NATIONAL CENTER FOR UNDERSTANDING FUTURE
TRAVEL BEHAVIOR AND DEMAND

Final Project Report

How Effective are Marker Variables at Predicting Attitudinal Factor Scores? An Out-of-sample Evaluation

BY

Jason Soria
Email: jsoria@cintra.us

Patricia L. Mokhtarian
Email: patmokh@gatech.edu

School of Civil & Environmental Engineering
Georgia Institute of Technology
790 Atlantic Drive
Atlanta, GA 30331

September 2024

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No. N/A	**2. Government Accession No.** N/A	**3. Recipient's Catalog No.** N/A	
4. Title and Subtitle How Effective are Marker Variables at Predicting Attitudinal Factor Scores? An Out-of-sample Evaluation		**5. Report Date** September 1, 2024	
		6. Performing Organization Code N/A	
7. Author(s) Jason Soria, Ph.D., https://orcid.org/0000-0003-1319-9197 Patricia L. Mokhtarian, Ph.D., https://orcid.org/0000-0001-7104-499X		**8. Performing Organization Report No.** N/A	
9. Performing Organization Name and Address School of Civil and Environmental Engineering Georgia Institute of Technology 790 Atlantic Drive, Atlanta, GA 30332		**10. Work Unit No. (TRAIS)** N/A	
		11. Contract or Grant No. 69A3552344815 and 69A3552348320	
12. Sponsoring Agency Name and Address U.S. Department of Transportation, University Transportation Centers Program, 1200 New Jersey Ave, SE, Washington, DC 20590		**13. Type of Report and Period Covered** Final Report, 2023-2024	
		14. Sponsoring Agency Code USDOT OST-R	
15. Supplementary Notes N/A			
16. Abstract This report is one in a series of studies designed to investigate the practicality of including attitudes as explanatory variables in practice-oriented travel demand forecasting models. In this study, we respectively applied random forest (RF) and elastic net regression (ENR) to 15 marker variables (MVs), to predict factor scores on four attitudes: pro-car ownership, pro-non-car alternatives, pro-suburban, and urbanite. We incorporated those four imputed attitudes into multinomial logit vehicle ownership (VO) models and compared the results to those of models including the original four factor scores, models containing only the four MVs most strongly associated with the same four attitudes, and models containing no attitudes. We created 1,000 random splits of a sample of 3,178 responses to a 2017 survey of Georgia adults, training the RF and ENR functions on each donor half-sample and applying those functions to the recipient half-sample. We reported results averaged over the 1,000 recipient half-samples. In the VO models, estimated coefficients for all sets of attitudes (original, ENR-imputed, RF-imputed, and MVs only) were by far most often both statistically significant and with the expected sign. Perhaps most importantly, the predictive power of the models markedly improved specifically for zero-car households whenever the attitudes were included. Using only the marker variables themselves gave results nearly as good as those associated with the more elaborate prediction of factor scores using machine learning methods.			
17. Key Words travel behavior surveys, attitudes, factor analysis, random forest, elastic net regression, vehicle ownership		**18. Distribution Statement** No restrictions.	
19. Security Classif.(of this report) Unclassified	**20. Security Classif.(of this page)** Unclassified	**21. No. of Pages** 37	**22. Price** N/A

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The report is funded, partially or entirely, under Grant No. 69A3552344815 and 69A3552348320 from the U.S. Department of Transportation's University Transportation Centers Program. The U.S. Government assumes no liability for the contents or use thereof.

ACKNOWLEDGMENTS

This research was partially supported by the National Center for Understanding Future Travel Behavior and Demand (TBD), a National University Transportation Center sponsored by the U.S. Department of Transportation (USDOT) under grant numbers 69A3552344815 and 69A3552348320. Collection of the data on which this study is based was funded by the Georgia Department of Transportation (DOT) (RP 16-31). Subsequent work was funded by the Center for Teaching Old Models New Tricks (TOMNET), a Tier l University Transportation Center sponsored by the U.S. Department of Transportation (USDOT) under grant number 69A3551747116. The authors would like to thank the TBD National Center, TOMNET, USDOT, and GDOT for their support of university-based research in transportation, particularly for the funding provided for this project.

TABLE OF CONTENTS

EXECUTIVE SUMMARY 1
1. INTRODUCTION 2
2. LITERATURE REVIEW 3
2.1 Attitudinal marker variables 3
2.2 Vehicle ownership modeling 4
3. ANALYSIS METHODS 5
3.1 Transfer process with random forest and elastic net regression 5
3.2 Data 10
3.3 Exploratory factor analysis and marker variable selection 12
4. RESULTS 13
4.1 Internal evaluation 14
4.2 Adjacent-sample external evaluation 14
5. DISCUSSION AND CONCLUSION 20
6. REFERENCES 22
7. APPENDIX 26

LIST OF TABLES

Table 1 Descriptive statistics (N = 3,178) 11

Table 2 Exploratory factor analysis results 12

Appendix Table 1 Vehicle ownership multinomial logit model with all observations and observed attitude factor scores (N = 3,178) 26

Appendix Table 2 Average estimated coefficients and p-values of multinomial logit model results for model with no attitudes (across 1,000 random splits, holding model specification constant) 27

Appendix Table 3 Average estimated coefficients and p-values of multinomial logit results with standardized attitude marker statements as explanatory variables (across 1,000 random splits, holding model specification constant) 28

Appendix Table 4 Average estimated coefficients and p-values of multinomial logit results with elastic net regression imputed attitude factor scores as explanatory variables (across 1,000 random splits, holding model specification constant) 29

Appendix Table 5 Average estimated coefficients and p-values of multinomial logit results with random forest imputed attitude factor scores as explanatory variables (across 1,000 random splits, holding model specification constant) 30

Appendix Table 6 Average estimated coefficients and p-values of multinomial logit results with observed attitude factor scores as explanatory variables (across 1,000 random splits, holding model specification constant) 31

LIST OF FIGURES

Figure 1 Transfer process, learning functions, and evaluation 7

Figure 2 Average (over 1,000 random splits) correlations between observed and imputed attitude scores 15

Figure 3 Average (over 1,000 random splits) root mean squared errors between observed and imputed attitude scores 15

Figure 4 Percent of 1,000 attitude coefficient results that are expected or unexpected and statistically significant or insignificant (at $\alpha = 0.05$) 18

Figure 5 Average (over 1,000 random splits) McFadden's $\rho 2$ (note: vertical axis does not begin at 0) 19

Figure 6 Average (over 1,000 random splits) success index for each alternative and different inclusions of attitudes 19

Appendix Figure 1 Share of non-attitude coefficient results that are expected or unexpected and statistically significant or insignificant 32

EXECUTIVE SUMMARY

This report is one in a series of studies designed to investigate the practicality of including attitudes as explanatory variables in practice-oriented travel demand forecasting models. In this study, we respectively applied random forest (RF) and elastic net regression (ENR) to 15 marker variables (MVs), to predict factor scores on four attitudes: pro-car ownership, pro-non-car alternatives, pro-suburban, and urbanite. We incorporated those four imputed attitudes into multinomial logit vehicle ownership (VO) models and compared the results to those of models including the original four factor scores, models containing only the four MVs most strongly associated with the same four attitudes, and models containing no attitudes. We created 1,000 random splits of a sample of 3,178 responses to a 2017 survey of Georgia adults, training the RF and ENR functions on each donor half-sample and applying those functions to the recipient half-sample. We reported results averaged over the 1,000 recipient half-samples. In the VO models, estimated coefficients for all sets of attitudes (original, ENR-imputed, RF-imputed, and MVs only) were by far most often both statistically significant and with the expected sign. Perhaps most importantly, the predictive power of the models markedly improved specifically for zero-car households whenever the attitudes were included. Using only the marker variables themselves gave results nearly as good as those associated with the more elaborate prediction of factor scores using machine learning methods.

1. INTRODUCTION

The explanation and prediction of travel demand is central to transportation planning, policy, and research. A variety of methods are available for estimating travel demand: at one extreme are simple, parsimonious models requiring little data, statistical knowledge, and computational power, while at the other extreme are complex, advanced, data-hungry models requiring substantial computational resources. Analysts must make tradeoffs to find the situationally-optimal balance between these two extremes, evaluating data availability, computational capabilities, and budgets. Particularly in the context of real-world travel demand forecasting, models are often simplified by limiting the number and kinds of variables that they include. Such a decision has the benefits of reducing the data collection burden and involving more straightforward models, at the cost of loss of realism and lack of flexibility for dealing with a rapidly-changing landscape.

Attitudes comprise one type of variable that is commonly omitted from practice-ready travel demand models (Ewing et al. 2019). Conventional wisdom holds that attitudes are difficult to measure, and perhaps impossible to forecast, rendering them unsuitable for inclusion. And yet, attitudes have long been a staple in travel behavior *research*, invariably adding considerable explanatory power to models including them (Bhagat-Conway et al. 2022). The present paper is the latest link in a chain of studies designed to address this dilemma (Malokin et al. 2019; Shaw 2021; Soria and Mokhtarian 2024; Mokhtarian 2024). Specifically, this research arc is exploring ways to make the inclusion of attitudinal variables into travel demand forecasting models more practical – focusing first on the measurement and inclusion challenge, while reserving the forecasting challenge for future studies. Earlier stages established the value of using machine learning methods to impute ("transfer") missing attitudinal variables into the household travel survey datasets that supply the data used to estimate regional travel demand models. In preliminary investigation Shaw (2021) obtained very promising results by using just a few attitudinal statements, called "marker variables" (MVs), to represent the information contained in a far larger set of statements.

The present paper continues that line of inquiry. Specifically, the goals of this paper are threefold: first, to see how well we can predict attitude factor scores (which were created from "numerous" attitudinal statements) using only a small number of MVs (the "internal evaluation"); second, to see how well the predicted factor scores perform in a model of household vehicle ownership; and third, to compare the performance of predicted factor scores to that of simply using the individual MVs as explanatory variables (the latter two goals constituting the "external evaluation"). To achieve these goals, we iterate through the following steps. A random half of the data (comprising the "donor sample") is used to train a "learning function" using MVs. This function is applied to the other half of the data (comprising the "recipient sample") to predict attitude factor scores, and we internally evaluate how well it predicts attitude factor scores for both halves. Household vehicle ownership (VO) models are estimated using the test sample, and we externally evaluate the models by measuring their predictive accuracy. These steps are iterated on multiple different random splits, to obtain a bootstrap distribution of attitude factor scores and VO model prediction accuracy measures. If we are able to (1) predict attitudes well from a small number of marker statements, and (2) obtain useful explanatory power from the predicted attitudes or even (3) from the marker statements themselves, it will provide another step forward in the path toward incorporating attitudinal variables into practice-ready travel demand forecasting models.

The rest of this paper is organized as follows. The following section briefly reviews the literature on attitudinal MVs, as well as providing an overview of VO models. Next, we present the data obtained from a 2017 statewide travel behavior survey, provide descriptive statistics for variables used in the VO model, and describe the exploratory factor analysis (EFA) results that produced the attitudinal factor scores used in the data imputation process. The methodology section discusses the MVs particular to this line of research, attitude imputation via elastic net regression (ENR) and random forest (RF) methods, and the multinomial logit (MNL) model applied to VO. Next, we present the results from the bootstrapped MNL VO models, including the VO prediction accuracy associated with different methods for predicting attitudes, with using the MVs themselves as explanatory variables, and with no attitudes in the model. We then conclude by discussing the implications, for practitioners and researchers alike, of *attitude* prediction accuracy and VO model improvements with the inclusion of attitudes.

2. LITERATURE REVIEW

2.1 Attitudinal marker variables

Travel behavior research incorporates attitudes in many ways. Many researchers use them directly as explanatory variables in behavioral models (Domarchi et al. 2008; Popuri et al. 2011; Wang et al. 2023). In other instances, they are used to identify/explain heterogeneity in the impacts of *other* explanatory variables on behavior, for example via a latent class model (e.g. Kim and Mokhtarian 2018). In both types of applications, attitudes are typically quantified via EFA; the attitudinal factor scores are calculated independently from the behavioral model. Other studies specify attitudinal constructs simultaneously with their behavioral models, for example using hybrid choice models (HCMs), also known as integrated choice and latent variable (ICLV) models (Soria et al. 2023; Tarabay and Abou-Zeid 2019). Bahamonde-Birke and Ortúzar (2021) explore methods of including both latent classes and latent variables by classifying individuals based on their estimated latent variables.

The methods in these research-oriented models, however, have typically not been adopted for practice-oriented travel demand forecasting models, as they involve either measuring a large number of attitudinal statements (which may fatigue survey respondents) or experimenting with complex models (which may require too much time and expertise to specify appropriately). Therefore, we are exploring ways to include attitudes in behavioral models that do not require substantially more data than conventional models do, and that involve approachable models.

The use of marker variables to create and validate a "short-form scale" is well-established in the psychology and health fields. A prominent example is using shortened versions of the Big Five personality inventory (Rammstedt and John 2007). Richins (2004) explores measuring the material values scale by reducing the number of items from the original 15 to 9, 6, and three with varying results. Hendrick et al. (1998) tests love attitude subscales by reducing each 7-item subscale down to 4-item and 3-item versions, finding that these short-form subscales had strong psychometric properties. In a physical and mental health survey, Ware et al. (1996) takes an already short-form 36-item survey and reduces it to 12 items. Using the scores from the 12-item factor analysis to predict the scores in the 36-item version, the authors achieve R^2 values greater than 0.90 for all factors.

Two methods to select items are seen in the literature. The first method is to use regression to identify items that predict factor or component scores most accurately. Ware et al. (1996) uses forward-stepwise regression to predict the original scores until 12 are chosen from the original 36 items. Slade (1997) takes a similar approach and reduces a 49-item questionnaire about oral health to 14 items by utilizing forward-stepwise regression. The author achieves an R^2 of 0.96 when using the 14 items to predict the single component score measuring oral ill-health, where even the first 7 items produced an R^2 of 0.92. The second method is to use items having the highest pattern loadings on the respective factors of interest. Kaufman et al. (2016) selects items for their short-form scale based on this criterion and finds that the long- and short-form versions have correlations of scores between 0.90 and 0.97. Tambs and Moum (1993) use this approach and then test the items' predictive power using forward-stepwise regression. They achieve an R^2 of 0.90 after 7 steps (the seventh item chosen from a pool of 25 candidates).

Within the field of transportation, there are fewer but pertinent examples. Cain et al. (2017) took a 120-item survey about pedestrian landscapes and created a 54-item version to measure physical activity. Comparing the measures obtained from the full and abbreviated surveys, a strong correlation of 0.94 was found. Cerin et al. (2009) developed an abbreviated version of the Neighborhood Walkability Scale using the correlation between the original scale and the Walk Score index (Hall and Ram 2018), which was later validated as an instrument for measuring neighborhood walkability (Silveira and Motl 2020). Green et al. (2021) considered more than 50 attitudinal items associated with five relevant scales in the psychometric literature, to model the decision to use managed lanes (i.e. tolled lanes on an expressway). In a multi-step process, they identified the 25 most promising items for inclusion in their survey, and after testing the inclusion of both composite scales and individual items in their models and finding that the items performed better, they retained six individual items in their final model. Within the present line of research and preceding our study, Shaw (2021) explores the efficacy of MVs in several models of mode usage, finding that the models were significantly enhanced with the addition of attitudinal factor scores predicted from MVs using machine learning methods. The attitudes improved the model fits for VO, ridehailing usage, vehicle-miles driven, public transit usage, and bicycle usage.

The methods we use in the present study blend elements from these prior approaches. Specifically, like Slade (1997) and others, one of our goals is to see how well a reduced set of items can predict the *original* scores (as opposed to re-doing the factor analysis on the reduced set and analyzing the correlations of the *two* sets of scores); like Kaufman et al. (2016) and others, we use highest pattern loadings (rather than stepwise regression) to select the reduced item set; like Shaw (2021), we use machine learning (rather than conventional regression) to predict the original factor scores from the reduced item set; and like Green et al. (2021), we also embrace the use of individual attitudinal items as direct explanatory variables (comparing their efficacy to that of composite factor scores).

2.2 Vehicle ownership modeling

Because the present study also models VO, we briefly review the literature on that subject. To model VO, analysts typically turn to count models such as negative binomial and Poisson models (Karlaftis and Golias 2002; Shay and Khattak 2007), or discrete choice models, whether ordered or unordered (Anowar, Eluru, et al. 2014). Ordinal models assume that the choices have a natural ordering, which certainly applies to different levels of VO, and are seen in many studies (Baldwin Hess and Ong 2002; Cao et al. 2007; Kim and Kim 2004; Ma and Srinivasan 2010). Unordered discrete choice models are often multinomial logit (MNL) or a variant (Anowar, Yasmin, et al. 2014; Dissanayake and Morikawa 2002; Guo 2013; Ryan and Han 1999). Anowar et al. (Anowar, Eluru, et al. 2014) provide a thorough review of several types of VO models. The variety of models appearing in the literature can be attributed to each formulation having its own pros and cons. Here, we follow Bhat and Pulugurta (1998), who found that MNL models of vehicle ownership were preferable to ordinal response models because they are less restrictive in nature.

The variables used to explain VO are just as important as the model form. Traditionally, VO is modeled using socio-economic variables (Bhat et al. 2013; Dargay and Gately 1997; Dargay et al. 2007). Some models have included built environment effects (Kim and Mokhtarian 2018; Li et al. 2010). Though attitudes have been used to model VO in the past (Kim and Mokhtarian 2018; Mokhtarian and Cao 2008), they have typically been quantified from factor-analyzing responses to a full set of attitudinal statements. When conducting an exploratory factor analysis (EFA), typically each attitudinal construct, or factor, is associated with three or more survey statements (items), thus entailing relatively larger data requirements than the MV approach we propose in this research. Here, we hypothesize that MVs can be used effectively in a VO model, and that the resulting model has better predictive accuracy than a model without attitudes altogether.

3. ANALYSIS METHODS

3.1 Transfer process with random forest and elastic net regression

For clarity, we include important definitions below. Importantly, this analysis within the MV line of research is based on a single dataset, in which all observations contain all variables (in particular, the full sample was used to conduct the EFA and compute factor scores, which for simplicity we will call the "observed" scores[1]). However, we simulate a scenario in which a random half of the data (the donor sample) contains all the variables, including observed attitude scores, while the other half (the recipient sample) contains only MVs and variables appearing in the VO multinomial logit (MNL) model. Following earlier work (Mokhtarian 2023), we use "internal evaluation" to refer to how well the learning function estimates the observed scores, and "external evaluation" to refer to how well the imputed scores perform in a travel behavior model. We use the term "adjacent-sample" to convey that the recipient dataset, although not overlapping the donor dataset, is not a fully separate sample (since the two halves were collected at the same time, over the same geography, using identical surveys and sampling methodologies), for which we would reserve the term "out-of-sample" (meaning, fully separate from the donor sample).

Donor dataset: data that are used to train the imputation functions (learning functions). This dataset minimally includes the MVs and the observed attitude scores computed from the EFA solution in Table 2 below.

[1] In reality, of course, they are only estimates of scores on latent (i.e. unobserved) factors.

Recipient dataset: data that (are treated as though they) only contain the MVs, as well as other variables appearing in the downstream travel behavior model (of vehicle ownership, in this case).

Enriched dataset: the recipient dataset augmented by the imputed attitude scores obtained by applying to the recipient sample the learning functions trained on the donor sample.

In-sample internal evaluation: assessing the Pearson correlation and root mean squared error (RMSE) between an *observed* attitude score and the *imputed* score, in the donor dataset. For a given factor, RMSE is computed as $\sqrt{\frac{\sum_{n=1}^{N}(I_n - O_n)^2}{N}}$, where I_n is the imputed score and O_n the observed score for person *n*, and *N* is the sample size.

Adjacent-sample internal evaluation: although the transfer process assumes that the recipient dataset lacks the observed attitude scores, since it actually has those scores in our split-sample approach we can still assess the Pearson correlation and RMSE of the *imputed* and *observed* attitude scores in the recipient/enriched dataset.

External evaluation: assessing the differences in model fit, interpretability, and prediction accuracy between models *without* attitudes, with the marker variables themselves, with *imputed* attitude scores, and with *observed* attitude scores (where available). For brevity, the VO models presented in this study are based only on the enriched dataset. Therefore, we do not present an in-sample external validation, using the donor dataset. Rather, we use the enriched dataset with *observed* attitude scores and conduct an *adjacent-sample* external evaluation (defined below).

Adjacent-sample external evaluation: assessing the differences in model fit, interpretability, and prediction accuracy between models without attitudes, with the marker variables themselves, with *imputed* attitude scores (in the enriched dataset), and (since the recipient dataset actually has these scores in our split-sample approach) with *observed* attitude scores.

The transfer process involves six steps. Figure 1 summarizes each of the steps, which we describe in greater detail below. Importantly, Steps 2 – 6 are repeated 1,000 times, each iteration using a different donor and recipient dataset obtained through random sampling. Case weights for the donor sample are incorporated by taking the previously-obtained weights for those cases (based on weighting over the entire sample), and rescaling them so that the sum of the rescaled weights is equal to the number of observations in the donor sample – and similarly for the recipient sample. By iterating through these steps numerous times, we are able to view the distribution of outcomes for all evaluation steps.

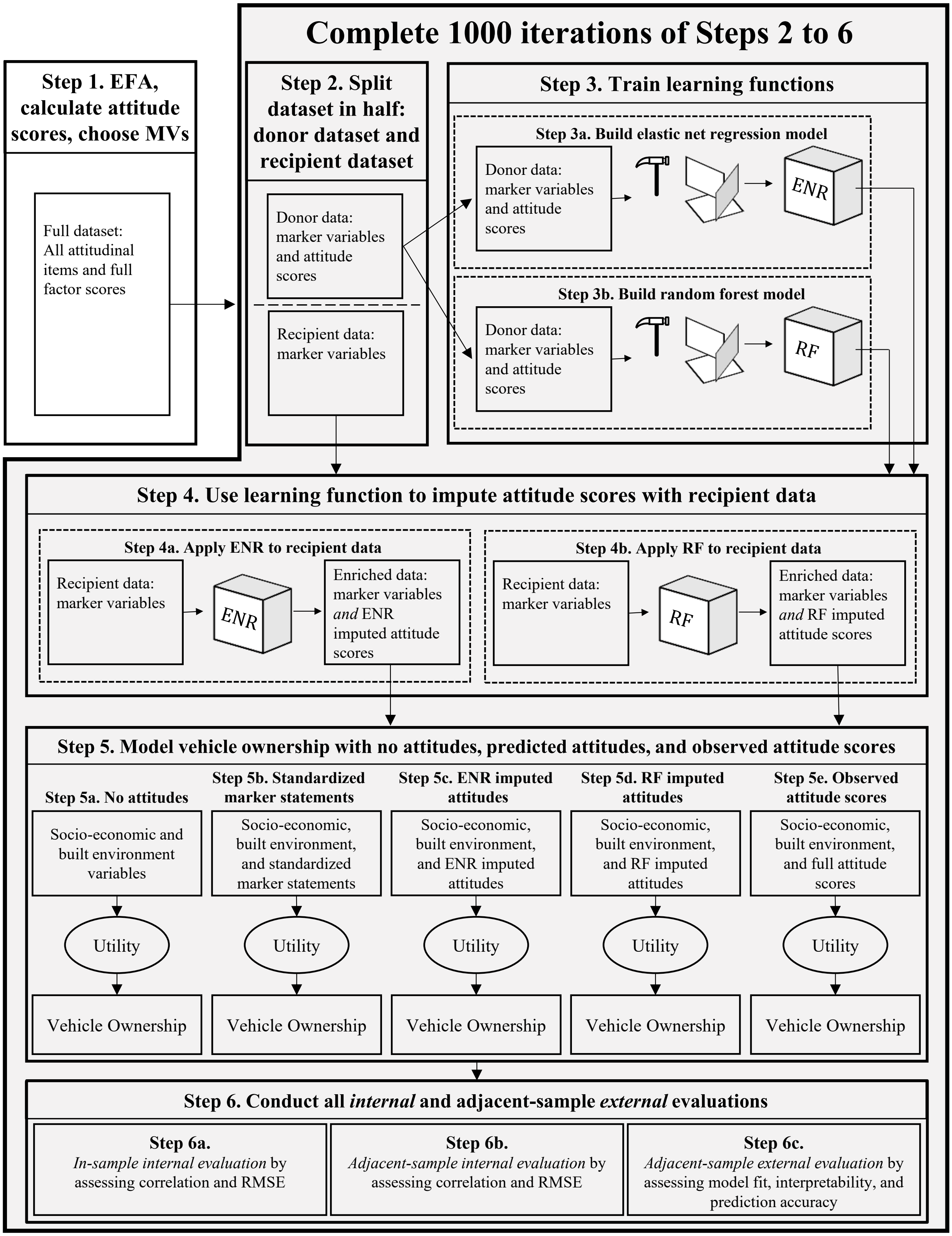


Figure 1 Transfer process, learning functions, and evaluation

Step 1. As discussed further in Section 3.3, factor analyze the attitudinal statements using the entire sample[2], with the resulting (Bartlett) scores treated as the "observed attitude scores". Identify the MVs as the highest-loading statement on each factor. A total of 15 MVs are chosen, one for each factor.

Step 2. Split the dataset in half by randomly sampling, *without* replacement, 1,579 observations. Treat this set as the donor dataset, with the remaining data used as the recipient dataset.

Step 3. Train the learning functions using the donor dataset.

Step 3a. Train an elastic net regression (ENR) model using the case-weighted donor dataset. All MVs are the features and observed attitude scores are the respective targets. Hyperparameter tuning (α and L_1 ratio) is done via 10-fold cross-validation, and the "best" parameters are chosen by using the values that produce the lowest RMSE for the test data within this cross-validation process. The hyperparameter values used in the grid search are $\alpha \in \{1x10^{-4}, 1x10^{-3}, 0.01, 0.1, 1, 10, 100\}$ and $L_1 \in \{0.05, 0.10, 0.15, \dots, 0.95, 1.00\}$. Each of the 1,000 iterations of the entire process undergoes its own ENR hyperparameter tuning. For more information about ENR, we refer readers to Zou and Hastie (2005).

Step 3b. Train a random forest (RF) model using the case-weighted donor dataset. All MVs are the features and observed attitude scores are the targets. Hyperparameter tuning (number of trees, n_{trees}, to grow) is done via 10-fold cross-validation, and the "best" parameter is chosen by using the value that produces the lowest RMSE for the test data within this cross-validation process. The tuning process explored values of $n_{trees} \in \{20, 40, 60, 80, 100\}$. Each iteration undergoes its own RF hyperparameter tuning. For more information about RF, we refer readers to Ho (1995).

Step 4. Use the trained learning functions to impute attitude scores into the recipient dataset, thereby obtaining the enriched dataset.

Step 4a. ENR learning function is applied to recipient (adjacent-sample) data.

Step 4b. RF learning function is applied to recipient (adjacent-sample) data.

Step 5. Estimate five VO models using the case-weighted enriched dataset. The alternatives in the MNL model are 0-vehicle, one-vehicle, two-vehicle, and three-plus-vehicle households. For more information about MNL, we refer readers to Train (2009). The five models differ in the way attitudes are measured (if at all). Specifically, the models respectively include:

Step 5a. No attitudes;

---

[2] A case can be made for only using the donor subsample to create the factor scores, since by using the full sample the recipient subsample is also influencing the solution. However, if only using the donor subsample: (a) Which one? We created 1,000 of them. (b) We would lose the ability to perform the adjacent-sample internal evaluation and to include the observed scores in one of the VO models compared in the adjacent-sample external evaluation. At this early stage of this line of inquiry, we consider the value of these analyses to outweigh the argument for using only the donor subsample, but eventually, of course, this approach should be tested in a true out-of-sample evaluation, for which observed attitudes will not be available. As a test on the stability of the factor solution, we investigated similarly structured EFA analyses (i.e. same oblimin rotation with the same items) using only the data from the donor samples and found that they also have very similar interpretations.

Step 5b. Only the standardized marker statement responses (SMS) associated with the attitudes of interest;

Step 5c. ENR-imputed attitude scores;

Step 5d. RF-imputed attitude scores; and

Step 5e. The observed attitude scores.

This array of models allows us to compare the efficacy of the first four alternatives against the maximally-informed approach of creating factor scores from a full set of attitudinal statements.

Step 6. Conduct all internal and adjacent-sample external evaluations. Specifically:

Step 6a. Conduct an *in-sample internal evaluation* with the donor dataset, by assessing the correlation and RMSE between observed and imputed attitude scores.

Step 6b. Conduct an *adjacent-sample internal evaluation* with the enriched dataset, also by assessing the correlation and RMSE between observed (which we had treated as unavailable in Steps 3 and 4) and imputed attitude scores.

Step 6c. On the enriched sample, conduct an *adjacent-sample external evaluation* by comparing the model fit, interpretability, and prediction accuracy of VO models without attitudes, with standardized marker statement responses, with (respectively ENR- and RF-) imputed attitude scores, and with observed attitude scores.

We estimate the household VO models with the same specifications throughout, for consistency of comparisons. The base specification is obtained from an MNL VO model estimated using *all* observations (N=3,178)[3] and observed attitudes, and it is inspired by the utility specification of a preliminary study which did *not* generate a bootstrap distribution of key model variables and statistics (Soria and Mokhtarian 2024). We include a table containing the VO MNL reference model using all observations with observed attitude scores in the appendix. The systematic utility V of $X \in \{0, 1, 2, 3+\}$ household vehicle ownership is a function of the number of drivers in a household, household income, activity density (employment + residential population per acre in the Census block group of the residential location), the number of stores within a mile of the residence, and attitudes. The coefficients are alternative-specific, and for identifiability we fix the values to 0 for the 0-vehicle alternative. Therefore, the model results are relative to this reference category (i.e. a negative coefficient reduces the probability of an alternative compared to the 0-vehicle alternative). Using this specification, we employ the enriched datasets to estimate the models mentioned in Step 5 of the process outlined above – for each of 1,000 iterations. A total of five VO models are estimated in this step: one model without attitudes and four others that include attitudes.

We evaluate the performance of the VO models in several ways, both quantitatively and qualitatively. Quantitatively, we compare McFadden ρ^2's, and perform likelihood ratio tests

[3] Again, we could have obtained a specification from a donor sample only, but (a) which one? and (b) in this case it is even less critical to do so than for the EFA solution, since in this case the specification is simply a template, for which coefficients are separately estimated for each recipient sample (unlike the EFA, which is a single solution applied to every case and remaining unchanged across splits).

between the attitude-less VO model and attitude-inclusive models. In addition, we calculate the success index for each alternative (McFadden 1976) as:

$$\frac{\left(\frac{\sum_n P_n(X) y_{nX}}{\sum_n P_n(X)}\right)}{\left(\frac{\sum_n y_{nX}}{N}\right)}. \tag{1}$$

To motivate the success index, we note that the (probability-weighted) number of cases *predicted* to choose alternative X is given by $\sum_n P_n(X)$ (where $P_n(X)$ is the probability of person n choosing alternative X); the number predicted to choose X who actually *do* choose X is given by $\sum_n P_n(X) y_{nX}$ (where y_{nX} = 1 if person *n* actually chooses alternative *X*, and 0 else); and the ratio $\frac{\sum_n P_n(X) y_{nX}}{\sum_n P_n(X)}$ is called the *success proportion* of alternative X, i.e. the *share* of cases predicted to choose X that actually do choose X. The *success index* of X is the success proportion normalized by the *observed share* of X, $\left(\frac{\sum_n y_{nX}}{N}\right)$. Thus, a success index less than 1 indicates that the model does not predict this alternative's adoption even as well as a naïve market-share prediction would, and the greater than 1 the success index is, the better the model predicts adoption of this alternative compared to the market-share prediction. To enable a qualitative evaluation of model interpretability, we calculate (and present in the appendix) summary statistics for the bootstrap distribution of model coefficients. Specifically, for each explanatory variable in each model we present the average coefficient and p-value, and the percent of models for which the coefficient is statistically significant at the $\alpha = 0.05$ level.

3.2 Data

The data consists of responses to the Georgia Department of Transportation (GDOT)-funded Emerging Technologies Survey conducted from September 2017 to January 2018. This section includes information relevant to the present study; more information can be found in Kim et al. (2019). The survey collected data on socio-economic characteristics of individuals and households, technology use, current and future expected travel behavior, and general attitudes and preferences.

Respondents were recruited through invitations mailed to two groups of individuals. Each invitation included a cover letter explaining the survey purpose as well as a paper copy of the survey, and provided individual-specific access codes to the online version of the survey. The first group of individuals was a geographically-stratified random address-based sample of 30,000 adults living in the 14 Metropolitan Planning Organization areas of the state of Georgia. Approximately 1,800 responses were collected from this group. The second group consisted of respondents to the Georgia subsample of the 2017 National Household Travel Survey (NHTS) who indicated that they were willing to be contacted about future surveys (Federal Highway Administration 2018). From this group of 5,000, approximately 1,500 responses were collected. Data cleaning removed observations based on incomplete responses, surveys that were completed too quickly, flatlining (repeated answers within blocks of attitudinal statements), and failed attention checks. The attention checks were questions that directed the survey taker to choose a specific response (e.g., "To confirm you're really reading this, please select 'strongly agree' here"). After cleaning the data, we retained 3,178 observations for this analysis. Table 1 includes descriptive statistics of variables used in the VO model. The few missing responses within this dataset were imputed using a meticulously tuned Random Forest algorithm. For example, of the 46 attitude statements a respondent saw, a missing response to a single statement was imputed using that individual's socio-economic variables and responses to other statements.

The explanatory variables used in the VO models were chosen because (except for the attitudes) they are usually available to practitioners via travel diary surveys. Table 1 provides unweighted and case-weighted descriptive statistics of these variables for the entire sample, including the Bartlett factor scores of multiple attitudes derived from the exploratory factor analysis described in the next subsection. Importantly, responses to the four marker statements shown are converted from categorical to numerical values as follows: -2 = "strongly disagree", -1 = "strongly agree", 0 = "neutral", 1 = "agree", 2 = "strongly agree".

Table 1 Descriptive statistics (N = 3,178)

Variable	Unweighted average (or share)	Unweighted SD[a]	Weighted average (or share)	Weighted SD[a]
Vehicle ownership				
Zero vehicles	2.5%		5.4%	
One vehicle	25.9%		33.0%	
Two vehicles	39.9%		33.5%	
Three or more vehicles	31.7%		28.1%	
Number of household drivers	1.92	0.84	1.84	0.86
HH income < $50K	31.3%		36.6%	
HH income $50K – $99,999	36.6%		31.6%	
HH income $100K or more	32.1%		31.9%	
Pop. + jobs per acre in Census block group of residence	4.06	6.07	5.08	7.74
Number of stores in 1-mi radius	10.04	8.30	11.04	8.41
Pro-car-owning Bartlett score	0.00	1.00	-0.03	1.09
Pro-suburban Bartlett score	0.00	1.00	-0.03	1.01
Urbanite Bartlett score	0.00	1.00	0.13	1.01
Pro-non-car alternatives Bartlett score	0.00	1.00	-0.01	1.03
Pro-car owning item[b]	1.34	0.88	1.31	0.93
Pro-suburban item[b]	0.08	1.14	0.07	1.18
Urbanite item[b]	0.31	1.25	0.43	1.26
Pro-non-car alternatives item[b]	-0.06	1.15	-0.08	1.20

[a] Standard deviation; [b]The MV for the associated factor – see the italicized statements in Table 2.

3.3 Exploratory factor analysis and marker variable selection

An exploratory factor analysis (EFA) condensed correlated responses to attitude statements on transportation mode preferences, residential location, technology, travel, time-use, and comfort with other people, into composite scores on a smaller number of factors. The results are shown in Table 2. Because the EFA results are readily available from an existing report and for brevity, we refer readers to Kim et al. (2019) for details. The rest of this subsection is dedicated to explaining the essence of the factor analysis and how it is applied to this MV research. The analysis resulted in a 15-factor solution based on principal axis factoring with an oblimin (oblique) rotation and using 38 statements (after discarding those not loading heavily on any factor, or comprising the only strongly-loading statement on a factor). Bartlett scores were calculated for each factor; these constitute the observed attitude scores. For this research one MV for each factor is identified (by italics) in Table 2, for a total of 15. We hypothesize that the ability of machine learning models to capture possible non-linear relationships of the item responses to the factor scores may lead to better imputation accuracy. Though not all factors are present in the VO model, we decided to use 15 items rather than only four because practitioners may want to study multiple factors related to a multitude of travel behaviors other than VO. Furthermore, each marker variable can potentially provide additional predictive value, even for scores on attitudes with which it is not strongly (linearly) associated.

Table 2 Exploratory factor analysis results

Factor	Statement [a]	Pattern matrix loading [b,c]
Pro-non-car alternatives	*I like the idea of walking as a means of travel for me.*	*0.666*
	I like the idea of bicycling as a means of travel for me.	0.628
	I like the idea of public transit as a means of travel for me.	0.336
Pro-car-owning	*I definitely want to own a car.*	*0.748*
	I am fine with not owning a car, as long as I can use/rent one any time I need it.	−0.576
	I like the idea of driving as a means of travel for me.	0.535
	As a general principle, I'd rather own things myself than rent or borrow them from someone else.	0.404
Pro-suburban	*I prefer to live in a spacious home, even if it's farther from public transportation or many places I go to.*	*0.609*
	I see myself living long-term in a suburban or rural setting.	0.387
Urbanite	*I like the idea of having stores, restaurants, and offices mixed among the homes in my neighborhood.*	*0.417*
	My phone is so important to me, it's almost part of my body.	0.350
Tech-savvy	*Learning how to use new technologies is often frustrating for me.*	*−0.866*
	I am confident in my ability to use modern technologies.	0.801
Commute benefit	*My commute is a useful transition between home and work (or school).*	*0.677*
	My travel to/from work (or school) is usually pleasant.	0.579
	I wish I could instantly be at work (or school)—the trip itself is a waste of time.	−0.428

Work-oriented	*At this stage of my life, having fun is more important to me than working hard.*	*−0.572*
	I'm too busy to have as much leisure time as I'd like.	0.527
	It's very important to me to achieve success in my work.	0.298
Materialistic	*I usually go for the basic ("no-frills") option rather than paying more money for extras.*	*−0.565*
	The functionality of a car is more important to me than the status of its brand.	−0.431
	I would/do enjoy having a lot of luxury things.	0.426
	I like to wait a while rather than being first to buy new products.	−0.357
	I prefer to minimize the amount of things I own.	−0.341
Polychronic	*I prefer to do one thing at a time.*	*−0.834*
	I like to juggle two or more activities at the same time.	0.697
Pro-environmental	*Cost or convenience takes priority over environmental impacts (e.g., pollution) when I make my daily choices.*	*−0.914*
	I am committed to an environmentally friendly lifestyle.	0.481
Pro-exercise	*The importance of exercise is overrated.*	*−0.669*
	I am committed to exercising regularly.	0.663
Family/friends-oriented	*Family/friends play a big role in how I schedule my time.*	*0.612*
	It's okay to give up a lot of time with family and friends to achieve other worthy goals.	−0.468
Waiting-tolerant	*Having to wait is an annoying waste of time.*	*−0.831*
	Having to wait can be a useful pause in a busy day.	0.533
Travel liking	*I generally enjoy the act of traveling itself.*	*0.618*
	I like exploring new places.	0.593
Sociable	*I consider myself to be a sociable person.*	*0.563*
	I'm uncomfortable being around people I don't know.	−0.507

Note: Among the four factors of interest, the highest-magnitude correlation is 0.41, between non-car alternatives and pro-car owning.

[a] Italicized statements are the marker variables; [b] Oblimin rotation; [c] Statements with loadings lower than 0.30 are suppressed (with an exception near that threshold for the work-oriented factor)

Several approaches to selecting MVs were tested, including choosing items to represent each factor based on highest communalities, choosing items that appear in multiple factors, and choosing items with the highest pattern loading of each factor. We found that using the items with the highest pattern loading on their respective factor produced the best in-sample internal validation (i.e. highest correlation between predicted and observed factor scores). This mirrors results from the literature (Kaufman et al. 2016; Tambs and Moum 1993), where the same method produced high correlations (greater than 0.90) between predicted and observed scores.

4. RESULTS

In this section we focus on the evaluation step (Step 6 in Figure 1). For each evaluation, we aggregate the results across the 1,000 iterations. Rather than making inferences about the MV approach from one snapshot, iterating through the process 1,000 times allows us to make stronger conclusions by basing them on a distribution of outcomes.

4.1 Internal evaluation

We begin with the internal evaluation of imputation accuracy for the donor and enriched datasets (Steps 6a and 6b), first through measuring the average correlation between observed and imputed attitude scores and calculating 95% confidence intervals. In addition to the imputed factor scores, we apply the same approach to the standardized marker statements (SMS). Figure 2 summarizes the results. Within the machine learning methods, both ENR and RF yield high correlations between the observed and imputed scores. RF produces markedly higher values for the donor dataset, but for the enriched dataset (the adjacent-sample internal evaluation), the correlations for the RF approach are not much better than those for ENR. Together, these results suggest RF overfitting on the donor dataset, likely an artefact of the automatic hyperparameter tuning process. Overall, Figure 2 still shows sizable correlations of 0.70 – 0.87 between observed and imputed attitude scores in the enriched dataset. The SMS method produces results almost as good, showing correlation coefficients in both donor and enriched datasets ranging from 0.59 to 0.86.

Next, still focused on internal evaluations, we do the same thing for root mean squared error (RMSE), and Figure 3 reveals a similar story for that metric of accuracy. Where a lower RMSE indicates better predictions, RF again shows better performance than ENR on the donor dataset. For the recipient sample, however, the adjacent-sample internal evaluation indicates that the RMSE differences between ENR and RF are nearly indiscernible, where the 95% confidence intervals show plenty of overlap for each attitude. Both figures show well-performing attitude imputation results. Additionally, in line with previous research, the accuracy of imputed attitude scores is related to the factor pattern loading of the respective MV. For instance, the urbanite MV has the lowest pattern loading (0.417), lowest correlation for the enriched dataset (0.702 with ENR), and highest RMSE (0.737 with RF), while MVs for the other attitudes have pattern loadings greater than 0.600, and consequently higher correlations and lower RMSE with the full scores.

4.2 Adjacent-sample external evaluation

Having found that the learning functions perform adequately, we now turn to the household VO model results, and conduct the adjacent-sample external evaluation (Step 6c). On the qualitative side, we consider model interpretability. Though the 1,000 versions produced for each of the five models were not examined individually, we include tables in the appendix for each of the five that provide the average coefficient estimates and p-values, as well as the percentage of the 1,000 runs for which each coefficient is statistically significant at the $\alpha = 0.05$ level. Not surprisingly for a model of vehicle ownership, the attitudes that are consistently statistically significant are pro-car-owning and pro-non-car-alternatives. In general, the average coefficients indicate that the directionality of effects follows our expectations. For example, pro-car owning attitudes are related to higher levels of vehicle ownership whereas pro-non-car-alternatives are related to lower levels of ownership. Because we only vary the data and not the model specification (aside from whether attitudes are included, and the alternative ways of measuring the attitudes), the tables in the appendix show that some attitudes are statistically insignificant more often than not. However, these results are dependent on the choice of reference alternative and other specification decisions.

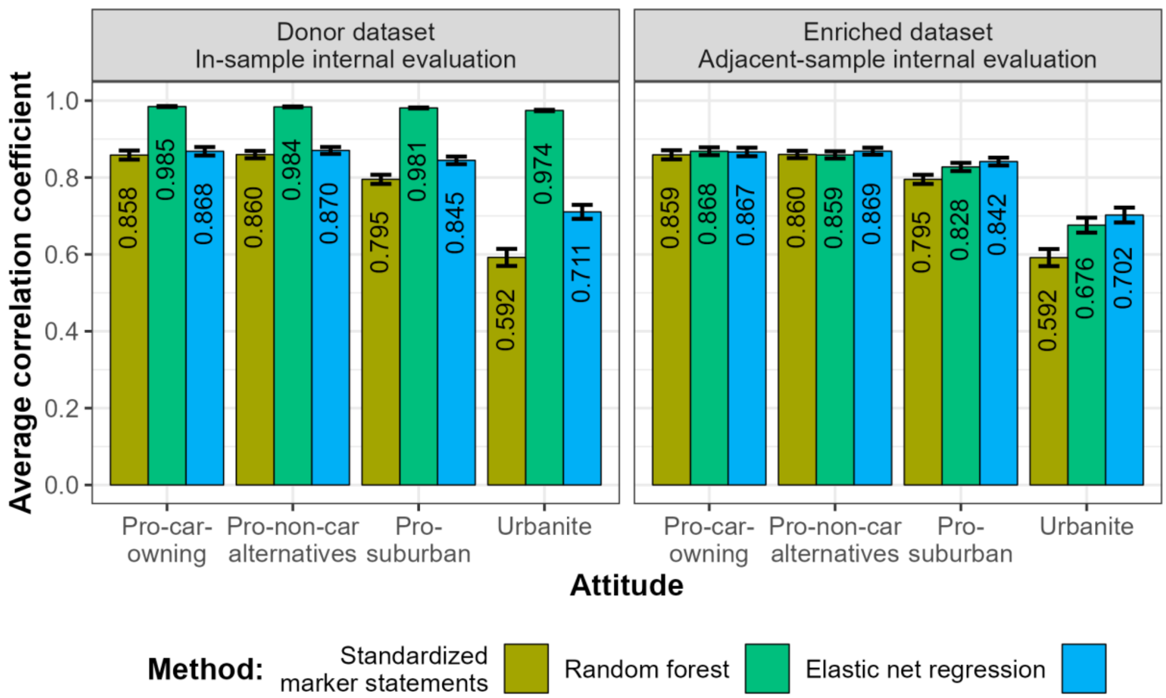


Figure 2 Average (over 1,000 random splits) correlations between observed and imputed attitude scores

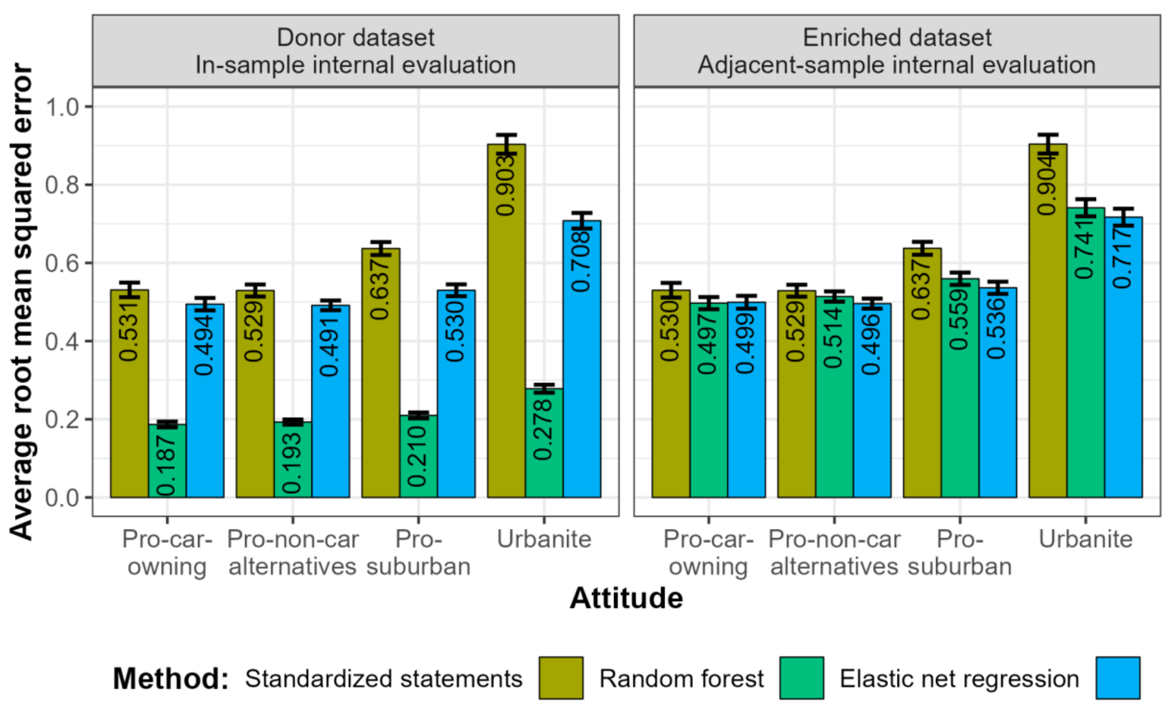


Figure 3 Average (over 1,000 random splits) root mean squared errors between observed and imputed attitude scores

To further investigate the qualitative interpretation of the VO MNL model results, we assess the extent to which the coefficients have the expected sign and are statistically significant (at the 0.05 level). Strongly identified attitudes (those with large-magnitude pattern loadings) tend to have the expected sign and are statistically significant more often. Figure 4 illustrates the attitude coefficient results based on these criteria (an appendix figure presents the non-attitude coefficient results). It shows that the coefficients of the pro-car-owning and pro-non-car alternatives attitudes have the expected sign and are statistically significant a majority of the time across all model types (i.e. an affinity for car ownership significantly increases the probability of owning vehicles and an affinity for non-car alternatives significantly decreases the probability of owning vehicles). In most iterations, the pro-car-owning, pro-suburban, and pro-non-car alternatives attitudes at least have the expected sign (though they may not be statistically significant), and in nearly all cases for which they have the wrong sign, they are insignificant. Interestingly, the pro-non-car alternatives attitude performs much better in the SMS and ENR- and RF-imputed attitudes models than it does in the observed factor score models. It may well be that the MV representing this attitude is more influential on vehicle ownership than the composite score built from multiple items. The urbanite attitude performs the worst of the four: it has the unexpected sign by far the most often, and such coefficients are often statistically significant (especially for the ENR- and RF- imputed versions of the attitude). This result may be attributed to the urbanite attitude's MV being the weakest in the EFA solution, having the smallest-magnitude pattern loading (as mentioned in Section 4.1).

On the quantitative side, the MNL models allow for several likelihood-based comparisons. First, we compare the McFadden ρ^2 of the models, as shown in Figure 5 which presents the average ρ^2 value based on the equally-likely and market-share loglikelihoods. We see that the models with attitudes perform better than the models with no attitudes, generally adding about 0.02 to the fit measure and with negligible differences among the four model types that involve attitudes. The adjusted ρ^2 measures, which penalize for adding variables, show similar improvement, indicating that the additional explanatory power brought by the attitudes is worth the loss of parsimony.

Next, we conduct likelihood ratio tests assessing whether the four attitudes, included for each of the three alternatives (1, 2, 3+ vehicles; thus, 12 degrees of freedom) are collectively statistically significant. We conduct four such tests for each set of five models, assessing the inclusion of attitudes measured four different ways against the corresponding model without attitudes, and thus 4,000 likelihood ratio tests are conducted in all. At $\alpha = 0.05$, there are no instances of an insignificant test. At $\alpha = 0.01$, there is only one instance among the 4,000 tests. There are no more instances of insignificant likelihood ratio tests until the value of α decreased to 0.0005 – i.e., at $\alpha = 0.0005$, only two tests out of 4,000 did not yield statistical significance. At $\alpha = 0.0001$ there are only 7 tests out of 4,000 that did not yield statistical significance. Thus, overall, the evidence is extremely strong that the attitudes bring significant explanatory power to the VO model, and that this is true even for the simplest way of measuring attitudes, the standardized marker statements.

As the final component of the adjacent-sample external evaluation, we calculate the success index, a measure of predictive performance, for each alternative of each model, and average the results from the 1,000 iterations. Figure 6 shows how the average success index for each alternative varies across the five models. The most noticeable difference regarding this measure is the improvement in 0-vehicle household predictions, where the average success index increases from 5.8 for the model with no attitudes, to 7.2 – 7.4 for the models containing attitudes. The tight range in the latter case shows that prediction of the 0-vehicle alternative using imputed attitudes is nearly as good as predictions using observed scores. Interestingly, however, for the 1-, 2-, and 3+-vehicle alternatives the inclusion of attitudes offers negligible improvement beyond what conventional explanatory variables already provide. Nevertheless, in view of the policy importance of the 0-vehicle household (for equity and social exclusion considerations on the one hand, and sustainability goals on the other hand), and also in view of its relatively small share of 5.4% of the weighted sample (which generally increases prediction difficulty), the marked improvement in the ability to predict that alternative is valuable for more enabling analysts to more accurately assess the effects of various policies.

In summary, the results from Figure 2 to Figure 6 illustrate the effectiveness of MVs at imputing attitude scores and improving the model fit and predictive accuracy of household VO models. From 15 MVs, we are able to train ENR and RF learning functions with reasonable accuracy according to the adjacent-sample internal evaluation based on correlation and RMSE metrics that compare imputed and observed attitude scores. For the adjacent-sample external evaluation of VO modeling, we observe improvements in model fit (i.e. McFadden's ρ^2) and statistically significant likelihood ratio tests with the inclusion of standardized marker statements alone, as well as with ENR- and RF- imputed attitude scores. Assessment of predictive ability based on the success index indicates that nearly all of the model improvement derives from better prediction of 0-vehicle households. Refreshingly, the SMS, RF-imputed attitudes, and ENR-imputed attitudes models perform nearly as well as the models with observed attitude scores. We discuss the broader impacts of this research in the next section.

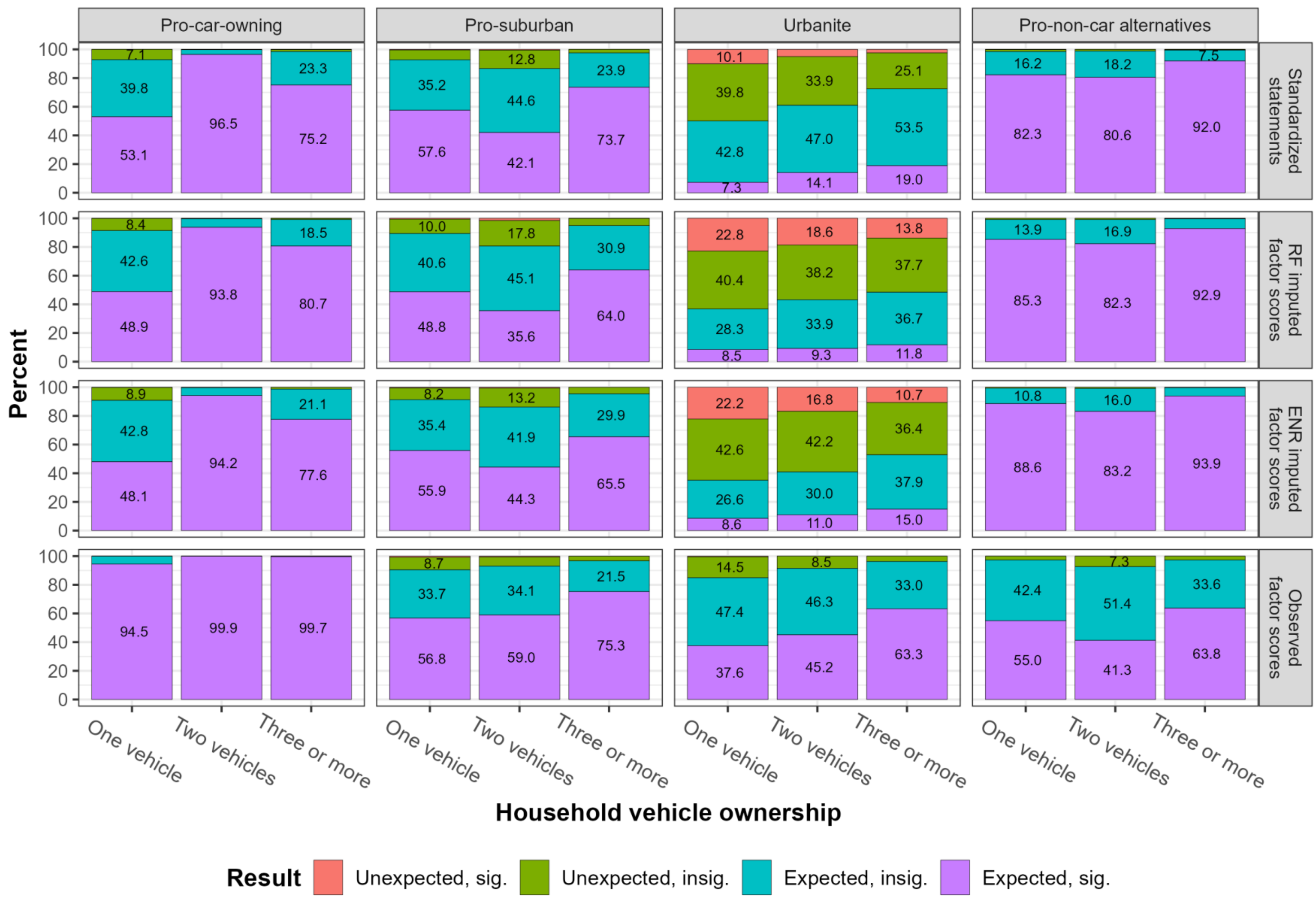


Figure 4 Percent of 1,000 attitude coefficient results that are expected or unexpected and statistically significant or insignificant (at α = 0.05)

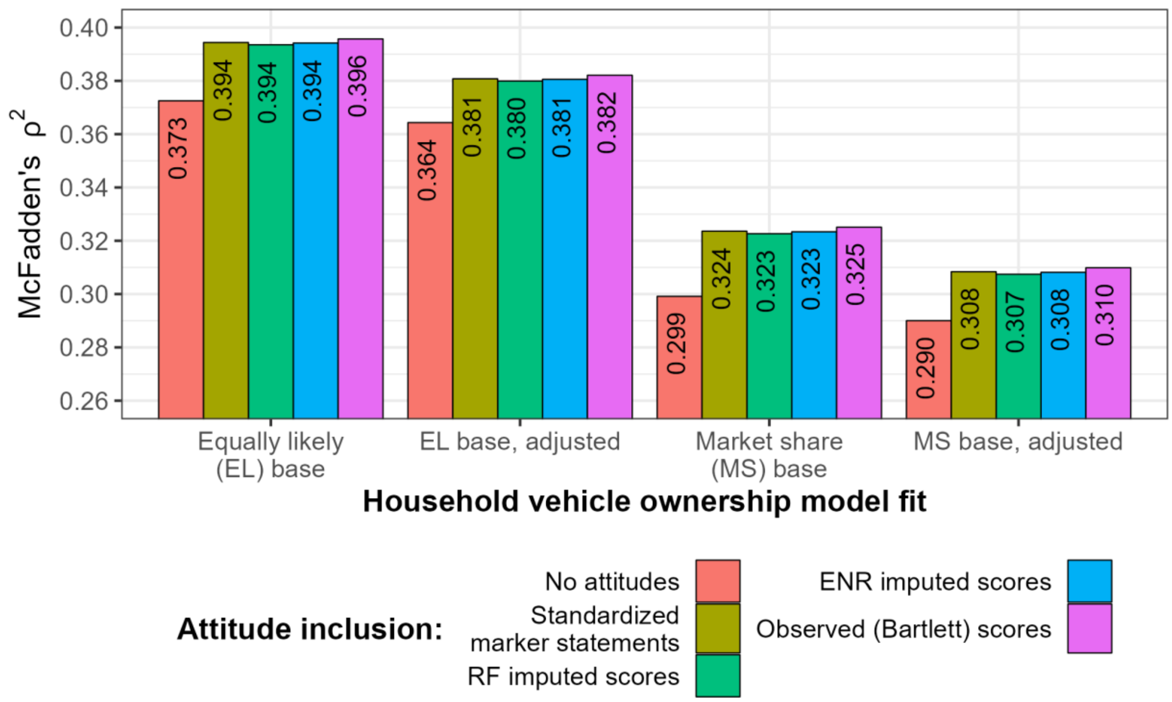


Figure 5 Average (over 1,000 random splits) McFadden's ρ^2 (note: vertical axis does not begin at 0)

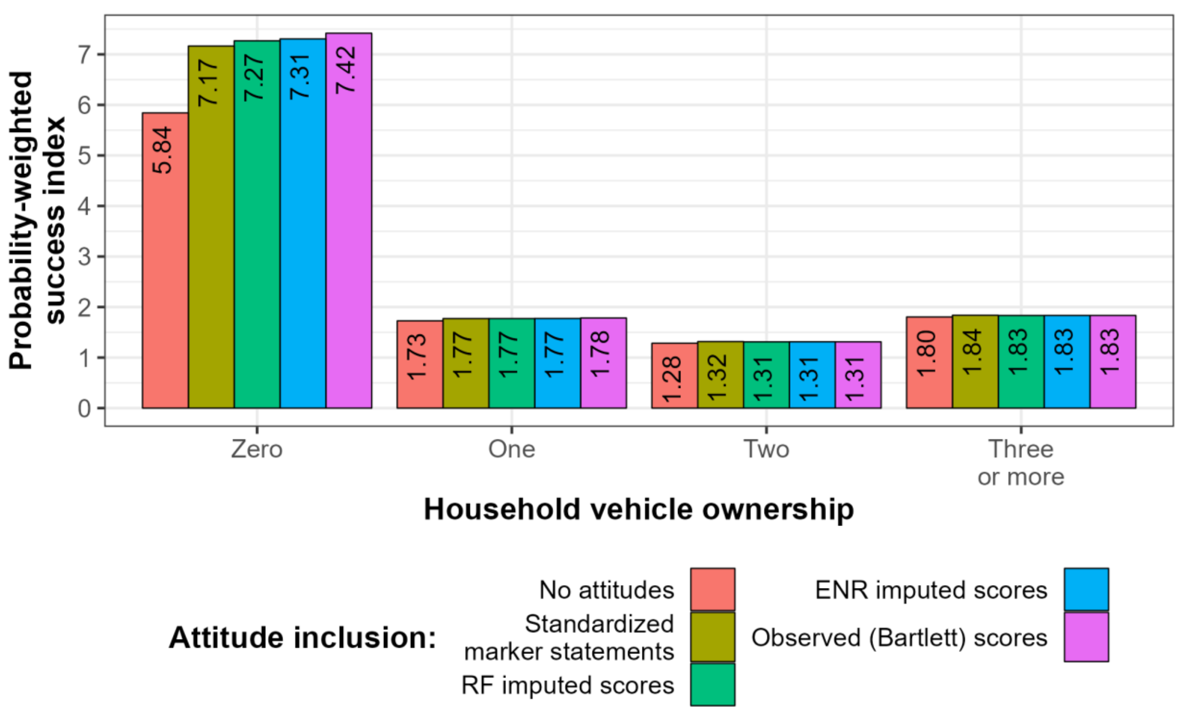


Figure 6 Average (over 1,000 random splits) success index for each alternative and different inclusions of attitudes

5. DISCUSSION AND CONCLUSION

This paper is one in a series of studies designed to investigate the practicality of including attitudes as explanatory variables in practice-oriented travel demand forecasting models. The specific goals of this study were to evaluate (1) how well we can predict attitudes using only a small number of marker variables (MVs); (2) to see how well the predicted attitudes perform in a model of household vehicle ownership (VO); and (3) to compare the performance of predicted factor scores to that of simply using the individual MVs as explanatory variables. In this study, we respectively applied random forest (RF) and elastic net regression (ENR) to 15 MVs, to predict factor scores on four attitudes: pro-car ownership, pro-non-car alternatives, pro-suburban, and urbanite. We incorporated those four imputed attitudes into multinomial logit VO models, and compared the results to those of models including the original four factor scores, models containing only the four MVs most strongly associated with the same four attitudes, and models containing no attitudes. We created 1,000 random splits of a sample of 3,178 responses to a 2017 survey of Georgia adults, training the RF and ENR functions on each donor half-sample and applying those functions to the recipient half-sample. We reported results averaged over the 1,000 recipient half-samples.

With respect to how well we can predict attitudes, we achieved average correlations between the original and predicted factor scores ranging between 0.68 and 0.87 in the recipient half-samples (0.59 to 0.86 for the singular MVs). With respect to how well the predicted attitudes performed in the VO model, we first of all found, consistent with numerous previous studies (Bhagat-Conway et al. 2022), that our results support the usefulness of attitudes for explaining travel behavior. In particular, all sets of attitudes (original, ENR-imputed, RF-imputed, and MVs only) only rarely (and in some cases never) had estimated coefficients that were both statistically significant *and* with the *un*expected sign (see Figure 4 and the accompanying text for details). By far most often, estimated coefficients were statistically significant and with the *expected* sign. All sets of attitudes generally increased McFadden's ρ^2 and adjusted ρ^2 by about 0.02. All sets of attitudes were collectively statistically significant based on likelihood ratio tests (in all but 7 out of 4,000 tests, significant at $\alpha = 0.0001$).

Perhaps most importantly, success indices show that the predictive power of the models markedly improves specifically for zero-car households whenever the attitudes are included (Figure 6). This signals that attitudes are useful for differentiating atypical behavior from regional norms. Specifically, in the North American context, households tend to have at least one vehicle, and 0-vehicle households are a mixture of those that are 0-vehicle by necessity (for which income is a good predictor) and those that are 0-vehicle by choice (for which attitudes are good predictors). Accordingly, this early finding suggests that attitudes may benefit other travel behavior models relevant to demand forecasting, beyond vehicle ownership.

With respect to how well the marker variables themselves performed relative to the more elaborate prediction of factor scores using machine learning methods, the short answer is, "very nearly as well". The ENR and RF transfer processes used 15 statements in the imputation phase. The standardized marker statement (SMS) approach showed that using only the four statements most strongly associated with the four attitudes included in the model was almost as good[4] – and

[4] Of course, this will only be true to the extent that the four factors associated with the retained MVs are essentially orthogonal to the remaining MVs. If, instead, the latter are also meaningfully correlated with the factors of interest, then including them can be expected to improve the imputation process.

clearly better than not including attitudes at all. The slightly larger average success indices of the zero-car household predictions for the RF- and ENR-imputed factor scores compared to that of the SMS method can be attributed to the SMS method not allowing much variation in attitude measurement since the item responses only take on five discrete values, whereas the factor scores predicted by machine learning methods are continuous-valued, affording finer-grained contributions to explaining VO. For practitioners, an implication of this result is that even the SMS method can lead to marked improvements, while the extra effort of training a "simple" machine learning model will likely yield results even closer to those of using the observed attitude scores.

These results have meaningful implications for the future design of surveys that provide the data used to build regional models. The first phase of this research required data on 38 attitudinal statements, which, using EFA, were condensed into scores on 15 factors. Selecting only 15 statements allowed us to predict VO with improved accuracy over a model that included no attitudes. Future practice-oriented surveys that wish to measure the same factors/attitudes may only need to collect responses on these 15 statements, or even on only the four key marker statements, rather than on all 38.[5]

Further, the methodology demonstrated in this study can be extended to identify multiple sets of marker variables (MVs) from different donor surveys, each set designed to predict different behaviors. An EFA or similar technique that produces a composite attitude score based on multiple statements can identify which statement is most representative of that attitude, and therefore likely to be most useful as a MV. Then, if a future survey designer wants to include 20 attitude statements, she may select 5 of the MVs identified in this study and 15 other MVs drawn from multiple other studies based on those studies' factor analysis results.

An important caveat to this line of argument, however, is to beware the reification fallacy, that is, "believing our own labels". For example, if a factor analysis identifies 10 factors from 40 statements, with the literature-advised (Fabrigar et al. 1999) 3 – 6 statements strongly associated with each factor but with all 40 statements contributing richness and texture to the resulting scores, it is all too easy to imbue a factor with the meaning that the analyst wants/expects it to have. It is even easier to do so when a factor score is imputed from a set of 10 MVs (only one or two of which will be strongly associated with it, but all of which are, again, contributing to the imputed score). From that perspective, going to the extreme of simply using individual MVs as explanatory variables may help avoid the reification fallacy trap, to the extent that a single statement is more likely to be taken at face value and not imbued with meaning that it does not possess.

As this research continues to expand, we acknowledge the limitations of the present study and plan the next steps. For example, this study uses only one dataset. In the future, it is important to evaluate this MV framework with two different datasets rather than artificially creating donor and recipient datasets from a single sample (i.e., to conduct true out-of-sample evaluations). Specifically, we plan to apply the MV framework to different donor and recipient datasets, and investigate the spatial and temporal transferability of attitudes. In the spatial sense, the formulas generating attitudinal factor scores, or even the highest-loading statements informing those scores, may not be constant across different geographic areas. For example, the order of statements that

[5] We hasten to note that longer surveys will still be important for research purposes, notably including the purpose of identifying which few attitudinal items, among potentially numerous candidates, will be most powerful at explaining the travel behaviors of interest.

load most strongly onto the pro-car owning attitude may not be the same between highly urbanized and highly rural areas. Similarly, they may not be the same over time. In support of the transferability investigation, additional research on factors associated with the spatial and temporal stability of attitudes (or lack thereof) would be highly desirable. Such research will be essential to tackling the *second* challenge of including attitudes in practice-ready travel demand models: namely, how to forecast them.

Another important line of research is to explicitly model the measurement error associated with using estimates of attitudes rather than their true values (likely involving ICLV and structural equation modeling approaches), and assess its impact on parameter estimates and resulting inferences. While we have focused on offering practitioners a relatively simple way of including attitudes into practice-oriented models, and keeping in mind the truism that "all models are wrong; some are useful", it is important to ascertain whether the wrongness-usefulness tradeoff of this particular approach is helpful on net.

While much remains to be learned, this approach appears to offer considerable promise for improving the behavioral content and predictive power of travel demand forecasting models, increasing their responsiveness to the rapidly changing transportation landscape. We look forward to the next steps along the way.

6. REFERENCES

Anowar, Sabreena, Naveen Eluru, and Luis F. Miranda-Moreno (2014) Alternative modeling approaches used for examining automobile ownership: a comprehensive review. *Transport Reviews* **34(4)**, 441-473.

Anowar, Sabreena, Shamsunnahar Yasmin, Naveen Eluru, and Luis F. Miranda-Moreno (2014) Analyzing car ownership in Quebec City: a comparison of traditional and latent class ordered and unordered models. *Transportation* **41**, 1013-1039.

Bahamonde-Birke, Francisco J., and Juan de Dios Ortúzar (2021) How to categorize individuals on the basis of underlying attitudes? A discussion on latent variables, latent classes and hybrid choice models. *Transportmetrica A: Transport Science* **17(4)**, 856-877 https://doi.org/10.1080/23249935.2020.1817171.

Baldwin Hess, Daniel, and Paul M. Ong (2002) Traditional neighborhoods and automobile ownership. *Transportation Research Record* **1805(1)**, 35-44.

Bhagat-Conway, Matthew Wigginton, Laura Mirtich, Deborah Salon, Nathan Harness, Alexis Consalvo, and Shuyao Hong (2022) Subjective variables in travel behavior models: a critical review and Standardized Transport Attitude Measurement Protocol (STAMP). *Transportation*.

Bhat, Chandra R., Rajesh Paleti, Ram M. Pendyala, Karen Lorenzini, and Karthik C. Konduri (2013) Accommodating immigration status and self-selection effects in a joint model of household auto ownership and residential location choice. *Transportation Research Record* **2382(1)**, 142-150.

Bhat, Chandra R., and Vamsi Pulugurta (1998) A comparison of two alternative behavioral choice mechanisms for household auto ownership decisions. *Transportation Research Part B: Methodological* **32(1)**, 61-75.

Cain, Kelli L., Kavita A. Gavand, Terry L. Conway, Carrie M. Geremia, Rachel A. Millstein, Lawrence D. Frank, Brian E. Saelens, Marc A. Adams, Karen Glanz, and Abby C. King (2017) Developing and validating an abbreviated version of the Microscale Audit for Pedestrian Streetscapes (MAPS-Abbreviated). *Journal of Transport & Health* **5**, 84-96.
Cao, Xinyu, Patricia L. Mokhtarian, and Susan L. Handy (2007) Do changes in neighborhood characteristics lead to changes in travel behavior? A structural equations modeling approach. *Transportation* **34**, 535-556.
Cerin, Ester, Terry L. Conway, Brian E. Saelens, Lawrence D. Frank, and James F. Sallis (2009) Cross-validation of the factorial structure of the Neighborhood Environment Walkability Scale (NEWS) and its abbreviated form (NEWS-A). *International Journal of Behavioral Nutrition and Physical Activity* **6(1)**, 1-10.
Dargay, Joyce, and Dermot Gately (1997) Vehicle ownership to 2015: implications for energy use and emissions. *Energy Policy* **25(14-15)**, 1121-1127.
Dargay, Joyce, Dermot Gately, and Martin Sommer (2007) Vehicle ownership and income growth, worldwide: 1960-2030. *The Energy Journal* **28(4)**.
Dissanayake, Dilum, and Takayuki Morikawa (2002) Household travel behavior in developing countries: Nested logit model of vehicle ownership, mode choice, and trip chaining. *Transportation Research Record* **1805(1)**, 45-52.
Domarchi, Cristian, Alejandro Tudela, and Angélica González (2008) Effect of attitudes, habit and affective appraisal on mode choice: an application to university workers. *Transportation* **35(5)**, 585-599 https://doi.org/10.1007/s11116-008-9168-6.
Ewing, Reid, Sadegh Sabouri, Keunhyun Park, Torrey Lyons, and Guang Tian (2019) Key Enhancements to the WFRC/MAG Four-Step Travel Demand Model. https://pdxscholar.library.pdx.edu/cgi/viewcontent.cgi?article=1212&context=trec_reports.
Fabrigar, Leandre R., Duane T. Wegener, Robert C. MacCallum, and Erin J. Strahan (1999) Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods* **4(3)**, 272.
Federal Highway Administration (2018) *Summary of Travel Trends 2017 National Household Travel Survey.* https://nhts.ornl.gov/assets/2017_nhts_summary_travel_trends.pdf.
Green, Lisa L., Mark W. Burris, David Florence, and Winfred Arthur (2021) Psychological items: a useful addition in modeling travel behavior on managed lanes. *Transportation* **48(1)**, 215-237. https://doi.org/10.1007/s11116-019-10049-z.
Guo, Zhan (2013) Does residential parking supply affect household car ownership? The case of New York City. *Journal of Transport Geography* **26**, 18-28.
Hall, C. Michael, and Yael Ram (2018) Walk score® and its potential contribution to the study of active transport and walkability: A critical and systematic review. *Transportation Research Part D: Transport and Environment* **61**, 310-324.
Hendrick, Clyde, Susan S. Hendrick, and Amy Dicke (1998) The Love Attitudes Scale: Short Form. *Journal of Social and Personal Relationships* **15(2)**, 147-159 https://doi.org/10.1177/0265407598152001.
Ho, Tin Kam (1995) "Random decision forests." Proceedings of 3rd International Conference on Document Analysis and Recognition. https://ieeexplore.ieee.org/document/598994.
Karlaftis, Matthew, and John Golias (2002) Automobile ownership, households without automobiles, and urban traffic parameters: are they related? *Transportation Research Record* **1792(1)**, 29-35.

Kaufman, Erin A., Mengya Xia, Gregory Fosco, Mona Yaptangco, Chloe R. Skidmore, and Sheila E. Crowell (2016) The Difficulties in Emotion Regulation Scale Short Form (DERS-SF): Validation and Replication in Adolescent and Adult Samples. *Journal of Psychopathology and Behavioral Assessment* **38(3)**, 443-455 https://doi.org/10.1007/s10862-015-9529-3.

Kim, Hong Sok, and Eungcheol Kim (2004) "Effects of public transit on automobile ownership and use in households of the USA." *Review of Urban & Regional Development Studies: Journal of the Applied Regional Science Conference*. https://doi.org/10.1111/j.1467-940X.2005.00090.x.

Kim, Sung Hoo, Patricia L. Mokhtarian, and Giovanni Circella (2019) The Impact of Emerging Technologies and Trends on Travel Demand in Georgia. https://rosap.ntl.bts.gov/view/dot/56095/dot_56095_DS1.pdf.

Kim, Sung Hoo, and Patricia L. Mokhtarian (2018) Taste heterogeneity as an alternative form of endogeneity bias: Investigating the attitude-moderated effects of built environment and socio-demographics on vehicle ownership using latent class modeling. *Transportation Research Part A: Policy and Practice* **116**, 130-150 https://www.sciencedirect.com/science/article/pii/S0965856417313721.

Li, Jieping, Joan L. Walker, Sumeeta Srinivasan, and William P. Anderson (2010) Modeling private car ownership in China: investigation of urban form impact across megacities. *Transportation Research Record* **2193(1)**, 76-84.

Ma, Lu, and Sivaramakrishnan Srinivasan (2010) Impact of individuals' immigrant status on household auto ownership. *Transportation Research Record* **2156(1)**, 36-46.

Malokin, Aliaksandr, Patricia L. Mokhtarian, and Giovanni Circella (2019) A Transfer Learning-Based Framework for Enriching National Household Travel Survey Data with Attitudinal Variables. https://doi.org/10.35090/gatech/66813.

Mokhtarian, Patricia L., and Xinyu Cao (2008) Examining the impacts of residential self-selection on travel behavior: A focus on methodologies. *Transportation Research Part B: Methodological* **42(3)**, 204-228 https://www.sciencedirect.com/science/article/pii/S0191261507000744.

Mokhtarian, Patricia L. (2024) Pursuing the impossible (?) dream: Incorporating attitudes into practice-ready travel demand forecasting models. *Transportation Research Part A*, in press.

Popuri, Yasasvi, Kimon Proussaloglou, Cemal Ayvalik, Frank Koppelman, and Aimee Lee (2011) Importance of traveler attitudes in the choice of public transportation to work: findings from the Regional Transportation Authority Attitudinal Survey. *Transportation* **38(4)**, 643-661 https://doi.org/10.1007/s11116-011-9336-y.

Rammstedt, Beatrice, and Oliver P. John (2007) Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality* **41(1)**, 203-212 https://www.sciencedirect.com/science/article/pii/S0092656606000195.

Richins, Marsha L. (2004) The Material Values Scale: Measurement Properties and Development of a Short Form. *Journal of Consumer Research* **31(1)**, 209-219 https://doi.org/10.1086/383436.

Ryan, James M., and Gregory Han (1999) Vehicle-ownership model using family structure and accessibility application to Honolulu, Hawaii. *Transportation Research Record* **1676(1)**, 1-10.

Shaw, Faaiqa Atiyya. "Methods for enriching transportation survey datasets: With sample applications using psychometric variables." PhD dissertation, Civil and Environmental Engineering, Georgia Institute of Technology, 2021. http://hdl.handle.net/1853/64640.

Shay, Elizabeth, and Asad J. Khattak (2007) Automobiles, trips, and neighborhood type: Comparing environmental measures. *Transportation Research Record* **2010(1)**, 73-82.

Silveira, Stephanie L., and Robert W. Motl (2020) Abbreviated Neighborhood Environment Walkability scale in persons with multiple sclerosis: Initial validation of score inferences. *Journal of Transport & Health* **19**, 100952 https://www.sciencedirect.com/science/article/pii/S2214140520301560.

Slade, Gary D. (1997) Derivation and validation of a short-form oral health impact profile. *Community Dentistry and Oral Epidemiology* **25(4)**, 284-290 https://doi.org/10.1111/j.1600-0528.1997.tb00941.x.

Soria, Jason, and Patricia L. Mokhtarian (2024) How Effective are Marker Variables at Imputing Attitudes? An External Evaluation Using Vehicle Ownership Models. Presentation #TRBAM-24-04220 at the 103rd Annual Meeting of the Transportation Research Board, Washington DC, January 2024. https://hdl.handle.net/1853/75405.

Soria, Jason, Shelly Etzioni, Yoram Shiftan, Eran Ben-Elia, and Amanda Stathopoulos (2023) Microtransit adoption in the wake of the COVID-19 pandemic: Evidence from a choice experiment with transit and car commuters. *Transportation Research Part C: Emerging Technologies* **157**, 104395 https://www.sciencedirect.com/science/article/pii/S0968090X23003856.

Tambs, K., and T. Moum (1993) How well can a few questionnaire items indicate anxiety and depression? *Acta Psychiatrica Scandinavica* **87(5)**, 364-367 https://doi.org/10.1111/j.1600-0447.1993.tb03388.x.

Tarabay, Rana, and Maya Abou-Zeid (2019) Modeling the choice to switch from traditional modes to ridesourcing services for social/recreational trips in Lebanon. *Transportation* https://doi.org/10.1007/s11116-019-09973-x.

Train, Kenneth. *Discrete Choice Methods with Simulation*. Cambridge University Press, 2009.

Wang, Xinyi, F. Atiyya Shaw, Patricia L. Mokhtarian, Giovanni Circella, and Kari E. Watkins (2023) Combining disparate surveys across time to study satisfaction with life: the effects of study context, sampling method, and transport attributes. *Transportation* **50(2)**, 513-543 https://doi.org/10.1007/s11116-021-10252-x.

Ware, John E., Mark Kosinski, and Susan D. Keller (1996) A 12-Item Short-Form Health Survey: Construction of Scales and Preliminary Tests of Reliability and Validity. *Medical Care* **34(3)** https://journals.lww.com/lww-medicalcare/fulltext/1996/03000/a_12_item_short_form_health_survey__construction.3.aspx

Zou, Hui, and Trevor Hastie (2005) Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **67(2)**, 301-320 https://doi.org/10.1111/j.1467-9868.2005.00503.x.

7. APPENDIX

Appendix Table 1 Vehicle ownership multinomial logit model with all observations and observed attitude factor scores (N = 3,178)

Reference alternative: Zero vehicles	One vehicle		Two vehicles		Three or more vehicles	
	Value	P-value	Value	P-value	Value	P-value
Alternative specific constant	0.232	0.562	-3.090	0.000	-5.819	0.000
Number of household drivers	1.982	0.000	4.074	0.000	5.220	0.000
HH income \$50K – \$99,999	2.217	0.000	3.092	0.000	3.542	0.000
HH income \$100K or more	2.453	0.000	4.106	0.000	4.573	0.000
Pop. + jobs per acre around residence	-0.018	0.047	-0.028	0.009	-0.051	0.004
Number of stores in 1-mi radius	-0.035	0.016	-0.083	0.000	-0.109	0.000
Att: pro-car-owning (obs.)	0.432	0.000	0.833	0.000	0.850	0.000
Att: pro-suburban (obs.)	0.275	0.009	0.334	0.004	0.467	0.000
Att: urbanite (obs.)	-0.198	0.053	-0.274	0.013	-0.393	0.001
Att: pro-non-car alternatives (obs.)	-0.301	0.001	-0.269	0.012	-0.394	0.000
Observed (unweighted) share	25.9%		33.9%		31.7%	
Observed (weighted) share	33.0%		33.0%		28.1%	
Equally-likely (EL) loglikelihood	-4405.643					
Market-share (MS) loglikelihood	-3944.424					
Final loglikelihood	-2702.211					
Adjusted $\rho^2_{EL\ base}$	0.380					
Adjusted $\rho^2_{MS\ base}$	0.307					

Appendix Table 2 Average estimated coefficients and p-values of multinomial logit model results for model with no attitudes (across 1,000 random splits, holding model specification constant)

Reference alternative: Zero vehicles	**One vehicle**			**Two vehicles**			**Three or more vehicles**		
	Value (SD[a])	P-value (SD[a])	% sig.[b]	Value (SD[a])	P-value (SD[a])	% sig.[b]	Value (SD[a])	P-value (SD[a])	% sig.[b]
Alternative specific constant	0.192 (0.849)	0.359 (0.302)	19.5%	-2.941 (0.192)	0.012 (0.077)	96.6%	-5.571 (0.979)	0.000 (0.003)	99.9%
Number of household drivers	2.136 (0.644)	0.000 (0.003)	99.9%	4.164 (0.667)	0.000 (0.000)	100.0%	0.683 (0.295)	0.000 (0.000)	100.0%
HH income $50K – $99,999	2.271 (0.845)	0.001 (0.016)	99.7%	3.179 (0.854)	0.000 (0.014)	99.9%	0.875 (0.513)	0.000 (0.013)	99.9%
HH income $100K or more	3.929 (3.237)	0.316 (0.205)	14.4%	5.421 (3.224)	0.172 (0.185)	51.0%	3.232 (0.000)	0.160 (0.174)	51.0%
Pop. + jobs per acre around residence	-0.024 (0.020)	0.221 (0.295)	51.3%	-0.046 (0.027)	0.110 (0.209)	67.4%	-0.087 (0.033)	0.024 (0.090)	90.0%
Number of stores in 1-mi radius	-0.056 (0.032)	0.094 (0.189)	69.3%	-0.107 (0.035)	0.005 (0.031)	98.0%	-0.136 (0.036)	0.001 (0.008)	99.8%

Appendix Table 3 Average estimated coefficients and p-values of multinomial logit results with standardized attitude marker statements as explanatory variables (across 1,000 random splits, holding model specification constant)

Reference alternative: Zero vehicles	One vehicle			Two vehicles			Three or more vehicles		
	Value (SD[a])	P-value (SD[a])	% sig.[b]	Value (SD[a])	P-value (SD[a])	% sig.[b]	Value (SD[a])	P-value (SD[a])	% sig.[b]
Constant	0.138 (0.936)	0.355 (0.298)	19.5%	-3.157 (1.002)	0.013 (0.086)	96.5%	-5.927 (1.065)	0.000 (0.002)	100.0%
Number of household drivers	2.319 (0.706)	0.000 (0.001)	100.0%	4.405 (0.725)	0.000 (0.000)	100.0%	5.561 (0.742)	0.000 (0.000)	100.0%
HH income \$50K – \$99,999	2.455 (0.863)	0.001 (0.012)	99.8%	3.373 (0.875)	0.000 (0.010)	99.9%	3.866 (0.893)	0.000 (0.009)	99.9%
HH income \$100K or more	4.462 (2.959)	0.208 (0.194)	42.1%	6.138 (2.953)	0.142 (0.152)	51.0%	6.618 (2.959)	0.130 (0.141)	51.0%
Pop. + jobs per acre around residence	-0.013 (0.027)	0.286 (0.309)	35.6%	-0.030 (0.030)	0.243 (0.302)	45.6%	-0.059 (0.035)	0.146 (0.243)	61.5%
Number of stores in 1-mi radius	-0.053 (0.036)	0.141 (0.231)	58.7%	-0.102 (0.038)	0.018 (0.090)	93.7%	-0.129 (0.039)	0.004 (0.036)	98.4%
Att: pro-car-owning (MV)	0.232 (0.156)	0.173 (0.250)	53.1%	0.602 (0.188)	0.010 (0.060)	96.5%	0.443 (0.223)	0.071 (0.162)	75.3%
Att: pro-suburban (MV)	0.314 (0.215)	0.156 (0.250)	58.0%	0.264 (0.234)	0.230 (0.282)	42.6%	0.480 (0.244)	0.085 (0.191)	73.7%
Att: urbanite (MV)	0.004 (0.251)	0.389 (0.308)	17.4%	-0.079 (0.267)	0.380 (0.310)	19.1%	-0.150 (0.275)	0.353 (0.309)	21.4%
Att: pro-non-car alternatives (MV)	-0.465 (0.220)	0.053 (0.146)	82.4%	-0.513 (0.237)	0.053 (0.142)	80.6%	-0.665 (0.254)	0.023 (0.096)	92.1%

Appendix Table 4 Average estimated coefficients and p-values of multinomial logit results with elastic net regression imputed attitude factor scores as explanatory variables (across 1,000 random splits, holding model specification constant)

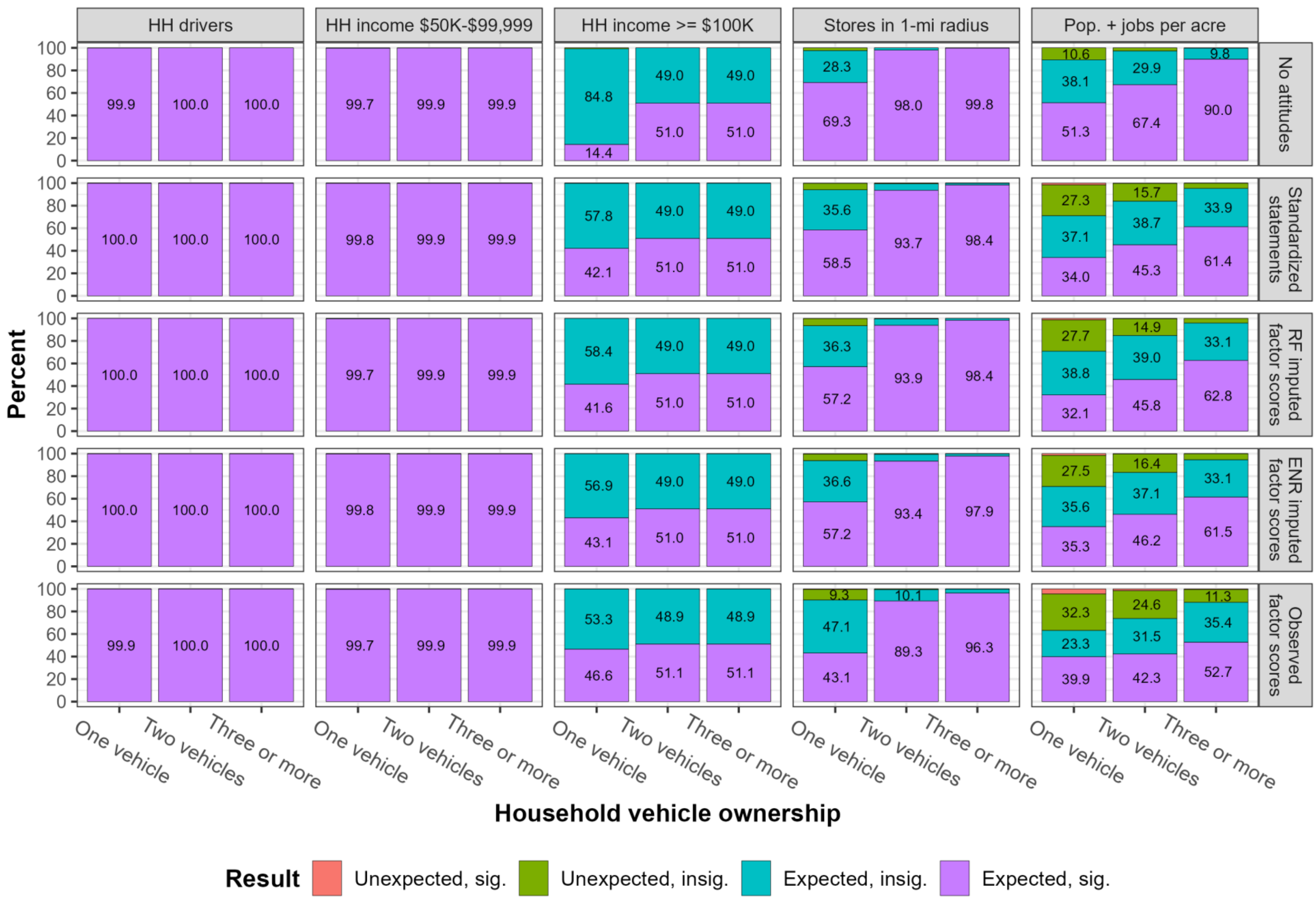
Reference alternative: Zero vehicles	One vehicle			Two vehicles			Three or more vehicles		
	Value (SD[a])	P-value (SD[a])	% sig.[b]	Value (SD[a])	P-value (SD[a])	% sig.[b]	Value (SD[a])	P-value (SD[a])	% sig.[b]
Alternative specific constant	0.089 (0.942)	0.359 (0.299)	18.6%	-3.189 (1.009)	0.012 (0.084)	96.8%	-5.962 (1.072)	0.000 (0.001)	100.0%
Number of household drivers	2.354 (0.729)	0.000 (0.002)	100.0%	4.440 (0.746)	0.000 (0.000)	100.0%	5.600 (0.764)	0.000 (0.000)	100.0%
HH income \$50K – \$99,999	2.475 (0.889)	0.001 (0.013)	99.8%	3.388 (0.901)	0.000 (0.011)	99.9%	3.889 (0.919)	0.000 (0.010)	99.9%
HH income \$100K or more	4.719 (3.169)	0.241 (0.228)	43.1%	6.376 (3.162)	0.182 (0.193)	51.0%	6.879 (3.170)	0.170 (0.181)	51.0%
Pop. + jobs per acre around residence	-0.012 (0.028)	0.280 (0.310)	36.9%	-0.030 (0.031)	0.246 (0.307)	46.5%	-0.059 (0.036)	0.150 (0.247)	61.6%
Number of stores in 1-mi radius	-0.051 (0.036)	0.149 (0.235)	57.6%	-0.101 (0.038)	0.020 (0.096)	93.4%	-0.127 (0.039)	0.005 (0.037)	97.9%
Att: pro-car-owning (ENR)	0.267 (0.206)	0.201 (0.270)	48.3%	0.713 (0.247)	0.015 (0.067)	94.2%	0.577 (0.287)	0.067 (0.161)	77.7%
Att: pro-suburban (ENR)	0.374 (0.278)	0.170 (0.258)	56.4%	0.341 (0.312)	0.222 (0.286)	44.9%	0.542 (0.330)	0.121 (0.224)	65.6%
Att: urbanite (ENR)	0.114 (0.418)	0.313 (0.310)	30.8%	0.064 (0.446)	0.325 (0.308)	27.8%	-0.051 (0.460)	0.344 (0.315)	25.7%
Att: pro-non-car alternatives (ENR)	-0.614 (0.250)	0.031 (0.111)	88.6%	-0.627 (0.272)	0.045 (0.123)	83.2%	-0.836 (0.293)	0.016 (0.077)	93.9%

Appendix Table 5 Average estimated coefficients and p-values of multinomial logit results with random forest imputed attitude factor scores as explanatory variables (across 1,000 random splits, holding model specification constant)

Reference alternative: Zero vehicles	One vehicle			Two vehicles			Three or more vehicles		
	Value (SD[a])	P-value (SD[a])	% sig.[b]	Value (SD[a])	P-value (SD[a])	% sig.[b]	Value (SD[a])	P-value (SD[a])	% sig.[b]
Alternative specific constant	0.086 (0.947)	0.356 (0.299)	19.7%	-3.189 (1.011)	0.012 (0.082)	96.5%	-5.952 (1.072)	0.000 (0.001)	100.0%
Number of household drivers	2.367 (0.744)	0.000 (0.002)	100.0%	4.459 (0.762)	0.000 (0.000)	100.0%	5.613 (0.778)	0.000 (0.000)	100.0%
HH income \$50K – \$99,999	2.471 (0.892)	0.001 (0.015)	99.8%	3.389 (0.903)	0.000 (0.014)	99.9%	3.883 (0.921)	0.000 (0.013)	99.9%
HH income \$100K or more	4.651 (3.200)	0.244 (0.228)	41.5%	6.312 (3.193)	0.182 (0.193)	51.0%	6.806 (3.201)	0.170 (0.181)	51.0%
Pop. + jobs per acre around residence	-0.012 (0.027)	0.289 (0.310)	33.5%	-0.031 (0.031)	0.245 (0.306)	46.1%	-0.060 (0.035)	0.144 (0.240)	62.4%
Number of stores in 1-mi radius	-0.052 (0.036)	0.140 (0.225)	57.8%	-0.102 (0.038)	0.018 (0.087)	93.8%	-0.128 (0.039)	0.004 (0.033)	98.2%
Att: pro-car-owning (RF)	0.271 (0.208)	0.190 (0.257)	50.1%	0.688 (0.236)	0.016 (0.070)	93.5%	0.581 (0.260)	0.051 (0.139)	80.4%
Att: pro-suburban (RF)	0.333 (0.284)	0.200 (0.274)	49.4%	0.273 (0.307)	0.262 (0.295)	37.0%	0.514 (0.323)	0.131 (0.231)	63.7%
Att: urbanite (RF)	0.134 (0.424)	0.308 (0.304)	30.1%	0.086 (0.444)	0.330 (0.315)	27.9%	0.026 (0.457)	0.339 (0.307)	25.7%
Att: pro-non-car alternatives (RF)	-0.584 (0.252)	0.038 (0.116)	86.2%	-0.621 (0.272)	0.045 (0.124)	82.7%	-0.797 (0.291)	0.019 (0.084)	93.0%

Appendix Table 6 Average estimated coefficients and p-values of multinomial logit results with observed attitude factor scores as explanatory variables (across 1,000 random splits, holding model specification constant)

Reference alternative: Zero vehicles	One vehicle			Two vehicles			Three or more vehicles		
	Value (SD[a])	P-value (SD[a])	% sig.[b]	Value (SD[a])	P-value (SD[a])	% sig.[b]	Value (SD[a])	P-value (SD[a])	% sig.[b]
Alternative specific constant	0.266 (0.940)	0.343 (0.304)	22.7%	-3.094 (1.006)	0.014 (0.085)	96.1%	-5.890 (1.074)	0.000 (0.002)	99.9%
Number of household drivers	2.118 (0.693)	0.000 (0.004)	99.9%	4.241 (0.712)	0.000 (0.000)	100.0%	5.409 (0.732)	0.000 (0.000)	100.0%
HH income \$50K – \$99,999	2.513 (0.949)	0.001 (0.012)	99.7%	3.403 (0.962)	0.000 (0.010)	99.9%	3.871 (0.980)	0.000 (0.009)	99.9%
HH income \$100K or more	4.796 (3.084)	0.192 (0.189)	46.6%	6.491 (3.078)	0.136 (0.147)	51.1%	6.986 (3.086)	0.125 (0.136)	51.1%
Pop. + jobs per acre around residence	-0.007 (0.034)	0.241 (0.295)	44.4%	-0.021 (0.037)	0.256 (0.308)	43.9%	-0.047 (0.040)	0.208 (0.293)	53.3%
Number of stores in 1-mi radius	-0.042 (0.036)	0.213 (0.278)	43.6%	-0.090 (0.038)	0.032 (0.114)	89.3%	-0.115 (0.039)	0.009 (0.055)	96.3%
Att: pro-car-owning (obs.)	0.474 (0.161)	0.014 (0.060)	94.5%	0.876 (0.187)	0.000 (0.003)	99.9%	0.900 (0.206)	0.000 (0.006)	99.7%
Att: pro-suburban (obs.)	0.309 (0.239)	0.159 (0.250)	57.6%	0.365 (0.251)	0.144 (0.237)	59.6%	0.499 (0.261)	0.082 (0.188)	75.4%
Att: urbanite (obs.)	-0.231 (0.226)	0.259 (0.295)	38.1%	-0.303 (0.234)	0.208 (0.270)	45.2%	-0.423 (0.244)	0.119 (0.212)	63.3%
Att: pro-non-car alternatives (obs.)	-0.331 (0.188)	0.147 (0.227)	55.0%	-0.291 (0.206)	0.223 (0.273)	41.3%	-0.422 (0.218)	0.113 (0.202)	63.8%



Appendix Figure 1 Share of non-attitude coefficient results that are expected or unexpected and statistically significant or insignificant