**Investigating objective and subjective factors influencing the adoption, frequency, and characteristics of ride-hailing trips**

**Patrícia S. Lavieri**
The University of Melbourne
Department of Infrastructure Engineering
Grattan Street, Parkville, Victoria, 3010, Australia
Tel: +61-3-9035-3274; Email: patricia.lavieri@unimelb.edu.au
and
The University of Texas at Austin, Austin, TX 78712, USA

**Chandra R. Bhat (corresponding author)**
The University of Texas at Austin
Department of Civil, Architectural and Environmental Engineering
301 E. Dean Keeton St. Stop C1761, Austin, TX 78712, USA
Tel: 1-512-471-4535; Email: bhat@mail.utexas.edu
and
The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

## ABSTRACT

Even as ride-hailing has become ubiquitous in most urban areas, its impacts on individual travel are still unclear. This includes limited knowledge of demand characteristics (especially for pooled rides), travel modes being substituted, types of activities being accessed, as well as possible trip induction effects. The current study contributes to this knowledge gap by investigating ride-hailing experience, frequency, and trip characteristics through two multi-dimensional models estimated using data from the Dallas-Fort Worth Metropolitan Area. Ride-hailing adoption and usage are modeled as functions of unobserved lifestyle stochastic latent constructs, observed transportation-related choices, and sociodemographic variables. The results point to low residential location density and people's privacy concerns as the main deterrents to pooled ride-hailing adoption, with non-Hispanic Whites being more privacy sensitive than individuals of other ethnicities. Further, our results suggest a need for policies that discourage the substitution of short-distance "walkable" trips by ride-hailing, and a need for low cost and well-integrated multi-modal systems to avoid substitution of transit trips by this mode.



*Keywords:* Ride-hailing, pooling, travel behavior, market adoption and use of disruptive mobility services, psycho-social latent variables, GHDM model.

# 1. INTRODUCTION

Ride-hailing services have experienced a rapid growth in the past few years and are currently operating in hundreds of cities around the world. In 2018, the two largest ride-hailing companies served together more than 45 million trips per day (Schlobach and Retzer, 2018), and the overall penetration rate of this service reached 9% of the global market (Statista, 2019).[1] Despite the considerable traction that ride-hailing is continually gaining in many urban areas, it is still unclear whether such services bring benefits or drawbacks to urban transportation. For example, on the positive side, pooled ride-hailing services have the potential to increase vehicle occupancy, while still offering convenience to its users.[2] This could result in an increase in vehicle occupancy and a reduction of vehicle miles traveled (VMT) in cities with current predominance of drive alone trips. Ride-hailing may also serve as an accessibility enhancer for those who cannot drive or do not own vehicles (see, for example, Leistner and Steiner, 2017 and Lavieri et al., 2018a). On the other hand, the convenience of hailing a ride through a smartphone app may reduce transit ridership and active travel or induce the generation of new trips that would not be undertaken otherwise, and consequently increase motorized travel (Rayle et al., 2016).

A decisive evaluation of ride-hailing effects on the behavior of travelers is still far from being achieved as the service and market are dynamically changing. More companies continue to join the market adopting different marketing, fare, and service strategies to attract customers, and the general population is still gaining familiarity with the service. It is no surprise then that the past years have seen multiple, even if limited, efforts in the literature focused on shedding light to the ride-hailing phenomenon, as discussed in the next section.

## 1.1 Ride-Hailing Adoption, Use and Impacts on Travel Behavior

The scarcity of publicly available data on ride-hailing is among the main reasons for the currently limited research on travel behavior considerations associated with ride-hailing. In particular, although ride-hailing companies collect user information and detailed trip characteristics, such data are usually not publicly released due to privacy concerns and proprietary value. Thus, the majority of the studies on ride-hailing demand rely on specialized user surveys (Rayle et al., 2016; Leistner and Steiner, 2017), online surveys (Smith, 2016; Clewlow and Mishra, 2017; Alemi et al., 2018a; Alemi et al., 2018b; Hampshire et al., 2018), the limited information available in large-scale household travel surveys (Dias et al., 2017; Lavieri et al., 2017), or aggregate time series data of transportation related characteristics (Li et al., 2017; Ward et al., 2018). Just recently, some studies have examined ride-hailing considerations and

[1] Ride-hailing refers to ride services provided by a loose collection of drivers to the wider public through a smartphone application.

[2] Ride-hailing services can be hired in a pooled mode, in which the user accepts to share a ride with strangers in exchange for a cheaper fare.

impacts based on big data (Kooti et al., 2017), or data from ride-hailing companies (Gerte et al., 2018; Komanduri et al., 2018; Lavieri et al., 2018a; Zheng et al., 2018), but these data sets have limited to no user information.

The aforementioned diversity of data sources has led to the use of a variety of analytic frameworks to investigate ride-hailing behavior and its impacts. Many earlier survey-based ride-hailing studies have been of an exploratory nature, relying on descriptive statistics, though a handful of studies have developed individual-level ordered-response models to examine ride-hailing frequency (see Dias et al., 2017 and Alemi et al., 2018a), and Lavieri et al. (2017) incorporated the binary choice representing ride-hailing experience as one of multiple travel behavior dimensions within a stochastic latent constructs approach. The earlier non-survey ride-hailing studies that do not have data at the individual user level have typically aggregated trip-level information to spatial units, and appended demographic information on the spatial units from Census data, to relate aggregate demand to spatial-level demographics. Examples of such studies include Gerte et al. (2018) and Lavieri et al. (2018a), with the former using a random effects linear panel model to evaluate the variation in generation of ride-hailing trips per taxi zone within a 49-week period and the latter using a spatially lagged bivariate count model to examine the average number of ride-hailing trips generated per traffic analysis zone on week and weekend days. In related non-survey studies, Li et al. (2017) and Ward et al. (2018) utilized a “difference in differences” regression to model aggregate time series data of transportation related variables, including vehicle ownership, miles traveled and congestion levels before and after ride-hailing services penetrated the market.

The many earlier studies listed above have provided important insights regarding individual-level and trip-level ride-hailing behavior. In terms of individual-level behavior, and notwithstanding the differences in the types of data and analysis techniques used in earlier studies, some common themes have emerged. For example, the adoption of ride-hailing seems to be more prevalent among young adults (18-30 years old) who live in urban areas, have a college education, and are in medium-to-high income segments (Smith, 2016; Clewlow and Mishra, 2017; Dias et al., 2017; Kooti et al., 2017). Kooti et al. (2017) further observed that, although older individuals use ride-hailing services less frequently than their younger counterparts, older individuals tend to make longer ride-hailing trips and choose services that are more expensive. Additionally, some studies have noted that the majority of ride-hailing users own personal vehicles (Smith, 2016; Dias et al., 2017), but vehicle ownership is lower among frequent users (Smith, 2016, Alemi et al., 2018a).

With regard to trip-level ride-hailing behavior, the few studies that have collected information on trip purpose have identified that social and recreational trips are the most common purposes for ride-hailing (Rayle et al., 2016; Hampshire et al., 2018). In terms of travel modes from which ride-hailing draws patronage, earlier studies have indicated that ride-hailing often substitutes for taxi trips. The ease of payment, the ease to call, the lower cost, and the shorter wait times are frequently cited by individuals as reasons to use ride-hailing relative to taxis. However, even if to a lesser degree, ride-hailing also draws away from public transit trips, active mode trips (that is, trips made by bicycling and walking), and solo-driving trips, at least as reported by survey respondents (Rayle et al., 2016; Alemi et al., 2018b; Alemi et al., 2018c; Zheng et al., 2018). Shorter travel times are identified as the primary reason to prefer ride-hailing over public transit and active modes, while limited parking at the destination and avoiding driving while intoxicated are the typical reasons provided for preferring ride-hailing over driving a private car. There also is some evidence of newly generated ride-hailing trips that would not have been made otherwise. In particular, two studies from California (U.S.), one focusing on San Francisco (see Rayle et al., 2016) and the other covering multiple metropolitan areas (see Alemi et al., 2018b), found that about 8.0% of users would not have made their most recent ride-hailing trip if this mode were not available.

Aggregate analyses comparing gasoline consumption, VMT, and traffic congestion before and after ride-hailing market penetration in U.S cities have not identified any overall increase in vehicle usage per-capita; on the contrary, they have observed a slight decrease (Li et al., 2017; Ward et al., 2018). Ward et al. (2018) also observed a decrease in per-capita vehicle registration rates over the past few years, coinciding with the period of substantial ride-hailing use increase. Still, localized effects of increase in congestion, especially due to empty ride-hailing cars (only with drivers), were observed in dense activity centers such as downtown Manhattan, NY (Schaller, 2017). Such differences in results suggest that impacts of ride-hailing on urban travel may be heterogeneous depending on built environment and population characteristics.

The overview above indicates that earlier studies have contributed to our understanding of ride-hailing demand, yet the overall knowledge base about ride-hailing behavior and impacts is still limited in many ways. First, people's propensity to choose pooled ride-hailing services has received relatively little attention, and possible positive or negative externalities associated with its adoption have not been sufficiently examined. Second, there has been little discussion directed to the examination of different uses of ride-hailing services and their associated implications. Third, from an analytic standpoint, there is a need for more multivariate analyses that simultaneously control for the effects of multiple variables, including lifestyle-related users' characteristics and perceptions. Fourth, it is likely that the effects of ride-hailing in transit-rich cities (the case of most areas investigated by existing studies) and other cities with high levels of car dominance differ. Thus, research efforts in multiple cities and regions are important. Finally, as ride-hailing continues to expand rapidly with more companies joining the market (adopting different marketing, fare, and service strategies to attract customers), and the general population gains more familiarity with the service, it is important that ride-hailing investigations be undertaken over time.

### 1.2 The Current Study

The current study contributes to the gaps identified earlier by modeling ride-hailing experience, frequency, and trip characteristics in the Dallas-Fort Worth-Arlington Metropolitan Area (DFW) of Texas, U.S. We develop two multi-dimensional models of ride-hailing behavior, one at an individual-level and the second at a trip-level. In the first individual-level model, we propose a close look into the characteristics of current users and non-users of ride-hailing and pooled ride-hailing services by jointly modeling ride-hailing experience and frequency as functions of unobserved lifestyle stochastic latent constructs, observed transportation-related choices, and sociodemographic variables. In the second trip-level model, four nominal dimensions of the individual's last ride-hailing trip are modeled simultaneously: trip purpose, time-of-day, companionship, and mode substituted. In combination, the results from the two multivariate models developed in this paper allow the identification of behavioral differences regarding ride-hailing use across population segments with varying lifecycle and lifestyle conditions, and contribute to the discussion of four important issues: (1) people's acceptance and use of pooled rides; (2) the use of ride-hailing as an accessibility versus a convenience mobility tool; (3) the relationship of this mode with transit and active travel; (4) and the potential latent demand and trip induction generated by this service.

The data used in the study is drawn from an online survey, developed and administered by the authors in the fall of 2017, of commuters in the Dallas Fort Worth (DFW) metro area. DFW is the largest metropolitan area in Texas in terms of population and the fourth largest in the U.S. It has more than 7.4 million inhabitants and is the fastest growing metropolitan area in the country (U.S. Census Bureau, 2018a). In contrast to the majority of cities investigated in previous studies (such as San Francisco by Rayle et al., 2016; Boston, Chicago, New York, San Francisco, Los Angeles, and Washington, D.C by Clewlow and Mishra, 2017; New York by Gerte et al., 2018; and Seattle by Dias et al., 2017 and Lavieri et al., 2017), DFW is a car-dominated urban area where more than 81% of commute trips are undertaken using the drive alone mode and another 10% are pursued by a private vehicle car even if not alone. Public transit accounts for only 2% of the overall commute mode share (U.S. Census Bureau, 2018b). Additionally, while the national average for zero-vehicle households is 9.0%, the DFW average is 5.0% (U.S. Census Bureau, 2018c). Overall, considering that there may be differences in the way ride-hailing may be incorporated within people's travel plans when in a car-dominated environment relative to a multimodal environment, the current study brings a new perspective to the ride-hailing literature.

The remainder of this paper is organized as follows. The next section contains a detailed description of the analytic framework and data, including the description of the conceptual and methodological aspects of the models. Section 3 and 4 describe the results of the individual-level experience and frequency model, and the trip-level characteristics model, respectively. Policy implications are discussed with the conclusions in the final section.

## 2. DATA AND ANALYTIC FRAMEWORK

### 2.1 Survey

The data used for the analysis was obtained through a web-based survey. The distribution was achieved through mailing lists held by multiple entities (local transportation planning organizations, universities, private transportation sector companies, non-profit organizations, and online social media), yielding a final clean convenience sample of 1,607 respondents. To focus on individuals with commute travel, the survey was confined to individuals who had their primary work place outside their homes.[3] Respondents were presented with the definition of ride-hailing as “Ride-hailing services use websites and mobile apps to pair passengers with drivers who provide passengers with transportation in the driver's non-commercial vehicle. Examples are Uber and Lyft.”, and then were asked if they had ever used this type of service. The sub-sample that answered positively was further presented with a definition of pooled ride-hailing (“In the carpooling option of ride-sourcing, additional passengers with similar routes get picked and dropped off in the middle of the customer's ride. Customers receive discounted rates when they choose this option”) and asked about the use of such a pooled ride-hailing service. Based on the responses to these questions, and as applicable, the respondents were asked to indicate their frequency of use, in the past 30 days, of private and pooled ride-hailing services. Also, all respondents who indicated the use of ride-hailing services at some point in their lives were asked to recall the details of their last ride-hailing trip and provide information on trip purpose, time of day of travel, companionship, and mode substituted. The survey also collected socio-demographic (see Section 2.2.1) and attitudinal information (see Section 2.4).

### 2.2 Individual-Level Experience and Frequency of Use Model

The individual level model focuses on two main endogenous outcomes, ride-hailing experience and frequency, and two additional endogenous variables, residential location and household vehicle availability. Ride-hailing experience is represented as a nominal dependent variable with three categories: (1) no experience with ride-hailing services, (2) experience only with private services (the individual traveled alone or with people s/he knew), and (3) experience with private

[3] The decision to focus on individuals with commute travel was guided by factors external to the current study. The authors acknowledge that understanding the preferences and ride-hailing travel behavior of non-commuters is also important from a transportation policy standpoint, and we encourage future studies to explore this limitation of our study.

and pooled services (the individual has, at least once, traveled with strangers for a cheaper fare). Ride-hailing frequency corresponds to the number of trips made by ride-hailing users within a one-month period prior to the date of the survey. This is modeled as an ordinal discrete variable with five possible values: zero trips, 1-3 trips, 4-5 trips, 6-10 trips, and more than 10 trips. Residential location is defined based on a survey item in which the respondents identified the type of neighborhood where they lived: (1) rural area, (2) small town, (3) neighborhood in the suburbs, (4) neighborhood in a central area but not downtown, and (5) downtown. Due to paucity of responses in the "small town" and "downtown" categories, we decided to regroup these five categories into the following three categories of residential location type: rural area or small town (referred to as rural area in the remainder of the paper), suburban area, and central area/downtown (referred to as urban area). Vehicle availability is characterized as the number of vehicles per worker in the household and is categorized in one of three ordinal levels: *less than one vehicle per worker*, *one vehicle per worker*, and *more than one vehicle per worker*. This definition is widely accepted in the literature as an indicator of vehicle availability or sufficiency for households with workers, because of the role that work schedules and commuting episodes play in shaping household activity schedules and task/vehicle allocation among household members (see, for example, Astroza et al., 2018). The last two co-endogenous variables (residential location and vehicle availability) are considered in our analysis to account for the possibility that residential location and vehicle availability, along with ride-hailing behavior, are determined as a choice bundle, and to accommodate for any self-selection effects in the influence of residential location and vehicle ownership on ride-hailing behavior (our expectation, though, is that these self-selection effects will be rather small, because ride-hailing is a relatively recent mobility option available within the past five years, while residential location and vehicle ownership decisions are typically made at longer time intervals than five years). Exogenous socio-demographic characteristics and four endogenous stochastic latent constructs representing attitudinal and lifestyle characteristics of the individual (privacy-sensitivity, technology-savviness, variety-seeking lifestyle propensity, and green lifestyle propensity, described in Section 2.2.1) are used as determinants of the four endogenous variables of interest.

The modeling methodology adopted is based on the Generalized Heterogeneous Data Model (GHDM) developed by Bhat (2015a), which allows for the joint estimation of multiple outcomes of different types (continuous, ordinal, count and nominal) by establishing a parsimonious dependence structure through stochastic latent variables (note that in the current application, there are only ordinal and nominal outcomes). The dependence structure, variable representation formulation, and endogeneity hierarchy of the model are presented in Figure 1 and discussed in detail in Section 2.2.2.

*2.2.1 Attitudinal and Lifestyle Latent Constructs*

Four attitudinal and lifestyle latent constructs are considered in our framework: privacy-sensitivity, technology-savviness, variety-seeking lifestyle propensity (VSLP), and green lifestyle propensity (GLP). These are identified based on earlier studies in transportation as well as in the ethnography field that recognize these psycho-social constructs as important determinants of travel-related and technology-use patterns. For instance, the first latent construct, privacy-sensitivity has been acknowledged and included in multiple transportation studies that investigate public transit use (Hunecke et al., 2010; Haustein, 2012; Spears et al., 2013). This is because one of the main aspects of the public transit mode that may discourage use is the presence of strangers in a shared space. Although ride-hailing is a car-based transportation mode, individuals travel with the driver. Hence, understanding how much individuals value being in private environments is a key element to predicting the adoption of ride-hailing, especially the use of pooled ride-hailing. Controlling for privacy-sensitivity is also important because concerns about sharing spaces with strangers influence people's residential location and vehicle availability (through ownership) choices as privacy is strongly related to spaciousness and exclusivity considerations, with individuals with a stronger privacy disposition locating in low to medium density neighborhoods and owning many vehicles (see, for example, Bhat et al., 2016 and Bhat, 2015b). Thus, including this construct is important to avoid the overestimation of any positive impacts of dense residential location and low vehicle ownership on ride-hailing use.

The second latent construct, tech-savviness, represents the individual's familiarity and affinity with technology, in our case, information and communication technologies (ICTs). This latent construct is relevant because, to hail a ride, the individual needs to use a smartphone app. Indeed, previous studies have found a significant and positive impact of tech-savviness on ride-hailing experience and smart phone use (Alemi et al., 2018b; Lavieri et al., 2017; Astroza et al., 2017). The third construct, variety-seeking lifestyle propensity (VSLP) represents the individual's interest in exploration, and his/her openness to new experiences and changes. This construct has also been used in a past ride-hailing study (Alemi et al., 2018b) and is important to capture intrinsic heterogeneity in the willingness to deviate from travel habits and mode inertia (Tudela et al., 2011; Rieser-Schüssler and Axhausen, 2012). The construct has been widely used within the theory of basic human values in the cultural-psychology field, and two of the indicators used in our survey to measure this construct are based on Schwartz's core value measures of openness to change (see Schwartz et al., 2001).

Finally, the green lifestyle propensity (GLP) construct is used to capture individuals' tendencies toward environmentally friendly behaviors such as reduced use of drive-alone modes, reduced car ownership, and increased preference for dense and walkable neighborhoods. This latent variable is probably the most commonly used lifestyle factor in travel behavior studies (see for example, Van Acker et al., 2014; Bhat, 2015b; Lavieri et al., 2017; Ye and Titheridge, 2017).

Similar to privacy-sensitivity, controlling for VSLP and GLP is fundamental to capture potential self-selection effects that could bias the impacts of residential density and vehicle ownership on ride-hailing behavior. A list of the indicators associated with each latent construct is presented at the bottom of Figure 1.

*2.2.2 Model Structure and the Generalized Heterogeneous Data Model (GHDM) Approach*

The GHDM approach (Bhat, 2015a) enables us to investigate the relationship between ride-hailing adoption and other transportation decisions, while controlling for observed and unobserved factors that simultaneously influence such decisions. There are two components to the GHDM model: (1) the latent variable structural equation model (SEM), and (2) the latent variable measurement equation model (MEM). As illustrated in Figure 1, the SEM component defines each latent construct (represented as ovals on the left side of the Figure) as a function of exogeneous socio-demographic variables and an unobserved error term. Each error term represents the effect of unobserved individual factors on a specific latent construct. The unobserved factors are denoted by $\eta_1, \eta_2, \eta_3,$ and $\eta_4,$ and collected, as shown toward the left side bottom of Figure 1, in a vector $\boldsymbol{\eta}$ assumed to be multivariate standard normal with a mean vector of $\mathbf{0}$ and a correlation matrix of $\boldsymbol{\Gamma}$ with six possible correlation elements (note that, due to identification considerations, the variances of the individual $\boldsymbol{\eta}$ elements need to be normalized to 1; see Bhat, 2015a). Of course, the latent constructs are stochastic because of the presence of the random elements, and, by definition, are not observed. Thus, the SEM model relationship between the socio-demographic variables and the latent constructs, as well as the correlation matrix elements of $\boldsymbol{\Gamma}$, are not directly estimable, but are estimated through observations on the latent construct indicators (not shown in Figure 1 to avoid clutter, but see Table 2 later and Section 2.4 for a discussion of these indicators) and the endogenous outcomes of interest (shown toward the right side of Figure 1).[4] The endogenous outcomes are discrete variables in our framework, and thus, for modeling, are considered to be based on underlying latent utilities (for the nominal outcomes of residential location and ride-hailing experience, each with three possible alternatives) and propensities (for the two remaining ordinal outcomes). These underlying latent utilities and propensities are sandwiched between the latent constructs and the observed endogenous outcomes in Figure 1 (these latent utilities/propensities also have additional error term effects, which are again suppressed in Figure 1 to avoid clutter). As shown in Figure 1, the MEM component relates the underlying latent utilities/propensities of the observed discrete endogenous outcomes to the stochastic latent constructs and exogeneous socio-

[4] As discussed at length in Bhat (2015a), the latent construct indicators are not needed for the GHDM model estimation, though their presence provides much needed additional stability and information to "pin" down the SEM relationship. In particular, the presence of these indicators play a role in model identification, as discussed in detail in Bhat (2015a). Suffice it to say that, in the current context, the presence of the indicators allows us to estimate all possible stochastic latent construct effects on each endogenous outcome.

demographic variables. Figure 1 indicates only those stochastic latent construct determinants for each underlying latent utility/propensity variable (shown by arrows originating from the stochastic latent constructs and terminating in specific underlying latent utility/propensity ovals) that turned out to be ultimately statistically significant in our empirical specification. Thus, for example, in our final specification, privacy sensitivity impacts the underlying latent propensity for vehicle availability and the ridesharing experience utilties, but not the other two endogenous outcomes.

The error term elements in the $\boldsymbol{\eta}$ vector of the SEM (which impact the stochastic latent constructs) permeate into the underlying latent utilities/propensities in the MEM, creating a parsimoious dependence structure among all endogenous variables. Thus, for example, consider the underlying latent propensity for ridesharing use and the two underlying utility functions for the "private only" and "pooled" ride-hailing experience categories (the latter two utility functions are contained within the oval labeled as "ridesharing experience utilities" in Figure 1). These three underlying variables are all impacted by the variety-seeking lifestyle propensity (VSLP) latent construct. Thus, they get correlated because of the common presence of the $\eta_3$ stochastic term embedded in VSLP, obviating the need for three separate correlation terms that would be needed otherwise. Similar parsimonious correlations are engendered (through other common stochastic latent constructs) across the latent propensities/utilities underlying the other endogenous outcomes. This ability to capture correlations across the many endogenous outcome dimensions is important for controlling for residential location and vehicle availability-based self-selection effects (when examining the effects of these two variables on ride-hailing experience and frequency) in an econometrically consistent fashion.

Also, to be noted is that, while the four endogenous outcome variables are all modeled jointly through the aforementioned correlation effects in the underlying latent utilities and propensities, recursive effects among the observed endogenous outcomes can also be accommodated in the GHDM (see Bhat, 2015a). Different recursive directionalities were tested in our model system, but the best data fit was obtained in the causal specification with residential location influencing vehicle availability, both of these then impacting ride-hailing experience, and all three finally influencing ride-hailing frequency (as shown in Figure 1 under "observed endogenous outcomes").[5] The full details of the GHDM and its estimation are available in Bhat (2015a).

[5] In joint limited-dependent variables systems in which one or more dependent variables are not observed on a continuous scale, such as the joint system considered in this paper that has ordinal and nominal discrete dependent variables, the structural effects of one limited-dependent variable on another can only be in a single direction. See Maddala (1983) and Bhat (2015a) for a more detailed explanation.

### 2.3 Trip-Level Ride-Hailing Attributes Multivariate Model

The trip-level model utilizes the subsample of individuals with ride-hailing experience and examines the four choice dimensions of trip purpose, time-of-day of trip, trip companionship, and the mode substituted by ride-hailing for the most recent ride-hailing trip undertaken by respondents. The four dimensions are modeled jointly using a multivariate multinomial probit (MMNP) modeling approach so that common unobserved individual-level factors that affect multiple trip characteristics are captured through error correlations and endogeneity relationships. The MMNP formulation is presented in Appendix A and the interested reader is also referred to Bhat et al. (2013) and Bhat (2011) for additional estimation details.

The first nominal variable, trip purpose, is captured in the four categories of airport trips, errand trips (including shopping, personal business, and family errand trips), recreation trips (including leisure, social activities and sports), and work trips (including education trips). The second dimension is time-of-day, which is characterized by four time windows of morning (6:00 am – 10:59 am), mid-day (11:00 am – 3:59 pm), evening (4:00 pm – 8:59 pm), and night (9:00 pm – 5:59 am). The third is companionship, formed by two categories, alone or with others.[6] The fourth dimension is the mode substituted by ride-hailing (based on the response to the question "if ride-hailing were not available, which mode would you have used for the trip"), in the four categories of (a) private vehicle, (b) taxi, (c) transit and/or active travel (walk/bicycle), and (d) no trip (that is, the trip would not have been made if ride-hailing were not available).

As already discussed earlier, only recursive effects among endogenous variables can be identified in joint limited-dependent variable systems such as in the MMNP model (after accommodating error covariance that engenders the jointness in the first place). We tested alternative recursive structures, but the one that provided the best fit was the recursive hierarchy in which trip purpose impacted the remaining three attributes (time-of-day, companionship, and mode substituted), time-of-day impacted the remaining two attributes (companionship and mode substituted), and companionship affected the "mode substituted" trip dimension.

This trip-level analysis is exploratory in nature, because we are modeling the attributes of an isolated trip outside the broader context of the individual's daily activity-travel schedule. In particular, it is difficult to disentangle whether the choices made for the most recent ride-hailing trip are a reflection of specifically choosing ride-hailing in the last trip or simply a manifestation of the totality of the activity-travel pattern of the individual. For example, if a student is more likely than a non-student to run errands in the last ride-hailing trip relative to traveling to the airport, it is not clear whether this implies that students are more likely than non-students to use

[6] As will be discussed in Section 2.4, for modeling purposes, it was necessary to merge trips with strangers and trips with acquaintances and define a generic "with others" category because the sample of trips with strangers was very low (only 13 observations). Also note that among individuals who traveled alone, we are not able to identify whether they initially called for a pooled ride but had to travel alone because of an unsuccessful match, or whether they intended to travel alone from the very beginning.

ride-hailing to run errands than to go to the airport, or whether this is simply an artifact of students rarely going to the airport in general relative to their non-student counterparts. We will not belabor over this point again when discussing the trip-level results, although all the results there should be viewed through this cautionary interpretive lens. Nonetheless, we use a multivariate modeling approach to study the different trip attributes jointly, allowing us to control for the effects of multiple endogenous variables systematically and simultaneously.

Unlike the individual-level model, residential location density, vehicle availability, whether or not the individual has experience with pooled ride-hailing, and ride-hailing frequency are considered as exogenous variables in this exploratory trip-level analysis.[7] These variables are treated as exogeneous in this case to reduce the model complexity considering the very limited sample size and the exploratory nature of this analysis. Similarly, we also include the latent constructs as characterized from the individual-level model as exogenous variables by developing an expected value for each latent variable (based on the SEM model estimates from the individual-level model) and each individual in the sample.[8]

### 2.4 Sample Description

In this section we describe the sample distribution regarding sociodemographic characteristics, attitudinal indicators, the four endogenous outcomes in the individual-level model and the four endogenous variables in the trip-level model. Table 1 presents the socio-demographic distribution of the sample. A comparison of the sample with the employed population of DFW (as characterized by the U.S. Census Bureau, 2018d) indicates that the survey has an overrepresentation of males (58.4% in the survey compared to 54.0% from the Census data), individuals between 45 and 64 years of age (53.2% compared to 35.8%), Non-Hispanic Whites (75.0% compared to 51.0%), and individuals with bachelor's or post-graduate degrees (75.6% compared to 33.7%). We also observe that the majority of the sample corresponds to non-students (94.2%) and full time-employees (81.6%). Finally, in terms of household income and household composition, we are unable to compare the statistics from our survey with the Census data, because the latter provides income and household composition data only for all households

[7] The last of the endogenous variables from the individual-level model, ride-hailing frequency, is introduced as a binary variable in the trip-level analysis, by classifying individuals as either frequent users (at least 4 rides in the past 30 days) or not.

[8] The choice to adopt this approach of treating the latent constructs as exogenous rather than endogenous for our trip-level model (instead of estimating another elaborate GHDM for the trip-level model) is based on two considerations. First, the dependence between the trip-level choice dimensions is likely more due to unobserved factors associated with the nature of activities and trips (for example, bars and pubs generally open at night, so recreational trips may be more likely at this time), rather than individuals' psychological and lifestyle factors. Second, we believe that the characterization of the latent attitudinal and life-style constructs would be better based on broad individual-level decisions rather than trip-level decisions. Of course, given the smaller sample available for this trip-level analysis, we also felt a simpler exploratory modeling approach relative to the GHDM would be more appropriate.

(while our survey is focused on households with at least one worker with a primary workplace outside home). However, the sample statistics do suggest a skew toward individuals from higher income households and multi-worker households. Overall, there are many possible reasons for the socio-demographic differences between our sample and the Census data. For example, the main topic of the survey was self-driving vehicles, which may be of more interest to highly educated males. Also, the survey was conducted strictly through an online platform and the largest mailing list used in the distribution was of toll-road users, who are likely to be individuals with higher values of time that then correlates with the specific characteristics of our sample. In any case, while the general descriptive statistics of ride-hailing experience and use cannot be generalized to the DFW population, the disaggregate models still provide important insights on the relationship between ride-hailing travel behavior and socio-demographic/lifestyle characteristics.

Regarding the endogenous variables in the individual-level model, the majority of survey respondents live in suburban areas (65.0%; n=1046), followed by central area/downtown (23.4%; n=375). Vehicle ownership rates are high, as only 14.7% of the household have less than one vehicle per worker (50.8% have one vehicle per worker and 34.5% more than one vehicle per worker). In terms of ride-hailing experience, about 56.4% of the sample (n=906) reported using ride-hailing services at least once in their lifetimes. The column at the far right in Table 1 shows the fraction of individuals with ride-hailing experience by socio-demographic group. We observe that men, young adults (18–44 years of age), individuals of Hispanic and Asian origin, individuals with graduate degrees and students, high income individuals, and individuals living alone and in the central city areas have a higher than average tendency of having used ride-hailing services. The specific distribution of ride-hailing experience according to the three nominal categories is: *no experience* (43.6%; n=701), *experience with private rides only* (46.6%; n=906-157=749), and *experience with pooled rides* (9.8%; n=157; note that this group may have had experience with private rides too). When asked about ride-hailing frequency specifically in the month prior to the survey, 33.7% of all respondents (n=542) reported at least one trip, suggesting that there is a considerable percentage of ride-hailing users (22.7%=56.4%-33.7%) who rely on ride-hailing on a one-off basis rather than on a monthly basis. It also is important to point out that ride-hailing frequency is relevant only if the individual has had ride-hailing experience (that is, only if the individual is not in the "no experience" category for the ride-hailing experience variable). Within the sub-sample of individuals with some ride-hailing experience (n=906), the frequency of trips in the past 30 days is grouped in one of the following five ordinal levels (the share of each level, as a percentage of 906 individuals with ride-hailing experience, is represented in parentheses: zero trips (40.2%; n=364), 1-3 trips (30.9%; n=280), 4-5 trips (12.6%; n=114), 6-10 trips (11.0%; n=100), and more than 10 trips (5.3%; n=48).

The indicators of each latent construct used in the GHDM model are presented in Table 2, together with their sample distributions. The sample shows a general tendency toward being privacy-sensitive, tech-savvy, and having a variety-seeking lifestyle. The concern with privacy during a trip is consistent with the level of car-dominance in DFW, and may possibly impact the adoption of ride-hailing, especially pooled ride-hailing (note that the first indicator for privacy sensitivity is actually a measure of privacy insensitivity as elicited in the survey, and so the response is introduced in a reversed scale in the analysis to capture privacy sensitivity). A clear familiarity with ICTs and a variety-seeking lifestyle in the sample is expected, considering that the sample is skewed toward high levels of education and income. Interestingly, the responses related to the last measure; green lifestyle; show that over 50% of the sample "somewhat" or "strongly" agree that factors other than environmental friendliness dictate their commute mode choices, while just a little over 11% of the sample "somewhat" or "strongly" agree that they do not give much thought to energy saving at home. These descriptive statistics suggest that, while most people are sensitive to energy conservation considerations at home, most people also believe that considerations other than their commute-related environmental footprint dictate their commute mode choices (note again that the two questions pertaining to green lifestyle measure non-green lifestyle in the way they are worded, and so are introduced in a reversed scale in the analysis to capture green lifestyle propensity).

Finally, the sample distributions of the four trip-level choice dimensions are presented in Table 3. The descriptive statistics corresponding to trip purpose indicate that ride-hailing is mostly being used to access airports and recreational activities (with each of these purposes accounting for about 40% of all ride-hailing trips). The time-of-day shares show a relatively even intensity of trips during the morning and mid-day periods, though there is a definitive spike in the intensity during the evening period (note that all the morning, mid-day, and evening periods are of five hours duration, as we have defined them). The intensity of ride-hailing trips is lower during the nine-hour night period, though this is to be expected given the overall lower intensity of travel during the night relative to the day periods. In terms of trip companionship, about two-fifths of all trips are made alone, while the remaining are with others (co-workers, friends, family, and strangers). The trips with strangers, while having more of a flavor of pooled ride-hailing trips than those with co-workers, friends, and family, amounted to only 13 in number, and so were combined with trips with other accompaniment types. Finally, the dimension of mode substituted from for the ride-hailing trips suggests that much of the draw is from a private vehicle or a taxi. It is also interesting to note that almost 6% of the sample would not have traveled if ride-hailing were not available.

## 3. INDIVIDUAL-LEVEL EXPERIENCE AND FREQUENCY OF USE MODEL RESULTS

This section presents a detailed discussion of the results of the individual-level ride-hailing experience and frequency model. The final model specification was obtained based on a systematic process of testing alternative combinations of explanatory variables and eliminating statistically insignificant ones. However, some variables that were not statistically significant at a 95% confidence level were still retained due to their intuitive interpretations and important empirical implications. In this regard, the GHDM methodology used involves the estimation of a large number of parameters, so the statistical insignificance of some coefficients may simply be a result of having only 1,607 respondents (and only 906 respondents for the ride-hailing frequency variable). Also, the effects from this analysis, even if not highly statistically significant, can inform specifications in future ride-hailing investigations with larger sample sizes.

In the next section, we discuss the results of the SEM model component of the GHDM, as well as the latent variables' correlations and loadings on the attitudinal and lifestyle indicators (which is one part of the MEM). In subsequent sections, we discuss the MEM relationships corresponding to the effects of socio-demographic characteristics and the latent variables on the four main outcomes of interest in the individual-level model (including endogenous effects among these four outcome variables).

### 3.1 Lifestyle and Attitudinal Latent Factors

The structural relationships between socio-demographic variables representing lifecycle stages and the latent constructs are presented in Table 4. Gender shows no significant effect on the individual's level of privacy-sensitivity and tech-savviness. Yet, women display lower levels of VSLP and higher levels of GLP. These results are consistent with the social psychology literature. Gender comparisons based on the Theory of Basic Human Values (Schwartz, 1992) identify that men tend to be more open to experiences and changes than women as men generally attribute more value to stimulation, self-direction and hedonism values (Schwartz and Rubel, 2005; Vianello et al., 2013). On the other hand, women are generally more oriented toward prosocial values than men (Liu et al., 2014; Gifford and Nilsson, 2014), which result in more environmentally conscious behaviors (Gilg et al., 2005; Bhat, 2015b).

Age presents generally significant effects on all latent constructs except privacy-sensitivity. In general, younger adults show higher levels of tech-savviness and VSLP than their older counterparts. It is well established that younger generations, through their early exposure to ICT in their formative childhood years, are naturally more familiar and adept with such technologies (Helsper and Eynon, 2010; Twenge, 2013), which contributes to their higher level of tech-savviness. In terms of VSLP, the human values and personality literature identifies that younger individuals are more open to new experiences and more likely to attribute high importance to stimulation values, seeking variety in their daily lives (Gutierrez et al., 2005; Milojev and Sibley, 2017). The marginally significant negative GLP among the youngest group of individuals (18-34 years of age) relative to their older peers is interesting, though not inconsistent with findings from recent studies that identify a decrease in the younger generation's environmental consciousness. For example, Liu et al. (2014) and Gifford and Nilsson (2014) suggest that this trend among the youngest generation of adults may be the result of an increase in the importance of material pleasures in the American society as well as with an increased level of optimism that technology will solve environmental problems.

Non-Hispanic White individuals tend to be more privacy-sensitive and exhibit a lower VSLP relative to other ethnicities, results that also align with the higher levels of drive-alone travel and vehicle ownership by this ethnic group (Giuliano, 2003; Klein et al., 2018). As expected, individuals who are more highly educated tend to be more green, consistent with results in the social-psychological literature (see, for example, Franzen and Vogl, 2013) that individuals with a higher education are more self-aware of the negative consequences of degrading the environment. Usually, education is also an important predictor of tech-savviness (Helsper and Eynon, 2010; Lavieri et al., 2017; Lavieri et al., 2018b). However, in our model, such a relationship is not statistically significant, probably because the majority of the sample has at least a bachelor's degree. Part-time employees are less tech-savvy than full-time and self-employed individuals. As Helsper and Eynon (2010) explain, familiarity and ability to use ICTs is largely explained by exposure and experience. In that sense, it is plausible that part-time employees are generally less exposed to technology in the workplace (due to the nature of part-time jobs, and the time spent at work) than full-time and self-employed individuals.

In terms of household demographics, household income contributes to an increase in privacy-sensitivity, tech-savviness and VSLP. The higher privacy-sensitivity among the wealthiest segment of individuals can be a direct result of having more access to private property and/or a need to signal exclusivity through separation and differentiation from others (Chevalier and Gutsatz, 2012; Bhat, 2015b). These individuals may also focus on privacy due to concerns associated with safety and preservation of material assets. Also, higher consumption power allows wealthy individuals early access to new technologies, increasing their exposure and use of technology. Indeed, multiple studies find this positive association between income level and technology use or technology-savviness (see, for example, Astroza et al., 2017; Lavieri et al., 2017; and Liu and Yu, 2017). The higher VSLP in the wealthiest segment of individuals is also reasonable, since this segment has more financial wherewithal to pursue a variety of different types of activities. Finally, compared to multi-worker and single individual (worker) households, individuals living in single-worker multi-person households have lower VSLP.

Two out of six correlations between latent variables are statistically significant (see bottom of Table 4, corresponding to $\gamma_{23}$ and $\gamma_{14}$ in Figure 1). Privacy-sensitivity is negatively associated with GLP, and tech-savviness is positively associated with VSLP. Both relationships are intuitive. For example, the second positive and reciprocal relationship between tech-savviness and VSLP is to be expected because (a) individuals who seek variety are more likely to experiment with new products and technology, and (b) ICT and internet use expand an individual's awareness and spatial cognition about activity options and opportunities.

The SEM estimation is made possible through the observations on the endogenous variables, which include the latent construct indicators and the four endogenous outcomes of interest (see Figure 1). As discussed earlier, the presence of the latent construct indicators is not essential, though they provide stability in the SEM estimation. To conserve on space, we relegate the loadings of the latent constructs on the underlying latent variables characterizing the construct indicators to Appendix B (Table B1). However, we will note that the loadings were all as expected.

### 3.2 Residential Location and Vehicle Availability

Residential location and vehicle availability are modeled as endogenous variables so that we can control for self-selection effects when analyzing the impacts of these variables on ride-hailing behavior. Interestingly, as shown in Table 5, after controlling for the latent variable effects, there were few other sociodemographic variables having a direct impact on residential location and vehicle availability (though sociodemographic variables have an indirect effect through their impacts on the latent variables).

In terms of latent variable impacts on residential density, individuals who are tech-savvy and pursue a green lifestyle appear to prefer to reside in higher density suburban and urban areas rather than in a rural area. Access to ICT is generally more limited in rural areas, which may explain the negative effect of tech-savviness on rural living. Also, GLP is measured in our study in terms of concern about transportation and energy footprint, which may not be a priority for rural dwellers. On the other hand, the results indicate that individuals with a high variety-seeking lifestyle propensity (VSLP) tend to be more likely to live in an urban area relative to other areas, presumably because urban areas offer easy access to a diverse portfolio of activities and products. In addition to the indirect sociodemographic effects through the latent variable effects just discussed, the direct sociodemographic effects on residential location choice reveal that the youngest segment of individuals prefer more urbanized living relative to their older peers, presumably a reflection of wanting to have a variety of activity opportunities in close proximity to satisfy a heightened need for social interactions. Part-time employees tend to be located in urban areas, while self-employed individuals are more likely to reside in rural and urban areas rather than in suburban neighborhoods. As expected, households with income above $150K dollars per year are less likely than those with lower incomes to be located in rural areas compared to suburbs and urban areas. Finally, individuals living alone show a higher propensity to locate in urban areas, consistent with the age effect discussed earlier.

Vehicle availability is positively impacted by privacy-sensitivity, which is expected since the automobile is the most private transportation mode. In contrast, tech-savviness has a negative effect on vehicle availably, plausibly because these lifestyle variables facilitate the use of, and draw toward, multi-modal travel options (Astroza et al., 2017). As anticipated, households with high incomes and with fewer workers have a higher vehicle availability, the first effect due to higher car ownership levels in households with high incomes and the second effect simply a manifestation of how we created the vehicle availability variable. Finally, households residing in the high-density urban areas of the DFW area have a lower vehicle availability, a reflection of the reduced need for vehicles in such areas because of good multi-modal transportation service as well as better access to activity opportunities within a compact geographic footprint. Importantly, this urban living effect is a "true" built environment effect after controlling for residential self-selection effects through the impacts of the latent attitudinal lifestyle variables on both residential location and vehicle availability.

### 3.3 Ride-Hailing Experience

The results of the ride-hailing experience model are presented in the third column of Table 5. The latent variable effects have the expected direction, with privacy-sensitive individuals less likely to have experience with pooled service and tech-savvy individuals most likely to have experience with private ride-hailing only. On the other hand, variety-seeking individuals are most likely to have the pooled service experience. Interestingly, GLP does not seem to play a role in ride-hailing adoption.

In addition to the indirect socio-demographic influences through the latent variable effects just discussed, there are quite a few direct socio-demographic effects on ride-hailing experience. This is unlike the case for residential location density and vehicle ownership where there are relatively fewer direct sociodemographic effects after controlling for latent variable effects. This disparity makes sense because ride-hailing is a relatively recent phenomenon and individuals are still in the process of exploring the many dimensions of this service. That is, ride-hailing preferences are still in a formative stage, with the impacts of attitudes and lifestyles not yet as deeply entrenched as for residential location density and vehicle availability (the latter choices have been available to individuals over a much longer period of time).[9] During these initial exploratory/formative stages of preference, it is the immediate demographic lifecycle

[9] Note that the attitudinal and lifestyle latent variables and indicators used in this study do not reflect individual's direct attitudes, beliefs and perceptions about ride-hailing services. Instead, they reflect more general lifestyle dimensions.

considerations that dictate and drive ride-hailing experience and frequency. Earlier studies in the social psychology literature support this notion that the effects of attitudes/lifestyle toward preference for a service/product take time to materialize and stabilize (see, for example, Hoeffler and Ariely, 1999; Amir and Levav, 2008).

Table 5 indicates that age has a direct negative effect on ride-hailing experience, with younger individuals more likely than their older counterparts to have used ride-hailing both in the private as well as pooled arrangements. While this is consistent with some earlier studies (Smith, 2016; Kooti et al., 2017), our study indicates that this effect is beyond the negative effect of age on ride-hailing experience through the tech-savviness and variety-seeking effects. This direct effect may be a result of younger individuals having more exposure to new services and products through larger social networks (English and Carstensen, 2014).

The results also show that non-Hispanic Whites are less likely to have used pooled services, even after accounting for indirect ethnicity effects through privacy-sensitivity and VSLP, and controlling for income effects. The reason behind this ethnicity effect is not clear and calls for more qualitative studies investigating the willingness to share rides. Higher education appears to increase the experience with pooled ride-hailing, and employment status shows significant direct effects on private ride-hailing experience but not on the pooled option. Specifically, part-time employees are less likely to have experienced private ride-hailing services relative to full-time employees. Similar results were observed by Dias et al. (2017).

In terms of household level variables, a higher household income increases experience with both private and pooled ride-hailing, beyond the positive effect of household income through tech-savviness and VSLP (and while individuals with a household income over $200,000 have a higher privacy sensitivity, and privacy sensitivity negatively impacts pooled ride-hailing experience, this indirect negative effect gets swamped by the magnitude of the positive direct effect in Table 5). Considering that attitudinal and lifestyle factors are being controlled for, the direct income effect is probably an indicator of higher consumption power, though there is still a distinct preference for private ride-hailing over pooled ride-hailing within this high income group. Individuals living alone are more likely to have used private ride-hailing service relative to individuals in other household types, while those in single-worker multi-person households are the least likely to have used both private and pooled services. Even after controlling for self-selection effects, individuals living in more urbanized locations are more likely than their counterparts in less urbanized locations to have used both private and pooled ride-hailing. A similar result holds for individuals in households with more than one vehicle per worker. The result that a higher private vehicle availability leads to a higher experience with ride-hailing suggests that, in an area such as DFW where almost all households own at least one vehicle, ride-hailing serves as more of a convenience feature for those one-off trips rather than being an accessibility facilitator for routine trips (though, as we will see in the next section, increasing vehicle availability has a negative effect on ride-hailing frequency). That is, there appears to be an overall "wealth effect" (living in urban areas and owning more than one vehicle per worker) that is separate from the direct income/education effect leading to more experience with ride-hailing.

### 3.4 Ride-Hailing Frequency

Our model, similar to that of Alemi et al. (2018a), shows that few variables have an impact on ride-hailing frequency. Among the latent variable effects, only VSLP has a significant impact. This effect may be a result of individuals with a high VSLP experimenting and exploring different travel options and different activity pursuits (see, for example, Rieser-Schüssler and Axhausen, 2012).

Among other demographic effects, individuals in households with very high income (above $200K dollars per year) have a high ride-hailing frequency propensity, as also observed by Dias et al. (2017). Although using ride-hailing is usually cheaper than calling a taxi, frequent use can incur significant costs that may be more easily afforded by those in the high income segments. Living in an urban area (relative to living in suburbs or rural areas) also contributes to a higher propensity associated with ride-hailing trip frequency, even after controlling for self-selection effects. There are at least three possible reasons for this result. First, urban areas have more parking restrictions, increasing the benefit of being dropped-off at a destination. Second, distances are shorter, compared to more spread-out suburbs and rural areas, limiting the costs of the trips. Third, in urban areas, the supply of drivers is higher, increasing the overall reliability of the service, which is possibly an essential condition for maintaining a demand of frequent users. As also observed by Alemi et al. (2018a), higher vehicle availability rates reduce the propensity underlying the frequency of ride-hailing usage. Combined with the earlier finding of the positive effect of vehicle availability on ride-hailing experience, the results perhaps suggest that individuals in households with high vehicle availability make generally many more out-of-home trips (including those one-off trips to the airport and other recreational sites) and so are more likely to have used ride-hailing at some point as a convenience mode. However, it still holds that higher vehicle availability reduces the overall ride-hailing dependence. Another endogenous effect is that users of pooled ride-hailing have higher frequency propensities. Pooled trips offer lower fares, which may be a key element for ride-hailing services to maintain regular users.

### 3.5 Model Fit Comparison

The improved data fit from jointly modeling the four choice dimensions in the individual-level model system may be assessed by comparing the GHDM model with an Independent Heterogeneous Data Model (IHDM) that does not consider the jointness in the four dimensions (that is, the covariances engendered by the stochastic latent constructs in the GHDM model are ignored). In this IHDM model, we introduce the exogenous variables (sociodemographic variables) used to explain the latent constructs as exogenous variables in the choice dimension equations. In this way, the contribution to the observed part of the utility due to sociodemographic variables is still maintained (and is allowed to vary relative to the GHDM to absorb, to the extent possible, the GHDM covariances due to unobserved effects). The resulting IHDM may be compared to the GHDM using the composite likelihood information criterion (CLIC) introduced by Varin and Vidoni (2005). The CLIC takes the following form (after replacing the composite marginal likelihood (CML) with the maximum approximate CML (MACML)):

$$\log L^{*}_{MACML}(\hat{\ddot{\boldsymbol{\theta}}}) = \log L_{MACML}(\hat{\ddot{\boldsymbol{\theta}}}) - tr\left[\hat{\boldsymbol{J}}(\hat{\ddot{\boldsymbol{\theta}}})\hat{\boldsymbol{H}}(\hat{\ddot{\boldsymbol{\theta}}})^{-1}\right] \tag{1}$$

The model that provides a higher value of CLIC is preferred. The $\log L_{MACML}(\hat{\ddot{\boldsymbol{\theta}}})$ values for the GHDM and IHDM models were estimated to be –394,131 and –398,801, respectively, with the corresponding CLIC statistic values of –395,982 and –400,229. These CLIC statistics clearly favor the GHDM over the IHDM.

The ordinal indicator variables used in the measurement equation are included solely for the purpose of model identification and do not serve any purpose in predicting the endogenous choice bundle of interest once the model is estimated. Therefore, we can also use the familiar non-nested likelihood ratio test to informally compare the two models. To do so, we evaluate a predictive log-likelihood value $L(\hat{\ddot{\boldsymbol{\theta}}})$ of both the GHDM and IHDM models using the parameter values at the GHDM convergent values by excluding the indicator variables and focusing only on the four endogenous variables of interest. Then, one can compute the adjusted likelihood ratio index of each model with respect to the log-likelihood with only the constants as follows:

$$\bar{\rho}^2 = 1 - \frac{L(\hat{\ddot{\boldsymbol{\theta}}}) - M}{L(c)}, \tag{2}$$

where $L(\hat{\ddot{\boldsymbol{\theta}}})$ and $L(c)$ are the predictive log-likelihood functions at convergence and at constants, respectively, and $M$ is the number of parameters (not including the constant(s) for each dimension and not including the ordinal indicators) estimated in the model. If the difference in the indices is $(\bar{\rho}_2^2 - \bar{\rho}_1^2) = \tau$, then the probability that this difference could have occurred by chance is no larger than $\Phi\left\{-[-2\tau L(c) + (M_2 - M_1)]^{0.5}\right\}$ in the asymptotic limit (however, this is only an informal test, because the use of the MACML inference approach rather than the traditional maximum likelihood approach changes the asymptotic properties). A small value for the probability of chance occurrence suggests that the difference is statistically significant and that the model with the higher value for the adjusted likelihood ratio index is to be preferred. The

$L(\hat{\hat{\theta}})$ values (number of parameters) for the GHDM and IHDM models were computed to be –2,728.85 (number of parameters=85) and –2,726.12 (number of parameters=94), respectively. The $L(c)$ value was –2,915.55. The non-nested adjusted likelihood ratio test (in its informal version use here) returns a value of $\Phi(-4.64)$, which is literally zero, reinforcing the result from the more formal CLIC statistic in rejecting the IHDM model in favor of the GHDM model and underscoring the importance of considering the stochastic latent constructs that engender covariation among the choice dimensions.

### 3.6 Average Treatment Effects

To compare the magnitudes of effects and identify of the most significant determinants of ride-hailing usage, we compute average treatment effects (ATEs) of the explanatory (exogenous, latent and endogenous) variables on ride-hailing experience and frequency.

In these ATE computations, we consider the latent psycho-social variables too as explicit determinant variables, rather than translating these latent variable effects into corresponding exogenous demographic variable effects through the structural equation model results of Table 4. That is, we do not combine the direct demographic effects and the indirect demographic effects (through the latent variables); rather, we compute the ATEs for the direct demographic effects and the ATEs for the latent variables. This is because, while the overall (indirect plus direct) demographic effects provide ride-hailing tendencies by demographic segment, they do not provide insights that may help in formulating policies. For example, one of the overall demographic effects is that non-Hispanic Whites are less likely to use pooled ride-hailing. However, this does not provide us additional insights on why this may be so. By including latent variables in the ATE computation, we may find, for example, that privacy sensitivity is one of the most important determinant variables in terms of the magnitude of effect on the use of pooled ride-hailing. If so, and because non-Hispanic Whites are likely to be more privacy sensitive relative to individuals of other ethnicity groups (according to our structural equation model results), it provides additional insights on how to position pooled ride-hailing information campaigns directed toward this segment of the population. One additional note regarding the computation of ATE effects for the latent variables. We compute these effects by examining the impact of changing each latent variable from its minimum value (the base) to its maximum value (that is, the continuous latent variable values are changed to two discrete values for the ATE computations; the minimum expected value representing the base category).

#### *3.6.1 Formulation and Computation*

The ATE measure for the ride-hailing experience variable (which is a nominal variable in our analysis) provides the expected difference in ride-hailing experience for a random individual if s/he were in a specific category $i$ of the explanatory variable as opposed to another configuration $k \neq i$. The ATE is estimated as follows for each explanatory variable:

$$\hat{\text{ATE}}_{ikj} = \frac{1}{Q}\sum_{q=1}^{Q}\left(\left[P(y_q = j \mid a_{qi} = 1) - P(y_q = j \mid a_{qk} = 1)\right]\right) \tag{3}$$

where $a_{qi}$ is the dummy variable for the category $i$ of the explanatory variable for the individual $q$ ($Q$=1607), $y_q$ stands for the ride-hailing experience nominal variable, and $j$ represents a specific nominal category of ride-hailing experience. Thus, $\hat{\text{ATE}}_{ikj}$ above represents the estimate of the expected value change in the nominal category $j$ of ride-hailing because of a change from category $k$ to $i$ of the explanatory variable. In computing this effect, we first assign the value of the base category for each individual in the sample (that is, we assign the value of $a_{qk} = 1$ to the determinant variable of each individual to compute $P(y_q = j \mid a_{qk} = 1)$) and then change the value of the variable to $a_{qi} = 1$ to compute $P(y_q = j \mid a_{qi} = 1)$).

The ATE measures may be computed for each nominal category $j$ of the ride-hailing experience variable as well as each combination of $i$ and $k$ for the explanatory variables. In our analysis, we compute the ATE measures for the nominal categories of “private only” and “pooled” ride-hailing experience, and for one combination of $i$ and $k$. For example, in the case of age, the base category is the “65 years or more” age group, while the changed category corresponds to the “18-34 years” age group. Similarly, for ethnicity, the base category is the “Other” ethnicity (including individuals of Hispanic and non-White ethnicities) and the changed category is the “Non-Hispanic White” ethnicity. Table 6, which provides the ATE values, shows the base category as well as the “changed category” for each determinant variable. As already indicated, in the case of the latent psychosocial variables, the base “category” corresponds to the minimum expected (that is, deterministically predicted) value of the variable, and the changed “category” corresponds to the maximum value of the variable.

For the ride-hailing frequency ordinal variable, we assign cardinal values to each of the frequency ordinal levels, and then compute the ATE of explanatory variables (in the same binary categorizations as discussed earlier for ride-hailing) on the expected total number of ride-hailing trips per month. The cardinal value assignments for the ordinal frequency levels in the model are as follows: (1) no ride-hailing trips: 0 trips in the past month, (2) 1-3 ride-hailing trips: 2 trips, (3) 4-5 ride-hailing trips: 4.5 trips, (4) 6-10 ride-hailing trips: 8 trips, and (5) more than 10 trips: 12 trips. With these assignments, the ATE corresponding to ride-hailing frequency for any determinant variable that is changed from category $k$ to category $i$ is computed as follows:

$$\hat{\text{ATE}}_{ik} = \frac{1}{Q}\sum_{q=1}^{Q}\left(\sum_{h=1}^{5} c_h \cdot \left[P(freq_q = h \mid a_{qi} = 1) - P(freq_q = h \mid a_{qk} = 1)\right]\right) \tag{4}$$

where $c_h$ is the cardinal value assignment corresponding to the ordinal ride-hailing frequency level $h$, and $freq_q$ corresponds to the ordinal ride-hailing frequency of individual $q$ in the 30 days prior to the survey ($Q$=906).

To calculate the ATE values in Equations (3) and (4), a realization of random draws is constructed by appropriately drawing from the sampling distribution of all the relevant parameters in the model. The ATE values are computed for 1000 different draws (for each individual) so that standard errors are obtained. The values of all dependent variables are calculated appropriately by following the chain of causal effects among the endogenous variables. All results are presented in Table 6.

*3.6.2 Results*

Among the latent variables, tech-savviness seems to be the strongest predictor of private ride-hailing experience with an ATE coefficient of 0.16. That is, if 100 random individuals increased their level of tech-savviness from the minimum to the maximum sample value, there would be 16 more individuals with private ride-hailing experience. In terms of pooled ride-hailing experience, privacy-sensitivity appears to be the most important deterrent, which suggests the need for concerted efforts to better understand the fundamental origins of high privacy-sensitivity, especially within the wealthiest population segment and non-Hispanic Whites (because these two groups have the highest privacy sensitivity). Another important insight from our results is the negative correlation between green lifestyle propensity (GLP) and privacy sensitivity, which suggests that targeting individuals with a high GLP (women, non-millennials, and individuals with a graduate degree) and positioning information campaigns about the environmental benefits of pooled compared to riding alone may be effective through the low privacy sensitivity prevalent in these population subgroups. While such campaigns should immediately increase pooled ride-hailing in women and in the group of individuals with a graduate degree (the second group is already pre-disposed toward pooled ride-hailing, as we will discuss later), our results suggest that information campaigns targeted toward non-millennials (and especially the oldest group of 65+ years) would be more effective if also combined with efforts to make this group of the population more tech-savvy, as discussed next.

The effects of the other latent variables in Table 6 indicate that tech-savviness and variety-seeking latent propensity (VSLP) have a positive impact on ride-hailing in general and pooled ride-hailing in particular. The positive impact on pooled ride-hailing adoption provides additional important policy insights. Tech-savviness levels in the population are generally increasing, thanks to information and communication technologies permeating into our routine daily lives. However, as evidenced in the results of Table 4, older and lower income segments seem to be falling behind and may need additional support to become “technologically-included”. This calls for informational campaigns targeted at these population segments on how ride-hailing services function and how to use smartphone apps. The positive impact of VSLP on pooled ride-hailing suggests that perhaps one other way to promote pooled ride-hailing would be to promote the notion of VSLP through the development of personalized trip plans that show multiple travel options, including pooled ride-hailing.

The ATEs corresponding to the direct impacts of socio-demographic variables and the other endogenous variables, when combined with the latent variable effects just discussed, point to millennials, individuals belonging to ethnicities other than the non-Hispanic White ethnicity with a graduate degree or higher, and those residing in urban areas as being the most likely to adopt pooled ride-hailing. In particular, the direct positive effect of being a millennial on pooled ride-hailing complements the indirect positive effect through the high tech-savviness and VSLP prevalent among millennials, while the direct positive effect of being of an ethnicity other than non-Hispanic White complements the low privacy sensitivity in ethnicities other than non-Hispanic White (as discussed earlier, privacy sensitivity appears to be the most important consideration in the use of pooled ride-hailing). Similarly, the direct positive effect of being a non-rural area resident complements the indirect positive effect through the high tech-savviness, VSLP and GLP among non-rural residents. The direct effects of income suggest that pooled ride-hailing is likely to be more adopted among individuals in low income households, which reinforces the positive indirect effect on pooled ride-hailing through the low privacy sensitivity in this low-income group; however, this low-income group also is less tech-savvy and has a low VSLP, both of which take away from the positive direct income effect. More generally, the positive direct effect of low income on pooled ride-hailing is likely a reflection of the cost of ride-hailing services, which are still high. After controlling for the latent variable effects, the number of monthly ride-hailing trips would increase by an average of 1.2 trips (over a 30-day period) if a random individual were transferred from the lowest to the highest household income category, which indicates that ride-hailing use by the overall employed population can increase quite substantially if ride-hailing costs significantly drop. In that sense, the introduction of self-driving ride-hailing fleets, which promise to reduce ride-hailing trip costs, may play an important role in increasing the demand for ride-hailing services in general, and pooled ride-hailing services in particular.

We also computed ATEs based on the IHDM model so that we can evaluate the magnitude of any self-selection effects of residential choice and vehicle availability on ride-hailing experience and frequency. As expected, ignoring these self-selection effects (as the IHDM model does) led to a higher magnitude of effect of urban living and vehicle availability on both private and pooled ride-hailing, as well as on ride-hailing frequency. Similarly, the effect of being a pooled ride-hailing user on ride-hailing frequency was also over-estimated in the IHDM model. However, as also anticipated in Section 2.2, these overestimations from the IHDM model were marginal and statistically insignificant. The important insight is that, at least at the current point in time, ride-hailing is a relatively new mobility option within the larger time scale at which residential choice and vehicle ownership decisions are made. Thus, at least in the very near term, studies may assume residential location choice and vehicle ownership decisions as being exogenous to ride-hailing choices, with reasonable confidence that, in doing so, the effects of residential location and vehicle ownership choices are still "true” causal effects. Of course, over time, this could change, with ride-hailing not just viewed as a travel mode, but as one element of a much broader lifestyle choice that includes residential choice and vehicle ownership. The analysis framework used in this study is thus very general, and can accommodate the more expansive lifestyle choice bundle context that is likely to unfold over time.

## 4. TRIP-LEVEL CHARACTERISTICS MODEL RESULTS

This section analyzes the model results corresponding to the four dimensions of the individual’s last ride-hailing trip: purpose, time-of-day, companionship, and mode substituted. In this trip-level analysis, we included the latent psycho-social constructs as exogenous variables. However, except for the VSLP construct (and that too only for the trip purpose dimension), no other latent variable turned out to be statistically significant in explaining trip-level ride-hailing choices. This result is consistent with our notion earlier that trip-level choices regarding ride-hailing are likely more affected by unobserved factors associated with the nature of activities and trips rather than individuals’ psychological and lifestyle factors.

The trip-level model in this paper is more of an exploratory nature, and thus the variable effects on the many dimensions of ride-hailing should be viewed with much more caution than for the individual-level model of the previous section. However, there are still some important insights from the results that we briefly summarize in this section.

### 4.1 Trip Purpose

The results of the model component representing trip purpose are presented in Table 7a. In the first category of latent constructs, only the VSLP variable influences trip purpose, with individuals with a higher VSLP more inclined to participate in recreation relative to other purposes. This is reasonable simply because recreation intrinsically captures a sense of variety and exploration relative to the other more sustenance and maintenance activity purposes.

Although women are usually responsible for more personal, family and shopping errands than men (Fan, 2017), being a woman is associated with a lower likelihood of using ride-hailing for these purposes, probably indicating that ride-hailing is not the preferred option when it comes to completing these routine commitments. By way of summarizing the effects of other socio-demographic effects, we observe that students and those with lower vehicle availability are more likely than their peers to have pursued errands in their last ride-hailing trip rather than other activity purposes, while millennials and those with lower vehicle availability are more likely to have pursued work-related travel rather than airport travel in their most recent ride-hailing trip. These results perhaps are indicative of the use of ride-hailing as an "accessibility mobility tool" to compensate for limited access to routine activities using other mobility options. On the other hand, millennials and non-Hispanic Whites are most likely to have pursued recreation (relative to all other activity purposes) in their last ride-hailing trip, presumably a reflection of the use of ride-hailing here as a "convenience mobility tool". The results also indicate that frequent ride-hailing users are more likely to have pursued work relative to other activity purposes. Finally, we observe that living in more urbanized areas decreases the probability of having pursued other activities compared to going to the airport in the last trip. Above all, this result shows that individuals living in rural areas do not use ride-hailing to go to the airport, probably because the associated costs are still higher than parking at an airport.

### 4.2 Time of Day

The earlier ride-hailing literature indicates that the peak period of ride-hailing trips occurs during the night and does not coincide with the commuting and traffic peak periods (see Kooti et al., 2017; Komanduri et al., 2018). However, our descriptive statistics indicate otherwise; as discussed earlier in Section 2.4, the evening period (which includes the afternoon commute period) is when the overall intensity of ride-hailing activity is highest. But the model results in Table 7a show that there are variations across individuals regarding when they are most likely to make a ride-hailing trip (at least based on their most recent trip). Not surprisingly, millennials (18-34 years of age) make most of their ride-hailing trips during the night period, consistent with this group more likely to socialize during this period (see Garikapati et al., 2016). High-income individuals, on the other hand, are the least likely to ride-hail during the evening and night periods. Individuals living in single-worker multi-person households (relative to those in other households) tend to ride-hail during the morning and evening periods, while those residing in suburbs and urban areas (relative to those residing in rural areas) appear to ride-hail more during the morning and mid-day periods, presumably due to the convenience to get to work by ride-hailing in dense areas. Frequent ride-hailing users appear to do so during the daytime. On the other hand, individuals who ride-hail during the nights appear to be from households with high vehicle availability and do so primarily for recreation, suggesting that these effects may be related to not wanting to drink and drive.

### 4.3 Companionship

The trip-level companionship results in Table 7b reveal some similarities with the individual-level pooled ride-hailing results, as one would expect. For example, middle-aged individuals are more likely than their peers to have had a companion on their most recent ride-hailing trip, while non-Hispanic Whites are more likely to have traveled alone. These are consistent with the results for pooled ride-hailing experience. Interestingly, a highly educated individual is more likely to have traveled alone during her/his last ride-hailing trip, though highly educated individuals are in general more likely to have had a pooled ridesharing experience. This perhaps is simply a reflection of highly educated individuals using a combination of private and pooled modes as they see best fit for specific trips, even if they are more open to pooled ride-hailing in general. Part-time employees and individuals from low-income households are more likely than their peers to have traveled with others. As expected, individuals who live alone, and individuals running errands or going to work are more likely to have traveled alone during their previous ride-hailing trip, while individuals pursuing recreation are more likely to have traveled with others during their previous ride-hailing trip. Finally, ride-hailing trips made during the morning peak serve mostly individuals traveling alone, which may have a negative implication on traffic congestion during this period.

**4.4 Mode Substituted by Ride-Hailing**

The results for this component of the trip-level model are presented in the last column of Table 7b. The base category is the "private car". Women, more than men, appear to substitute active travel or transit usage by ride-hailing (at least based on the most recent ride-hailing trip). Non-Hispanic Whites, those with graduate-level education, students, part-time employees, and individuals living in medium and high income households have a higher tendency than their peers to substitute ride-hailing for taxi trips, while millennials, self-employed individuals, individuals living in non-rural locations, individuals in households with one vehicle per worker, and individuals making their trip in the evening period are the least likely to substitute ride-hailing for taxi trips. In the context of active/public transportation (APT) modes, individuals younger than 65 years of age, those with a bachelor's degree or higher, and individuals with experience with pooled ride-hailing tend to replace APT modes with ride-hailing (see also Alemi et al., 2018b), while high income individuals and frequent ride-hailing users are not very likely to replace their APT travel with ride-hailing. Obviously, while one can explain these results in more ways than one, there is a clear need to investigate these effects in much more detail in future studies within the context of overall activity-travel patterns. However, the result regarding age effects does suggest that one potential detrimental effect of ride-hailing is a reduction in public health benefits, due to the substitution of active forms of transportation by ride-hailing among the substantial fraction of the population that is below the age of 65 years.

The last sub-column of the "Mode substituted by ride-hailing" corresponds to "no trips", which essentially implies that ride-hailing generated a new trip that would not have occurred otherwise. The demographic effects specific to this alternative indicate that young adults (18-44 years of age) are more likely than their older peers to have generated a new trip in their most recent ride-hailing experience, although it is more likely that these adults (relative to senior adults over the age of 65 years) switched to ride-hailing from active/public transportation. Also, part-time employees, self-employed individuals and those that live in multi-worker households appear to generate new ride-hailing trips more so than individuals in other households, perhaps a reflection of the added convenience to pursue activities due to ride-hailing. New trips are also more likely to occur among those living in non-rural areas. The generation of new trips in dense areas can, in the long term, intensify traffic congestion problems due to increased automobile usage. The new generated trips seem to be for the purposes of running errands and pursuing recreational activities, and are more likely to happen during the non-evening periods. The implied newly generated ride-hailing trips during the morning commute needs to be investigated more carefully, because the trips may add to traffic congestion as well as traffic crashes (the morning commute period is a traffic crash-prone period of the day due the combination of traffic congestion as well as the need to get to work on-time, which leads to aggressive driving during this period; see Paleti et al., 2010).

### 4.5 Dependence between Alternatives and Choice-Dimensions

The estimated covariance matrix (presented as an addendum in Appendix B, Table B2) corresponds to the differenced error terms for each dimension (the error term of the utility of an alternative minus that of the utility of the base alternative for that dimension; see Appendix A). In our analysis, we could not reject the hypothesis that the error terms for each of the trip-level dimensions were independently and identically distributed (in differenced error terms, we could not reject the hypothesis that all the diagonal terms in the covariance matrix of the differenced error terms were 1.0 and all the off-diagonal elements were 0.5). However, there were two covariances statistically different from zero across dimensions, both associated with the taxi alternative in the "mode substituted by ride-hailing" dimension, corresponding to (1) the morning alternative in the time-of-day dimension and the taxi alternative (= –0.181, t-stat of –3.14), and (2) the "with others" alternative in the companionship dimension and the taxi alternative (= 0.147, t-stat of 2.49). If we assume that the error term in the base alternative in each dimension is independent of the error terms of all other alternatives in other dimensions, the implication of the first covariance is that unobserved factors that increase taxi substitution also decrease the likelihood of the ride-hailing trip occurring during the morning, while the second covariance factor suggests that unobserved factors that increase taxi substitution increase the propensity to travel with others. The first effect may be related to the overall lower share of taxi trips in the morning compared to other modes (especially drive alone), while the second effect may be a consequence of the reduced costs in using ride-hailing relative to a taxi, especially in the pooled form of ride-hailing.

### 4.6 Model Fit Comparison

The statistically significant covariance effects, even if only two in number, point to the importance of developing a joint model at the trip-level. To further examine model fit, we compare the log-likelihood of the final model (= –10,747.52), and that of the model which ignores the two covariances discussed in the previous section (= 10,751.23). The log-likelihood ratio test statistic of comparison between the two nested models is 7.42. This value is greater than the table chi-squared value with two degrees of freedom at even a 0.025 level of significance.

## 5. POLICY IMPLICATIONS AND CONCLUSION

In the previous sections, we described a comprehensive analysis of ride-hailing travel behavior. The analysis involved the specification of multivariate models of ride-hailing experience, frequency and trip characteristics as functions of lifecycle and lifestyle variables. The results are now used as inputs to four broader discussions with implications to transportation policy-making and planning: (1) guiding people's acceptance and use of pooled rides; (2) the use of ride-hailing as an accessibility or a convenience mobility tool; (3) the relationship of ride-hailing with transit and active travel; (4) and the potential latent demand and trip induction generated by this service.

### 5.1 Acceptance and Usage of Pooled Rides

In terms of vehicle occupancy and VMT, private ride-hailing trips may be considered analogous to the drive alone mode, or even more problematic, as empty travel is generated by drivers relocating and waiting to pick up a next passenger. On the other hand, by serving multiple travelers that share similar routes, pooled ride-hailing can reduce VMT in car-centric environments. More generally, pooled ride-hailing has the potential to become a stepping stone toward multi-modal and service-based transportation in cities with less-than-adequate public transportation systems (such as in DFW and many metropolitan areas in the U.S.). However, our results show that individual's privacy sensitivity may be a significant barrier to choosing pooled rather than private rides. Privacy sensitivity seems to be a concern especially within the wealthiest population segment and non-Hispanic Whites. High income is also among the strongest determinants of frequent ride-hailing usage, which suggests the need for qualitative research (such as focus groups) to identify how individuals may be steered toward being less sensitive to the presence of strangers in a ride-hailing trip. For instance, based on the prejudice literature within the social psychology field (see for example, Zebrowitz et al., 2008; Barlow et al., 2012), greater exposure may reduce people's aversion to strangers as long as experiences are positive. Thus, breaking the inertia barrier and encouraging people to experiment with pooled services even if only temporarily (through substantial cost incentives or convenience incentives) may naturally reduce privacy concerns and have a snow-balling effect on the use of future pooled ride-sharing.

While pooled ride-hailing shows immediate potential to serve individuals in lower income segments, as they demonstrate lower levels of privacy sensitivity, a decrease in service fares and an increase in technology awareness among this group might still be necessary (as their lower levels of tech-savviness appear as an important deterrent to overall ride-hailing adoption). Similarly, educational campaigns targeting the increase of tech-savviness among older population segments would also make ride-hailing and pooled ride-hailing more accessible to this segment and thus contribute to an improvement in transportation equity.

Policies that have the result of increasing the number of individuals who have experienced pooled ride-hailing seem to also have the effect of increasing ride-hailing frequency. According to the results in Table 6, a pooled ride-hailing user is likely to make about two more monthly ride-hailing trips than an individual who has had no experience with pooled ride-hailing. Thus, our results suggest that getting an individual to try pooled ride-hailing that one time can have a lasting impact on the frequency of pooled ride-hailing over the longer term. It is also important to note that policies targeting variety-seeking behavior (such as personalized trip plans in mobility-as-a-service apps that offer multiple travel choices and include pooled ride-hailing) may be specially welcomed by young adults (18-44 years of age) and result in an increase in both pooled ride-hailing experience and overall ride-hailing frequency. However, this age group is also the segment that is more likely to substitute active or transit trips by ride-hailing and to generate new trips. Thus, the development of multi-modal apps needs to be especially careful to avoid the proliferation of unintended behaviors (such as the reduction of active travel/transit ridership and the generation of new car-based trips).

Another interesting result pertaining to ride-hailing as an opportunity to increase vehicle occupancy relates to commuting. Commuting encompasses a significant share of daily trips that, in areas such as DFW, are predominantly made by the drive alone mode (see Section 1.2) and could potentially be accommodated by pooled rides. Despite the lower numbers of work trips captured in our sample (compared to trips to the airport and trips to recreational activities), the model results show that frequent users are likely to use ride-hailing for work trips (from the trip purpose model), and work trips by ride-hailing are typically made alone (based on the trip companion model) during the morning and evening periods (as per the time-of-day model). The net result is that many ride-hailing trips for work during the morning and evening are undertaken in private ride-hailing mode as opposed to pooled ride-hailing mode. There is substantial opportunity for ride-hailing services as well as employers to work together to increase vehicle occupancy during the commute periods, through low cost pooled ride-hailing services (such as Uber's most recently introduced "Express Pool" service) and subsidizing the use of such services. Also, appealing to the range of co-travelers one has the possibility to meet, alongside campaigns to reduce privacy-sensitivity among individuals of non-White Hispanic ethnicity and high-income individuals, may be additional policy instruments available to promote pooled ride-hailing. Compared to traditional car-pooling arrangements that typically have scheduled times of arrival and departure, pooled ride-hailing would offer more time flexibility for workers and allow for them to use travel time productively.

### 5.2 Accessibility versus Convenience Mobility Tool

Ride-hailing can provide more access to activity opportunities for individuals who do not own vehicles and/or those with limited driving capabilities. Our model results provide initial evidence for this, as we observe that students, individuals with low vehicle availability, and individuals from low-income households are generally more likely than their peers to use ride-hailing to run errands. Thus, ride-hailing can assume a welfare role, but fares would need to be revisited to fit the needs of these more financially challenged segments of our society. On the other hand, the result that a higher private vehicle availability leads to a higher experience with ride-hailing suggests that, in an area such as DFW where almost all households own at least one vehicle, ride-hailing serves as more of a convenience feature for those one-off trips rather than being an accessibility facilitator for routine trips. Indeed, we also observe that vehicle availability has a negative effect on ride-hailing frequency.

An observation from the trip purpose results is that women are rather unlikely to use ride-hailing for routine errand trips, even though the women in a household are primarily responsible for personal, family, and shopping errands. At the same time, the "mode substituted" model results reveal that many of the new trips generated by the availability of ride-hailing (and that would not have been made otherwise) are for running errands. The implication is that, while ride-hailing provides more access to activity opportunities, it is also not the most convenient for running errands. This is perhaps because running errands typically involves chaining of multiple activities in the same sojourn from home and/or involves carrying and storing food and other perishable goods during the trip, and ride-hailing is not the most convenient because it is more of a pure trip-based consumption service as opposed to a broader transportation option that allows a cost-effective time-based consumption service (in which the same vehicle is available to pursue multiple activities and over an extended period of time). Perhaps ride-hailing providers need to be thinking about providing a time-based option too, which effectively would combine today's ride-hailing and car-sharing services into one service. As the mobility landscape moves more toward automated vehicles, this integration of trip-based and time-based consumption options may become even easier to implement.

### 5.3 Relationship with Transit and Active Travel

The results relative to substitution of transit and active travel by ride-hailing trips provide relevant insights for the planning of multi-modal urban transportation systems. First, the results reveal that people of age younger than 65 years of age are more likely than those 65 years of age or older to substitute active travel/transit by ride-hailing. This can further reduce the physical activity levels of individuals, and pose additional public health problems given that regular physical activity levels have now been proven to be effective as preventive medicine for a number of obesity-related diseases (Ku et al., 2018; Stamatakis et al., 2018). This is particularly of concern, since about 85.5% of the US population is under the age of 65 years (U.S. Census Bureau, 2018e) and obesity-related problems earlier on in life do lead to poor health outcomes later on in life (Cheng et al., 2016; Zheng et al., 2017). Thus, from both a traffic congestion standpoint as well as a health standpoint, policies that discourage the substitution of short-distance "walkable" trips by ride-hailing (such as a pricing scheme that more heavily prices the first mile, except if the patron is mobility-challenged) would be particularly valuable. Second, pooled ride-hailing users are more likely to have been drawn away from transit and active travel. The door-to-door travel convenience and relatively low-cost differential (between pooled ride-hailing and transit) appears to lead to the substitution of transit by ride-hailing (see also Clewlow and Mishra, 2017). Of course, to increase the efficiency and sustainability of a multi-modal transportation system, the relationship between (pooled) ride-hailing and transit should be one of complementarity rather than substitution. Yet, it is reasonable to expect that a service that can be used for door-to-door trips will not be used for first- and last-mile connectivity to transit hubs, unless low cost and well-integrated systems are designed.

### 5.4 Latent Demand and Trip Induction

Finally, our results point that the ride-hailing induced trips (trips that would not have been pursue if ride-hailing were not available) observed in our sample may be a significant point of concern for future planning. This is because they are being generated by household with already high vehicle availability in suburban and urban areas, serve a single passenger, and occur in the morning commute period as well as the mid-day and night periods (see the last column of Table 7b). In other words, ride-hailing is generating more "drive alone" trips in the already-congested suburban and urban areas of the DFW area to serve individuals that already have the resources to undertake drive-alone trips, contrary to the main transportation and environmental goals of sustainable and equitable urban systems. This is a serious concern in an era of dwindling real estate and financial assets to build new roads, along with increasing urban populations. In this sense, there is a need for more consideration of congestion pricing schemes that discourage private ride-hailing (especially in the morning commute period), as well as a need to re-visit the criteria and fee structure for the use of managed lanes (for example, a high-occupancy vehicle may have to be defined as 3+ individuals in the vehicle as opposed to 2+).

## 5.5 Conclusion

This study provided relevant insights on current ride-hailing adoption and usage as well as possible aspects that should be considered when planning to incorporate these services within the bigger picture of urban transportation. Nevertheless, the current study also contains many limitations that need to be stressed and potentially addressed by future research. First, the sample investigated is not representative of the DFW population and excludes non-commuters. Understanding the ride-hailing travel behavior of the population as a whole, including all segments, is extremely important from a transportation policy standpoint and thus, more comprehensive samples should be considered by new studies. Second, with the still recent growth of ride-hailing, it is plausible that many of the currently ride-hailing users are early adopters and the observed behaviors reflect an adaptation to the new service. Therefore, as ride-hailing becomes a reality to most of the population, the overall characteristics of ride-hailing users and their trip characteristics may significantly differ from the current observations. Third, the trip-level model in this paper is more of an exploratory nature, and thus, the variable effects on the many dimensions of ride-hailing should be viewed with much more caution than for the individual-level model. As discussed earlier, the impact of socio-demographics on general activity-travel patterns may be confounded with ride-hailing trip characteristics in some of the observed effects. To generate predictive models, future studies should seek to measure ride-hailing behavior within the broader picture of activity scheduling and mode choice (including mode and trip attributes), while also measuring built-environment characteristics. Indeed, our result that the key ingredient to pooled ride-hailing use is still urban density, even after accounting for self-selection effects and considering that our area of analysis has a high share of suburban land (DFW), points to the importance of including more detailed built-environment variables in future models. Finally, it is important for planning agencies to collect data on ride-hailing and incorporate ride-hailing behavior within their activity-travel modeling systems. Doing so is not only important for forecasting activity-travel patterns, but also to design good multi-modal systems through an understanding of how ride-hailing may be integrated with other travel modes.

## ACKNOWLEDGEMENTS

This research was partially supported by Cintra and the U.S. Department of Transportation through the Data-Supported Transportation Operations and Planning (D-STOP) Tier 1 University Transportation Center. The first author acknowledges funding support from CAPES and the Brazilian International Mobility Program "Ciência sem Fronteiras". The authors are grateful to

Lisa Macias for her help in formatting this document. Two anonymous reviewers provided useful comments on an earlier version of this paper.

## REFERENCES

Alemi, F., Circella, G., Mokhtarian, P., Handy, S., 2018a. On-demand ride services in California: Investigating the factors affecting the frequency of use of Uber/Lyft. Presented at the 97th Annual Meeting of the Transportation Research Board, Washington, D.C., January, Paper No. 18-05563.

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

Alemi, F., Circella, G., Mokhtarian, P., Handy, S., 2018c. Exploring the latent constructs behind the use of ridehailing in California. *Journal of Choice Modelling* 29, 47-62.

Amir, O., Levav, J., 2008. Choice construction versus preference construction: The instability of preferences learned in context. *Journal of Marketing Research* 45(2), 145-158.

Astroza, S., Garikapati, V.M., Bhat, C.R., Pendyala, R.M., Lavieri, P.S., Dias, F.F., 2017. Analysis of the impact of technology use on multimodality and activity travel characteristics. *Transportation Research Record: Journal of the Transportation Research Board* 2666, 19-28.

Astroza, S., Bhat, P.C., Bhat, C.R., Pendyala, R.M., Garikapati, V.M., 2018. Understanding activity engagement across weekdays and weekend days: A multivariate multiple discrete-continuous modeling approach. *Journal of Choice Modelling* 28, 56-70.

Barlow, F.K., Paolini, S., Pedersen, A., Hornsey, M.J., Radke, H.R., Harwood, J., Rubin, M., Sibley, C.G., 2012. The contact caveat: Negative contact predicts increased prejudice more than positive contact predicts reduced prejudice. *Personality and Social Psychology Bulletin* 38(12), 1629-1643.

Bhat, C.R., 2011. The maximum approximate composite marginal likelihood (MACML) estimation of multinomial probit-based unordered response choice models. *Transportation Research Part B* 45(7), 923-939.

Bhat, C.R., 2015a. A new generalized heterogeneous data model (GHDM) to jointly model mixed types of dependent variables. *Transportation Research Part B* 79, 50-77.

Bhat, C.R., 2015b. A comprehensive dwelling unit choice model accommodating psychological constructs within a search strategy for consideration set formation. *Transportation Research Part B* 79, 161-188.

Bhat, C., Paleti, R., Pendyala, R., Lorenzini, K., Konduri, K., 2013. Accommodating immigration status and self-selection effects in a joint model of household auto ownership and residential location choice. *Transportation Research Record: Journal of the Transportation Research Board* 2382, 142-150.

Bhat, C.R., Astroza, S., Bhat, A.C., Nagel, K., 2016. Incorporating a multiple discrete-continuous outcome in the generalized heterogeneous data model: Application to residential self-selection effects analysis in an activity time-use behavior model. *Transportation Research Part B* 91, 52-76.

Cheng, H.L., Medlow, S., Steinbeck, K., 2016. The health consequences of obesity in young adulthood. *Current Obesity Reports* 5(1), 30-37.

Chevalier, M., Gutsatz, M., 2012. *Luxury Retail Management: How the World's Top Brands Provide Quality Product and Service Support*. John Wiley & Sons, Singapore.

Clewlow, R., Mishra, G.S., 2017. Shared mobility: Current adoption, use, and potential impacts on travel behavior. Presented at the 96th Annual Meeting of the Transportation Research Board, Washington, D.C., January, Paper No. 17-05729.

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

English, T., Carstensen, L.L., 2014. Selective narrowing of social networks across adulthood is associated with improved emotional experience in daily life. *International Journal of Behavioral Development* 38(2), 195-202.

Fan, Y., 2017. Household structure and gender differences in travel time: spouse/partner presence, parenthood, and breadwinner status. *Transportation* 44(2), 271-291.

Franzen, A., Vogl, D., 2013. Two decades of measuring environmental attitudes: A comparative analysis of 33 countries. *Global Environmental Change* 23(5), 1001-1008.

Garikapati, V.M., Pendyala, R.M., Morris, E.A., Mokhtarian, P.L., McDonald, N., 2016. Activity patterns, time use, and travel of millennials: A generation in transition? *Transport Reviews* 36(5), 558-584.

Gerte, R., Konduri, K.C., Eluru, N., 2018. Is there a limit to adoption of dynamic ridesharing systems? Evidence from analysis of Uber demand data from New York. Presented at the 97th Annual Meeting of the Transportation Research Board, Washington, D.C., January, Paper No. 18-03265.

Gifford, R., Nilsson, A., 2014. Personal and social factors that influence pro-environmental concern and behaviour: A review. *International Journal of Psychology* 49(3), 141-157.

Gilg, A., Barr, S., Ford, N., 2005. Green consumption or sustainable lifestyles? Identifying the sustainable consumer. *Futures* 37(6), 481-504.

Giuliano, G., 2003. Travel, location and race/ethnicity. *Transportation Research Part A* 37(4), 351-372.

Gutiérrez, J.L.G., Jiménez, B.M., Hernández, E.G., Pcn, C., 2005. Personality and subjective well-being: Big five correlates and demographic variables. *Personality and Individual Differences* 38(7), 1561-1569.

Hampshire, R.C., Simek, C., Fabusuyi, T., Di, X., Chen, X., 2018. Measuring the impact of an unanticipated suspension of ride-sourcing in Austin, Texas. Presented at the 97th Annual Meeting of the Transportation Research Board, Washington, D.C., January, Paper No. 18-03105.

Haustein, S., 2012. Mobility behavior of the elderly: an attitude-based segmentation approach for a heterogeneous target group. *Transportation* 39(6), 1079-1103.

Helsper, E.J., Eynon, R., 2010. Digital natives: Where is the evidence? *British Educational Research Journal* 36(3), 503-520.

Hoeffler, S., Ariely, D., 1999. Constructing stable preferences: A look into dimensions of experience and their impact on preference stability. *Journal of Consumer Psychology* 8(2), 113-139.

Hunecke, M., Haustein, S., Böhler, S., Grischkat, S., 2010. Attitude-based target groups to reduce the ecological impact of daily mobility behavior. *Environment and Behavior* 42(1), 3-43.

Klein, N.J., Guerra, E., Smart, M.J., 2018. The Philadelphia story: Age, race, gender and changing travel trends. *Journal of Transport Geography* 69, 19-25.

Komanduri, A., Wafa, Z., Proussaloglou, K., 2018. Assessing the impact of app-based ride share systems in an urban context: Findings from Austin. Presented at the 97th Annual Meeting of the Transportation Research Board, Washington, D.C., January, Paper No. 18-01356.

Kooti, F., Grbovic, M., Aiello, L.M., Djuric, N., Radosavljevic, V., Lerman, K., 2017. Analyzing Uber's ride-sharing economy. *Proceedings of the 26th International Conference on World Wide Web Companion*, pp. 574-582. International World Wide Web Conferences Steering Committee.

Ku, P.W., Chen, L.J., Fox, K.R., Chen, Y.H., Liao, Y., Lin, C.H., 2018. Leisure-time, domestic, and work-related physical activity and their prospective associations with all-cause mortality in patients with cardiovascular disease. *American Journal of Cardiology* 121(2), 177-181.

Lavieri, P.S., Garikapati, V.M., Bhat, C.R., Pendyala, R.M., Astroza, S., Dias, F.F., 2017. Modeling individual preferences for ownership and sharing of autonomous vehicle technologies. *Transportation Research Record: Journal of the Transportation Research Board* 2665, 1-10.

Lavieri, P.S., Dias, F.F., Juri, N.R., Kuhr, J., Bhat, C.R., 2018a. A model of ridesourcing demand generation and distribution. *Transportation Research Record: Journal of the Transportation Research Board* 2672(46), 31-40.

Lavieri, P.S., Dai, Q., Bhat, C.R., 2018b. Using virtual accessibility and physical accessibility as joint predictors of activity-travel behavior. *Transportation Research Part A* 118, 527-544.

Leistner, D.L., Steiner, R.L., 2017. Uber for seniors? Exploring transportation options for the future. *Transportation Research Record: Journal of the Transportation Research Board* 2660, 22-29.

Li, Z., Hong, Y., Zhang, Z., 2017. An empirical analysis of on-demand ride sharing and traffic congestion. Presented at the 2016 International Conference on Information Systems (ICIS 2016), Dublin, September.

Liu, N., Yu, R., 2017. Identifying design feature factors critical to acceptance and usage behavior of smartphones. *Computers in Human Behavior* 70, 131-142.

Liu, X., Vedlitz, A., Shi, L., 2014. Examining the determinants of public environmental concern: Evidence from national public surveys. *Environmental Science & Policy* 39, 77-94.

Maddala, G.S., 1983. *Limited-Dependent and Qualitative Variables in Econometrics* (No. 3). Cambridge University Press, Cambridge, MA.

Milojev, P., Sibley, C.G., 2017. Normative personality trait development in adulthood: A 6-year cohort-sequential growth model. *Journal of Personality and Social Psychology* 112(3), 510.

Paleti, R., Eluru, N., Bhat, C.R., 2010. Examining the influence of aggressive driving behavior on driver injury severity in traffic crashes. *Accident Analysis & Prevention* 42(6), 1839-1854.

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

Rieser-Schüssler, N., Axhausen, K., 2012. Investigating the influence of environmentalism and variety seeking on mode choice. *Transportation Research Record: Journal of the Transportation Research Board* 2322, 31-41.

Schaller, B., 2017. Empty seats, full streets: Fixing Manhattan's traffic problem. Schaller Consulting, 1(3). Available at: http://www.schallerconsult.com/rideservices/emptyseats.htm [Accessed: 02/22/2019]

Schlobach, M., Retzer, S., 2018. Didi Chuxing – How China's ride-hailing leader aims to transform the future of mobility. Sustainable Transport China [online]. https://www.sustainabletransport.org/archives/6317 [Accessed: 02/22/2019]

Schwartz, S.H., 1992. Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. *Advances in Experimental Social Psychology* 25, 1-65.

Schwartz, S.H., Rubel, T., 2005. Sex differences in value priorities: Cross-cultural and multimethod studies. *Journal of Personality and Social Psychology* 89(6), 1010-1028.

Schwartz, S.H., Melech, G., Lehmann, A., Burgess, S., Harris, M., Owens, V., 2001. Extending the cross-cultural validity of the theory of basic human values with a different method of measurement. *Journal of Cross-Cultural Psychology* 32(5), 519-542.

Smith, A., 2016. Shared, collaborative and on demand: The new digital economy. Pew Research Center, Washington, D.C. Available at: http://www.pewinternet.org/2016/05/19/the-new-digital-economy/ [Accessed: 02/06/2018].

Spears, S., Houston, D., Boarnet, M.G., 2013. Illuminating the unseen in transit use: A framework for examining the effect of attitudes and perceptions on travel behavior. *Transportation Research Part A* 58, 40-53.

Stamatakis, E., de Rezende, L.F., Rey-López, J.P., 2018. Sedentary behaviour and cardiovascular disease. In Leitzmann, M.F., Jochem, C., Schmid, D. (Eds.) *Sedentary Behaviour Epidemiology*, pp. 215-243, Springer.

Statista, 2019. Ride-hailing worldwide: Penetration rate chart [online]. Available at: https://www.statista.com/outlook/368/100/ride-hailing/worldwide#market-revenue [Accessed: 02/22/2019]

Tudela, A., Habib, K.M.N., Carrasco, J.A., Osman, A.O., 2011. Incorporating the explicit role of psychological factors on mode choice: A hybrid mode choice model by using data from an innovative psychometric survey. Presented at the 2nd International Choice Modelling Conference, Leeds.

Twenge, J.M., 2013. Does online social media lead to social connection or social disconnection? *Journal of College and Character* 14(1), 11-20.

U.S. Census Bureau, 2018a. New Census Bureau Population Estimates Show Dallas-Fort Worth-Arlington Has Largest Growth in the United States. Release Number: CB18-50. Available at: https://www.census.gov/newsroom/press-releases/2018/popest-metro-county.html [Accessed: 03/23/2018].

U.S. Census Bureau, 2018b. Means of Transportation to Work by Selected Characteristics, Table S0802. 2012-2016 American Community Survey 5-Year Estimates. Available at: https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_16_5YR_S0802&prodType=table [Accessed: 03/23/2018].

U.S. Census Bureau, 2018c. Household Size by Vehicles Available, Table B08201. 2012-2016 American Community Survey 5-Year Estimates. Available at: https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_16_5YR_B08201&prodType=table [Accessed: 04/18/2018].

U.S. Census Bureau, 2018d. Employment Status, Table S2301. 2012-2016 American Community Survey 5-Year Estimates. Available at: https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_16_5YR_S2301&prodType=table [Accessed: 03/23/2018].

U.S. Census Bureau, 2018e. Age and Sex, Table S0101. 2012-2016 American Community Survey 5-Year Estimates. Available at: https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_16_5YR_S0101&prodType=table [Accessed: 05/16/2018].

Van Acker, V., Mokhtarian, P.L., Witlox, F., 2014. Car availability explained by the structural relationships between lifestyles, residential location, and underlying residential and travel attitudes. *Transport Policy* 35, 88-99.

Varin, C., Vidoni, P., 2005. A note on composite likelihood inference and model selection. *Biometrika* 92(3), 519-528.

Vianello, M., Schnabel, K., Sriram, N., Nosek, B., 2013. Gender differences in implicit and explicit personality traits. *Personality and Individual Differences* 55(8), 994-999.

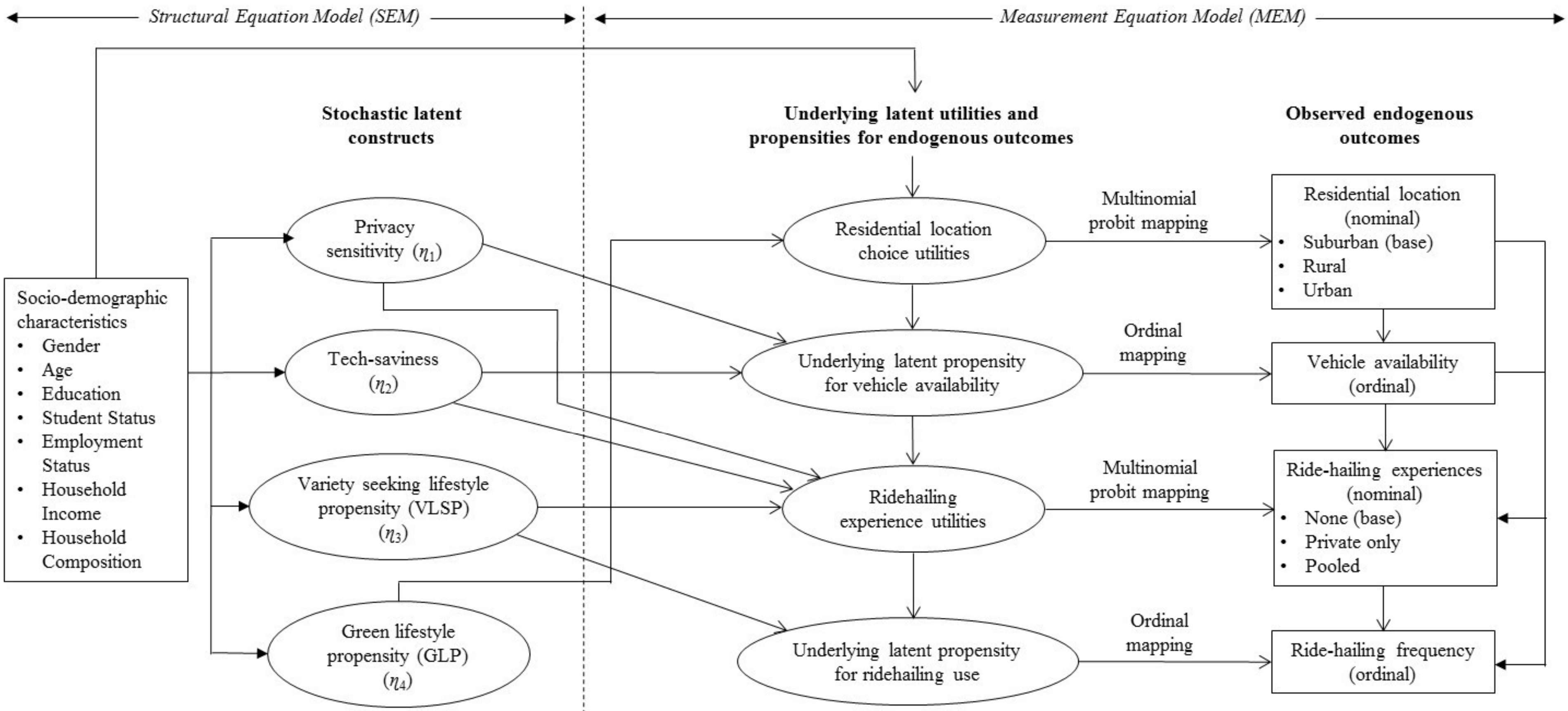
Ward, J.W., Michalek, J.J., Azevedo, I.L., Samaras, C., Ferreira, P., 2018. On-demand ridesourcing has reduced per-capita vehicle registrations and gasoline use in US States. Presented at the 97th Annual Meeting of the Transportation Research Board, Washington, D.C., January, Paper No. 18-05185.

Ye, R., Titheridge, H., 2017. Satisfaction with the commute: The role of travel mode choice, built environment and attitudes. *Transportation Research Part D* 52, 535-547.

Zebrowitz, L.A., White, B., Wieneke, K., 2008. Mere exposure and racial prejudice: Exposure to other-race faces increases liking for strangers of that race. *Social Cognition* 26(3), 259-275.

Zheng, Y., Manson, J.E., Yuan, C., Liang, M.H., Grodstein, F., Stampfer, M.J., Willett, W.C., Hu, F.B., 2017. Associations of weight gain from early to middle adulthood with major health outcomes later in life. *JAMA* 318(3), 255-269.

Zheng, H., Chen, X., Chen, X.M., 2018. How does on-demand ridesplitting influence vehicle use and ownership? A case study in Hangzhou, China. Presented at the 97th Annual Meeting of the Transportation Research Board, Washington, D.C., January, Paper No. 18-04327.




$$\boldsymbol{\eta} = \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \end{bmatrix} \sim MVN_4(\mathbf{0}, \boldsymbol{\Gamma}),\ \boldsymbol{\Gamma} = \begin{bmatrix} 1 & & & \\ \gamma_{12} & 1 & & \\ \gamma_{13} & \gamma_{23} & 1 & \\ \gamma_{14} & \gamma_{24} & \gamma_{34} & 1 \end{bmatrix}$$

**Figure 1. Diagrammatic Representation of the GHDM Model**

**Table 1. Sample distribution of socio-demographic characteristics**

| Variable | Count | % | % Ride-hailing experience |
|---|---|---|---|
| **Gender** | | | |
| Female | 668 | 41.57 | 54.04 |
| Male | 939 | 58.43 | 58.04 |
| **Age** | | | |
| 18-34 | 261 | 16.24 | 75.48 |
| 35-44 | 360 | 22.40 | 65.28 |
| 45-54 | 432 | 26.88 | 54.63 |
| 55-64 | 423 | 26.32 | 45.86 |
| 65 or more | 131 | 8.16 | 33.59 |
| **Ethnicity** | | | |
| Non-Hispanic White | 1205 | 74.98 | 55.19 |
| Non-Hispanic Black | 102 | 6.35 | 55.88 |
| Hispanic | 109 | 6.78 | 62.39 |
| Asian/Pacific Islander | 101 | 6.29 | 65.35 |
| Other | 90 | 5.60 | 55.55 |
| **Education** | | | |
| Completed high-school | 238 | 14.82 | 42.44 |
| Completed technical school/associates' degree | 154 | 9.58 | 59.74 |
| Completed undergraduate degree | 724 | 45.05 | 56.22 |
| Completed graduate degree | 491 | 30.55 | 62.32 |
| **Student (attending institution in person)** | | | |
| Yes | 93 | 5.79 | 65.59 |
| No | 1514 | 94.21 | 55.81 |
| **Employment type** | | | |
| Full-time employee | 1312 | 81.64 | 57.39 |
| Part-time employee | 138 | 8.59 | 51.45 |
| Self-employed | 157 | 9.77 | 52.23 |
| **Household income** | | | |
| Under $50,000 | 184 | 11.45 | 51.09 |
| $50,000-$99,999 | 443 | 27.57 | 46.50 |
| $100,000-$149,999 | 496 | 30.86 | 54.64 |
| $150,000-$199,999 | 269 | 16.74 | 63.94 |
| $200,000 or more | 215 | 13.38 | 75.81 |
| **Household composition** | | | |
| Single person household | 191 | 11.89 | 62.30 |
| Single worker multi-person household | 265 | 16.49 | 44.91 |
| Multi-worker household | 1151 | 71.62 | 58.04 |

**Table 2. Sample distribution of attitudinal and behavioral indicators (n=1607)**

| **Privacy-sensitivity** | | | | | |
|---|---|---|---|---|---|
| | Strongly disagree | Somewhat disagree | Neither | Somewhat agree | Strongly agree |
| I don't mind sharing a ride with strangers if it reduces my costs | 13.44% | 22.15% | 20.41% | 35.53% | 8.46% |
| Having privacy is important to me when I make a trip | 2.80% | 10.52% | 22.84% | 41.19% | 22.65% |
| I feel uncomfortable sitting close to strangers | 8.59% | 22.53% | 27.88% | 29.12% | 11.89% |
| **Tech-savviness** | | | | | |
| | Does not describe me at all | Describes me slightly well | Describes me moderately well | Describes me very well | Describes me extremely well |
| I frequently use online banking services | 2.43% | 3.42% | 6.41% | 18.67% | 69.07% |
| I frequently purchase products online | 1.24% | 7.28% | 14.87% | 23.58% | 53.02% |
| Learning how to use new smartphone apps is easy for me | 2.49% | 5.48% | 16.68% | 27.13% | 48.23% |
| **Variety-seeking lifestyle propensity (VSLP)** | | | | | |
| | Does not describe me at all | Describes me slightly well | Describes me moderately well | Describes me very well | Describes me extremely well |
| I think it is important to have all sorts of new experiences and I am always trying new things | 3.48% | 12.62% | 29.33% | 34.12% | 20.45% |
| Looking for adventures and taking risks is important to me | 13.36% | 24.98% | 33.25% | 21.81% | 6.59% |
| I love to try new products before anyone else | 6.90% | 15.91% | 28.28% | 30.33% | 18.58% |
| **Green lifestyle propensity (GLP)** | | | | | |
| | Strongly disagree | Somewhat disagree | Neither | Somewhat agree | Strongly agree |
| When choosing my commute mode, there are many things that are more important than being environmentally friendly | 4.60% | 15.93% | 28.38% | 34.85% | 16.24% |
| I don't give much thought to saving energy at home | 39.33% | 37.59% | 11.64% | 8.59% | 2.86% |

**Table 3. Sample distribution of trip characteristics (n=906)**

| Variable | Count | % |
|---|---|---|
| **Trip purpose** | | |
| Airport | 359 | 39.62 |
| Shopping, personal, or family errands | 86 | 9.49 |
| Recreational and leisure activities | 362 | 39.96 |
| Work or education | 99 | 10.93 |
| **Time-of-day** | | |
| Moring (6:00am-10:59am) | 191 | 21.08 |
| Mid-day (11:00am-3:59pm | 183 | 20.20 |
| Evening (4:00pm-8:59pm) | 305 | 33.66 |
| Night (9:00pm-5:59am) | 227 | 25.06 |
| **Companion** | | |
| Alone | 370 | 40.84 |
| With friends, family or co-workers | 523 | 57.73 |
| With a stranger (pooled ride) | 13 | 1.43 |
| **Mode substituted** | | |
| My own vehicle | 419 | 46.25 |
| Taxi | 347 | 38.30 |
| Transit, bicycle or walk | 87 | 9.60 |
| Would not have traveled | 53 | 5.85 |

**Table 4. Determinants of latent constructs**

| Variables (base category) | Structural Equations Model Component Results | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Privacy-sensitivity | | Tech-savviness | | VSLP | | GLP | |
| | Coeff. | t-stat | Coeff. | t-stat | Coeff. | t-stat | Coeff. | t-stat |
| **Gender (male)** | | | | | | | | |
| Female | -- | -- | -- | -- | -0.270 | -3.24 | 0.426 | 1.89 |
| **Age (55 years or more)** | | | | | | | | |
| 18-34 | -- | -- | 1.144 | 11.28 | 0.480 | 4.89 | -1.174 | -1.86 |
| 35-44 | -- | -- | 0.899 | 10.14 | 0.287 | 3.51 | -- | -- |
| 45-54 | -- | -- | 0.441 | 5.58 | -- | -- | -- | -- |
| **Ethnicity (other ethnicities)** | | | | | | | | |
| Non-Hispanic White | 0.187 | 1.98 | -- | -- | -0.177 | -3.34 | -- | -- |
| **Education (less than undergraduate degree)** | | | | | | | | |
| Graduate degree | -- | -- | -- | -- | -- | -- | 0.859 | 2.52 |
| **Employment (full-time)** | | | | | | | | |
| Part-time employee | -- | -- | -0.395 | -3.29 | -- | -- | -- | -- |
| Self-employed | -- | -- | -- | -- | -- | -- | -- | -- |
| **Household income** | | | | | | | | |
| **(Under $50,000)** | | | | | | | | |
| $50,000-$99,999 | -- | -- | 0.283 | 2.55 | -- | -- | -- | -- |
| $100,000-$149,999 | -- | -- | 0.446 | 3.94 | -- | -- | -- | -- |
| $150,000-$199,999 | -- | -- | 0.668 | 5.27 | -- | -- | -- | -- |
| $200,000 or more | 0.259 | 2.55 | 0.803 | 5.98 | 0.257 | 2.61 | -- | -- |
| **Household composition (multi-worker and single person)** | | | | | | | | |
| Single worker multi-person | -- | -- | -- | -- | -0.209 | -2.07 | -- | -- |
| **Correlations between latent variables** | | | | | | | | |
| Privacy-sensitivity | 1.000 | n/a | | | | | | |
| Tech-savviness | -- | -- | 1.000 | n/a | | | | |
| VSLP | -- | -- | 0.360 | 2.48 | 1.000 | n/a | | |
| GLP | -0.465 | -2.01 | -- | -- | -- | -- | 1.000 | n/a |

"--" = not statistically significantly different from zero at the 90% level of confidence and removed from the specification.
"n/a" = not applicable

## Table 5. Results of the residential location, vehicle availability, ride-hailing experience, and ride-hailing frequency model components

| Variables (base category) | Residential location (base: Suburban) | | | | Vehicle availability | | Ride-hailing experience (base: none) | | | | Ride-hailing frequency | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Rural | | Urban | | (ordinal) | | Private only | | Pooled | | (ordinal) | |
| | Coeff. | t-stat | Coeff. | t-stat | Coeff. | t-stat | Coeff. | t-stat | Coeff. | t-stat | Coeff. | t-stat |
| *Latent variables* | | | | | | | | | | | | |
| Privacy-sensitivity | -- | -- | -- | -- | 0.405 | 2.25 | -- | -- | -0.422 | -1.95 | -- | -- |
| Tech-savviness | -0.142 | -3.51 | -- | -- | -0.124 | -2.54 | 0.251 | 2.61 | 0.165 | 1.80 | -- | -- |
| VSLP | -- | -- | 0.081 | 1.79 | -- | -- | 0.068 | 1.68 | 0.271 | 1.83 | 0.166 | 2.17 |
| GLP | -0.376 | -1.89 | 0.203 | 1.85 | -- | -- | -- | -- | -- | -- | -- | -- |
| *Exogenous effects* | | | | | | | | | | | | |
| **Age (65 years or more)** | | | | | | | | | | | | |
| 18-34 | -0.789 | -4.54 | 0.657 | 4.57 | -- | -- | 0.739 | 6.97 | 0.677 | 2.72 | -- | -- |
| 35-44 | -- | -- | -- | -- | -- | -- | 0.508 | 7.93 | 0.432 | 1.90 | -- | -- |
| 45-54 | -- | -- | -- | -- | -- | -- | 0.213 | 3.83 | 0.326 | 1.85 | -- | -- |
| 55-64 | -- | -- | -- | -- | -- | -- | 0.161 | 2.99 | -- | -- | -- | -- |
| **Ethnicity (other ethnicities)** | | | | | | | | | | | | |
| Non-Hispanic White | -- | -- | -- | -- | -- | -- | -- | -- | -0.148 | -1.87 | -- | -- |
| **Education (less than undergraduate degree)** | | | | | | | | | | | | |
| Graduate degree | -- | -- | -- | -- | -- | -- | -- | -- | 0.186 | 4.54 | -- | -- |
| **Employment (full-time)** | | | | | | | | | | | | |
| Part-time employee | -- | -- | 0.369 | 9799 | -- | -- | -0.135 | -2.71 | -- | -- | -- | -- |
| Self-employed | 0.188 | 3.04 | 0.242 | 7.42 | -- | -- | -- | -- | -- | -- | -- | -- |
| **Household income (Under $100,000)** | | | | | | | | | | | | |
| $100,000-$149,999 | -- | -- | -- | -- | 0.519 | 6.16 | 0.326 | 6.67 | -- | -- | -- | -- |
| $150,000-$199,999 | -0.106 | -2.87 | -- | -- | 0.519 | 6.16 | 0.546 | 11.39 | 0.146 | 1.85 | -- | -- |
| $200,000 or more | -0.106 | -2.87 | -- | -- | 0.883 | 7.27 | 0.913 | 15.35 | 0.434 | 1.96 | 0.427 | 3.37 |
| **Household composition (multi-worker)** | | | | | | | | | | | | |
| Single person | -0.106 | -2.52 | 0.189 | 3.74 | 0.532 | 6.22 | 0.386 | 8.50 | -- | -- | -- | -- |
| Single worker multi-person | -- | -- | -- | -- | 1.638 | 15.94 | -0.176 | -2.94 | -0.243 | -2.25 | -- | -- |
| *Endogenous effects* | | | | | | | | | | | | |
| **Residential location (rural)** | | | | | | | | | | | | |
| Suburban | n/a | n/a | n/a | n/a | -- | -- | 0.332 | 2.03 | 0.392 | 1.93 | -- | -- |
| Urban | n/a | n/a | n/a | n/a | -0.175 | -2.23 | 0.668 | 4.24 | 0.777 | 4.50 | 0.190 | 1.74 |
| **Vehicle availability (less than 1 per worker)** | | | | | | | | | | | | |
| 1 per worker | n/a | n/a | n/a | n/a | n/a | n/a | -- | -- | -- | -- | -0.239 | -1.79 |
| More than 1 per worker | n/a | n/a | n/a | n/a | n/a | n/a | 0.084 | 2.30 | 0.183 | 4.70 | -0.239 | -1.79 |
| **Pooled ride-hailing user (no)** | | | | | | | | | | | | |
| Yes | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | 0.655 | 4.80 |
| **Constant** | -0.759 | -2.24 | -0.876 | -4.47 | 0.680 | 7.88 | -1.172 | -4.73 | -1.702 | -8.26 | 0.246 | 1.6 |
| *Thresholds* | | | | | | | | | | | | |
| Threshold 2 | n/a | n/a | n/a | n/a | 1.688 | 28.91 | n/a | n/a | n/a | n/a | 0.870 | 13.97 |
| Threshold 3 | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | 1.338 | 17.13 |
| Threshold 4 | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | 2.037 | 17.51 |

"--" = not statistically significantly different from zero at the 90% level of confidence and removed from the specification.
"n/a" = not applicable

**Table 6. Treatment effects of different variables on ride-hailing adoption and frequency**

| Variable | Categories compared (base versus changed) | Private only | | Pooled | | Frequency | |
|---|---|---|---|---|---|---|---|
| | | Est. | St. err | Est. | St. err | Est. | St. err. |
| *Latent variables* | | | | | | | |
| Privacy-sensitivity | Min vs. Max | 0.000 | | -0.038 | (0.021) | 0.000 | |
| Tech-savviness | Min vs. Max | 0.160 | (0.016) | 0.029 | (0.018) | 0.000 | |
| VSLP | Min vs. Max | 0.007 | (0.007) | 0.028 | (0.019) | 0.626 | (0.151) |
| GLP | Min vs. Max | 0.000 | | 0.000 | | 0.000 | |
| *Sociodemographic variables* | | | | | | | |
| Gender | Male vs. female | 0.000 | | 0.000 | | 0.000 | |
| Age | 65 or more vs. 18-34 | 0.223 | (0.012) | 0.034 | (0.013) | 0.000 | |
| Ethnicity | Other vs. Non-Hispanic White | 0.000 | | -0.020 | (0.012) | 0.000 | |
| Education | < Bachelor's vs. graduate | 0.000 | | 0.029 | (0.007) | 0.000 | |
| Employment type | Full-time vs. part-time | -0.040 | (0.017) | 0.000 | | 0.000 | |
| Income | Under $50,000 vs. $200,000 or more | 0.290 | (0.028) | -0.021 | (0.014) | 1.194 | (0.134) |
| Household composition | Multi-worker vs. single-worker multi-person | -0.032 | (0.011) | -0.012 | (0.007) | 0.000 | |
| *Endogenous variables* | | | | | | | |
| Residential location | Rural vs. urban | 0.160 | (0.037) | 0.067 | (0.020) | 0.580 | (0.210) |
| Vehicle availability | Less than 1 vs. more than 1 per worker | 0.011 | (0.004) | 0.023 | (0.006) | -0.087 | (0.022) |
| Pooled ride-hailing user | No vs. yes | 0.000 | | 0.000 | | 2.062 | (0.234) |

## Table 7a. Trip characteristics model results

| Variables (base category) | Purpose (base: airport) | | | | | | Time (base: mid-day) | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Errands | | Recreation | | Work | | Morning | | Evening | | Night | |
| | Coeff. | t-stat | Coeff. | t-stat | Coeff. | t-stat | Coeff. | t-stat | Coeff. | t-stat | Coeff. | t-stat |
| **Latent variables** | | | | | | | | | | | | |
| VSLP | -- | -- | 0.279 | 4.23 | -- | -- | -- | -- | -- | -- | -- | -- |
| *Exogenous effects* | | | | | | | | | | | | |
| **Gender (male)** | | | | | | | | | | | | |
| Female | -0.139 | -3.43 | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- |
| **Age (35 years or more)** | | | | | | | | | | | | |
| 18-34 | -- | -- | 0.390 | 8.82 | 0.221 | 4.72 | -0.113 | -2.54 | -- | -- | 0.461 | 12.56 |
| **Ethnicity (other ethnicities)** | | | | | | | | | | | | |
| Non-Hispanic White | -- | -- | 0.342 | 10.10 | -- | -- | -- | -- | -- | -- | -- | -- |
| **Education (less than undergraduate degree)** | | | | | | | | | | | | |
| Undergraduate degree | -0.294 | -6.20 | -0.180 | -4.83 | -- | -- | -- | -- | -- | -- | -- | -- |
| Graduate degree | -0.294 | -6.20 | -0.238 | -5.88 | -0.092 | -2.21 | -- | -- | -- | -- | -- | -- |
| **Student (no)** | | | | | | | | | | | | |
| Yes | 0.562 | 8.73 | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- |
| **Employment (full-time)** | | | | | | | | | | | | |
| Part-time employee | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- |
| Self-employed | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- |
| **Household income (Under $50,000)** | | | | | | | | | | | | |
| $50,000-$99,999 | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- |
| $100,000-$149,999 | -0.236 | -5.14 | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- |
| $150,000-$199,999 | -0.403 | -6.86 | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- |
| $200,000 or more | -- | -- | -- | -- | -0.155 | -3.06 | -- | -- | -0.082 | -2.07 | -0.247 | -5.25 |
| **Household composition (multi-worker)** | | | | | | | | | | | | |
| Single person | -- | -- | -- | -- | -0.303 | -4.85 | -- | -- | -- | -- | -- | -- |
| Single worker multi-person | -- | -- | -- | -- | -- | -- | 0.217 | 5.54 | 0.217 | 5.54 | -- | -- |
| **Residential location (rural)** | | | | | | | | | | | | |
| Suburban | -0.361 | -4.85 | -0.330 | -5.62 | -0.373 | -5.18 | -- | -- | -0.323 | -5.32 | -0.215 | -3.09 |
| Urban | -0.361 | -4.85 | -0.163 | -2.63 | -0.373 | -5.18 | -- | -- | -0.399 | -6.20 | -0.358 | -4.85 |
| **Vehicle availability (Less than 1 per worker)** | | | | | | | | | | | | |
| 1 per worker | -0.248 | -4.71 | -- | -- | -0.162 | -3.17 | -- | -- | -- | -- | -- | -- |
| More than 1 per worker | -0.248 | -4.71 | -- | -- | -0.162 | -3.17 | -- | -- | -- | -- | 0.214 | 6.26 |
| **Ride-hailing frequent user (no)** | | | | | | | | | | | | |
| Yes | -0.283 | -5.42 | -- | -- | 0.264 | 6.69 | -- | -- | -- | -- | -0.109 | -3.15 |
| **Pooled ride-hailing user (no)** | | | | | | | | | | | | |
| Yes | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- |
| *Endogenous Effects* | | | | | | | | | | | | |
| **Trip purpose (airport)** | | | | | | | | | | | | |
| Errands | | | | | | | -- | -- | -- | -- | -- | -- |
| Recreational | | | | | | | -- | -- | 1.033 | 29.86 | 1.268 | 34.24 |
| Work | | | | | | | 0.284 | 5.61 | 0.284 | 5.61 | -0.209 | -2.83 |
| **Constant** | 0.251 | 2.78 | 0.101 | 1.54 | -0.202 | -2.48 | -0.058 | -2.50 | 0.193 | 3.23 | -0.250 | -3.57 |

"--" = not statistically significantly different from zero at the 90% level of confidence and removed from the specification.

**Table 7b. Trip characteristics model results (cont.)**

| Variables (base category) | Trip companion (base: alone) | | Mode substituted (base: own car) | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Not alone | | Taxi | | Active or transit | | No trip | |
| | Coeff. | t-stat | Coeff. | t-stat | Coeff. | t-stat | Coeff. | t-stat |
| *Exogenous effects* | | | | | | | | |
| **Gender (male)** | | | | | | | | |
| Female | 0.093 | 2.85 | -- | -- | -0.255 | -5.73 | -- | -- |
| **Age (65 years or more)** | | | | | | | | |
| 18-34 | -- | -- | -0.273 | -6.74 | 0.639 | 4.61 | 0.280 | 5.60 |
| 35-44 | 0.077 | 2.41 | -- | -- | 0.862 | 6.30 | 0.280 | 5.60 |
| 45-54 | 0.077 | 2.41 | -- | -- | 0.557 | 4.08 | -- | -- |
| 55-64 | -- | -- | -- | -- | 0.557 | 4.08 | -- | -- |
| **Ethnicity (other ethnicities)** | | | | | | | | |
| Non-Hispanic White | -0.225 | -6.17 | 0.092 | 2.79 | -- | -- | -- | -- |
| **Education (less than undergraduate degree)** | | | | | | | | |
| Completed undergraduate degree | -- | -- | -- | -- | 0.135 | 2.64 | -- | -- |
| Completed graduate degree | -0.153 | -4.46 | 0.089 | 2.97 | 0.135 | 2.64 | -- | -- |
| **Student (no)** | | | | | | | | |
| Yes | -- | -- | 0.201 | 2.62 | -- | -- | -- | -- |
| **Employment type (full-time)** | | | | | | | | |
| Part-time employee | 0.291 | 4.26 | 0.550 | 8.39 | 0.581 | 8.20 | 0.180 | 2.78 |
| Self-employed | -- | -- | -0.133 | -2.60 | -- | -- | 0.180 | 2.78 |
| **Household income (Under $50,000)** | | | | | | | | |
| $50,000-$99,999 | 0.123 | 3.69 | 0.383 | 5.96 | -- | -- | -- | -- |
| $100,000-$149,999 | -- | -- | 0.383 | 5.96 | -0.331 | -7.38 | -- | -- |
| $150,000 or more | -- | -- | 0.497 | 7.33 | -0.331 | -7.38 | -- | -- |
| **Household composition (multi-worker)** | | | | | | | | |
| Single person | -0.618 | -13.35 | -- | -- | -- | -- | -0.295 | -3.52 |
| Single worker multi-person | -- | -- | -- | -- | -- | -- | -0.316 | -4.22 |
| **Residential location (rural)** | | | | | | | | |
| Suburban | -- | -- | -0.176 | -3.28 | -- | -- | 0.600 | 4.19 |
| Urban | -- | -- | -0.269 | -4.65 | -- | -- | 0.600 | 4.19 |
| **Vehicle availability (Less than 1 per worker)** | | | | | | | | |
| 1 per worker | -- | -- | -0.111 | -3.95 | -- | -- | -- | -- |
| More than 1 per worker | -- | -- | -- | -- | -- | -- | 0.366 | 6.40 |
| **Ride-hailing frequent user (no)** | | | | | | | | |
| Yes | -- | -- | -- | -- | -0.332 | -6.65 | -- | -- |
| **Pooled ride-hailing user (no)** | | | | | | | | |
| Yes | -- | -- | -- | -- | 0.182 | 3.72 | -- | -- |
| *Endogenous Effects* | | | | | | | | |
| **Trip purpose (airport)** | | | | | | | | |
| Errands | -0.484 | -7.65 | -- | -- | -- | -- | 1.061 | 15.88 |
| Recreational | 0.928 | 23.94 | -- | -- | -- | -- | 0.527 | 7.52 |
| Work | -0.671 | -12.36 | -- | -- | -- | -- | -- | -- |
| **Trip time (mid-day)** | | | | | | | | |
| Morning | -0.144 | -3.41 | -- | -- | -- | -- | -- | -- |
| Evening | -- | -- | -0.138 | -4.12 | -0.140 | -2.85 | -0.254 | -4.38 |
| Night | -- | -- | -- | -- | -0.295 | -5.42 | -- | -- |
| **Trip companion (alone)** | | | | | | | | |
| Not alone | | | -- | -- | 0.218 | 5.00 | -0.366 | -5.50 |
| **Constant** | 0.194 | 4.01 | -0.292 | -3.60 | -1.286 | -9.31 | -1.974 | -13.25 |

"--" = not statistically significantly different from zero at the 90% level of confidence and removed from the specification.

## Appendix A

**The Multivariate Multinomial Probit (MMNP) Approach**

The model adopted for the analysis of trip-level attributes is the MMNP that allows flexible covariances due to unobserved elements within the utilities of each trip dimension's alternatives, and also allows covariances across the utilities of different trip dimensions. Let $g$ be the index for the nominal variables ($g$ = 1, 2,…, $G$), in the current empirical context, $G$=4. Also, let $I_g$ be the number of alternatives corresponding to the $g^{\text{th}}$ nominal variable ($I_g \geq 2$) and let $i_g$ be the corresponding index ($i_g$ = 1, 2, 3,…, $I_g$). In the current empirical context, $I_1$ = 4 (four trip purpose categories), $I_2$ = 4 (four time of day categories), $I_3$ = 2 (two companionship categories, and $I_4$ = 4 (four mode substituted categories). Under the usual assumptions of random utility theory, the utility associated with alternative $i_g$ of the $g^{th}$ nominal variable may be written as:

$$U_{gi_g} = \boldsymbol{\beta}'_g \boldsymbol{x}_{gi_g} + \varepsilon_{gi_g}, \tag{1}$$

where $\boldsymbol{x}_{gi_g}$ is a ($K_g$×1)-column vector of exogenous and endogenous explanatory variables, following the assumptions and endogeneity hierarchy discussed Section 2.3 of the paper. $\boldsymbol{\beta}_g$ is a column vector of corresponding coefficients, and $\varepsilon_{gi_g}$ is a normal error term. Let the variance-covariance matrix of the vertically stacked vector of errors $\varepsilon_g[=(\varepsilon_{g1}, \varepsilon_{g2}, ..., \varepsilon_{gI_g})']$ be $\boldsymbol{\Omega}_g$. Assuming that the household chooses the alternative $m_g$, all the utility differences with respect to this chosen alternative $m_g$ must be less than zero. These conditions can be denoted numerically as:

$$U_{gi_g} - U_{gm_g} < 0 \; \forall \, i_g \neq m_g \tag{2}$$

Let $y^*_{gi_gm_g} = U_{gi_g} - U_{gm_g} \, (i_g \neq m_g)$, and $\boldsymbol{y}^*_g$ be the stacked vector of the latent utility differentials $\boldsymbol{y}^*_g = \left[\left(y^*_{g1m_g}, y^*_{g2m_g}, ..., y^*_{gI_gm_g}\right)'; \; i_g \neq m_g\right]$. $\boldsymbol{y}^*_g$ has a mean vector of $\boldsymbol{B}_g \, (\boldsymbol{\beta}'_1\boldsymbol{z}_{g1m_g}, \boldsymbol{\beta}'_1\boldsymbol{z}_{g2m_g}, ..., \boldsymbol{\beta}'_1\boldsymbol{z}_{gI_gm_g})'$, where $\boldsymbol{z}_{gi_gm_g} = \boldsymbol{x}_{gi_g} - \boldsymbol{x}_{gm_g}, i_g = 1, 2, ..., I_g; \; i_g \neq m_g$ and a covariance matrix given by $\boldsymbol{\Sigma}^*_g = \boldsymbol{M}_g \boldsymbol{\Omega}_g \boldsymbol{M}'_g$, where $\boldsymbol{M}_g$ is an $(I_g - 1) \times I_g$ matrix that corresponds to an $(I_g - 1)$ identity matrix with an extra column of –1's added as the $m_g^{th}$ column. Thus, we can write

$$\boldsymbol{y}^*_g \sim N\left(\boldsymbol{B}_g, \boldsymbol{\Sigma}^*_g\right), \tag{3}$$

Now, for the four nominal variables, consider the stacked $\tilde{G} = (I_1 + I_2 + I_3 + I_4 - 4)$ vector $\boldsymbol{y}^* = \left[\left(\boldsymbol{y}^{*'}_1, \boldsymbol{y}^{*'}_2, \boldsymbol{y}^{*'}_3, \boldsymbol{y}^{*'}_4\right)'\right]$, each of whose element vectors is formed by differencing utilities of alternatives from the chosen alternative $m_g$ for the $g^{th}$ nominal variable. Then, we may write: $\boldsymbol{y}^* \sim N(\boldsymbol{B}, \boldsymbol{\Sigma}^*)$, where $\boldsymbol{B} = \left(\boldsymbol{B}'_1, \boldsymbol{B}'_2, \boldsymbol{B}'_3, \boldsymbol{B}'_4\right)'$ and $\boldsymbol{\Sigma}^*$ is a $\tilde{G} \times \tilde{G}$ matrix as follows:

$$\mathbf{\Sigma}^* = \begin{bmatrix} \mathbf{\Sigma}_1^* & & & \\ \mathbf{\Sigma}_{21}^* & \mathbf{\Sigma}_2^* & & \\ \mathbf{\Sigma}_{31}^* & \mathbf{\Sigma}_{32}^* & \mathbf{\Sigma}_3^* & \\ \mathbf{\Sigma}_{41}^* & \mathbf{\Sigma}_{42}^* & \mathbf{\Sigma}_{43}^* & \mathbf{\Sigma}_4^* \end{bmatrix} \qquad (4)$$

The off-diagonal elements in $\mathbf{\Sigma}^*$ capture the dependencies across the utility differentials of the four nominal variables, the differential being taken with respect to the chosen alternative for each nominal variable. Let $\boldsymbol{\theta}$ be the collection of parameters to be estimated: $\boldsymbol{\theta} = [\boldsymbol{\beta}_1, \boldsymbol{\beta}_2; \text{Vech}(\mathbf{\Sigma}^*)]$, where Vech($\mathbf{\Sigma}$) represents the vector of lower triangle elements of $\mathbf{\Sigma}$ (as represented in Eq. 4). Then the likelihood function for a household may be written as:

$$L(\boldsymbol{\theta}) = F_{\tilde{G}}\,[-\tilde{\boldsymbol{B}}\,, \mathbf{\Sigma}^*]\,, \qquad (5)$$

where $F_{\tilde{G}}(.,.)$ is the $\tilde{G} = (I_1 + I_2 + I_3 + I_4 - 4)$-dimensional normal cumulative distribution function.

The above likelihood function involves the evaluation of a $\tilde{G} = (I_1 + I_2 + I_3 + I_4 - 4)$-dimensional integral for each individual, which can be very computationally expensive if each nominal variable can take a large number of values. Therefore, in this study, the Maximum Approximate Composite Marginal Likelihood (MACML) approach (Bhat, 2011) is used. In this approach, the likelihood function only involves the computation of univariate and bivariate cumulative distributive functions. We refer the reader to Bhat (2011) for more details regarding identification considerations and estimation.

## Appendix B

### Table B1. Thresholds and constants of indicators and loadings of latent variables on indicators

| Attitudinal and lifestyle indicators | Threshold 2 | | Threshold 3 | | Threshold 4 | | Constant | | Latent variable loading | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Coeff. | t-stat | Coeff. | t-stat | Coeff. | t-stat | Coeff. | t-stat | Coeff. | t-stat |
| **Privacy-sensitivity** | | | | | | | | | | |
| I don't mind sharing a ride with strangers if it reduces my costs (inverse scale) | 2.523 | 19.85 | 3.598 | 21.06 | 5.123 | 19.08 | 2.504 | 12.84 | 1.792 | 14.09 |
| Having privacy is important to me when I make a trip | 0.922 | 12.13 | 1.799 | 22.17 | 3.076 | 33.69 | 2.101 | 23.01 | 0.575 | 16.21 |
| I feel uncomfortable sitting close to strangers | 0.954 | 17.55 | 1.737 | 25.04 | 2.777 | 25.44 | 1.409 | 22.24 | 0.427 | 6.19 |
| **Tech-savviness** | | | | | | | | | | |
| I frequently use online banking services | 1.133 | 8.67 | 2.606 | 18.136 | 4.099 | 28.56 | 2.559 | 12.83 | 1.601 | 55.44 |
| I frequently purchase products online | 0.506 | 6.475 | 1.017 | 11.17 | 1.849 | 19.27 | 1.861 | 14.69 | 0.681 | 26.15 |
| Learning how to use new smartphone apps is easy for me | 1.138 | 9.685 | 1.993 | 16.22 | 2.859 | 23.18 | 2.255 | 15.08 | 0.787 | 30.61 |
| **Variety-seeking lifestyle propensity (VSLP)** | | | | | | | | | | |
| I think it is important to have all sorts of new experiences and I am always trying new things | 1.159 | 13.78 | 2.374 | 26.41 | 3.676 | 35.80 | 2.631 | 19.33 | 0.930 | 22.40 |
| Looking for adventures and taking risks is important to me | 1.195 | 2.45 | 2.468 | 2.33 | 3.834 | 2.17 | 1.739 | 2.67 | 1.033 | 23.83 |
| I love to try new products before anyone else | 0.910 | 6.67 | 1.859 | 7.37 | 2.934 | 7.38 | 1.908 | 6.88 | 0.704 | 2.69 |
| **Green lifestyle propensity (GLP)** | | | | | | | | | | |
| When choosing my commute mode, there are many things that are more important than being environmentally friendly (inverse scale) | 1.045 | 15.37 | 1.860 | 16.49 | 2.746 | 15.00 | 0.988 | 12.66 | 0.158 | 1.84 |
| I don't give much thought to saving energy at home (inverse scale) | 0.708 | 10.87 | 1.182 | 16.44 | 2.203 | 25.18 | 1.910 | 21.34 | 0.132 | 1.80 |

**Table B2. Dependence between choice dimensions and choice alternatives**

| | | Purpose | | | Time | | | Mode substituted | | | Companion |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Errands | Recreation | Work | Morning peak | Afternoon peak | Evening/ night | Taxi | Active/ transit | No trip | Not alone |
| Purpose | Errands | 1.0 | | | | | | | | | |
| | Recreation | 0.5 | 1.0 | | | | | | | | |
| | Work | 0.5 | 0.5 | 1.0 | | | | | | | |
| Time | Morning peak | 0.0 | 0.0 | 0.0 | 1.0 | | | | | | |
| | Afternoon peak | 0.0 | 0.0 | 0.0 | 0.5 | 1.0 | | | | | |
| | Evening/ night | 0.0 | 0.0 | 0.0 | 0.5 | 0.5 | 1.0 | | | | |
| Mode substituted | Taxi | 0.0 | 0.0 | 0.0 | -0.181 (-3.14) | 0.0 | 0.0 | 1.0 | | | |
| | Active or transit | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.5 | 1.0 | | |
| | No trip | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.5 | 0.5 | 1.0 | |
| Companion | With others | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.147 (2.49) | 0.0 | 0.0 | 1.0 |

*Note*: Covariance parameters are represented in gray cells and correlation parameters in white cells. T-statistics are represented between parentheses when applicable.