A Joint Behavioral Choice Model of Carpool Formation and Frequency

Vivek Verma
The University of Texas at Austin, Dept of Civil, Architectural and Environmental Engineering
301 E. Dean Keeton St. Stop C1761, Austin TX 78712, USA
Email: vivek.verma@utexas.edu

Aupal Mondal
The University of Texas at Austin, Dept of Civil, Architectural and Environmental Engineering
301 E. Dean Keeton St. Stop C1761, Austin TX 78712, USA
Email: aupal.mondal@utexas.edu

Katherine E. Asmussen
The University of Texas at Austin, Dept of Civil, Architectural and Environmental Engineering
301 E. Dean Keeton St. Stop C1761, Austin TX 78712, USA
Email: kasmussen29@utexas.edu

Irfan Batur
Arizona State University, School of Sustainable Engineering and the Built Environment
660 S. College Avenue, Tempe, AZ 85287-3005
Email: ibatur@asu.edu

Tassio B. Magassy
Arizona State University, School of Sustainable Engineering and the Built Environment
660 S. College Avenue, Tempe, AZ 85287-3005
Email: tmagassy@asu.edu

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

Ram M. Pendyala
Arizona State University, School of Sustainable Engineering and the Built Environment
660 S. College Avenue, Tempe, AZ 85287-3005
Email: ram.pendyala@asu.edu

ABSTRACT
The future of transportation is often characterized by a vision of shared mobility in which multiple individuals ride in the same vehicle together. The most prevalent form of such shared personal mobility is carpooling. Despite decades of efforts to increase carpool mode shares, the share of carpooling for most travel, and especially work travel, has decreased. There is a need for a deeper understanding of the phenomenon and the drivers that influence carpool choice behavior. To this end, we use a novel data set and a holistic framework to focus on three critical dimensions: frequency of carpooling, choice of companion for carpooling, and the choice of platform or method for making the carpool arrangement. Results show that individuals do not embrace carpooling with strangers and do not use formal carpool programs to arrange their carpool arrangements. Model results show that a host of socio-demographics and built environment and workplace characteristics affect all three dimensions of carpool behavior. Insights from this study would help in identifying policies and technological platforms that would promote carpooling for disparate population subgroups.

Keywords: carpool formation, carpool accompaniment, carpool frequency, joint model of carpool choices, managed lanes.

1. INTRODUCTION

Vehicular travel demand and associated environmental externalities have been an endless issue for transportation planners, policymakers, and the general public. The transportation sector contributes the most (29 percent) to total greenhouse gas emissions in the United States (*1*). Growing population, urbanization, and vehicular travel demand, and increased attention to global climate change, have motivated communities to develop and foster sustainable travel options to reduce vehicle miles of travel (VMT). Among various travel demand management (TDM) strategies aimed at curbing vehicle use, "shared mobility" has emerged as an important option. Shared mobility modes and platforms aim to reduce 'drive alone' or single-occupancy vehicle travel by enabling multiple travelers to share a vehicle either concurrently (carpool or rideshare) or sequentially (carshare). Carpooling is a form of shared mobility in which two or more persons with similar spatial and temporal characteristics share a ride (*2*).

Carpooling is different from ridehailing (which is also referred to as ridesharing or ridesplitting, and offered by companies such as Didi, Uber and Lyft), where individuals pay for mobility as a service. In carpooling, an individual agrees to share a ride with other passengers, generally for free or by splitting the travel costs. In contrast to other TDM strategies that involve alternative mode use, carpooling allows individuals to travel more economically without significantly compromising their accessibility or mobility and requires no additional infrastructure investments. In addition to providing time and cost savings to carpoolers, carpooling offers significant benefits to the environment and society; carpooling helps to reduce vehicle miles of travel, decrease energy consumption and emissions, control traffic congestion, and limit parking infrastructure demand. Less frequently discussed are the social benefits of carpooling, which include increased accessibility to destinations for low-income people, reduced driving stress and increased social interaction. Furthermore, in many regions, carpoolers are able to take advantage of high occupancy vehicle (HOV) lanes (see for example, Burris, et. al. (*3*)).

However, since its heightened adoption in the 1970s due to the energy crisis, carpooling has been experiencing continuous decline in mode share. According to the U.S. Census Bureau (*4*), in a typical week, only 8.9 percent of workers carpool to work on most days, as compared to 76 percent who typically drive alone. The 2017 National Household Travel Survey (*5*) reported average vehicle occupancy for commute trips as 1.18. Well-known reasons stated by individuals for not choosing to carpool include reduced independence and flexibility, inability to find and coordinate with people with similar schedules and origin-destination patterns and lack of trust, safety, and certainty (*6*). COVID-19 has dealt the mode share of carpooling a further blow, as people prioritize health and safety, and commuting patterns have become more flexible and unpredictable (*7*).

Despite these headwinds being experienced by carpooling, many have been articulating a utopian vision of the future of transportation characterized by *sharing (pooling)*, besides connectivity, electrification, and automation. If the future is to truly realize a higher share of carpooling than today, a far deeper understanding of carpooling behaviors and phenomena is needed. The current understanding of carpooling may be considered deficient because all of the attempts to increase carpool mode share have largely been a dismal failure to date. Any inquiry into carpooling must explore two realms: *first*, to reveal why people are and are not carpooling today, and *second*, to explore what would make individuals carpool more in the future, especially with the proliferation of new mobility on demand services and app-based platforms/technology (that may not necessarily work in favor of boosting carpooling).

This paper focuses on jointly modeling the choices that comprise the phenomenon of carpooling. In particular, this study considers three dimensions of carpooling in an integrated joint modeling framework. The *first* is the decision to carpool or not. The second is the choice of *who* to carpool with, and the *third* is the manner in which the carpool is formed. Several papers have addressed these dimensions of interest, including the analysis of accessibility implications of carpooling, the impact of carpool app registration on the frequency of driving alone, and the examination of factors that affect carpooling frequency (see for example, Gheorghiu and Delhomme (*8*)). In all of these studies, the analysis sample has been limited to carpoolers, thus limiting the ability to identify carpooling deterrents for current non-carpoolers. In terms of the mechanism of finding carpool partners, earlier studies examined various ride-matching processes, from household-based carpool formation to co-worker carpools to slugging with strangers, as well as investigated differences between different types of carpool arrangements (e.g., Chan and Shaheen (*9*)). But most of these studies have been rather descriptive in nature and have not modeled carpool companion as an explicit choice. Finally, when it comes to carpool formation, there has been little investigation on methods for forming carpools and underlying preferences. Researchers have explored the potential impacts of, and user interactions with, various carpooling and travel apps (e.g., Shaheen et al. (*10*)); however, the use of such apps requires critical investment to ensure wide participation. The impact of carpool programs and clubs has been discussed in the literature as well (e.g., Correia and Viegas (*11*)). Once again, none of these studies actually consider the method of carpool formation as an explicit *choice* that needs to be modeled in a holistic behavioral framework.

This paper aims to make a substantial contribution to the literature by modeling the decisions of *whether* to carpool, *who* to carpool with, and *how* to form the carpool jointly within a comprehensive choice modeling framework. The data for the study is derived from a special purpose HOV and carpooling survey conducted in the Washington, D.C. and Charlotte, North Carolina regions of the United States. The survey provided rich information about carpooling behaviors suitable for a modeling effort of the nature undertaken in this paper.

The remainder of the paper is organized as follows. The next section presents a detailed description of the data. The third section offers a presentation of the modeling framework and methodology. The fourth section is dedicated to presenting model estimation results, while the fifth and final section offers a discussion of the study implications and concluding remarks.

2. DATA DESCRIPTION

The data for this study is derived from a High Occupancy Vehicle (HOV) and carpool survey, deployed in the areas around Washington, D.C. (mostly in Virginia) and Charlotte, North Carolina during March – May 2021. The survey was specifically aimed at collecting information about various aspects of carpooling behavior and HOV lane usage. The survey was distributed via e-mail to a random sample of residents in the two regions of interest; in addition, information about the survey was disseminated via media outlets and social channels with a view to boost response rates. A total of 2,735 responses were obtained; of this sample, 1,382 were workers who had a workplace away from home and commuted on a regular basis. The survey gathered detailed information about individual socio-economic and demographic attributes, geographic locations of residence and workplaces (at the zip code level), and commute travel. The geographic location information was used to merge a host of built environment (BE) variables to the data set using Smart Location Database of the U.S. Environmental Protection Agency (*12*). These BE variables include population density, with a trinary classification of low density (<5 people/acre), medium density

(5 to 10 people/acre, and high density (>10 people/acre); employment density, also with a three-way classification of low (<2 jobs/acre), medium (2 to 5 jobs/acre), and high (>5 jobs/acre); and auto ownership density classified as low (<1.5 households per acre with two or more vehicles) and high (≥1.5 households per acre with two or more vehicles).

Of particular interest in the context of this paper are workplace characteristics. About 60 percent of the respondents indicate that they have an express lane on commute route. Also, 79.3 percent indicate that they are not required to pay for parking at work, suggesting that a vast majority of respondents are afforded free employer-provided parking at the workplace.

The respondents were specifically asked to answer questions related to carpooling behavior in a pre-COVID era (i.e., carpool behavior in 2019) to control for any pandemic effects. In particular, the questions about carpool behavior included the number of days that respondents carpool to work in a typical month, the companion(s) they primarily choose to carpool with (choices included: family member, friend, co-worker, or a stranger), and the method/platform they primarily use to organize their carpool (choices included: website/application, phone calls/texts, in-person conversations, or company/school/neighborhood/agency carpool or rideshare program).

The data set was extensively cleaned and filtered to ensure that only valid and complete records were retained for analysis. After removing records that have missing data, particularly for questions related to carpool behavior, a total of 1,067 observations remained in the final analysis sample. Among the dependent variables, Figure 1 depicts the frequency distribution of carpool frequency for the sample of respondents. It is found that 315 respondents (29.5 percent) do not engage in carpooling, while the remainder engaged in at least one instance of carpooling in the past 30 days. An examination of the frequency distribution suggests that there is a good representation of individuals in all of the frequency bins, with the high frequency bins of 25-29 and 30+ instances of carpooling in the last 30 days showing relatively modest numbers. It should be noted that this sample is not representative of the population in terms of carpool frequency; this is a sample of workers specifically targeted to study carpool behavior and hence the nature of the sample recruitment process was such that carpoolers would naturally be over-represented (substantially) in the respondent sample. This was necessary to ensure that carpooling-related behavioral choices could be analyzed and modeled effectively.

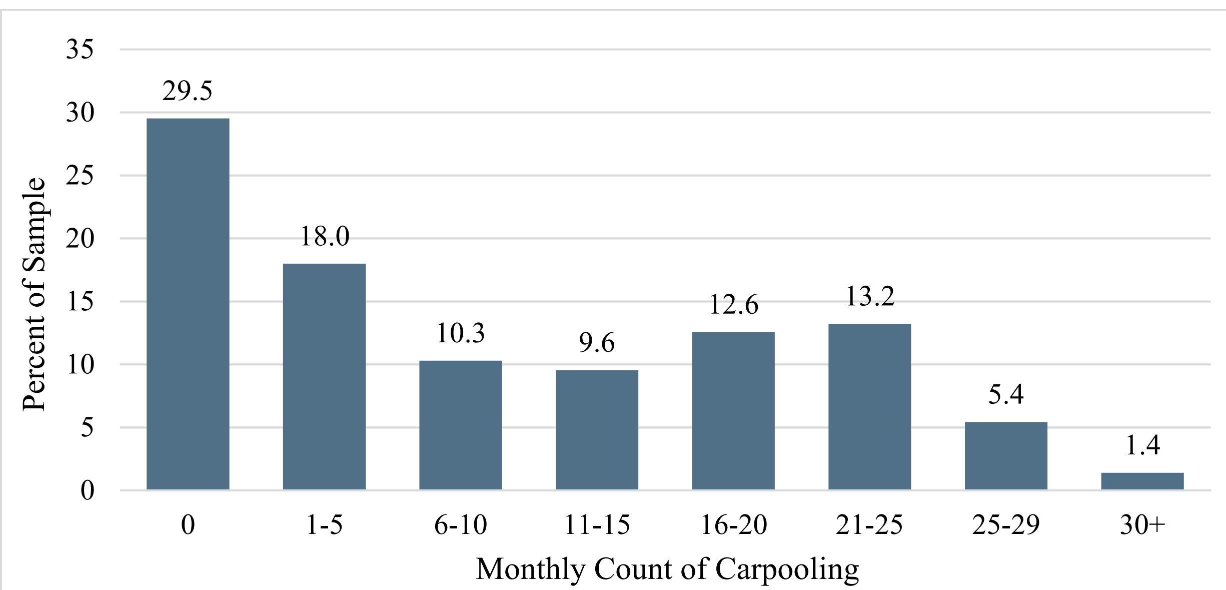

Figure 1 Distribution of frequency of carpooling

Table 1 shows a bivariate crosstabulation of carpool formation method versus carpool companion. The vast majority of carpooling is done with family and friends, although carpooling with co-workers is quite prevalent as well. Carpooling with strangers is quite modest, when compared with other categories. The percent of respondents using carpool programs (to form carpools) is quite low; carpools seem to be formed largely through personal communications (either telecommunications or in-person) and web/applications. The bivariate distribution, however, reveals an interesting finding in that carpooling with strangers is facilitated to a significant degree through the use of web/applications, much more so than other carpool companion arrangements. This is consistent with expectations as web/applications facilitate matching of strangers who share similar commute spatio-temporal characteristics. When it comes to family, friends, and co-workers, carpool formation tends to happen largely via personal communications (in-person or telecommunications). Overall, 90.8 percent of carpoolers do so with family, friends, and co-workers, suggesting that carpooling with strangers remains a choice not embraced by most. This generally does not bode well for a future of *shared* mobility, but such a discussion is beyond the scope of this paper.

TABLE 1 Bivariate Crosstabulation of Carpool Formation and Companion

Carpool Formation Method		Carpool Companion					Total	
		N/A	Family	Friends	Coworkers	Strangers	N	%
N/A		315	0	0	0	0	315	29.5
Website/App		0	69	36	40	54	199	18.7
Phone Call/Texts		0	75	119	70	11	275	25.8
In Person Conversations		0	107	61	48	3	219	20.5
Carpool Programs		0	20	7	31	1	59	5.5
Total	N	315	271	223	189	69	1067	100%
	%	29.5	25.4	20.9	17.7	6.5		

Note: N/A = not applicable

3. MODEL FRAMEWORK AND METHODOLOGY

As discussed earlier, the three dependent outcomes of interest in this study are: (1) the monthly count of carpooling to work (including the possibility of a zero) treated as a *count* variable, (2) if carpooling, the *nominal* outcome of who the individual primarily carpools with, (with the choice set comprised of family, friends, co-workers, or strangers; and (3) if carpooling, the *nominal* outcome of how the commuter typically forms their carpool, with the choice set comprised of website/applications, text/call, in-person conversation, or organized carpool programs. In this study, the three dependent outcomes are modeled jointly while explicitly considering the error correlations among the three dimensions. As with any data set, there are likely to be many unobserved factors (such as attitudes, perceptions, values, and preferences) that influence these choice dimensions. A joint model specification is capable of accounting for correlated unobserved factors that simultaneously affect multiple choice dimensions of interest. The mathematical formulation for the methodology used in this study is described below.

For the monthly carpooling frequency dimension, the outcome variable takes the form of a count model recast as a special case of a Generalized Ordered-Response Probit (GORP) that allows for more flexibility compared to the ordinary Poisson model. Following the GORP framework for count models (and dropping the subscript for individual q for ease of presentation), the count model expression for the monthly carpooling count dimension can be written as follows:

$$y^* = \varepsilon \;,\; y \;= k \;\text{ if } \psi_{k-1} < y^* < \psi_k, \tag{1}$$

$$\text{with } \psi_k = f_k(\boldsymbol{z}) = \Phi^{-1}\left(e^{-\lambda} \sum_{l=0}^{k} \frac{\lambda^l}{l!} \right) + \sum_{l=0}^{k} \alpha_l \text{, where } \lambda = e^{\gamma' z} \;,\; k \in \{0,\ 1,\ 2,\ ...\}. \tag{2}$$

In the above equation, y^* is a latent continuous stochastic propensity variable associated with the count that maps into the observed count k through the $\boldsymbol{\psi}$ vector (which is a vertically stacked column vector of thresholds $(\psi_{-1}, \psi_0, \psi_1,\ \psi_2, ...\psi_K)'$, where K is the max threshold level determined based on the empirical sample. ε is a standard normal random error term. $\boldsymbol{\gamma}$ is a column vector $[A \times 1]$ of coefficients corresponding to the vector $\boldsymbol{z}$. Φ^{-1} in the threshold function of Equation (2) is the inverse function of the univariate cumulative standard normal. The α_l terms $\left(\boldsymbol{\alpha} = \alpha_0, \alpha_1, \alpha_2, ..., \alpha_L\right)'$ in the thresholds are parameters to be estimated to accommodate high or low probability masses (spikes and dips) for specific count outcomes without the need for using zero-inflated or related mechanisms in multi-dimensional model systems (where L is the highest count for which α_l is specified).

Now, let there be G nominal outcome variables for an individual, and let g be the index for these variables $(g = 1, 2, 3, ..., G)$. For our analysis, G=2 (specific to the “who to carpool” and “how to carpool” dimensions). Also, let I_g be the number of alternatives corresponding to the g^{th} variable ($I_g \geq 3$) and let i_g be the corresponding index $(i_g = 1, 2, 3, ..., I_g)$. In our analysis, I_g =4 for all $g = 1,\ 2$ since all the variables have four alterntives each. Consider the g^{th} variable and assume the usual random utility structure for each alternative i_g:

$$U_{gi_g} = \boldsymbol{b}'_{gi_g} \boldsymbol{x} + \varsigma_{gi_g}, \tag{3}$$

where $\boldsymbol{x}$ is an $(A \times 1)$ vector of exogenous variable (including a constant), $\boldsymbol{b}_{gi_g}$ is an $(A \times 1)$ column vector of corresponding coefficients, and ς_{gi_g} is a normal error term. Let $\boldsymbol{\varsigma}_g = (\varsigma_{g1}, \varsigma_{g2}, ...\varsigma_{gI_g})'$ ($I_g \times 1$ vector), and $\boldsymbol{\varsigma}_g \sim MVN_{I_g}(\mathbf{0}, \boldsymbol{\Lambda}_g)$. Taking the difference with respect to the first alternative, the only estimable elements are found in the covariance matrix $\breve{\boldsymbol{\Lambda}}_g$ of the error differences, $\breve{\boldsymbol{\varsigma}}_g = (\breve{\varsigma}_{g2}, \breve{\varsigma}_{g3}, ..., \breve{\varsigma}_{gI_g})$ (where $\breve{\varsigma}_{gi} = \varsigma_{gi} - \varsigma_{g1}, i \neq 1$). Further, the variance term at the top left diagonal of $\breve{\boldsymbol{\Lambda}}_g$ $(g = 1, 2, ..., G)$ is set to 1 to account for scale invariance. However, beyond identification, for further simplicity and ease of estimation, we restrict all the diagonal elements of $\breve{\boldsymbol{\Lambda}}_g$ to be one, implying a correlation matrix. $\boldsymbol{\Lambda}_g$ is constructed from $\breve{\boldsymbol{\Lambda}}_g$ by adding a row on top and a column to the left. All elements of this additional row and column are filled with values of zero. In addition, the usual identification restriction is imposed such that one of the alternatives serves as the base when introducing alternative-specific constants and variables that do not vary across alternatives To proceed, define $\boldsymbol{U}_g = (U_{g1}, U_{g2}, ..., U_{gI_g})'$ ($I_g \times 1$ vector),

$\boldsymbol{b}_g = (\boldsymbol{b}_{g1}, \boldsymbol{b}_{g2}, \boldsymbol{b}_{g3}, ..., \boldsymbol{b}_{gI_g})'$ $(I_g \times A$ matrix), $\vec{G} = \sum_{g=1}^{G} I_g$, $\tilde{G} = \sum_{g=1}^{G} (I_g - 1)$, $\boldsymbol{U} = (\boldsymbol{U}_1', \boldsymbol{U}_2', ..., \boldsymbol{U}_G')'$ $(\vec{G} \times 1$ vector), $\boldsymbol{\varsigma} = (\varsigma_1, \varsigma_2, ..\varsigma_G)'$ $(\vec{G} \times 1$ vector), $\boldsymbol{b} = (\boldsymbol{b}_1', \boldsymbol{b}_2', ..., \boldsymbol{b}_G')'$ $(\vec{G} \times A$ matrix). Then, in matrix form, we may write Equation (1) as:

$$\boldsymbol{U} = \boldsymbol{bx} + \boldsymbol{\varsigma}, \qquad \text{where } \boldsymbol{\varsigma} \sim MVN_{\vec{G}}(\mathbf{0}_{\vec{G}}, \mathbf{\Lambda}). \tag{4}$$

$$\mathbf{\Lambda} = \begin{bmatrix} \mathbf{\Lambda}_1 & \mathbf{\Lambda}_{12} & \mathbf{\Lambda}_{13} & \mathbf{\Lambda}_{14} & \cdots \mathbf{\Lambda}_{1G} \\ \mathbf{0} & \mathbf{\Lambda}_2 & \mathbf{\Lambda}_{23} & \mathbf{\Lambda}_{24} & \cdots \mathbf{\Lambda}_{2G} \\ \mathbf{0} & \mathbf{0} & \mathbf{\Lambda}_3 & \mathbf{\Lambda}_{34} & \cdots \mathbf{\Lambda}_{3G} \\ \vdots & \vdots & \vdots & \vdots & \ldots \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \ldots \ldots \mathbf{\Lambda}_G \end{bmatrix} (\vec{\boldsymbol{G}} \times \vec{\boldsymbol{G}} \text{ matrix}).$$

The off-diagonal elements of the $\mathbf{\Lambda}$ matrix capture the correlations of the unobserved factors across the alternatives of the various nominal variables.

Now, consider the $[(\vec{G}+1) \times 1]$ vector $\tilde{\boldsymbol{y}} = \begin{pmatrix} \boldsymbol{U} \\ y^* \end{pmatrix}$ Next, define

$$\boldsymbol{B} = \begin{bmatrix} \boldsymbol{bx} \\ 0 \end{bmatrix} \text{ and } \mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Lambda} & \mathbf{\Omega} \\ \mathbf{\Omega}' & 1 \end{bmatrix}. \ (\vec{G}+1) \times (\vec{G}+1) \text{ matrix}. \tag{5}$$

The off-diagonal block matrix $\mathbf{\Omega}$ constitute the correlations between the multinomial dimensions and the count dimension. Then $\tilde{\boldsymbol{y}} \sim \boldsymbol{MVN}_{(\vec{G}+1)}(\boldsymbol{B}, \mathbf{\Sigma})$ is the multivariate joint distribution of the main outcomes of "how to carpool", "who to carpool" and carpooling frequency. Note that the "who to carpool" and "how to carpool" dimensions only come to effect for non-zero frequency of carpool; for zero carpooling frequency, our model collapses to a univariate count model.

For model estimation, define a matrix $\mathbf{M}$ of size $[\tilde{G}+1] \times [\vec{G}+1]$. Fill this matrix with values of zero. Then, in the first $(I_1 - 1)$ rows I_1 columns, insert an identity matrix of size $(I_1 - 1)$ after supplementing with a column of '-1' values in the column corresponding to the chosen alternative in the first nominal outcome. . Next, rows I_1 through $I_1 + I_2 - 2$ and columns $I_1 + 1$ through $I_1 + I_2$ correspond to the second nominal variable. Again position an identity matrix of size $(I_2 - 1)$ after supplementing with a column of '-1' values in the column corresponding to the chosen alternative for the second nominal variable. Continue this procedure for all G nominal variables (G=2 in our case). Then, at position $\tilde{G}+1$ and $\vec{G}+1$, enter a value one, corresponding to the count outcome. With the matrix $\mathbf{M}$ as defined, we can write $\ddot{\boldsymbol{y}} \sim \boldsymbol{MVN}_{\tilde{G}+1}(\tilde{\boldsymbol{B}}, \tilde{\mathbf{\Sigma}})$ where $\tilde{\boldsymbol{B}} = \mathbf{M}\boldsymbol{B}$ and $\tilde{\mathbf{\Sigma}} = \mathbf{M\Sigma M}'$.

Let $\boldsymbol{\delta}$ be the collection of parameters to be estimated: $\boldsymbol{\delta} = \left(\left[\text{Vech}(\boldsymbol{b})\right]', \boldsymbol{\gamma}, \boldsymbol{\alpha}, \left[\text{Vechup}(\mathbf{\Sigma})\right]' \right)'$, where the operator "Vech(.)" row-vectorizes all the non-zero elements of the matrix/vector on which it operates, and the operator Vechup(.) row-vectorizes the non-zero upper diagonal elements of a matrix.

Finally, define the thresholds, $\breve{\boldsymbol{\psi}}_{low} = \left[\left(-\boldsymbol{\infty}_{\tilde{G}} \right)', \psi_{k-1} \right]'$ $[(\tilde{G}+1)\times 1]$ vector and $\breve{\boldsymbol{\psi}}_{high} = \left[\left(\mathbf{0}_{\tilde{G}} \right)', \psi_{k} \right]'$ another $[(\tilde{G}+1)\times 1]$ vector.

Then the likelihood function may be written as:

$$L(\boldsymbol{\delta}) = \Pr\left[\breve{\boldsymbol{\psi}}_{low} \le \breve{\boldsymbol{y}} \le \breve{\boldsymbol{\psi}}_{high} \right], \tag{6}$$

$$= \int_{D_r} f_{(\tilde{G}+1)}(\boldsymbol{r} \mid \tilde{\boldsymbol{B}}, \tilde{\Sigma}) dr,$$

where the integration domain $D_r = \{ \boldsymbol{r} : \breve{\boldsymbol{\psi}}_{low} \le \boldsymbol{r} \le \breve{\boldsymbol{\psi}}_{high} \}$ is simply the multivariate region of the elements of the $\breve{\boldsymbol{y}}$ vector determined by the observed count and the the utility differences taken with respect to the utility of the chosen alternative for the multinomial outcomes. The likelihood function for a sample of Q decision-makers is obtained as the product of the individual-level likelihood functions. We use the analytical approximation based methods proposed by Bhat (*13*) for approximating this integral.

4. MODEL ESTIMATION RESULTS

This section presents the model estimation results in detail. Initial model specifications were informed by evidence in the literature, while the final model specification presented in the paper was determined based on a number of iterative trials and carefully examining the statistical significance and behavioral intuitiveness of estimated parameters. Table 2 presents estimation results for the joint model of carpool frequency, carpool companion, and carpool formation method. Note that some variables with marginal statistical significance are retained due to their intuitive meaning and important implications, and because the rather moderately sized sample may be artificially diminishing statistical significance of variables.

In joint limited dependent variable models, such as the one estimated in the current paper, only recursive effects of an endogenous observed variable on the underlying propensities of other variables are allowed, due to logical consistency considerations (see Bhat (*14*) for a detailed discussion). As part of the joint model system estimated in this paper, many different recursive endogenous effects across the dependent variables were explored. For example, tests were undertaken to examine the potential endogenous effects of “who to carpool with” on “how to carpool” (such as the decision to carpool with family leading to the use of personal communications, or the choice of carpooling with strangers leading to the use of web apps). However, all such tests showed that, after accounting for the unobserved correlations that influence the three decisions, no statistically significant (or even marginally significant) endogenous effects were found. This suggests that the three choice dimensions are contemporaneous in nature (for example, the decisions of carpooling with family and use of personal communications happen as one package decision synchronously).

4.1. Carpool Frequency

The carpool frequency estimation results show that younger individuals are more likely to embrace carpooling. This is entirely consistent with expectations and prior literature (e.g., Abrahamese and Keall (*15*)) and stems from several possible phenomena at play. Younger individuals may not have the same level of car ownership as older individuals, either by choice or due to income constraints.

Carpooling frequency is positively associated with higher levels of education, possibly because educational attainment affects pro-environmental attitudes, further engendering ridesharing behavior. Relative to students, workers are found to carpool more, presumably due to regular work schedules that facilitate carpooling with co-workers. Students generally use an array of modes of travel, including walk, bicycle, and transit. Vehicle ownership is found to have a statistically significant impact on carpool frequency, with individuals residing in households with more vehicles choosing to carpool less (as expected).

Household structure plays an important role in shaping carpool frequency choice. Single women are less likely to carpool than single men, presumably due to safety considerations (*16*). Couples and households with multiple adults are less likely to carpool, when compared with single men (the base category). Single men generally have greater flexibility and less household obligations that constrain their schedules; in general, travel behavior of people in families are more family-oriented and constrained with respect to degrees of freedom and adaptability (*17*). As expected, presence of children further diminishes carpool frequency due to increased need to coordinate child schedules and take care of household obligations. On the other hand, a greater number of workers in the household engenders a higher level of carpooling, consistent with the notion that the most prevalent type of carpooling is "fampooling" in which household members carpool with each other (*18*).

The findings clearly show the impact of transport policies and infrastructure. The provision of an HOV lane on the commute route has a positive influence on carpool frequency, a finding also reported by Bento et al. (*19*). Similarly, if an individual has to pay for parking at work, that contributes positively towards carpool frequency as individuals value cost-savings and fewer automobiles are driven to work when employees have to pay. These findings clearly point to the need to invest in infrastructure (HOV lanes) where it makes sense to do so (not all locations may warrant an HOV lane, especially if congestion levels are low), and to implement workplace (parking) policies that would incentivize carpooling (and disincentivize driving alone).

Built environment attributes are found to significantly affect carpool frequency, with individuals in dense areas exhibiting a higher frequency of carpooling, presumably because it is easier to find carpool companions in high density neighborhoods. On the other hand, residing in an area with high employment density tends to diminish carpooling, largely because dense environments are often characterized by the availability of alternative modes of transportation (*20*) and the need to carpool is diminished when commute distances are smaller. When auto ownership is high in the area, carpooling frequency tends to be higher due to greater availability of (and potential access to) vehicles.

4.2. Carpool Companion

The carpool companion model shows that older people tend to carpool with strangers; this is somewhat counter to expectations as one would expect young people to be more open to participating in the sharing economy and carpooling with strangers. However, there is some evidence that suggests older individuals have smaller social circles (*21*) and greater access to more formal rideshare programs through employers and other organizations. Those with a high degree of education are likely to be more time-pressured than others (*22*), thus rendering carpooling with friends and strangers more difficult. For such individuals, carpooling with family and co-workers is likely to be easier due to ease of schedule coordination. Cultural differences are likely to explain the higher propensity to carpool with family members among Whites (*23*).

An interesting finding is that higher income individuals are more likely to carpool with strangers (if they do carpool). It is likely that these individuals are time pressured (*22*) and they rely more on convenient third-party websites/apps for carpool arrangements (which naturally favor carpooling with strangers). When residing in households with multiple vehicles, multiple adults (couple), or children, the tendency is to engage in family-oriented carpooling. In such households, there are generally more opportunities to form "fampools" and there is a greater need to coordinate intra-household schedules and fulfilment of household and childcare obligations. As such, it is found that variables corresponding to high vehicle ownership, presence of a couple in the household, and presence of children in the household, all contribute positively to family-based carpooling as opposed to any other carpool companionship.

The presence of HOV lanes on the commute route appears to engender a higher level of carpooling with friends. Although it is admittedly difficult to explain this result readily, it is likely that the cost sharing/time savings (accrued through the use of HOV facilities) would motivate friends to band together and forge carpool arrangements. In areas with higher employment densities, people are more likely to choose to carpool with co-workers as compared to areas with lower densities. This is logically consistent with the notion that workers residing in such areas are more likely to find a co-worker as a carpool companion more readily. Although the alternative specific constants do not have a clear and natural interpretation, they do provide a sense of the general market tendency. It is found that the sample as a whole is least inclined to pool with strangers and more inclined to pool with co-workers and friends.

4.3. Carpool Formation Method

Finally, the table shows the estimation results for the carpool formation method component of the joint model system. The "in-person conversations" alternative is the overall base alternative for this specific model component (and hence there is no explicit column showing this alternative). It is found that higher educated individuals (graduate degree) tend to use websites/applications and more formal carpool programs to participate in carpools. People in this higher education group are generally more technology-savvy, live in urban environments, and open to adopting technology-based solutions for their transport needs (e.g., Alemi et al. (*24*)). At the same time, they are also likely to be experiencing greater time pressure (*22*), and hence they turn to apps and employer-based carpool programs that automatically coordinate and form carpools for them, rather than their having to personally coordinate via personal communications.

It is found that Blacks are less likely to use website/applications and formal carpool programs, presumably because Blacks trail Whites in internet and technology use, and this digital divide may be contributing to their lower level of use of such carpool formation tools; moreover, there is evidence of implicit bias which discourages Blacks from using carpool program platforms (*25*). Individuals with higher income and higher vehicle ownership are less likely to use carpool programs to organize their carpools, presumably of their diminished need to do so. Individuals with higher income are more likely to use websites or applications, presumably because high income individuals have more access to internet and are technology-savvy and are therefore able to use websites/applications easily. Individuals in households with high vehicle ownership do not need to rely on formal carpool programs because access to a vehicle is not a barrier to carpooling. Single women are found to use phone calls and texts to arrange carpools (if they carpool at all). As mentioned previously, single women may have heightened security and safety concerns that motivate them to use personal communications to vet a carpool companion (rather than relying on third party applications and programs to do this for them).

TABLE 2 Model Estimation Results

Exogenous variables (base shown in parenthesis)		Carpool Frequency (count)		Carpool Companion (MNP)								Carpool Method (MNP)					
				Family		Friends		Co-workers		Strangers		Website/ Application		Phone Calls/Texts		Carpool Programs	
		Coef.	t-stat	Coef.	t-stat	Coef.	t-stat	Coef.	t-stat	Coef.	t-stat	Coef.	t-stat	Coef.	t-stat	Coef.	t-stat
Individual level characteristics																	
Age (<30)	30 or older	-0.01	-2.82							0.53	2.78						
Education (bachelor's or less)	Grad. degree(s)	0.01	4.12			-0.52	-3.84			-0.49	-2.56	0.32	1.81			0.73	3.01
Race (*)	White			0.28	2.49												
	Black											-0.32	-2.08			-0.38	-1.27
Occupation (student only, non-worker)	Worker	0.01	1.12														
	Student & worker	0.02	2.93														
Household characteristics																	
Income (<$100,000)	≥$100,000									0.24	1.66	0.26	1.92			-0.20	-0.91
Number of vehicles (one)	Two or more	-0.05	-7.08	0.46	4.30											-0.37	-1.28
Household structure (*)	Single woman	-0.06	-8.43											0.26	2.58		
	Couple only	-0.02	-3.25	0.32	3.04												
	Multi-adult	-0.01	-1.05														
Number of kids (none)	One or more	-0.01	-1.75	0.15	1.55												
Number of workers (one)	Two or more	0.03	9.12														
Workplace characteristics																	
HOV lane on commute (none)	Present	0.01	1.90			0.63	4.26										
Parking at work (free)	Paid	0.01	1.41														
Built environment characteristics																	
Population density (*)	Medium	0.01	1.82														
	High	0.03	6.14									-0.34	-2.08				
Employment density (low)	Medium/high	-0.02	-6.33					0.12	1.22			0.23	1.69	0.14	1.28	0.28	1.74
Auto density (low)	High	0.02	3.80														
Constant		6.53	37.46			0.21	0.83	0.32	1.31	-0.44	-0.97	-0.32	-1.26	-0.07	-0.40	-0.93	-1.73
Threshold at zero		35.64	46.15														

*Base category is all other omitted categories

People residing in areas of lower population density are more likely to use websites or applications for carpooling. This is because it is generally difficult to find someone in close proximity to readily organize a carpool. On the other hand, people residing in areas of higher employment density are more likely to use websites/applications, calls and texts, and carpool programs, all of which are quicker and more convenient methods to organize carpools among co-workers (than in-person conversations). Recall that individuals in higher employment density areas are more likely to carpool with co-workers, if they choose to carpool at all.

Finally, although the alternative specific constants do not have a clear and natural interpretation, they do provide a sense of the general market tendency. It is found that the sample as a whole is least inclined to use carpool programs and most inclined to form carpools through personal communications.

4.4. Error Correlations and Goodness-of-Fit Measures

As noted earlier, the jointness in the decision-making process related to carpool behavior is captured through the explicit accounting for error covariances across the choice dimensions included in the model system. Table 3 presents the error correlation matrix together with goodness-of-fit statistics for the model. The values of the log-likelihood together with the likelihood ratio test reveal that the joint model (which incorporates error correlations) provides a statistically superior fit than the independent model which would ignore error correlations and essentially treat the three dimensions as separate and independent carpool choice dimensions. This is a clear indication that a joint model that brings these choice dimensions together into a "choice bundle" is preferred when investigating carpool choice behavior).

TABLE 3 Correlation Matrix and Goodness-of-fit Statistics

Attribute dimensions	Friend	Co-worker	Stranger	Website/ App	Call/text	Carpool Program	Carpooling Count
	(base: family)			*(base: in-person conversations)*			
Friend	1.000	0.510*	0.517*	0.183*	0.384*	0.092	0.018
Co-worker	-	1.000	0.431*	0.062	0.185	0.351*	0.051*
Stranger	-	-	1.000	0.594*	0.120	0.021	-0.024
Website/App	-	-	-	1.000	0.347	0.244	-0.060
Call/text	-	-	-	-	1.000	0.398	-0.018
Carpool Program	-	-	-	-	-	1.000	-0.037
Carpooling Count	-	-	-	-	-	-	1.000

Model-Fit Statistics

Loglikelihood of proposed model = -4948.56; Loglikelihood of independent model = -5012.70
Loglikelihood of constants only model = -9609.94
Likelihood ratio test = -2(4948.56-(-5012.70))=128.28
Critical $\chi^2_{21,0.001}$ = 46.797

*Note: Same base applies on the dimensions given in rows; *Statistically significant at 85% confidence level.*

Several error correlations are statistically significant, suggesting that there are correlated unobserved factors that simultaneously affect multiple carpool choice dimensions. There are significant positive correlations between the choice alternatives of friends, co-workers and strangers. This means that there are unobserved factors which enhance the propensity of an individual to carpool with all of these potential companions *simultaneously*. Household (family-based) carpools are known to be different than other carpools that involve non-household members (*18*, *26*), and hence this finding of significant error correlations across the non-household member alternatives is intuitive. There is a positive and significant correlation between the "Friends" and "Calls/Text" dimensions, which means that there are unobserved factors which simultaneously increase or decrease the propensity of both dimensions. Factors that contribute to increased carpooling with friends also contribute to elevated use of calls and text messages in organizing the carpool. On the other hand, there is a positive significant correlation between "Stranger" and "Website/App", suggesting that factors that contribute to carpooling with strangers also contribute to using website/app platforms to do so (e.g., being more technology-savvy, more open to new adventures or arrangements). For co-worker category, however, the significant positive correlations are with "carpool program" and "carpool count". This is entirely consistent with expectations in that carpooling with co-workers is generally facilitated through employer-based carpool programs and associated with a higher frequency of carpooling (due to regularity in schedules among matched co-workers).

5. DISCUSSION AND CONCLUSIONS

This study has presented a joint model of carpool choice behavior with a focus on several key dimensions that comprise carpooling behavior. The proposed joint model includes a count model of monthly carpool frequency, and multinomial probit model components of carpool companion (family, friends, co-workers, strangers) and carpool formation method (website/app, phone calls/texts, in-person conversations, carpool programs). The model system is estimated on a data set derived from a survey of commuters in the Washington, D.C. (Virginia) and Charlotte, North Carolina areas, and includes 1,067 commuters – 30 percent of whom do not carpool at all. A host of explanatory variables, including socio-economic, demographic, workplace, and built environment attributes are included in the model specification.

The findings have important implications for the future of carpooling. The history of carpooling has not been particularly encouraging with the mode share for carpooling seeing a steady decline over the past several decades. Carpooling is often viewed as inconvenient, reducing degrees of freedom and flexibility, and adding a time penalty due to the need to coordinate, wait, and go out of the way to drop off and pick up passengers. Carpooling programs have largely met with little success; despite this history, many envision a future of *shared mobility* that will lead to a more sustainable transport system in the years ahead. Given the evidence to date, it is difficult to see how such a future may be realized in practice.

The data set used in this study shows that people are least likely to carpool with strangers; individuals are most likely to carpool with family, friends, and co-workers (in that order) with strangers coming a very distant fourth in order. With respect to carpool formation method, the most popular methods involve personal communications (whether telecommunications or in-person conversations) with the use of newer website/apps coming a reasonably close third. However, the use of carpool programs (typically implemented by employers or transport agencies) come a very distant fourth in usage. In other words, this study suggests that there are considerable barriers to carpooling with strangers and to the use of formal carpool programs – findings that do not bode well for a future of *shared mobility*. However, model results show that websites and mobile applications that facilitate automated matching of ride(r)s for carpooling purposes hold considerable promise; both public and private sector investments in such technology-based platforms could potentially yield positive results in boosting carpooling. However, the development and deployment of such apps must be sensitive to the notion of inclusiveness and overcoming any digital divide, as model estimation results in this study showed that Blacks are less likely to use such platforms for facilitating carpool arrangements.

Policy instruments do exhibit the potential to make a difference when it comes to carpooling. Model estimation results show that the presence of HOV lanes does contribute to a higher frequency of ridesharing and a greater potential for carpooling with friends. Despite some of the concerns regarding the potential effectiveness of HOV lanes, this study suggests that the investments in (and designation of) such lanes would provide significant positive benefits, particularly in congested metropolitan contexts. Similarly, parking policies at the workplace (parking pricing in particular) would bring about shifts in carpool mode share. Similarly, while carpool programs may not work in all contexts, they appear to show promise in the context of engendering carpool arrangements between co-workers. As such, workplace-based carpool programs are of value, if implemented effectively (perhaps coupled with strong incentives) and communicated broadly to raise awareness about benefits. Agencies, on the other hand, may be better served by deploying website/applications to enable carpooling (particularly between strangers).

Efforts to promote carpooling must recognize the socio-demographic aspects that play a critical role. Single women will need systems that ensure trustworthy, safe, and secure transport to adopt carpooling in large numbers. Families will struggle to embrace carpooling in the wake of having to fulfill myriad household obligations, childcare responsibilities, and other chores through the course of a day. They treasure and value freedom and flexibility, particularly in the context of leading lives that are already constrained by work schedules, school schedules, after-school extra-curricular activity schedules, and business schedules (store, restaurant, and establishment opening and closing hours). Adding modal constraints (in the form of carpool arrangements that need coordination and matching) removes degrees of freedom, not to mention an element of privacy, that families are loathed to sacrifice. Public agencies and private employers need to band together to come up with a slate of services that serve households and their varied needs, from the standpoint of chauffeuring children, picking up goods and services, and providing flexibility in times of need.

Transportation models need to be enhanced to incorporate the multitude of dimensions that characterize, and the multitude of considerations that affect, carpooling behavior. Current transportation demand forecasting models simply include carpool as a single modal option in mode choice models, with little consideration for other choice dimensions that comprise carpooling (the questions of: with who and how). Transportation models do not account for the plethora of technology platforms and tools that now enable carpooling and ride-matching in real time. The future of mobility application platforms is likely to be one that is increasingly gamified, with incentives, prizes, and frequent-user reward points to bring about behavioral change. Once again, transportation demand models are ill-equipped to incorporate the effects of such platforms and their features. Unless there is a gamification of such platforms, it is unlikely that there will be any amplification of desirable sustainable modal choice behaviors (based on historical evidence to date), particularly in a post-COVID era. Future surveys should incorporate a battery of attitudinal questions to captures attitudes, values, perceptions, and preferences; by explicitly incorporating and considering such variables, it will be possible to identify (typically unobserved) barriers to carpooling and tailor future programs and platforms accordingly.

ACKNOWLEDGEMENT

This research was partially supported by the Center for Teaching Old Models New Tricks (TOMNET) Tier 1 University Transportation Centers sponsored by the U.S. Department of

Transportation under grant 69A3551747116. The authors would also like to acknowledge Cintra support, which helped the survey efforts undertaken as part of this study.

AUTHOR CONTRIBUTIONS

The authors confirm contribution to the paper as follows: study conception and design: V. Verma, A. Mondal, K.E. Asmussen, C.R. Bhat; data collection: V. Verma, A. Mondal, K.E. Asmussen, C.R. Bhat; analysis and interpretation of results: V. Verma, A. Mondal, K.E. Asmussen, C.R. Bhat, I. Batur; draft manuscript preparation: V. Verma, A. Mondal, K.E. Asmussen, R.M. Pendyala, C.R. Bhat, T.B. Magassy, I. Batur. All authors reviewed the results and approved the final version of the manuscript.

REFERENCES

1. United States Environmental Protection Agency. Global Greenhouse Gas Emissions Data, 2021. https://www.epa.gov/transportation-air-pollution-and-climate-change/carbon-pollution-transportation
2. Shaheen, S., and A. Cohen. Shared ride services in North America: Definitions, impacts, and the future of pooling. *Transport Reviews*, 2019. 39(4):427–442.
3. Burris, M., N. Alemazkoor, R. Benz, and N. S. Wood. The impact of HOT lanes on carpools. *Research in Transportation Economics*, 2014. 44:43–51.
4. Burrows, M., C. Burd, and B. McKenzie. Commuting by Public Transportation in the United States: 2019. *American Community Survey Reports*, ACS-48, U.S. Census Bureau, 2021.
5. National Household Travel Survey. U.S. Department of Transportation, 2017.
6. Park, Y., N. Chen, and G. Akar. Who is interested in carpooling and why: The importance of individual characteristics, role preferences and carpool markets. *Transportation Research Record: Journal of the Transportation Research Board*, 2018. 2672(8):708–718.
7. De Vos, J. The effect of COVID-19 and subsequent social distancing on travel behavior. *Transportation Research Interdisciplinary Perspectives*, 2020. 5:100121.
8. Gheorghiu, A., and P. Delhomme. For which types of trips do French drivers carpool? Motivations underlying carpooling for different types of trips. *Transportation Research Part A*, 2018. 113:460–475.
9. Chan, N. D., and S. A. Shaheen. Ridesharing in North America: Past, present, and future. *Transport Reviews*, 2012. 32(1):93–112.
10. Shaheen, S., A. Stocker, and M. Mundler. Online and app-based carpooling in France: Analyzing users and practices—A study of BlaBlaCar. In *Disrupting Mobility* (G. Meyer, S. Shaheen, eds), Springer, Cham, 2017, pp. 181–196.
11. Correia, G., and J. M. Viegas. Carpooling and carpool clubs: Clarifying concepts and assessing value enhancement possibilities through a Stated Preference web survey in Lisbon, Portugal. *Transportation Research Part A*, 2011. 45(2):81–90.
12. Ramsey, K., and A. Bell. Smart Location Database, Version 2.0 User Guide. 2014. Available at: https://www.epa.gov/sites/production/files/2014-03/documents/sld_userguide.pdf.

13. Bhat, C. R. New matrix-based methods for the analytic evaluation of the multivariate cumulative normal distribution function. *Transportation Research Part B*, 2018. 109:238–256.

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

15. Abrahamese, W., and M. Keall. Effectiveness of a web-based intervention to encourage carpooling to work: A case study of Wellington, New Zealand. *Transport Policy*, 2012. 21:45–51.

16. Meshram, A., P. Choudhary, and N. R. Velaga. Assessing and modelling perceived safety and comfort of women during ridesharing. *Transportation Research Procedia*, 2020. 48:2852–2869.

17. Javid, M. A., T. Mehmood, H. M. Asif, A. U. Vaince, and M. Raza. Travelers' attitudes toward carpooling in Lahore: Motives and constraints. *Journal of Modern Transportation*, 2017. 25(4):268–278.

18. Matsuo, M. Carpooling and drivers without household vehicles: Gender disparity in automobility among Hispanics and non-Hispanics in the US. *Transportation*, 2020. 47(4):1631–1663.

19. Bento, A. M., J. E. Hughes, and D. Kaffine. Carpooling and driver responses to fuel price changes: Evidence from traffic flows in Los Angeles. *Journal of Urban Economics*, 2013. 77:41–56.

20. Blumenberg, E., and M. Smart. Brother can you spare a ride? Carpooling in immigrant neighbourhoods. *Urban Studies*, 2014. 51(9):1871–1890.

21. Tang, F., and Y. Lee. Social support networks and expectations for aging in place and moving. *Research on Aging*, 2011. 33(4):444–464.

22. Batur, I., S. Sharda, T. Kim, S. Khoeini, R. M. Pendyala, and C. R. Bhat. Mobility, time poverty, and well-being: How are they connected and how much does mobility matter? Technical paper, Arizona State University, 2019.

23. Vasquez-Tokos, J., and P. Yamin. The racialization of privacy: racial formation as a family affair. *Theory and Society*, 2021. 50(5):717–740.

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

25. Uhlmann, E. L., V. L. Brescoll, and E. Machery. The motives underlying stereotype-based discrimination against members of stigmatized groups. *Social Justice Research*, 2010. 23(1):1–16.

26. Blumenberg, E., and M. Smart. Getting by with a little help from my friends… and family: immigrants and carpooling. *Transportation*, 2010. 37(3):429–446.