or bike-share services. An examination of the dependent variable of interest shows that 81 percent of the sample has never used car-share or ride-sourcing services in the past 30 days. This is consistent with the notion that shared mobility services are relatively new entrants in the transportation landscape.

Table 2.1: Descriptive Statistics of Sample

Person Variables			Household Variables		
Variable	Count	%	Variable	Count	%
Age			Vehicle Ownership		
18-24	55	2.53%	0	265	12.21%
25-34	396	18.25%	1	877	40.41%
35-44	370	17.05%	2 or more	1028	47.37%
45-54	354	16.31%	Residence Type		
55-64	487	22.44%			
65-74	338	15.58%	One HH – Detached Unit	1206	55.58%
75-84	139	6.41%	One HH – Attached Unit	115	5.30%
85 or older	31	1.43%	Multiple HH – Apt/Dorms	849	39.12%
Gender			Residential Density		
Male	943	43.46%	Upto 5000 HH per sq.mi	1519	70.00%
Female	1227	56.54%	Above 5000 HH per sq.mi	651	30.00%
Employment Status			Family Structure		
Employed full-time	1061	48.89%	Single Person HH	423	19.49%
Employed part-time	183	8.43%	Single Parent HH	46	2.12%
Self-employed	141	6.50%	Couple HH	775	35.71%
Unemployed	785	36.18%	Nuclear Family HH	441	20.32%
			Other (joint-families) HH	485	22.35%
Student			Number of kids		
Yes	134	6.18%	0	1772	81.66%
No	2036	93.82%	1	207	9.54%
			2 or more	191	8.80%
Driving License			Annual income		
Yes	2042	94.10%	Under $25,000	303	13.96%
No	128	5.90%	$25,000-$49,999	442	20.37%
Owns a smart-phone			$50,000-$74,999	353	16.27%
Yes	1519	70.00%	$75,000-$99,999	328	15.12%
No	651	30.00%	$100,000 or more	744	34.29%
Education Level			Car / Bike-share membership		
Less than Bachelor Degree	689	31.75%	Yes	226	10.41%
Bachelor Degree	828	38.16%	No	1944	89.59%
Any Graduate Degree	653	30.09%			

Chapter 3: Behavioral and Methodological Frameworks

This section offers a detailed description of the behavioral and methodological frameworks adopted in this study.

3.1 Linking Latent Constructs with Usage Patterns – Behavioral Framework

The behavioral framework adopted in this study is shown in Figure 3.1. Latent constructs that describe a decision maker's innate attitudes and lifestyle preferences are linked to the proclivity to adopt and use shared mobility services in this framework. Latent attitudinal constructs are modeled as functions of exogenous variables and manifest themselves in the data set as indicator variables (specifically, binary, ordinal frequency, and ordinal attitudinal indicator variables) that represent observed travel and locational choice behavior as well as responses to attitudinal questions. Instead of explicitly modeling the impacts of these latent constructs on shared mobility service usage, the latent constructs are used to induce dependency effects over a latent social network of decision makers who are proximally located in attitudinal space over and above the dependency effects attributed to spatial proximity.

Latent factors considered in this study include a decision maker's "neo-urban lifestyle propensity" and "pro-environmental attitude", both of which have surfaced repeatedly in the literature as determinants of activity-travel choices, especially in the context of shared mobility service usage (Lavieri et al, 2017; Astroza et al, 2017).

A pro-environmental attitude has been found to be significantly associated with shared mobility use (e.g., Efthymiou et al, 2013; Burkhardt and Millard-Ball, 2006).

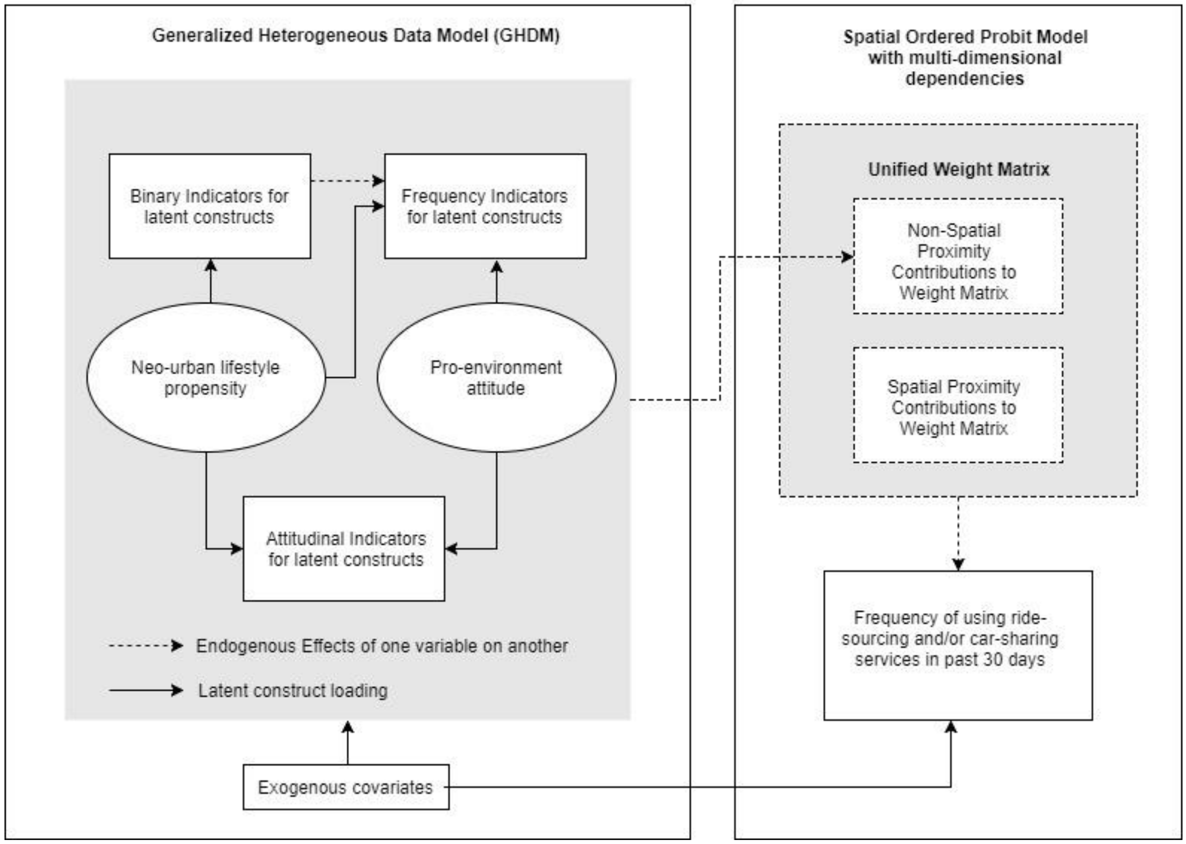

Figure 3.1: Overview of Behavioral Framework

It has been shown in these studies that pro-environmental decision makers eschew use of personal vehicles in favor of the use of transit and non-motorized modes and exhibit a higher affinity towards use of ride-sourcing and car-sharing services. In this study, two ordinal attitudinal variables and two ordinal frequency variables in the data set are considered representative of a pro-environmental attitude:

- Importance of residing close to transit (measured on a five-point scale: very unimportant to very important)

- Importance of residing in a walkable neighborhood with access to local activities located nearby (measured on a five-point scale: very unimportant to very important)
- Frequency of bicycling episodes (more than 15 minutes) in past 30 days (measured on a four-point scale: never, I do – but not in past 30 days, more than once in past 30 days – but at most one day per week, and two or more days per week)
- Frequency of walking episodes (more than 15 minutes) in past 30 days (measured on the same four-point scale as frequency of bicycling episodes)

The neo-urban lifestyle propensity is comprised of three unique features – use of technology to access travel-related information, proclivity for shared-space and collaborative ownership (i.e., proclivity to participate in the shared economy), and level of importance attached to residing in locations close to work and social-recreational activities. Previous studies have shown that these three attitudinal traits are significantly associated with the use of car-share and ride-sourcing services (Astroza et al, 2017; Montgomery, 2015). Since GPS-based technology, especially incarnated in the form of smart-phones, has paved the way for the rising popularity of mobility-on-demand services, dependency on technology certainly plays a pivotal role in adoption and continued usage of these services. Smart-phones allow decision makers to have a greater spatial-temporal control over how they plan and allocate time to activities – decision makers can now make decisions-on-the fly and participate in complex tours with multiple stops (Astroza et al., 2017). The second feature is embodied in an emerging trend in urban agglomerates. Attitudes and preferences associated with renting, borrowing and leasing have ushered in a new era of “disownership", flipping the consumerism idea of ownership. Over 52% of respondents in a survey in the US by Sunrun (2013) eschewed ownership and chose to borrow or lease traditionally-owned items. When it comes down to vehicle ownership, millennials are more willing to embrace shared vehicles than decision makers from previous generations (Montgomery, 2015). Individuals with such lifestyle preferences have also shown interest in car-sharing systems with fully autonomous vehicles (Lavieri, 2017). Neo-urban lifestyles are also characterized by preferences for dense neighborhoods, shorter commutes to work and proximity to social and recreational places (New Urbanite Study, 2016). In this study, one ordinal frequency indicator, one binary indicator, and three ordinal attitudinal/interest indicator variables are tested as indicators of a neo-urban lifestyle propensity:

- Frequency of using technology-based platforms (smartphone apps and/or websites) for travel information in past 30 days (measured on same four-point scale as frequency of walking and bicycling episodes)
- Smartphone ownership (binary indicator)
- Level of interest in participating in an autonomous vehicle car-share system (measured on a five-point scale: not at all interested to very interested)
- Importance of residing in a home location close to highways or major roads (measured on a five-point scale: very unimportant to very important)
- Importance of living within a 30-minute work commute (measured on a five scale: very unimportant to very important)

Table 3.1 presents a summary of the indicator variables for the analysis sample. Being close to highways and major roads is generally considered less important than being within a 30-minute work commute and having a walkable neighborhood with local activities nearby. Availability of public transit is also considered an important criterion in determining residential location. A majority of the sample is not at all interested in using an autonomous car-share system for daily travel. The frequency of walking is substantially larger than the frequency of bicycling, with over 66 percent respondents engaging in walking trips two or more times a week. About 22 percent own a smartphone, but never use apps for travel information. About 30 percent own a smartphone and use apps one or more days per week for travel information. About 31 percent of the sample never uses technology platforms for travel information. On the other hand, 23 percent do so two or more times per week. The statistics in the table show that there is considerable heterogeneity in the population with respect to residential location preferences, interest in autonomous car-share adoption, and use of technology platforms for travel information.

Table 3.1: Descriptive Statistics of Indicator Variables

Attitudinal (Ordinal) Indicator Variables					
	Response Distribution				
Importance of factor in choosing home location	Very Unimportant 1	Unimportant 2	Neutral 3	Important 4	Very Important 5
Close to major roads/highways	14.8%	16.0%	22.2%	34.6%	12.4%
Being within 30-minute commute to work	11.0%	6.1%	17.7%	20.4%	44.8%
Being close to public transit	15.4%	10.4%	17.8%	25.3%	31.1%
Having a walkable neighborhood and being near local activities	5.3%	6.7%	10.3%	33.2%	44.4%
	Response Distribution				
Level of interest in use of…	Not at all interested	Somewhat uninterested	Neutral	Somewhat interested	Very Interested
Autonomous car-share system for daily travel	55.4%	6.7%	11.9%	14.0%	11.7%

Frequency (Ordinal) Indicator Variables				
	Response Distribution			
Frequency of participating in…	Never	I do, but not in the past 30 days	More than once in past 30 days but at most 1 day/week	Two or more days/week
Bicycling (15 min or more)	62.7%	20.7%	8.4%	8.2%
Walking (15 min or more)	8.3%	6.0%	18.7%	66.9%

Frequency of…	Smartphone ownership and app use for travel info	*Frequency of:*	Technology-based platforms for travel info
Don't own smartphone	30.0%	Never	31.1%
Own smartphone but never use apps for travel info	21.8%	Less than one day per week	33.2%
Own smartphone and use apps less than one day per week for travel info	18.9%	One day per week	12.9%
Own smartphone and use apps one or more days per week for travel info	29.4%	Two or more times per week	22.8%

3.2. MODELLING FRAMEWORK

The modeling framework consists of two primary components, namely, the Generalized Heterogeneous Data Model (GHDM) and the spatially lagged ordinal response model with a composite weight matrix that includes both spatial and aspatial (attitudinal) components. Within the GHDM, there are two submodels – a latent structural equation model (SEM) and a latent measurement equation model (MEM). In the latent SEM, the latent psychological constructs are represented as linear functions of exogenous variables with the usual stochastic error terms. In the latent MEM component, psychometric indicators along with observed travel behavior indicators are posited as functions of latent constructs, exogenous variables, and other endogenous outcomes. The SEM and MEM sub-models are estimated jointly in a simultaneous equations modeling framework. The second component of modelling framework is the spatially lagged ordinal response model with a composite weight matrix, populated using expected latent variable scores from the GHDM model, that embodies the multi-dimensional dependency in spatial-attitudinal space.

In the following discussion, consider a sample of Q decision makers denoted by index $q(1,2,3...,Q)$ and L latent variables denoted by index l (L=2 in this study). Let there be a total of N ordinal indicators and G nominal indicators (binary or multinomial outcomes) for the MEM submodel of GHDM.

3.2.1. Latent Structural Equation Model (SEM)

For a given decision maker q, the L latent constructs can be compactly written as a vector $\boldsymbol{z_q}$ $(L\times 1)$, which in turn, is a function of observed covariates and a vector of stochastic error components, as specified below.

$$\boldsymbol{z}_q = \boldsymbol{\alpha}\boldsymbol{w}_q + \boldsymbol{\eta}_q \tag{3.1}$$

where, $\boldsymbol{\alpha}$ $(L \times F)$ is the vector of coefficients associated with the covariates (excluding a constant) given by $\boldsymbol{w}_q$ $(F \times 1)$ and $\boldsymbol{\eta}_q$ $(L \times 1)$ is the vector of stochastic error terms. $\boldsymbol{\eta}_q$ follows a multivariate normal (MVN) distribution which allows for a correlation structure to accommodate interactions amongst latent variables i.e. $\boldsymbol{\eta}_q \sim \mathbf{MVN}\left[\boldsymbol{0}, \boldsymbol{\Gamma}\right]$ where $\boldsymbol{0}$ $(L \times 1)$ is a zero vector and $\boldsymbol{\Gamma}$ $(L \times L)$ denotes the correlation matrix. We assume $\boldsymbol{\eta}_q$ is independent across decision makers i.e. $\mathrm{Cov}(\boldsymbol{\eta}_q, \boldsymbol{\eta}_{q'}) = 0\ \forall q \neq q'$.

3.2.2. Latent Measurement Equation Model (MEM)

The measurement equation system (MEM) component is expressed in matrix form for all indicators. Let G be the number of nominal indicators with I_g alternatives in nominal variable g. Let the total number of alternatives across all nominal indicators be given by $\ddot{G} = \sum_{g=1}^{G} I_g$. Then for decision maker q,

$$\ddot{\boldsymbol{y}}_q = \ddot{\boldsymbol{\gamma}}\boldsymbol{x}_q + \ddot{\boldsymbol{d}}\boldsymbol{z}_q + \ddot{\boldsymbol{\varepsilon}}_q \tag{3.2}$$

$$\boldsymbol{U}_q = \boldsymbol{b}\boldsymbol{x}_q + \boldsymbol{\varpi}\boldsymbol{z}_q + \boldsymbol{\varsigma}_q \tag{3.3}$$

where $\boldsymbol{y}$ $(N \times 1)$ is a vector of underlying latent continuous variables (onto which the actual ordinal outcomes are mapped), $\ddot{\boldsymbol{\gamma}}$ $(N \times A)$ captures the effects of exogenous variables expressed as vector $\boldsymbol{x}_q$ $(A \times 1)$ and $\ddot{\boldsymbol{d}}$ $(N \times L)$ is a vector of loadings associated with the latent variables vector $\boldsymbol{z}_q$. Let $\ddot{\boldsymbol{\varepsilon}}_q$ $(N \times 1)$ be specified as the vector of respective error terms and we allow a multi-variate structure such that $\ddot{\boldsymbol{\varepsilon}}_q \sim \mathbf{MVN}\left[\boldsymbol{0}, \boldsymbol{IDEN}_N\right]$, where $\boldsymbol{0}$ $(N \times 1)$ is a zero vector and $\boldsymbol{IDEN}_N$ is an identity vector of dimension N. $\boldsymbol{U}_q$ is the $(\ddot{G} \times 1)$ vector of utilities for the nominal outcomes, $\boldsymbol{b}$ $(\ddot{G} \times A)$ is the matrix of exogenous covariates and ς_q is the error term with zero mean and $\mathbf{\Lambda}$ correlation matrix (refer to Bhat, 2015).

Collect all parameters to be estimated in the SEM and MEM models as an ($R \times 1$) vector $\boldsymbol{\theta}_1$, where $R = (N+F)+(N+A)+(N+L)+(\ddot{G}+A)+(\ddot{G}+L)+\ddot{G}$. The model system (SEM and MEM) is jointly estimated using the Maximum Approximate Composite Marginal Likelihood (MACML) approach (Bhat, 2011).

3.2.3. Capturing Dependency Effects Using a Spatial Lag Structure

This section describes the approach to model the ordinal variable of interest with spatial-attitudinal dependency effects. The ordinal variable has three levels corresponding to usage of shared mobility services: never, occasionally but not in the past 30 days, and one or more times in the past 30 days. The use of a spatial lag structure allows choice behavior of a decision maker to be influenced by that of peers in the geographic-attitudinal space. While proximity in geographic space is derived using spatial distances between residence locations of decision makers, the proximity in attitudinal space is based on a latent social network defined by similarities in attitudes and lifestyle preferences. These are captured by the two latent constructs considered in the study: pro-environment attitude and neo-urban lifestyle propensity. The dependency effects due to each dimension of proximity are disentangled using separate coefficients for each proximity measure.

With the sample of Q decision makers denoted by index $q(1,2,3...,Q)$ and L latent variables denoted by index l (L=2), collect all of the constructs for latent variable l

across all decision makers in the vector $\boldsymbol{z_l} = (z_{l1}, z_{l2}, ..., z_{lQ})'$. Also, let the expected value of this vector, as obtained from the GHDM, be $\boldsymbol{\hat{z}_l}$. The ordinal variable of interest for decision maker q, in the spatial lag structure, is specified in terms of exogenous covariates as follows,

$$\tilde{y}_q = \rho \sum_{q'=1}^{Q} w_{qq'} \tilde{y}_{q'} + \boldsymbol{\tilde{\gamma}' x_q} + \xi_q \ , \ y_q = k \ \text{ if } \psi_{q,k-1} < \tilde{y}_q < \psi_{q,k} \tag{3.4}$$

where $\tilde{y}_q$ is the underlying continuous latent response variable whose partitioning relates to the K levels of the ordinal variable, and $\boldsymbol{\tilde{\gamma}}$ $(A \times 1)$ is the vector of coefficients associated with the $\boldsymbol{x_q}$ $(A \times 1)$ vector of exogenous covariates (excluding the constant). Let the idiosyncratic error term ξ_q be standard normally distributed and independently and identically distributed across decision makers. Let $w_{qq'}$ be the (q,q') element of the row-normalized multi-dimensional weight matrix $\boldsymbol{W}$ $(Q \times Q)$ with zeros on the diagonal ($w_{qq} = 0$, $\sum_{q \neq q'}^{Q} w_{qq'} = 1$) and $\rho\,(0 < \rho < 1)$ be the auto-regressive parameter. In vector notation, the consolidated formulation for all individuals Q is given as,

$$\boldsymbol{\tilde{y}} = \rho \boldsymbol{W\tilde{y}} + \boldsymbol{x\tilde{\gamma}} + \boldsymbol{\xi} \tag{3.5}$$

where $\boldsymbol{\tilde{y}} = (\tilde{y}_1, \tilde{y}_2, ..., \tilde{y}_Q)'$ and $\boldsymbol{\xi} = (\xi_1, \xi_2, ..., \xi_Q)'$ are $(Q \times 1)$ vectors, $\boldsymbol{x}$ is $(Q \times A)$ matrix of exogenous variables for individuals. Through a simple matrix operation, the equation can now be rewritten as:

$$\boldsymbol{\tilde{y}} = \boldsymbol{Tx\tilde{\gamma}} + \boldsymbol{T\xi} \tag{3.6}$$

$$\boldsymbol{T} = (\boldsymbol{I_Q} - \rho \boldsymbol{W})^{-1} \tag{3.7}$$

where $\boldsymbol{I_Q}$ is an identity matrix of size Q. The vector $\boldsymbol{\tilde{y}}$ is multivariate normally distributed with mean $\boldsymbol{Tx\tilde{\gamma}}$ and covariance matrix $\boldsymbol{TT}'$, i.e., $\boldsymbol{\tilde{y}} \sim \boldsymbol{MVN_Q}(\boldsymbol{Tx\tilde{\gamma}}, \boldsymbol{TT}')$.

The crux of this paper lies in the formulation of the composite weight matrix $\boldsymbol{W}$, which engenders the interdependencies amongst decision makers in geographic and attitudinal space. The composite weight matrix is a combination of spatial and non-spatial (one corresponding to each latent construct) weight matrices. Unlike previous formulations (e.g., Yang and Allenby, 2003), the number of constituent weight matrices does not explode with an increasing number of non-spatial measures of proximity. Instead the non-spatial proximity (in attitudinal space) is parsimoniously expressed using a reduced number of latent variable - distance matrices. The composite weight matrix $\boldsymbol{W}$ $(Q \times Q)$ is specified as follows,

$$\boldsymbol{W} = \boldsymbol{exp}(-(\boldsymbol{D}_{spatial} + \sum_{l=1}^{L} \kappa_l \boldsymbol{D}^{l}_{non-spatial})) \tag{3.8}$$

where, $\boldsymbol{D}_{spatial}$ is a $(Q \times Q)$ spatial distance matrix that is derived using latitude-longitude coordinates of decision makers' residential locations. $\boldsymbol{D}^{l}_{non-spatial}$ is the $(Q \times Q)$ non-spatial distance matrix, based on attitudinal proximity on latent variable $l\,(1,2,...,L)$. Further, $\{\kappa_1, \kappa_2, ..., \kappa_L\}$ are coefficients associated with the non-spatial proximity measures derived from each of the L latent variables. The element-by-element exponentiation operator allows for negative values for kappa while still ensuring non-negativity of the final weights. The coefficient associated with spatial distance is fixed to unity to ensure econometric identification.

The non-spatial distance matrix $\boldsymbol{D}^{l}_{non-spatial}$, associated with latent variable l, is populated using a $(Q \times Q)$ matrix $\hat{\mathbb{Z}}_l$ that is expressed as the Kronecker product of $\hat{\boldsymbol{z}}_l$ $(Q \times 1)$ vector of predicted values for latent variable l and a $(1 \times Q)$ row vector of ones. Due to the non-directionality of differences in latent lifestyles and preferences across decision makers, the absolute difference of $\hat{\mathbb{Z}}_l$ with its transpose $\left(\hat{\mathbb{Z}}_l\right)'$ is taken. This results in a $(Q \times Q)$ distance matrix of attitudinal proximity on latent variable l given by,

$$\hat{\mathbb{Z}}_l = \hat{\boldsymbol{z}}_l \otimes ones(1,Q) = \begin{bmatrix} \hat{z}_{1,l} \\ \hat{z}_{2,l} \\ \dots \\ \hat{z}_{Q,l} \end{bmatrix} \otimes \begin{bmatrix} 1 & 1 & \dots & 1 \end{bmatrix}_{1\times Q} \tag{3.9}$$

$$\boldsymbol{D}^{l}_{non-spatial} = |\hat{\mathbb{Z}}_l - (\hat{\mathbb{Z}}_l)'| = \begin{bmatrix} 0 & |\hat{\boldsymbol{z}}_{1,l} - \hat{\boldsymbol{z}}_{2,l}| & \dots & |\hat{\boldsymbol{z}}_{1,l} - \hat{\boldsymbol{z}}_{Q,l}| \\ |\hat{\boldsymbol{z}}_{2,l} - \hat{\boldsymbol{z}}_{1,l}| & 0 & \dots & |\hat{\boldsymbol{z}}_{2,l} - \hat{\boldsymbol{z}}_{Q,l}| \\ \dots & \dots & \dots & \dots \\ |\hat{\boldsymbol{z}}_{Q,l} - \hat{\boldsymbol{z}}_{1,l}| & |\hat{\boldsymbol{z}}_{Q,l} - \hat{\boldsymbol{z}}_{2,l}| & \dots & 0 \end{bmatrix}_{Q\times Q} \tag{3.10}$$

An important note here is that the non-spatial proximity measures among decision agents, as constructed above, are based on the expected values of the latent constructs as opposed to their actual values. The main reason for this formulation is that the sample is but a random fraction of the population of interest. It is impossible to represent every individual in spatial or social space, and therefore more appropriate to consider a sampled neighbor in spatial or social space as representative of many others in the population who may be in that space. It may then be intrinsically more appropriate to consider the *expected value* of a sampled neighbor’s latent construct (representing the larger set of individuals in the population with the same observed characteristics that impact the latent variable of the sampled neighbor), and examine the distance of this expected value from the expected value of the sampled individual in question.

From a methodological standpoint, applying the stochastic values of latent constructs for individuals, in lieu of the expected values, would entail specifying the joint-distribution of latent constructs and underlying propensities for the ordinal outcome of interest. However, hypothesizing a joint-distribution for such a complex system of models may be impractical or difficult to associate with theoretical underpinnings (Murphy and Topel, 2002). Additionally, joint-estimation of a model system with stochastic weights can also be computationally taxing, which motivates the use of a two-step modeling procedure illustrated in this thesis.

The spatial distances matrix ($\boldsymbol{D}_{spatial}$) and the non-spatial distance matrices ($\boldsymbol{D}^{l}_{non-spatial}$) are normalized (divided by the maximum value) before they enter Equation (6) to adjust for scale differences. Prior to feeding the composite weight matrix $\boldsymbol{W}$ into the SORP model (Equation 2), the diagonal elements of $\boldsymbol{W}$ are set to zero and $\boldsymbol{W}$ is row-normalized to ensure that each decision maker gets the same net influence from all other decision makers.

The parameters to be estimated in the ordered probit model with spatial and non-spatial dependencies are the vector of exogenous coefficients $\tilde{\boldsymbol{\gamma}}$, the auto-correlation parameter ρ, $(M-1)$ thresholds of the ordinal variable ($\psi_0 = -\infty, \psi_K = \infty, -\infty < \psi_1 < \psi_2 ... < \psi_{K-1} < \infty$), and κ $\{\kappa_1, \kappa_2, ..., \kappa_L\}$ coefficients associated with the non-spatial weight matrices. The likelihood function $L(\boldsymbol{\theta})$ for the model takes the following form,

$$L(\boldsymbol{\theta}) = P(\boldsymbol{y} = \boldsymbol{m}) = \int_{D_{\tilde{y}}} \boldsymbol{F_Q}(\tilde{\boldsymbol{y}} \,|\, \boldsymbol{Tx}\tilde{\boldsymbol{\gamma}}, \boldsymbol{TT}') d\tilde{\boldsymbol{y}} \qquad (3.11)$$

where $\boldsymbol{\theta} = (\tilde{\boldsymbol{\gamma}}', \rho, \kappa_1, \kappa_2, ..., \kappa_L, \psi_1, \psi_2, ..., \psi_{K-1})'$ is the $((A+L+K)\times 1)$ vector of coefficients to be estimated, $\boldsymbol{y} = \left(y_1, y_2, ..., y_Q\right)$, $\boldsymbol{m} = \left(m_1, m_2, ..., m_Q\right)$ is the $(Q\times 1)$ vector of actual observed level of frequency of using car-sharing and/or ride-sourcing. $D_{\tilde{y}}$ is the domain of integration defined as $D_{\tilde{y}} = \left\{\tilde{y} : \psi_{q,m-1} < \tilde{y}_q < \psi_{q,m},\ \forall q = 1, 2, ..., Q\right\}$. $\boldsymbol{F_Q}(.)$ is the Q-

variate normal cumulative function with mean $\boldsymbol{Tx\tilde{\gamma}}$ and correlation matrix $\boldsymbol{TT}'$. The autoregressive parameter ρ is reparametrized as $\rho = \dfrac{exp(\breve{\rho})}{1 + exp(\breve{\rho})}$ to ensure that $0 < \rho < 1$ and the likelihood function is maximized with respect to $\breve{\rho}$. The true value of ρ can be easily extracted after the estimation process. The likelihood function is maximized using a pair-wise composite marginal likelihood (CML) approach (Bhat, 2011). Dependency effects dilute very quickly as distance between observations increases (Castro et al, 2013). Based on statistical tests discussed in Bhat (2011), a distance threshold of eight miles is adopted and only those pairs of observations falling within this distance band are included in the CML function.

3.2.5. Standard Error Corrections

The two-stage estimation procedure allows for easy implementation of the composite marginal likelihood (CML) inference approach, which is backed by the well-established asymptotic properties of applying exogenous weight matrices in spatial models. The application of expected values of latent constructs, in lieu of actual stochastic values, makes the composite weight matrix essentially exogenous. For such two-stage models, which entail inclusion of predicted values of variables from one model into another, standard errors need to be corrected because the first-stage parameters are themselves estimated with sampling error.

The CML estimators in the second-stage model can be argued to be consistent. The first step estimators, as well as the covariance matrix, from GHDM are asymptotically consistent (Bhat, 2011) and this implies that under the usual regularity conditions, the CML estimators in the second-stage are asymptotically consistent too.

This is because the sampling error for the first-stage estimators vanishes in the limit (as sample size increases). However, the covariance matrix and the associated inference statistics for the second-stage estimators are biased, even with large samples. The standard errors for the second-stage parameter estimates are corrected using the procedure suggested by Murphy and Topel (2012). This procedure exploits the limiting distribution of the sampling error for first-stage estimators to consistently estimate the variances of the second-stage estimators. In this thesis, the GHDM model serves as the first-stage model (auxiliary model) and SORP with multi-dimensional dependencies is subsequently referred to as the second-stage model (model of interest).

Continuing the notations from previous sections, parameters estimated in the first-stage GHDM model are compactly expressed using vector $\boldsymbol{\theta}_1$ ($R \times 1$), and parameters estimated in the second-stage SORP model are represented with vector $\boldsymbol{\theta}_2\ (S \times 1)$, where $S = (A + L + K)$. Let the composite marginal likelihood (CML) values for individual $q\,(1,2,...,Q)$ in the two models be denoted using functions $L_{1,q}(\boldsymbol{\theta}_1)$ and $L_{2,q}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$. Parameters from the first-stage model are reflected in the likelihood function of the second-stage model by virtue of imputed variables (expected values of latent constructs in composite weight matrix) that are exogenously introduced in the second-stage model. Then the two-step CML estimators satisfy the following equations,

$$\sum_{q=1}^{Q} \frac{\partial \log L_{1,q}(\hat{\boldsymbol{\theta}}_1)}{\partial \boldsymbol{\theta}_1} = 0 \tag{3.12}$$

$$\sum_{q=1}^{Q} \frac{\partial \log L_{2,q}(\hat{\boldsymbol{\theta}}_1, \hat{\boldsymbol{\theta}}_2)}{\partial \boldsymbol{\theta}_2} = 0 \tag{3.13}$$

The next step in the process focuses on deriving the asymptotic joint distribution of the two vectors of parameters, using the central limit theorem and the law of large numbers. A priori define four matrices $\boldsymbol{M_1}(R \times R)$, $\boldsymbol{M_2}(S \times S)$, $\boldsymbol{M_3}(R \times S)$ and $\boldsymbol{M_4}(R \times S)$ as follows,

$$\boldsymbol{M_1} = -\mathrm{E}\frac{\partial^2 L_1}{\partial \boldsymbol{\theta_1} \boldsymbol{\theta_1}'} = \mathrm{Hess}(\boldsymbol{\theta_1}) \tag{3.14}$$

$$\boldsymbol{M_2} = -\mathrm{E}\frac{\partial^2 L_2}{\partial \boldsymbol{\theta}_2 \boldsymbol{\theta}_2'} = \mathrm{Hess}(\boldsymbol{\theta}_2) \tag{3.15}$$

$$\boldsymbol{M_3} = \mathrm{E}\frac{\partial L_2}{\partial \boldsymbol{\theta}_1}\left(\frac{\partial L_2}{\partial \boldsymbol{\theta}_2}\right)' = \frac{\partial L_2}{\partial \boldsymbol{\theta}_1}\left(\frac{\partial L_2}{\partial \boldsymbol{\theta}_2}\right)'_{\max} \tag{3.16}$$

$$\boldsymbol{M_4} = \mathrm{E}\frac{\partial L_1}{\partial \boldsymbol{\theta}_1}\left(\frac{\partial L_2}{\partial \boldsymbol{\theta}_2}\right)' = \frac{\partial L_1}{\partial \boldsymbol{\theta}_1}\left(\frac{\partial L_2}{\partial \boldsymbol{\theta}_2}\right)'_{\max} \tag{3.17}$$

Expected values of these matrices can be substituted with matrices evaluated with parameter values at convergence (denoted by subscript max). While $\boldsymbol{M_1}$ and $\boldsymbol{M_2}$ are the hessian matrices of the parameter estimates from the two models, $\boldsymbol{M_3}$ and $\boldsymbol{M_4}$ are obtained through manipulation of gradient functions at convergence. The vector $\frac{\partial L_2}{\partial \boldsymbol{\theta}_1}$ in $\boldsymbol{M_3}$ is vector ($R \times 1$) that takes non-zero gradient values (from second-stage model at convergence) for the parameters $\dot{\boldsymbol{\theta}}_1$ present in both models and zero for parameters $\ddot{\boldsymbol{\theta}}_1$ present only in first-stage model. Note that $\boldsymbol{\theta}_1 = (\dot{\boldsymbol{\theta}}_1, \ddot{\boldsymbol{\theta}}_1)$.

Denoting the true parameter values for first-stage and second-stage parameters as $\boldsymbol{\theta}_1^*$ and $\boldsymbol{\theta}_2^*$ respectively, the central limit theorem can be used to approximate the distribution of the first-order partial derivatives of the two-step log (CML) functions,

$$\begin{bmatrix} \dfrac{1}{\sqrt{Q}} \displaystyle\sum_{q=1}^{Q} \dfrac{\partial \log L_{1,q}(\boldsymbol{\theta}_1^*)}{\partial \boldsymbol{\theta_1}} \\ \dfrac{1}{\sqrt{Q}} \displaystyle\sum_{q=1}^{Q} \dfrac{\partial \log L_{2,q}(\boldsymbol{\theta}_1^*, \boldsymbol{\theta}_2^*)}{\partial \boldsymbol{\theta}_2} \end{bmatrix} \overset{A}{\sim} N(\boldsymbol{0}, \Theta) \tag{3.18}$$

$$\Theta = \begin{bmatrix} \boldsymbol{M_1} & \boldsymbol{M}_4 \\ \boldsymbol{M}_4' & \boldsymbol{M}_2 \end{bmatrix} \tag{3.19}$$

where $\boldsymbol{0}$ is a zero vector $((R+S)\times 1)$ and Θ is $((R+S)\times(R+S))$ covariance matrix. Under the standard assumptions of maximum likelihood estimation, the asymptotic distribution of first-stage and second-stage parameters can be expressed as:

$$\sqrt{Q}\,(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_1^*) \overset{A}{=} -\boldsymbol{M_1}^{-1} \frac{1}{\sqrt{Q}} \sum_{q=1}^{Q} \frac{\partial \log L_{1,q}(\boldsymbol{\theta}_1^*)}{\partial \boldsymbol{\theta_1}} \tag{3.20}$$

$$\sqrt{Q}\,(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_2^*) \overset{A}{=} -\boldsymbol{M}_2^{-1} \frac{1}{\sqrt{Q}} \sum_{q=1}^{Q} \frac{\partial \log L_{2,q}(\boldsymbol{\theta}_1^*, \boldsymbol{\theta}_2^*)}{\partial \boldsymbol{\theta}_2} + \boldsymbol{M}_2^{-1} \boldsymbol{M}_3' \boldsymbol{M}_1^{-1} \frac{1}{\sqrt{Q}} \sum_{q=1}^{Q} \frac{\partial \log L_{1,q}(\boldsymbol{\theta}_1^*)}{\partial \boldsymbol{\theta_1}} \tag{3.21}$$

Using the joint distribution stated in equation (3.18), the asymptotic distribution of the second-stage parameters can be written as,

$$\sqrt{Q}\,(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_1^*) \overset{A}{\sim} N(\boldsymbol{0}, \boldsymbol{\Sigma}) \tag{3.22}$$

$$\boldsymbol{\Sigma} = \boldsymbol{M}_2^{-1} + \boldsymbol{M}_2^{-1}[\boldsymbol{M}_3' \boldsymbol{M}_1^{-1} \boldsymbol{M}_3 - \boldsymbol{M}_4' \boldsymbol{M}_1^{-1} \boldsymbol{M}_3 - \boldsymbol{M}_3' \boldsymbol{M}_1^{-1} \boldsymbol{M}_4] \boldsymbol{M}_2^{-1} \tag{3.23}$$

where $\boldsymbol{0}$ is a zero vector $(S \times 1)$ and $\boldsymbol{\Sigma}$ is $(S \times S)$ corrected asymptotic covariance matrix for the second-stage parameters.

Chapter 4: Model Estimation Results

This section presents a detailed discussion of the model estimation results of the GHDM and Spatial Ordered Response Probit (SORP) model components with various forms of dependency effects among decision-makers. The final model specification was adopted, after testing an extensive number of alternative specifications, based on a combination of behavioral interpretation and statistical significance. In the following sections, results for the SEM and MEM submodel of the GHDM (which determines the latent constructs) and the SORP model component are presented. The MEM submodel is not of primary importance; it simply serves as the vehicle to estimate the SEM submodel by establishing correspondence between latent constructs and their observed indicators.

4.1 Structural Equation Model (SEM) Component of GHDM

Table 4.1 presents estimation results for the SEM component of the GHDM. In general, results are behaviorally intuitive and consistent with expectations. Young people are more likely to be pro-environment and show a proclivity for neo-urban lifestyles. Young adults are more sensitive to the environment as compared to their older counterparts. This is consistent with other studies (Garikapati et al. 2016, Lavieri et al. 2017) that found out young individuals are more likely to use alternative travel modes (transit and non-motorized modes). The impact of age on neo-urban lifestyle propensity is even more profound. Young adults, growing up in realm of ubiquitous presence of technology and emerging sharing economy that eschews ownership and promulgates leasing and renting, and attaching a higher importance to social contact through participation in social and recreational activities are more likely to conform to a neo-urban lifestyle, as compared to their older peers.

Females exhibit a greater sensitivity to the environment, a finding consistent with previous research (Kalof et al, 2002; McCright 2010). Income is strongly related to pro-environmental attitudes, with decision makers in lower income households exhibiting greater levels of the pro-environmental attitude. Lack of wherewithal also restricts the overall level of consumption and induces a higher dependence on alternative modes. Surprisingly, income effects do not significantly influence the likelihood of a decision maker exhibiting a preference for neo-urban lifestyles. A possible explanation maybe counter-acting effects of income in explaining different dimensions of a neo-urban lifestyle – lower incomes are associated with lower levels of technology-ownership and usage (Astroza et al. 2017) but higher inclination to participate in car and bike share programs owing to lower levels of car-ownership. Decision makers with a college education are likely to be pro-environmental and favor active neo-urban lifestyles, consistent with the notion that they are likely to have greater awareness of the ill-effects of pollution. Households with children are more likely to reside in suburban locations in larger homes; consistent with such a lifestyle, individuals in these households express lower levels of the pro-environmental attitude or preference for a neo-urban lifestyle. Also, the flexibility afforded by personal vehicles to chauffeur kids and freedom to undertake complex trips makes them less affine to transit and other active modes (Nolan, 2010).

An interesting finding is that the correlation between error terms is insignificant. The model specification may have captured all key effects, or it is possible that positive and negative correlations due to unobserved effects canceled out.

Table 4.1: Estimation Results for Structural Equation Model of GHDM

Structural Equation Component	**Pro-environment attitude**		**Neo-urban lifestyle propensity**	
Variable	**Coefficient**	**t-stat**	**Coefficient**	**t-stat**
Age (base: 55 + years old)				
18 to 24 years old	0.565	3.12	1.648	4.22
25 to 34 years old	0.374	4.31	1.396	4.55
35 to 44 years old	0.423	4.35	1.208	4.65
45 to 54 years old	0.183	1.99	--	--
Female (base: male)	0.137	2.13	--	--
Education (base: lower than Bachelor's)				
Bachelor's Degree	0.432	5.64	0.489	4.75
Graduate Degree	0.678	7.84	0.500	4.65
Income (base: $75,000 or more per year)				
Less than $24,999 per year	0.552	4.94	--	--
$25,000 - $49,999 per year	0.110	1.34	--	--
$50,000 - $74,999 per year	0.104	1.27	--	--
Employment Status (base: Unemployed)				
Full-time, part-time or self-employed	0.164	2.33	1.032	4.73
Household Structure (base: no kids)				
Atleast 1 kid (0-17 years)	-0.325	-3.79	-0.306	-1.78
Correlation between latent variables	--			

4.2 Measurement Equation Model (MEM) Component of GHDM

Table 4.2 and Table 4.3 provide the results of the measurement equation component of GHDM associated with nominal and ordinal indicator variables respectively. The two nominal variables include smart-phone ownership and membership in car and/or bike-share programs. The ordinal variables include a mix of attitudinal indicators (e.g. importance of being close of transit) and frequency variables that capture the degree of participation in certain activities (e.g. frequency of walking) or usage of facilities (e.g. frequency of using technology for travel information).

Affluent decision makers are more likely to own smart-phones, as reflected in the negative coefficients associated with the lower income groups. Younger decision makers have a higher tendency to own smart-phones as they embrace new technology more readily as compared to their older counterparts (Astroza et al, 2017). Smart-phone owners are more reliant on apps and websites to derive travel information which is evidenced in the positive and significant coefficient of the endogenous effect of smart-phone ownership on frequency of using technology platforms for travel information, consistent with findings for "technophiles" by Seebauer et al (2015). Living in a high-density locality further increases the dependency on such media to plan travel. Prior information about travel times and prevalent traffic conditions can impact travel decisions in multifarious ways – change the tour start and end times, mode opted or number/type of stops on the tour.

Young decision makers are also likely to make decisions on the fly and access to technology can help in planning such impromptu trips and convey the same to others via social media (Astroza et al, 2017). Females are more likely to own smart-phones, as access to smart-phones makes the travel more reliable and brings down uncertainties associated with travelling, especially when travelling alone. A preference for neo-urban lifestyle is an important predictor of smart-phone ownership, as evident in the positively significant loading of the latent construct.

Membership in car and bike-share programs is strongly related to vehicle ownership patterns – decision makers in households without vehicles are more likely to resort to such programs that allow them to experience the convenience and flexibility afforded by vehicle (or bike) ownership sans the burden of ownership, maintenance, insurance etc. These results are akin to those reported by Coll et al. (2014) and Clewlow (2016). Such programs draw more interest from patrons living in dense developments, presumably due to better accessibility to such services and costs associated with vehicle usage (for e.g. parking costs). Students show a greater interest in such programs. An intrinsic preference for a neo-urban lifestyle to eschew ownership and opt for mobility services when needed may attribute to their interest in enrolling for such programs.

Attitudinal indicators are strongly related to the latent lifestyle constructs posited in this study. Pro-environment attitude is associated with a high importance attached to living in a walkable neighborhood with proximity to transit, as can be observed from the positive factor loadings of the pro-environment latent construct. Closely related to these are the higher frequency of walking and bicycling episodes. On the other hand, decision makers conforming to a neo-urban lifestyle gravitate towards residential locations that offer short commutes and instant connectivity to locations of interest. Neo-urbanists, as the more technology affine folk, exhibit more interest in embracing future mobility systems that integrate autonomous driving in a sharing economy setup.

Table 4.2: Estimation Results for Nominal Measurement Equations

Measurement Equation Component for Nominal Indicators					
Latent Variable	**Indicators**	**Constant (t-stat)**		**Factor Loading (t-stat)**	
Neo-urban lifestyle propensity	Smart-phone ownership	0.017	0.61	0.556	5.98
	Car/Bike-share membership	-2.057	-21.30	0.641	6.33

Exogenous variables - smart-phone ownership			**Exogenous variables – car/bike-share membership**		
	Coefficient	***t-stat***		***Coefficient***	***t-stat***
Age (base: 55+ years)			*Vehicles owned (base:0)*		
18 to 34 years old	1.099	7.62	One vehicle	-0.817	-15.82
35 to 44 years old	0.265	4.55	Two or more vehicles	-1.211	-18.02
Income (base: $75,000+)			*Residential Density*		
Below $24,999	-0.939	-18.86	More than 5000 hh/mi^2	0.471	14.55
$25,000 - $49,999	-0.665	-17.91	*Student (base: not a student)*	0.166	3.76
$50,000 - $74,999	-0.430	-14.05			
Female (base: male)	0.290	12.40			

Table 4.3: Estimation Results for Non-Nominal Measurement Equations

Latent Variable	**Indicators**	**Constant (t-stat)**		**Factor Loading (t-stat)**	
Pro-environment attitude	Importance of being close to transit	0.692	6.06	1.066	17.80
	Importance of having walkable neighborhood	1.500	13.78	0.870	33.60
	Frequency of bicycling	-0.617	-3.16	0.343	2.42
	Frequency of walking	1.171	16.00	0.503	16.38
Neo-urban lifestyle propensity	Importance of staying close to major roads/ highways	1.326	18.85	0.105	-2.67
	Importance of being within 30 minutes of commute to work	0.903	12.71	0.281	5.13
	Interest in participating in car-share with fully AV	-0.809	-3.41	0.434	2.36
	Frequency of using technology based platforms for travel info	-0.187	-2.18	0.215	3.51

Endogenous effects on frequency of using technology-based platforms for travel info	**Coefficient**	**t-stat**
Smart-phone ownership (base: no smart-phone)	0.412	4.44
Exogenous effects on frequency of using technology-based platforms for travel info	**Coefficient**	**t-stat**
High Residential Density (more than 5000 hh/mi^2)	0.256	3.93
Age (base: above 35 years old)		
18 to 24 years old	0.650	2.47
25 to 34 years old	0.242	3.93

4.3 Spatial Ordered Response Probit (SORP) Model with Dependency Effects

Table 4.4 presents estimation results for the SORP model with spatial and non-spatial (attitudinal) dependencies. The dependent variable is the frequency of using shared mobility services. For comparison purposes, models with no dependency effects and only spatial dependency effects (autocorrelation) are also presented alongside the SORP model that incorporates multi-dimensional spatial and non-spatial dependencies. The last column indicates the t-stat values obtained through the post-estimation correction procedure stated in section 3.2.5.

In general, more frequent users of these services are young, more educated workers living in high density locations where mobility-on-demand services have higher penetration rates. All of these indications are consistent with findings reported elsewhere in the literature (e.g., Smith 2016, Dias et al 2017). However, decision makers in households with young kids engage in complex tours with multiple stops, making it challenging to effectively manage time, cost and incorporate uncertainties in travel plans through relying on mobility-on-demand services (see Dias et al. 2017). Thus, the presence of kids diminishes the positive effect of staying in high density areas on frequent use of such services. Those who own smartphones are more likely to use shared mobility services; this is presumably because the use of shared mobility services often requires the ownership of a smartphone. Female smartphone owners who use apps fairly regularly for travel information are less likely to use shared mobility services, possibly due to safety considerations and the consistent finding reported in the literature that females carry a greater burden of chauffeuring and household maintenance activities, thus engendering greater levels of trip chaining and joint travel (Garikapati et al, 2014). Such travel patterns are not as conducive to shared mobility service usage. Higher levels of vehicle ownership are associated with lower levels of shared mobility service use frequency, a finding that is consistent with expectations and prior literature (Coll et al, 2014).

What is particularly noteworthy is that the model coefficients differ in magnitude among the model forms. This suggests that the use of models that do not account for dependencies may offer erroneous forecasts and estimates of policy impacts. Also, the auto-correlation term, ρ, is statistically significant in both models. In addition, parameters representing social dependency arising from proximity in the attitudinal space are also statistically significant for both attitudinal constructs considered in this thesis. They are positive in value, suggesting that diffusion effects are at play. Taken together, and comparing coefficients on the non-spatial proximity contributions with the normalized value of one for the spatial proximity contribution, the net result is that both social and spatial proximity contributions are important, statistically significant, and of the same order of magnitude in diffusion effects. As more people use shared mobility services, the more visible they become to the rest of the population – both from a spatial perspective and a social (attitudinal and lifestyle) perspective.

4.4 Model Fit

Model selection procedures allow for statistical comparison of model fit between different model specifications, typically between nested models where one model is a restricted version of the other. Analogous to the log-likelihood ratio test (LRT) statistic used for comparing models using ordinary maximum likelihood estimation, the composite likelihood ratio test (CLRT) statistic is an appealing statistic for models estimated using composite marginal likelihood (CML) approach. The CLRT statistic can be expressed as,

$$CLRT = 2[logL_{CML}(\hat{\boldsymbol{\theta}}) - logL_{CML}(\hat{\boldsymbol{\theta}}_{\boldsymbol{0}})] \tag{4.1}$$

where $\boldsymbol{L_{CML}}(\hat{\boldsymbol{\theta}})$ and $\boldsymbol{L_{CML}}(\hat{\boldsymbol{\theta}}_{\mathbf{0}})$ are the CML values for the restricted and unrestricted models respectively. $\hat{\boldsymbol{\theta}}$ is the CML estimator for the unrestricted model and $\hat{\boldsymbol{\theta}}_{\mathbf{0}}$ is the CML estimator for the restricted model. Although the construction and calculation of the CLRT statistic is straight-forward, it does not have a standard chi-squared asymptotic distribution. This can be traced to the fact that the CML function does not correspond to the parametric model from which the data originates. Pace et al (2011) proposed a parameterization invariant adjustment to the CLRT statistic that yields a statistic that is asymptotically chi-squared distributed with known degrees of freedom. The present study uses the adjusted CLRT statistic to compare the fit between the SORP with multi-dimensional dependencies with its restrictive versions (only spatial dependency and no dependency).

Table 4.5 presents the results of the data fit comparisons between the different models. The number of parameters and composite log-likelihood (CLL) values for each of the three models are provided in the second and third rows, while the fourth row shows the results of the ADCLRT tests which clearly shows the superior performance of the SORP with multi-dimensional dependencies over its restrictive versions. The ADCLRT computations yield χ^2 statistics that are statistically significant at any level of confidence, demonstrating the importance of accounting for multi-dimensional dependency effects in activity-travel choice models.

Table 4.4: SORP Model with Spatial and Non-Spatial Dependencies

Exogenous effects on frequency of using ride-sourcing and/or car-sharing in past 30 days	Aspatial ORP		SORP with Spatial Dependencies Only		SORP with Spatial & Non-Spatial Dependencies		
	Coefficient	t-stat	Coefficient	t-stat	Coefficient	t-stat	t-stat (corrected)
Age (base: 45 or more years)							
18 to 24 years old	0.881	4.49	0.906	10.97	0.598	5.26	5.31
25 to 34 years old	0.661	6.90	0.777	16.11	0.492	5.04	5.11
35 to 44 years old	0.527	5.41	0.573	12.05	0.336	3.77	3.85
Work Status							
Full-time, part-time or self-employed *(base: unemployed)*	0.381	3.95	0.432	9.02	0.227	3.37	3.49
Student *(base: not a student)*	0.253	1.93	0.172	2.94	0.254	12.95	12.24
Income (base: above $100,000)							
Below $25,000	-0.684	-4.57	-0.445	-6.57	-0.741	-19.34	-17.58
$25,000 - $49,999	-0.581	-5.10	-0.513	-8.49	-0.736	-23.31	-20.33
$50,000 - $74,999	-0.366	-3.26	-0.194	-3.30	-0.385	-19.21	-17.25
$75,000 - $99,999	-0.397	-3.51	-0.156	-2.59	-0.318	-17.27	-16.17
Educational attainment (base: less than a bachelor's degree)							
Bachelor's degree	0.386	3.78	0.217	4.10	0.184	4.87	4.96
Graduate degree	0.430	4.00	0.249	4.66	0.182	4.13	4.21
Smart-phone ownership and frequency of usage for travel information in past 30 days (base: don't own a smart-phone)							
Own smart-phone but never use apps	0.881	6.42	0.923	14.25	0.959	22.52	19.65
Own smart-phone and use apps less than once a week	0.800	5.77	0.799	12.22	0.834	22.81	20.08
Own smart-phone and use apps once or more a week	1.080	7.62	1.079	16.49	1.116	23.31	20.26
Own smart-phone and use apps once or more a week x Female	-0.249	-2.28	-0.222	-4.27	-0.263	-14.16	-13.37
Residential Location Density (base: Low Density)							
High Density	0.694	7.64	0.246	5.83	0.497	25.40	22.61
High Density x Presence of atleast one kid	-0.416	-2.15	-0.293	-2.95	-0.300	-11.25	-10.82
Vehicle Ownership and Residence Type (base: no vehicles)							
One Vehicle and single-family residence	-0.505	-3.51	-0.344	-4.85	-0.606	-21.45	-18.99
Two or more Vehicles and single-family residence	-1.207	-8.32	-0.892	-12.58	-1.259	-26.21	-21.85
One Vehicle and multi-family residence	-0.653	-5.43	-0.564	-10.57	-0.648	-23.19	-20.01
Two or more Vehicles and multi-family residence	-0.608	-3.79	-0.383	-5.41	-0.597	-20.69	-18.27
ρ	--	--	0.562	2.24	0.895	2.02	1.96
κ_1 (pro-environment attitude)	--	--	--	--	0.883	2.79	2.66
κ_2 (neo-urban lifestyle propensity)	--	--	--	--	1.151	2.53	2.50

Table 4.5: Measures of fit

Summary Statistic	Aspatial ORP	SORP with Spatial Dependencies Only	SORP with Spatial & Non-Spatial Dependencies
Number of observations	2170		
Number of parameters	23	24	26
Composite log-likelihood (CLL) at convergence	-837,319	-680,959	-637,788
Adjusted composite likelihood ratio test (ADCLRT) between SORP with Spatial & Non-Spatial Dependencies and corresponding model	581.99 > Chi-Squared statistics with 3 degrees of freedom at any reasonable level of significance	125.92 > Chi-Squared statistics with 2 degrees of freedom at any reasonable level of significance	Not Applicable

Chapter 5: Discussion and Conclusions

Decision makers interact with one another as an inevitable part of living in a society. People observe what others do, interact and exchange information with others, and modify their own behaviors, choices, attitudes, and goals in response to societal forces. Yet, many travel models continue to ignore the forces of inter-dependency when simulating activity-travel choices. Models (largely in the research domain) that recognize inter-dependency are often limited to accounting for intra-household interactions among family members. Models that purport to capture influences beyond the immediate confines of the household do so through spatial dependency effects that are purely based on measures of geographic proximity. For example, people may purchase environmentally friendly vehicles, bicycle and walk, use transit, or let their children walk to school in response to observing what their neighbors do and interacting with them.

However, in an era of social media platforms and ubiquitous connectivity, inter-dependencies may no longer be solely influenced by geographic proximity. Rather, the strength of association among decision makers may be influenced by attitudes, values, preferences, and perceptions. Those with similar attitudes and lifestyle preferences may interact more closely (for example, in online communities and forums), thus enhancing social dependency effects among such decision makers who share comparable perspectives.

This thesis makes a fundamental contribution to the literature by proposing an econometric methodology that is capable of simultaneously accounting for both spatial and non-spatial (attitudinal) dependency effects. The model system takes the form of a simultaneous equations model system with latent constructs that describe decision maker attitudes and lifestyle preferences as a function of measured indicators in survey data. The proximity among decision makers with respect to the latent constructs is explicitly incorporated (along with spatial measures of separation) into the weight matrix that captures the strength of association across observations. The formulation is able to disentangle the strength of the inter-dependency due to attitudinal proximity from that due to spatial proximity.

The model system is applied to the study of the frequency of use of shared mobility services, including car-sharing and ride-sourcing services. Two latent constructs, representing pro-environmental attitude and preference for a neo-urban lifestyle, are used to account for non-spatial dependency effects. A spatially ordered response model (SORP) is estimated within a larger Generalized Heterogeneous Data Model (GHDM) framework to examine the dependency effects. It is found that both spatial and non-spatial (attitudinal) dependency effects are significant in explaining the use of emerging shared mobility services and that both of these effects are comparable in magnitude. The model that accounted for both sources of dependency offered statistically better goodness-of-fit than models that ignored one or both sources.

The model system shows that diffusion effects are at play, not just based on distance but also based on non-spatial attitudinal and lifestyle variables. Such models can help in developing estimates of market adoption of emerging transportation technologies as they capture the diffusion effects engendered by multiple sources. Policy strategies aimed at enhancing shared mobility service usage can be better informed via models that capture various inter-dependency effects. Agencies interested in seeing greater adoption of these services could identify virtual groups and forums that may be targeted for information campaigns, incentives and rebates, and seeking assistance in spreading the word. Through such mechanisms, agencies may be able to realize significant change in behavior in response to various strategies by leveraging the power of diffusion effects that influence people's activity-travel choice behaviors.

References

Ajzen, I. (1991) The Theory of Planned Behaviour. *Organizational Behavior and Human Decision Processes*, 50, 179-211.

Anable, J. (2005) 'Complacent Car Addicts' or 'Aspiring Environmentalists'? Identifying Travel Behaviour Segments Using Attitude Theory. *Transport Policy*, 12(1), 65-78.

Arentze, T.A. and H.J.P. Timmermans (2008) Social Network, Social Interaction, and Activity-Travel Behavior: A Framework for Microsimulation. *Environment and Planning B,* 35(6), 1012-1027.

Arentze, T.A. and H.J.P.Timmermans (2009) A Need-Based Model of Multi-Day, Multi-Person Activity Generation. *Transportation Research Part B*, 43(2), 251-265.

Astroza, S., V.M. Garikapati, C.R. Bhat, R.M. Pendyala, P.S. Lavieri, and F. Dias (2017) Analysis of the Impact of Technology Use on Multi-Modality and Activity-Travel Characteristics. *Transportation Research Record*, *Journal of the Transportation Research Board*, (*in press*).

Axhausen, K.W. (2008) Social Networks, Mobility Biographies, and Travel: Survey Challenges. *Environment and Planning B*, 35(6), 981-996.

Bagley, M.N. and P.L. Mokhtarian (2002) The Impact of Residential Neighborhood Type on Travel Behavior: A Structural Equation Modeling Approach. *Annals of Regional Science*, 36, 279–297.

Baptista, P., S. Melo, and C. Rolim (2014) Energy, Environmental and Mobility Impacts of Car-Sharing Systems. Empirical Results from Lisbon, Portugal. *Procedia-Social and Behavioral Sciences*, 111, 28-37.

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 New Spatial (Social) Interaction Discrete Choice Model Accommodating for Unobserved Effects due to Endogenous Network Formation, *Transportation*, 42(5), 879-914

Bhat, C.R., A.R. Pinjari, S.K. Dubey, and A.S. Hamdi (2016) On Accommodating Spatial Interactions in a Generalized Heterogeneous Data Model (GHDM) of Mixed Types of Dependent Variables. *Transportation Research Part B*, 94, 240-263.

Blume, L. and S. Durlauf (2003) Equilibrium Concepts for Social Interaction Models. *International Game Theory Review*, 5(3), 193-209.

Brock, W. and S. Durlauf (2001) Discrete Choice with Social Interactions. *Review of Economic Studies*, 68, 235-260.

Burkhardt, J. and A. Millard-Ball (2006) Who's Attracted to Car-Sharing?. *Transportation Research Record: Journal of the Transportation Research Board*, 1986, 98-105.

Carrasco, J.A., B. Hogan, B. Wellman, and E.J. Miller (2008) Collecting Social Network Data to Study Social Activity-Travel Behavior: An Egocentric Approach. *Environment and Planning B*, 35(6), 961-980.

Castro, M., R. Paleti, and C.R. Bhat (2013) A Spatial Generalized Ordered Response Model to Examine Highway Crash Injury Severity. *Accident Analysis and Prevention*, 52, 188-203.

Clewlow, R.R. (2016) Carsharing and Sustainable Travel Behavior: Results from the San Francisco Bay Area. *Transport Policy*, 51, 158-164.

Coll, M.H., M.H. Vandersmissen, and M. Thériault (2014) Modeling Spatio-Temporal Diffusion of Carsharing Membership in Québec City. *Journal of Transport Geography*, 38, 22-37.

Costain, C., C. Ardron, and K. N. Habib. (2012) Synopsis of Users' Behaviour of a Carsharing Program: A Case Study in Toronto. *Transportation Research Part A*, Vol. 46, No. 3, pp. 421- 434.

Dias, F.F., P.S. Lavieri, V.M. Garikapati, S. Astroza, R.M. Pendyala, and C.R. Bhat (2017) A Behavioral Choice Model of the Use of Car-Sharing and Ride-Sourcing Services. *Transportation (in press).*

Dugundji, E.R. and J.L.Walker (2005) Discrete Choice with Social and Spatial Network Interdependencies. *Transportation Research Record: Journal of the Transportation Research Board*, 1921, 70-78.

Efthymiou, D., C. Antoniou, and P. Waddell (2013) Factors Affecting the Adoption of Vehicle Sharing Systems by Young Drivers. *Transport Policy*, 29, 64-73.

Firnkorn, J., and M. Müller (2011). What Will be the Environmental Effects of New Free-Floating CarSharing Systems? The Case of Car2go in Ulm. *Ecological Economics*, Vol. 70, No. 8, pp. 1519-1528

Fishbein, M. (1980) A Theory of Reasoned Action: Some Applications and Implications. In H. Horve, Jr. (ed.) *Nebraska Symposium on Motivation*, 65-116.

Garikapati, V. M., R. M. Pendyala, E. A. Morris, P. L. Mokhtarian, and N. McDonald (2016). Activity Patterns, Time Use, and Travel of Millennials: A Generation in Transition? *Transport Reviews*, Vol. 36, No. 3, pp. 558-584.

Garikapati, V.M., D. You, R.M. Pendyala, P. Vovsha, V. Livshits, and K. Jeon (2014) Multiple Discrete-Continuous Model of Activity Participation and Time Allocation for Home-Based Work Tours. *Transportation Research Record: Journal of the Transportation Research Board*, 2429, 90-98.

Hackney, J.K. and K.W. Axhausen (2006) An Agent Model of Social Network and Travel Behavior Interdependence. Presented at the 11th International Conference on Travel Behaviour Research, Kyoto, August.

Hannon, E., C. McKerracher, I. Orlandi, and S. Ramkumar (2016) An Integrated Perspective on the Future of Mobility. McKinsey & Company, October, http://www.mckinsey.com/business-functions/sustainability-and-resource-productivity/our-insights/an-integrated-perspective-on-the-future-of-mobility. Accessed July 27, 2017.

Kalof, L., T. Dietz, G. Guagnano, and P.C. Stern (2002) Race, Gender and Environmentalism: The Atypical Values and Beliefs of White Men. *Race, Gender and Class*, 9(2), 112-130.

Kim, K. (2015) Can Carsharing Meet the Mobility Needs for the Low-Income Neighborhoods? Lessons from Carsharing Usage Patterns in New York City. *Transportation Research A*, 77, 249-260.

Kitamura, R., P.L. Mokhtarian, and L. Laidet (1997) A Micro-Analysis of Land Use and Travel in Five Neighborhoods in the San Francisco Bay Area. *Transportation*, 24, 125-158.

Lavieri, P., V.M. Garikapati, C.R. Bhat, R.M. Pendyala, S. Astroza, and F.F. Dias (2017) Modeling Individual Preferences for Ownership and Sharing of Autonomous Vehicle Technologies. *Transportation Research Record: Journal of the Transportation Research Board (in press)*.

Lyft (2017) http://www.businessinsider.com/lyft-tripled-its-rides-in-2016-2017-Accessed May 15th, 2017

McCright, A.M. (2010) The Effects of Gender on Climate Change Knowledge and Concern in the American Public. *Population and Environment*, 32(1), 66-87.

McPherson, M., L. Smith-Lovin, and J.M. Cook (2001) Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, 27(1), 415-444.

Metcalfe, G., and J. Warburg (2012) A Policy Agenda for the Sharing Economy. *The Urbanist*, October. http://www.spur.org/publications/urbanist-article/2012-10-09/policy-agenda-sharing-economy. Accessed August 1, 2017.

Mokhtarian, P.L., Cao, X. (2008). Examining the impacts of residential self-selection on travel behavior: a focus on methodologies. *Transportation Research Part B* 42(3) 204-208.

Montgomery, M. (2015) Millennials Insist the Future Will be Shared. *TechCrunch*, March 2015. https://techcrunch.com/2015/03/11/millennials-the-future-will-be-shared/. Accessed July 27, 2017.

Murphy, K.M. and Topel R.H. (2002) Estimation and Inference in Two-Step Econometric Models. *Journal of Business & Economic Statistics*, 20:1, 88-97.

New Urbanite Study (2016) A study of city dwellers in America and their feelings on everything from the cost of their rent to the character of their cities, http://www.zipcar.com/research/the-new-urbanite-and-city-living. Accessed July 29, 2017

Nolan, A. (2010) A Dynamic Analysis of Household Car Ownership. *Transportation Research Part A*, Vol. 44, No. 6, 2010, pp. 446-455

Pace, L., Salvan, A., Sartori, N., 2011. Adjusting composite likelihood ratio statistics. *Statistica Sinica* 21(1), 129-148.

Paleti, R., C.R. Bhat, R.M. Pendyala, and K.G. Goulias (2013) Modeling of Household Vehicle Type Choice Accommodating Spatial Dependence Effects. *Transportation Research Record: Journal of the Transportation Research Board*, 2343, 86-94.

Pew Research Center (2014) The Web at 25. http://www.pewinternet.org/2014/02/25/the-web-at-25-in-the-u-s. Accessed August 1, 2017.

PSRC (2015) *Puget Sound Regional Travel Study*: *Household Travel Survey*, http://www.psrc.org/assets/14106/2015-Household-Survey-Tech-Report.pdf. Accessed May 12, 2017.

Rayle, L., D. Dai, N. Chan, R. Cervero, and S. Shaheen (2016) Just a Better Taxi? A Survey-Based Comparison of Taxis, Transit, and Ridesourcing Services in San Francisco. *Transport Policy*, 45, 168-178.

Ronis, D.L., J.F. Yates, and J.P. Kirscht (1989) Attitudes, Decisions, and Habits as Determinants of Repeated Behaviour. In A.R. Pratkanis, S. J. Breckler, and A.G. Greenwald (eds) *Attitude Structure and Function*, Lawrence Erlbaum, Hillsdale, N.J., 213-239.

Sabatini, J. (2014). SF exploring ways to regulate ride services like Uber, Lyft. *San Francisco Examiner*

Seebauer, S., R. Stolz, and M. Berger (2015) Technophilia as a Driver for Using Advanced Traveler Information Systems. *Transportation Research Part C*, Vol. 60, 2015, pp. 498-510

Shaheen, S. A. (2016). Mobility and the Sharing Economy. *Transport Policy*, Vol. 51, pp. 141- 142

Shaheen, S., A. Cohen, and M. Chung (2009) North American Carsharing: 10-year Retrospective. *Transportation Research Record: Journal of the Transportation Research Board*, 2110, 35-44.

Silver, N. and R. Fischer-Baum (2015) Public Transit Should be Uber's New Best Friend. *FiveThirtyEight*, August. https://fivethirtyeight.com/features/public-transit-should-be-ubers-new-best-friend/. Accessed August 1, 2017.

Smith, A. (2016) Shared, Collaborative and On Demand: The New Digital Economy. *Pew Reserach Center*, May. http://www.pewinternet.org/2016/05/19/the-new-digital-economy/. Accessed July 15, 2017.

Sunrun (2013) New Survey Reveals 'Disownership' is the New Normal https://www.sunrun.com/why-sunrun/about/news/press-releases/new-survey-reveals-disownership-new-normal . Accessed August 1, 2017.

Uber (2017) https://newsroom.uber.com/page/2/ Accessed May 15th, 2017

Van Acker, V., B. Van Wee, and F. Witlox (2010) When Transport Geography Meets Social Psychology: Toward a Conceptual Model of Travel Behaviour. *Transport Reviews*, 30(2), 219-240.

van Wee, B., Holwerda, H., van Baren, R. (2002). Preferences for modes, residential location and travel behaviour: the relevance for land-use impacts on mobility. *European Journal of Transport and Infrastructure Research* 2 (3/4), 305–316.

Yang, S. and G.M. Allenby (2003) Modeling Interdependent Consumer Preferences. *Journal of Marketing Research*, 40(3), 282-294.