TRR
JOURNAL OF THE TRANSPORTATION RESEARCH BOARD

Research Article

Transportation Research Record
1–14

DOI: 10.1177/0361198119857953
journals.sagepub.com/home/trr
SAGE

Attitudes on Autonomous Vehicle Adoption using Interpretable Gradient Boosting Machine

Dongwoo Lee[1], John Mulrow[1], Chana Joanne Haboucha[2], Sybil Derrible[1], and Yoram Shiftan[2]

Abstract

This article applies machine learning (ML) to develop a choice model on three choice alternatives related to autonomous vehicles (AV): regular vehicle (REG), private AV (PAV), and shared AV (SAV). The learned model is used to examine users' preferences and behaviors on AV uptake by car commuters. Specifically, this study applies gradient boosting machine (GBM) to stated preference (SP) survey data (i.e., panel data). GBM notably possesses more interpretable features than other ML methods as well as high predictive performance for panel data. The prediction performance of GBM is evaluated by conducting a 5-fold cross-validation and shows around 80% accuracy. To interpret users' behaviors, variable importance (VI) and partial dependence (PD) were measured. The results of VI indicate that trip cost, purchase cost, and subscription cost are the most influential variables in selecting an alternative. Moreover, the attitudinal variables Pro-AV Sentiment and Environmental Concern are also shown to be significant. The article also examines the sensitivity of choice by using the PD of the log-odds on selected important factors. The results inform both the modeling of transportation technology uptake and the configuration and interpretation of GBM that can be applied for policy analysis.

The effort to make cities smarter, more sustainable, and more resilient has been accompanied by a wave of advances in artificial intelligence (AI) (*1*, *2*). One of the most disruptive AI technologies in transportation will be autonomous vehicles (AVs) that can drive themselves with less to no human intervention. The impacts of AVs are hard to estimate as they have the potential to dramatically change both lifestyles and urban systems (*1–3*). More precisely, AVs will fundamentally influence many aspects of transportation systems, including user behaviors, infrastructure (e.g., roads, traffic signals), safety, mobility (e.g., vehicle miles traveled), and environmental impacts. AVs notably allow drivers to replace driving time with more desirable activities such as reading, working, and sleeping (*1*, *4*). AVs will tend to improve mobility for children, the elderly, and individuals with disabilities. Moreover, some argue that AVs may reduce household vehicle ownership in the future since they can be more easily shared throughout a given day among users (*3*, *4*). AV system design also implies that urban areas may not require as many parking facilities in the near future. In addition, AVs' trained driving algorithms are arguably going to outperform humans in driving ability and frequency of errors. Thus, some argue that they will be able to reduce congestion and vehicle accidents while improving fuel efficiency (*4*, *5*). Despite these immediate benefits, others argue that these behavioral changes and technological benefits may, in the long run, increase vehicle miles traveled (VMT) and even exacerbate overall congestion on the roadways, resulting in negative environmental impacts (*4*, *6*).

When considering the propagation of AVs, two major platforms (i.e., markets) dominate: privately owned autonomous vehicles (PAVs) and shared autonomous vehicles (SAVs). SAVs receive particular attention because they will likely provide cost structures that facilitate the initial growth of use and exposure to AVs. Nonetheless, there remains a host of questions about the utility and effects of widespread AV use. In this context, it is essential to understand the behavioral motivations and individual attitudes regarding AVs. This improved

[1]Department of Civil and Materials Engineering, University of Illinois at Chicago, Chicago, IL
[2]Department of Civil and Environmental Engineering, Technion, Israel Institute of Technology

Corresponding Author:
Address correspondence to Dongwoo Lee: dlee226@uic.edu

understanding will also inform policy interventions and urban infrastructure investment decisions. To demystify the impacts of these disruptive future changes, researchers have recently faced a technical dilemma when modeling discrete choice behaviors: which modeling approach is appropriate? In this context, this article aims to provide useful insights into modeling discrete choice behaviors, especially for measuring individuals' AV preferences, while adopting a reliable and interpretable predictive modeling method.

In general, there is trade-off between modeling performance (e.g., prediction accuracy) and interpretability when selecting a modeling approach. Some models, such as parametric models, are often considered more easily interpretable than nonparametric models that use complex (a.k.a., black-box) algorithms. In predicting and modeling discrete choice problems, parametric approaches (e.g., the family of logit models) have been predominantly used since they are intuitive and easy to interpret based on strong theoretical backgrounds (random utility theory). Specifically, these approaches yield parameters that may be used to evaluate the impact of policies, economic changes, and technological adaptation. Recently, algorithm-based nonparametric statistical learning with machines [a.k.a., machine learning (ML)] techniques show relatively high performance to analyze a wide variety of modeling tasks compared with parametric approaches (*7–10*). In part, it is because of the fact that ML models possess fewer predetermined assumptions than the parametric models while adopting complex algorithms and myriads of sub-models (e.g., local nonparametric estimators) thanks to significant advancement in computational ability. Despite a high degree of predictive power, the general criticism about many ML techniques is their lack of interpretability because of their reliance on machine-based repetitive computation. Future models should try to possess the advantages of both modeling approaches—that is, ML and parametric approaches.

In addition, discrete choice models that estimate future behavioral changes are generally learned from stated preference (SP) and revealed preference (RP) survey data. Accordingly, this article also uses SP survey information that is collected from a total of 721 individuals and includes 4260 usable observations. This type of data is considered as *panel data* because the respondents' choices cannot be treated as independent of each other. To deal with such multi-level data, models incorporating random effects (i.e., mixed logit) have been used in many studies. These models capture unobserved heterogeneity across individuals and can account for sequentially correlated choices by single individuals. Although such hierarchical parametric modeling can control detrimental effects (e.g., fixed, random, autocorrelation), domain knowledge is still required to adjust the configuration of a model and take into account the underlying relationships between variables. As an alternative, advanced ML algorithms can be applied to address similar issues of error and correlation among observations.

This main goal of this article is to fill the knowledge gaps in modeling discrete choice behavior on AVs by applying a ML modeling method that is capable of providing reliable and interpretable information from multilevel data. Thus, the estimated (learned) model must be usable and adaptable by researchers in the future. In this context, this study makes use of a boosting method that is built on a rule-based, additive, and hierarchical learning algorithm (e.g., decision tree). Boosting methods have shown significantly high performance in predictive learning and benefit from a greater degree of interpretability than other machine learning approaches (*11–13*). In particular, boosting methods have an inherent ability to handle mixed-type data sets (*11*, *13*). They are designed to control both bias and variance. Furthermore, the learned model and interpretive statistics give useful insights for AV planning, policy making, and data collection efforts. Thus, this study contributes to the progress of machine learning in transportation engineering applications and to the body of literature aiming to characterize the future of AV technology adoption. Consequently, the specific objectives of this article are to answer the following research questions:

- How does machine-based predictive learning (i.e., ML models) perform in predicting discrete choice activities?
- How can ML models be interpreted to explain users' behaviors regarding AVs?
- How do individual and system-level characteristics influence SPs of future AV adoption?
- Which attributes of mode choice decisions are most sensitive?

After the introduction, existing studies about the adoption of AVs are reviewed and about predictive modeling methods that can be applied to discrete choice data. In the methodology section, the algorithmic properties of gradient boosting machine (GBM) and its interpretable measures are covered. Data descriptions and latent variables on attitudinal variables are then briefly explained in the third section. Finally, model results, interpretation, and conclusions are covered in the final sections.

Literature Review

Data science methods are increasingly being applied to problems of urban infrastructure analysis and design (*7,* *9, 14–19*). In particular, a wide variety of statistical learning approaches have been applied to predictive learning tasks within the realm of urban research (*7*, *20*). Each method tends to be tailored to specific problems and situations (*11*, *12*). In particular, ML and data mining techniques have been applied to travel survey data to model and predict many aspects of transportation (*21*). For instance, using data with known features, ML algorithms are trained to recognize patterns that correlate well with a target classification or attribute value (*8*, *22*). This recognition capability can then be applied to new data, yielding classifications or predictions. Thus, these techniques are especially well suited to examining the revealed mode choice preferences of individuals and societies (*7*). As Lu and Kawamura note, data science methods "view the travel mode choice as a pattern recognition problem whereby the travel choices can be identified by a combination of explanatory variables" (*23*). Lee et al. compared the multinomial logit (MNL) model with four types of artificial neural networks in predicting mode choice and showed that neural networks can outperform MNL in predictive power (*9*). In studies that compare machine learning algorithms, MNL is typically included as a baseline predictor, with mixed results. There is no clear answer to which algorithm type outperforms others, and the outcome is often dependent on how input variables are structured (*24*, *25*). Predicting transportation mode choice using decision tree classification (the method used in this study) has been studied, notably using Random Forest (i.e., tree-based ensemble learning methods) (*26*, *27*).

Gradient boosting of decision trees has become common in ML but it has not been widely applied to transportation modeling problems to date. It is, however gaining increasing attention. Semanjski and Gautama explored gradient boosting as a method for predicting mode choice from crowdsourced travel behavior data (*28*). They suggested that gradient boosting could become an essential predicting method in smart transportation. Wang and Ross found that a GBM had higher mode choice prediction accuracy than an MNL model when applied to household travel survey data (*10*). The utility of using gradient boosting for traffic flow and mode choice prediction has been similarly demonstrated by Zhang et al. and Lee and Min (*29*, *30*).

The study of AV adoption potential has also increased, anticipating the deployment of AVs for household transportation in coming years (*31*). Attempts to model the adoption of AVs based on passenger characteristics such as current driving habits, economic factors, and behavioral data have been conducted but not using decision tree-based methods such as gradient boosting (*6, 32–34*).

Methodology

This article applies a nonparametric model with machine-based repetitive and complex algorithms (also previously called ML models) to investigate users' behaviors on the adoption of AVs. Specifically, this ML model must be reliable and interpretable, while effectively controlling unobserved detrimental effects. Towards this goal, gradient boosting machine (GBM) is used.

Rule-based modeling is a nonparametric statistical model that can be applied to both regression and classification (i.e., discrete choice) problems. In particular, a decision tree used for classification is analogous to a discrete choice model (*5*, *6*). This nonparametric model is estimated by fitting models locally and then aggregating them together (i.e., additive learning). Specifically, the decision tree algorithm partitions the entire data space (i.e., a set of independent variables) into regions having the most homogeneous responses (i.e., dependent variables) to predictors (i.e., independent variables) yielding a hierarchical "tree" structure (*6*, *7*). Because of its nonparametric and rule-based properties, these have been widely used in choice prediction problems with multilevel data as is the case here (*11, 12, 35–37*). Nevertheless, single-tree models tend to have high prediction error and the results have lower prediction accuracy when the number of variables is increased. Specifically, single-tree models tend to have high variance and result in increasing the overall mean squared error (MSE) because of the additive learning process (*11*, *38*, *39*). Ensemble methods, such as boosting, are therefore preferred to enhance modeling performance.

Nonetheless, similar to other ML models, the inner processes of predictions of boosting machine can be difficult to understand—that is, the relationships between predictors and responses. There are two possible approaches to address this interpretability issue: (1) model-specific and (2) model-agnostic. The formal approaches are limited to a specific model. For example, the parameters in Random Utility Models (RUMs) are model-specific measures, which mostly require model-specific predetermined assumptions. However, this approach is not a feasible in ML since ML models are not tied to specific assumptions. Therefore, model-agnostic approaches are widely used to interpret ML models as an alternative (*40–42*). Specifically, these approaches can be applied to any ML model since the goal of learning a model is to minimize errors, including biases and variances. In the next section, the algorithmic characteristics and model-agnostic features of GBM are discussed.

Gradient Boosting Machine (GBM)

As a generalized rule-based approach, boosting offers robust and high predictive power by simultaneously controlling bias and variance. In addition, it is algorithmically well-suited to handle mixed-type data that often show mixed-effects (e.g., heterogeneity across individuals) since it learns patterns in the data based on a hierarchy (i.e., top-down approach). For instance, GBM can handle panel data by regarding predefined sub-group identifiers (i.e., respondents) as considering one of the criteria in constructing hierarchical structures (*11*, *12*). Thus, it can alleviate omitted bias issues. Moreover, in contrast to other rule-based methods, boosting algorithms sequentially forgive poorly learned samples in a single-tree structure by using multiple estimators.

As rule-based hierarchical learning, GBM also partitions the entire data space (i.e., all variables) into disjoint sub-regions that become siblings of parent nodes (also called terminal nodes). The fundamental learning process of GBM is to reconstruct the function dependence between x and y, while minimizing the following loss function to estimate ($\hat{f}(x)$, estimator) (*43*):

$$\hat{f}(x) = \arg\min_{f(x)} L(y,f(x)) = \arg\min_{f(x)} E_x[E_y L[y,f(x)]|x] \tag{1}$$

In general, the target variable (y) shows different distributions (e.g., binomial, multinomial); it is therefore required to be defined beforehand. Moreover, additional parameters are needed to estimate $\hat{f}(x)$. The equation can be expressed as follows:

$$\hat{f}(x) = f(x, \hat{\theta}) \tag{2}$$

$$\hat{\theta}(x) = \arg\min_{\theta} E_x[E_y L[y,f(x,\theta)]|x] \tag{3}$$

Equation 3 is typically non-tractable to estimate parameters; thus, the following iterative (i.e., incremental) optimization process is applied:

(1) Initially, set the average of a dependent variable, $\hat{f}(x) = 0$, and assume residual (r_i) and dependent variable (y_i) are same for all i in the data.

(2) For every sub-region (tree), repeat following steps:

- Compute the residual (negative gradient, r) for each observation
- Fitting $\hat{f}^b$ to the data (x,r); b indicates a single tree
- Update new dependent values $\hat{f}$ by adding in a shrunken version of the new tree,

$$\hat{f}(x) \leftarrow \hat{f}(x) + \lambda \hat{f}^b(x)$$

- Update the residuals,

$$r_i \leftarrow r_i - \lambda \hat{f}^b(x_i)$$

(3) Finally, the general boosting process is the sum of sequential trees (B), which predicts y given x,

$$f_B(x) = \sum_{b=1}^{N} \lambda \hat{f}^b(x)$$

GBM also possesses some limitations, especially with respect to overfitting. This can be exacerbated when the number of boosting iterations and the depth of the tree are large. When this happens, the model accuracy may be artificially inflated. One possible way to solve this problem of overfitting is to impose regularization (e.g., constraints on boosting iterations, tree depth, and other hyperparameters). In this study, the optimal model specification is estimated that minimizes MSE. This estimation process (i.e., tuning process) is conducted by testing a wide range of regularization terms and the optimal hyperparameters are chosen by using a 5-fold cross-validation technique (*9*–*11*).

Interpretation of GBM

As mentioned above, the modeling performance (i.e., prediction accuracy) of GBM tends to be high thanks to machine-based repetitive algorithms. Nonetheless, a single measurement such as prediction accuracy offers an incomplete description to address discrete choice problems (*44*). In particular, it is imperative to know the inner processes of a model to explain users' attitudes and behaviors. In this case, predictive modeling also requires an understanding of why a user makes a certain decision and how this decision is sensitive to the changes in determinant factors.

Variable Importance (VI) for Gradient Boosting Algorithm. In predictive modeling, the influence of predictor variables (i.e., independent variables) are different from one another. In general, a few variables have substantial impacts on the response. To gain useful information about the relative importance of each predictors, the reduction in error is measured. This is usually referred to as the "variable importance" (VI) (a.k.a., feature importance) (*11*, *13*). This metric serves to enhance the interpretability of ML. VI indicates which variables are the most significant in the decision tree's predictive hierarchy of decision rules. Specifically, it compares all predictors in the data and ranks the variables that contribute to the reduction in overall variances [or sum of squared errors (SSE)] (*5*, *12*). As the MSE criteria are applied to generate the tree, the reduction in the SSE is aggregated for each independent variable during the estimation process (*7*). The sum of squared residuals (SSR), which is the numerator of a weighted SSE in statistics, is defined as:

$$\mathrm{SSR} = \sum\nolimits_{c \in \mathrm{leaves}(T)} \sum\nolimits_{i \in c} (y_i - m_c)^2 \tag{4}$$

The change of SSR between the variables indicates the VI x_j at a certain node and it can be measured as follows:

$$\mathrm{VI}_d(x_j) = \Delta_d = \mathrm{SSR}_d - \sum\nolimits_i \mathrm{SSR}_i^d \tag{5}$$

where d represents a node, i is the child of node d, SSR_d is a terminal node of node d (leaf node), and SSR_i^d is an internal node (i.e., the split). For the entire tree, the VI score for each variable is calculated as the mean importance over all nodes in a tree (*12*, *13*). It can be expressed as follows where D is the total number of nodes:

$$\mathrm{VI}(x_j) = \frac{\sum_{d=1}^{D} \mathrm{VI}_d(x_j)}{\mathrm{n\_nodes}} \tag{6}$$

In general, the VI scores are applied with the standardized metric values (ranging from 0 to 1) (*12*). For instance, if a variable has no "contribution," the VI score becomes zero (i.e., $\mathrm{VI}_d(x_j) = 0$).

VI in this article is used to detect the variables that are most influential in causing an individual's final vehicle type choice. This interpretation is straightforward and gives useful insights into policy, planning, business, and regulatory factors that may have the most interplay with developing AV markets and infrastructure. GBM has rule-based and additive modeling features allowing for robust estimation of determinant factors affecting choice behaviors. Accordingly, it implies that variables with higher VI scores play significant roles in making a decision among alternatives.

Partial Dependence (PD). In addition to VI, the marginal effects of predictors (e.g., sensitivity) are further interpreted. These explain the relationships between the predictors and the response. Specifically, partial dependence (PD) provides the marginal effect of selective features on the predicted response of a learned model (*11*). Thus, PD plots are computed to provide a causal interpretation by measuring the marginal average response in relation to different values of the predictor, which can be valuable interpretation in analyzing future scenarios.

The way to measure the PD of a predictor is as follows:

$$\hat{f}_{x_s}(x_s) = E_{x_s}\left[\hat{f}(x_s, x_c)\right] \tag{7}$$

where x_s is the chosen predictor, x_c is its complete set containing the other predictors, and both x_s and x_c are used to learn the model (i.e., GBM), $\hat{f}$. The PD of x_s on y can be calculated by marginalizing the predicted values over the other predictors:

$$\hat{f}_{x_s}(x_s) = \int^{\hat{f}} (x_s, x_c)\mathrm{dP}(x_c) \tag{8}$$

For different numbers of cross-validation datasets (see details in the section for model specification), the PD is simulated for a given number of observations (a.k.a., instances) using a traditional Monte Carlo method.

When it comes to a choice modeling, PD shows the probability of a certain alternative given different values for predictors x_s. For a K -class choice model, each class (l) can be related to the following probabilities:

$$\hat{f}_k(x_s) = \mathrm{logp}_k\left(\hat{f}_{x_s}(x_s)\right) - \frac{1}{K}\sum\nolimits_{l=1}^{K} \mathrm{logp}_l\left(\hat{f}_{x_s}(x_s)\right) \tag{9}$$

where $f_k(x_s)$ is the logarithm of a monotonically increasing function on its respective probability (*12*). Thus, PD plots of each $f_k(x_s)$ on a chosen predictor can indicate how the class depends on the factors (i.e., realization of log-odds).

Data

The data used for this study was collected via online survey from September to November 2014. The survey was given only to individuals who currently drive a car for their daily commute to work or school, and was distributed in Israel and North America.

The survey was designed to investigate individuals' likelihood of choosing their future vehicle. After removing some individuals with missing values, 721 individuals completed the survey, giving a total of 4260 observations. Respondents were asked a variety of questions about their socioeconomic status, family structure (e.g., number of children), car ownership level, and their current travel habits as well as SP questions. In particular, the survey also included a series of 30 attitudinal statements to which respondents indicated their level of agreement on a 5-point Likert scale. These statements were used to assign a latent variable score to individuals under the categories Technology Interest, Public Transit Attitude, Pro-AV Sentiments, and Environmental Concern (described in Table 1).

In relation to SP questions, survey respondents were asked to state their vehicle choice between private REG, PAV, and SAV in a series of six SP scenarios of various

Table 1. Attitudinal Variables

Variable	Variable name	General description of high-scoring individual
Pro-AV sentiments	ATT_AUTONOMOUS	Responded positively to normative statements about the importance and utility of AVs.
Environmental concern	ENVIRO	Indicated a concern about global warming, desire to adopt environmentally friendly behaviors, or both.
Technology interest	TECH	Indicated an interest in and excitement about new technologies and products.
Public transit attitude	PT	Reported frequent usage of public transportation and relative comfort choosing public transit over private vehicle transportation.

Note: AV = autonomous vehicle.

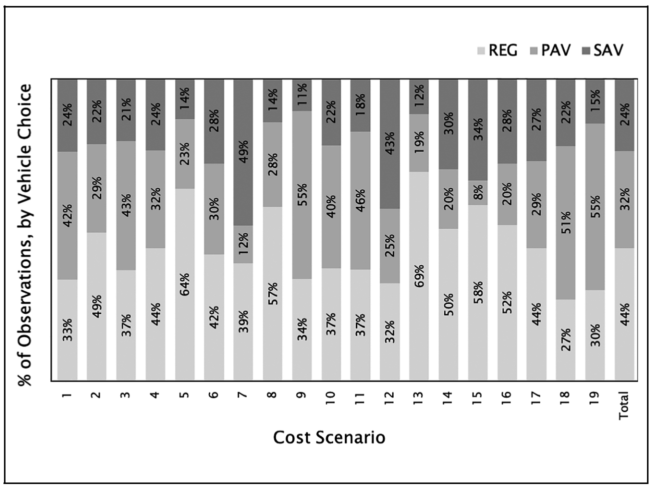

Figure 1. The percentage of observations yielding REG, PAV or SAV vehicle choices, for each cost scenario.

cost structures. Each scenario presented three cost factors—capital cost (i.e., purchase and membership cost), traveling cost, and parking cost—for each of the three vehicle types. For example, REG in a certain scenario is hypothesized by following the combination: 30,000 dollars for purchasing cost, 1.50 dollars per trip for traveling cost, and 4 dollars for parking cost. Therefore, the respondents would make different choices, when faced with the presented cost structures. To control the variation in cost scenarios provided to the respondents, a partial orthogonal balanced design was applied, and this resulted in 19 combinations of cost scenarios (see Figure 1) (*6*). Based on this finding, 6 scenarios out of 19 combinations were randomly assigned to each respondent. Thus, respondents were provided with one choice for each scenario, resulting in up to six observations per individual respondent. The target variable in each observation is simply the respondents' final choice of vehicle type: Current REG, PAV, or SAV. Overall, 44% of observations are classified as REG, 32% as PAV, and 24% as SAV, indicating that there is an acceptable stratification of target variable results (see Figure 1). Table 2 shows the descriptive statistics of each variable used in this study consisting of continuous and categorical variables.

Model Specification

Normalization

The SP survey information contains various types of variables that vary in scale and range. Most nonparametric approaches are sensitive to the scale of variables, and different scales and ranges may result in biased estimation. Therefore, all independent variables used in GBMs are standardized, except for the person and question identifiers. The min-max normalization method is applied:

$$x_i' = \frac{x_i - \min x}{\max x - \min x} \tag{7}$$

where x_i' presents the scaled value of sample i, and $\max x$ and $\min x$ indicate the maximum and minimum value of vector x.

Model Evaluation and Validation

This article aims to gain insights into users' attitudes and motivations for choosing AVs and to develop a generalized model that can be used for long-term planning and policy. Specifically, it is necessary to check and validate how well the learned model explains the data and predicts unseen instances, which is required to get reliable interpretation and minimize future risks in application.

In general, machine-based learning approaches generate a model that is discovered or constructed from the given data itself. The learned model is then used to predict a response on unseen data (i.e., test data), which evaluates the capability of the model to explain the unseen instances from the learned structure (also called generalization). To parsimoniously evaluate the model, this study uses both the hold-out method and the 5-fold

Table 2. Description of Statistics

Variables	Description	Mean	SD	Min.	Max.
Dependent choice alternatives					
CHOICE	1: REG 2: PAV 3: SAV				
Independent variable (continuous)					
DPP_REG	Purchasing cost for REG (thousand dollars)	47.39	37.98	5.00	195.00
DPP_PAV	Purchasing cost for PAV (thousand dollars)	50.31	41.76	4.00	254.00
DPP_SAV	Membership cost for SAV (thousand dollars)	1.74	2.54	0.00	7.5
DTC_REG	Trip cost for REG (in dollars)	13.13	14.08	0.30	50.00
DTC_PAV	Trip cost for PAV (in dollars)	11.75	13.38	0.17	60.00
DTC_SAV	Trip cost for SAV (in dollars)	24.53	34.19	0.00	166.67
DPARK_REG	Parking cost for REG (in dollars)	1.85	5.70	0.00	90.00
DPARK_PAV	Parking cost for PAV (in dollars)	0.76	3.23	0.00	90.00
DISTANCE	Commuting distance (in kilometers)	20.11	15.31	1.00	50.00
TIME	Commuting time (in minutes)	29.27	18.47	2.00	75.00
Independent variable (dummy and categorical)					
FT	Full-time job is 1, otherwise is 0			0	1
PT	Part-time job is 1, otherwise is 0			0	1
STUDENT	Student is 1, otherwise is 0			0	1
CARSHARING	User of car sharing is 1 otherwise is 0			0	1
TIME_OF_DEP	Time of departure (e.g., 1: before 6:00 a.m., 7: it varies from day to day)			1	7
FREQ_OF_ERRANDS	Frequency of errand or stops on the way to work or during the day (Answers ranging from daily to less than once a week)			1	4
ITEMS_LARGE	Frequency of leaving large items in the parked cars (Answers ranging from daily to never)			1	5
REGION	1: U.S., 2: Canada, 3: Israel, 4: Europe, 5: other			1	5
AGE	Age of respondent			1	4
GENDER	Gender of respondent			1	2
INCOME	Income of respondent			1	5
EDUCATION	Education level of respondent			1	5
ADULTS_HHSIZE	Number of adults in household			1	5
KIDS_HHSIZE	Number of children in household			1	7
DRIVERS_HH	Number of drivers in household			1	4
CARS_HH	Number of vehicles in household			1	5
TECH[a]	Attitude on new technologies			0	17
ENVIRO[a]	Attitude on environment			0	12
ATT_AUTONOMOUS[a]	Attitude on AVs			0	13

Note: Total number of observations = 4299. REG = regular vehicle; PAV = private autonomous vehicle; SAV = shared autonomous vehicle; AV = autonomous vehicle; SD = standard deviation; min. = minimum; max. = maximum.
[a]Latent attitudinal variables (see details in Table 1).

cross-validation (CV) method that are known to be statistically robust. The hold-out method separates the original data set into a training (70%) and a test (30%) data set. In the CV method, the training data set is also spilt into 5 partitions, and one left-out fold is used as a test data to evaluate the model that is learned from 4 out of 5 partitions. This evaluation process provides 5 trained models and their performance (i.e., mean absolute errors), from which the average and standard deviations are measured.

The modeling parameters for GBM are optimized from the CV process. The model accuracy (i.e., the number of times the correct class was predicted) and VI are also mainly estimated from the CV results. The test set is then used to identify the associations and relations between user preferences and important variables (e.g., PD).

Results

Model Accuracy

The prediction (i.e., choice) results of the GBM using 5-fold cross-validation are presented in Figure 2. As mentioned above, hyperparameters (e.g., depth of trees, number of boosting iterations) are also cross-validated to find the optimal parameters while minimizing the errors (i.e., maximizing the likelihood). The results show that the overall prediction accuracy of GBM is around 80% for the given choice alternatives: REG, PAV, or SAV. Based

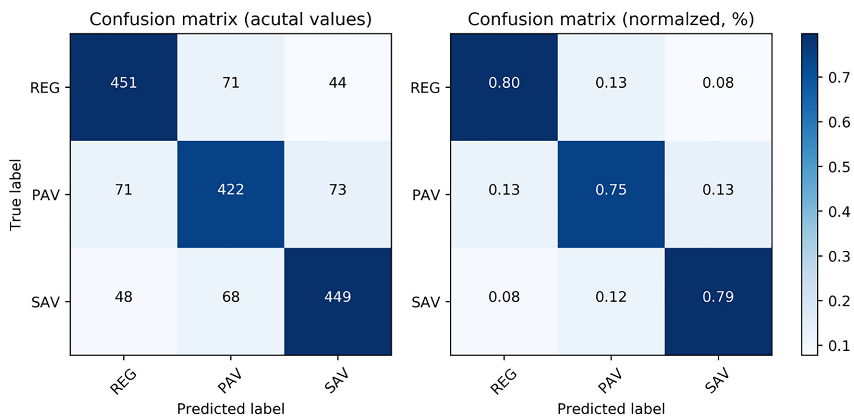

Figure 2. Accuracy of prediction in confusion matrix.

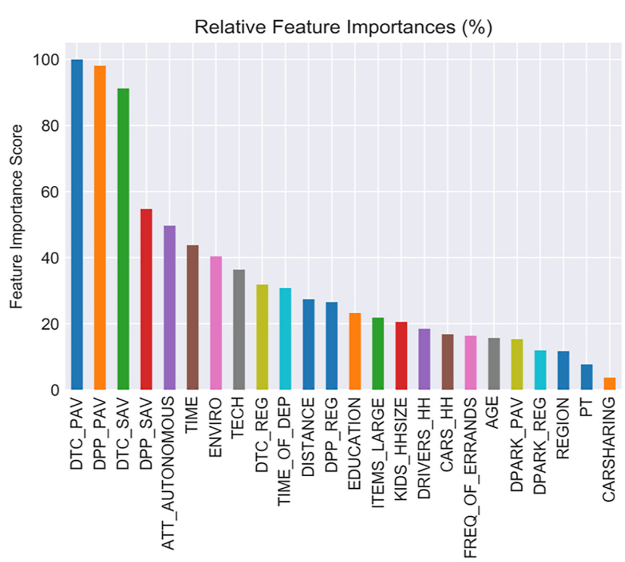

Figure 3. VI choosing three alternatives (REG, PAVs, SAVs).

on the learned and validated configurations, the following analysis focuses on interpreting users' attitudinal preferences for choosing AVs.

Interpretation of Important Variables

Figures 3 and 4 present the top 24 most important variables from a GBM trained on the survey data. Specifically, Figure 3 presents the base choice among REG, PAV, and SAV, and it is observed that DTC_PAV, DPP_PAV, DTC_SAV (AVs operating and purchasing cost in dollar value, see details in Table 2) are almost twice as important as the next important set of features. Latent attitudinal variables are prominent in the next set of important features, specifically the Pro-AV Sentiments, Environmental Concern, and Technology Interest features. Table 1 explains these latent variables further.

It is important to note that VI does not indicate the direction of correlation between variables. Rather, it is a measure of the magnitude of the reduction of variance in vehicle choice by partitioning on the given variable. Thus, for example, it is shown that the trip and purchase prices of a personal AV are more influential in driving individuals toward their final vehicle choice than any of their other attributes. The VI may also be interpreted as the variable's predictive power. For example, it is observed that the individual's Pro-AV attitude score is more than twice as likely to predict the individual's final vehicle choice, as compared with their education level. See the *Interpretation of GBM* section for further detail on interpreting VI.

Figure 4, charts *a* and *b* present feature importance rankings on two different choice groupings. In Figure 4, chart *a*, the algorithm was trained to predict whether the individual would choose their REG or AVs of either type (private or shared). This configuration provides insight into features that are most important to this binary choice level. In Figure 4, chart *b*, the data is subset to only those binary cases that yielded either PAV or SAV as the final vehicle choice. On this data, a GBM was trained to predict the decision of private or shared AV. The results indicate that AV purchase and trip cost expenses (for either PAV or SAV) are still the most important features. These binary tests show some marked differences from the three-variable predictor VI. For example, Pro-AV Sentiment, Commute Time, and Environmental Concern grow significantly in their influence over AV or non-AV decisions. And Commute Time

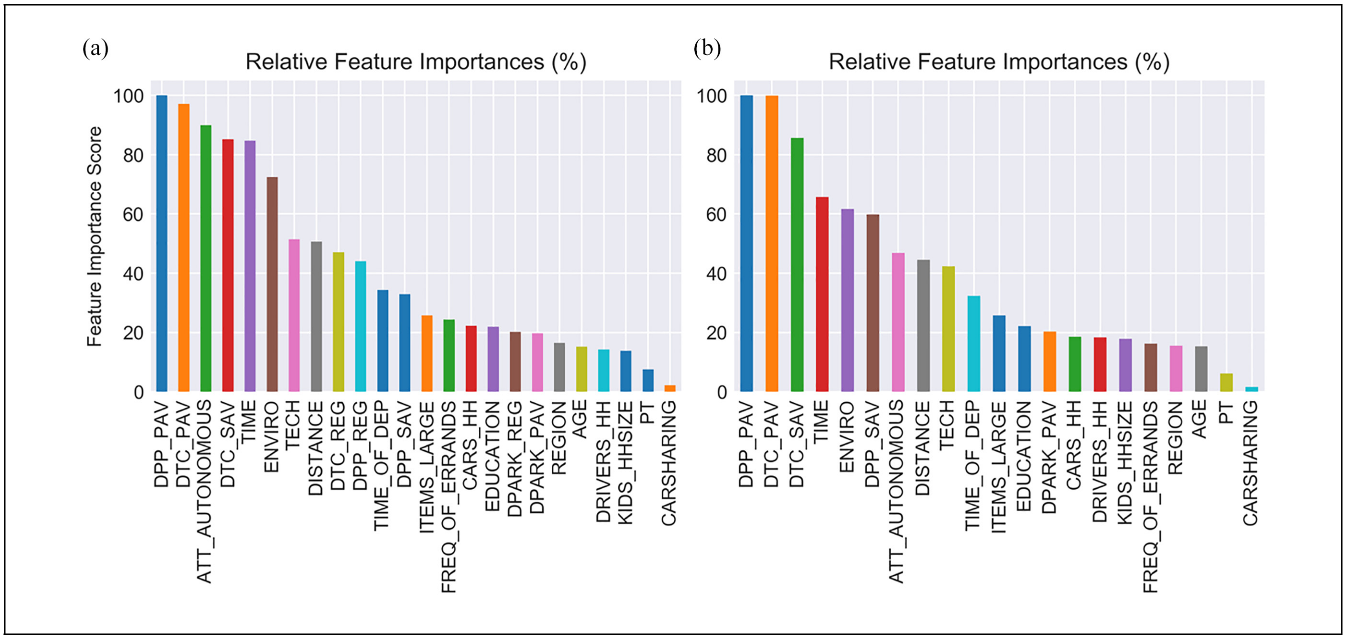

Figure 4. (*a*) VI on choice between REG and AVs and (*b*) VI on choice between PAV and SAV.

is the leading non-cost variable in determining private or shared AV choices.

Interpret Sensitivity of Vehicle-Specific Variables

The VI results suggest that important factors are more likely to affect choice behaviors than others. Specifically, the important factors include the purchasing and operating costs of vehicles and the latent attitudinal variables (Table 1). Based on this information, it is possible to further focus on interpreting users' choice behaviors on vehicle choices. Figure 5 shows the PD of the log-odds on the selected important factors in keeping one's current vehicle rather than transitioning to an AV. Specifically, the y-axis in Figure 5 indicates the log-odds of a current vehicle choice over AVs, and the x-axis correspond to the normalized value of importance factors. Markedly, the costs of AVs are positively associated with the choice of current vehicle, and a higher cost of current vehicle ownership has a negative impact on keeping choosing to keep one's current car. Moreover, qualitative latent variables such as Environmental Concern and Pro-AV Sentiments are negatively associated with keeping a current vehicle.

In addition, Figure 6 compares the PD of selected variables in determining PAV and SAV. Expectedly, someone with a positive attitude towards AVs (ATT_AUTONOMOUS) is more likely to prefer PAV to SAV. Moreover, a person with a higher level of environmental concern (ENVIRO) is also more likely to choose SAV rather than PAV. Moreover, there is also a marked abrupt decrease in the likelihood of choosing REG over AV options when the commute time (TIME) reaches a certain threshold.

Interpret Sensitivity of Household Attributes

We also interpret the sensitivity of the combination of two household attributes on future vehicle choice behaviors by using three-dimensional partial dependence (3D-PD) plots (Figure 7). Specifically, 3D-PD plots provide interrelated impacts of variables on the given choice scenarios. The top panel (Figure 7*a*) comparing REG and AV shows that respondents are likely to keep their regular vehicles when they are older and less educated. Moreover, a household with more vehicles has a tendency to choose AVs, and this tendency is more distinguishable in age groups such that elder people prefer to keep their current vehicle even if they have higher vehicle ownership. In addition, respondents in Israel (region = 0.5, see details in Table 2) show a greater preference for AVs than other regions.

Figure 7, section *b*, focuses on the choice between PAV and SAV. The results show that respondents who are older prefer PAV to SAV. Moreover, people prefer PAV when the frequency of trips is higher and when they are older. When it comes to identifying the effect of age and region, elders living in the U.S. are likely to prefer PAV to SAV. In addition, respondents have a tendency to choose PAV when the frequency of trips is higher and the departure time of work is earlier than others.

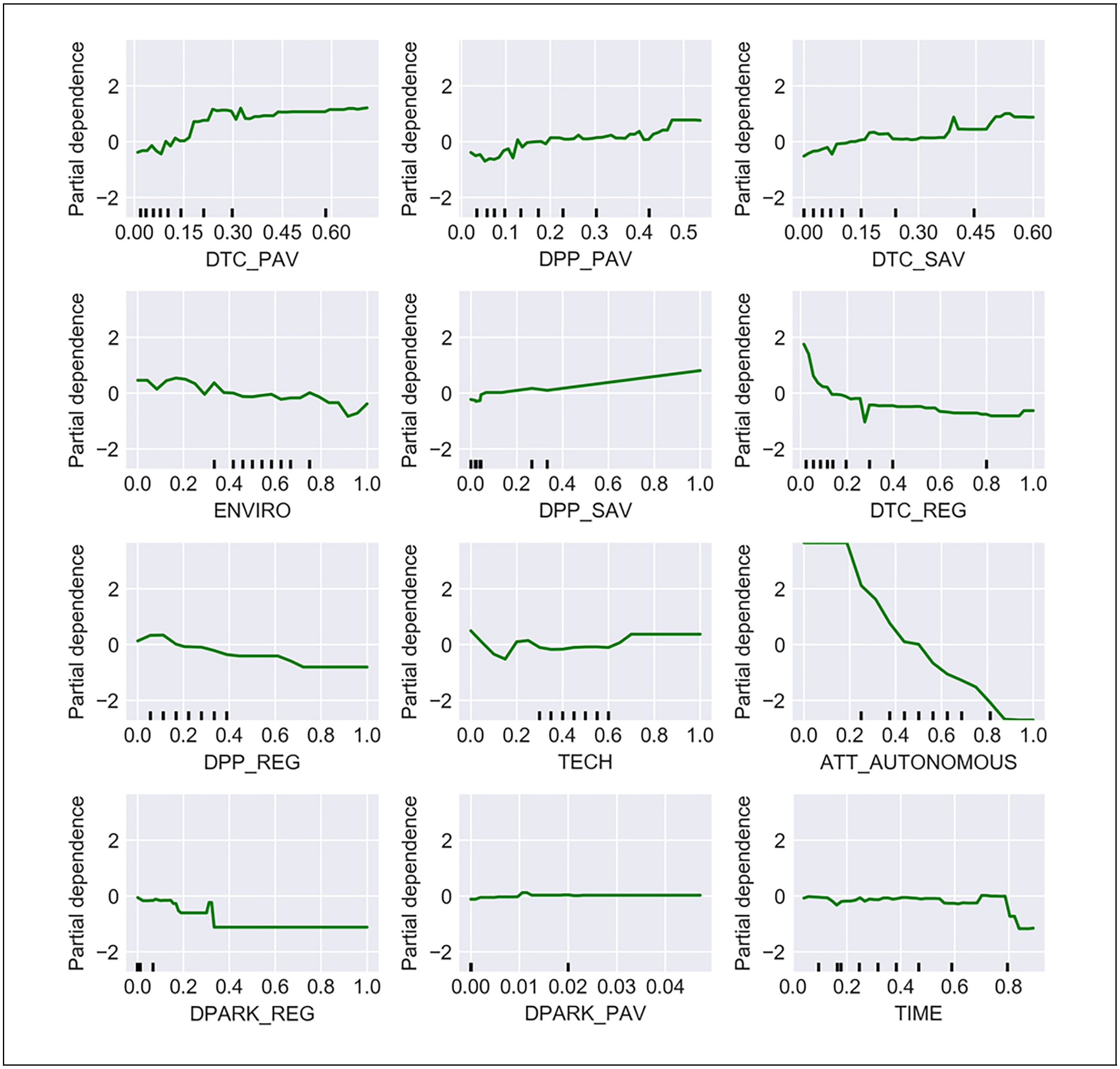

Figure 5. Partial dependencies of log-odds (positive or negative association between X and Y) in choosing current REG over PAV or SAV.

Conclusion

This article investigated AV adoption choice behaviors based on two main variables affecting users' behaviors: individual characteristics and system characteristics related to AVs. Specifically, ML techniques (i.e., GBM) were applied to investigate discrete choice behaviors, which have traditionally been considered as less "interpretable" than parametric approaches but which can achieve higher prediction performances. In particular, model-agnostic methods were used to interpret the estimated model using GBM, which can be applied to any ML algorithms. Through its findings, this article can fill the knowledge gaps in trade-off between prediction accuracy and interpretability when selecting a modeling approach for discrete choice problems.

The prediction performance of GBM is evaluated by conducting 5-fold cross-validation method and presented with a confusion matrix. The result of cross-validation showed that the overall prediction accuracy of GBM was acceptable (around 80%) to assess the impact of the

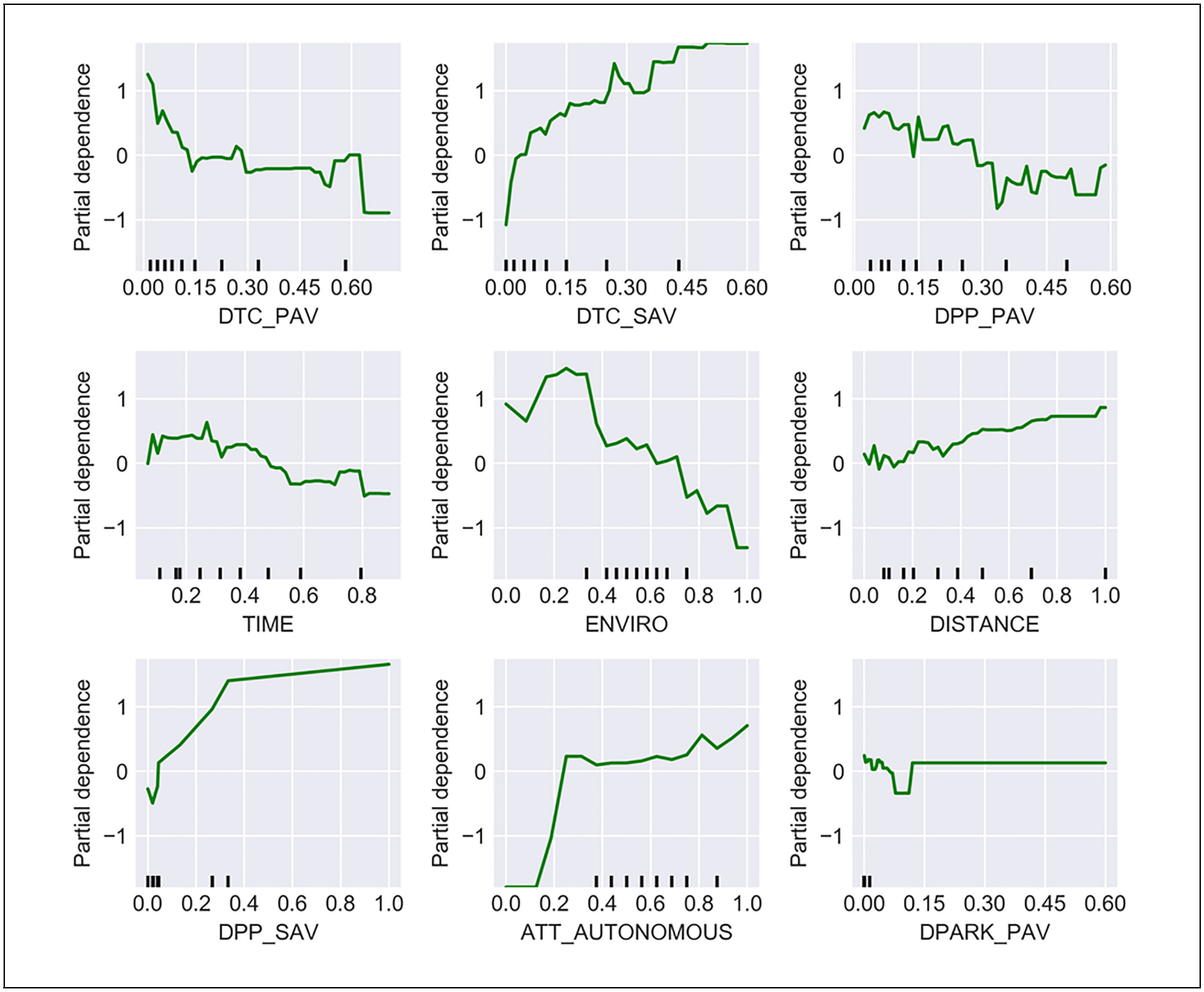

Figure 6. Partial dependencies of log-odds in choosing PAV over SAV.

variables on AVs. To interpret users' behaviors, VI and PD were measured. VI was used to identify the magnitude of variables (i.e., the reduction of variance) that have greater impacts on vehicle choice. The results of VI showed that the costs of vehicles, latent attitudinal factors, and time play significant roles in choosing a vehicle. Furthermore, the sensitivity of choice was also examined by using the PD of the log-odds on the selected important factors. The results from PD provided the expected positive or negative associations between the likelihood of alternative and the important variables.

While the most important property of statistical learning is to produce accurate predictions, the second most important may be how is it possible to further use the results. In other words, how interpretable a model is. When models adopt a more complex structure, their structures become less interpretable. This is particularly relevant to ML models that often face this issue because of their complex and unknown structures, although ML has particularly powerful prediction power and ability to capture complex patterns in data with its relatively easy applicability. In this context, this article aims to suggests the interpretability of ML that is often assimilated to a "black box." Thus, interpretable ML models have the potential to be promising statistical learning methods in addressing transportation behaviors.

For future works, the SP survey data used in this study was originally applied to mixed logit model in (*6*). This successfully identified the users' preferences on AVs. Thus, if this study compares the results and findings from GBM and the those of mixed logit model, it is possible to further validate the model. In addition, every

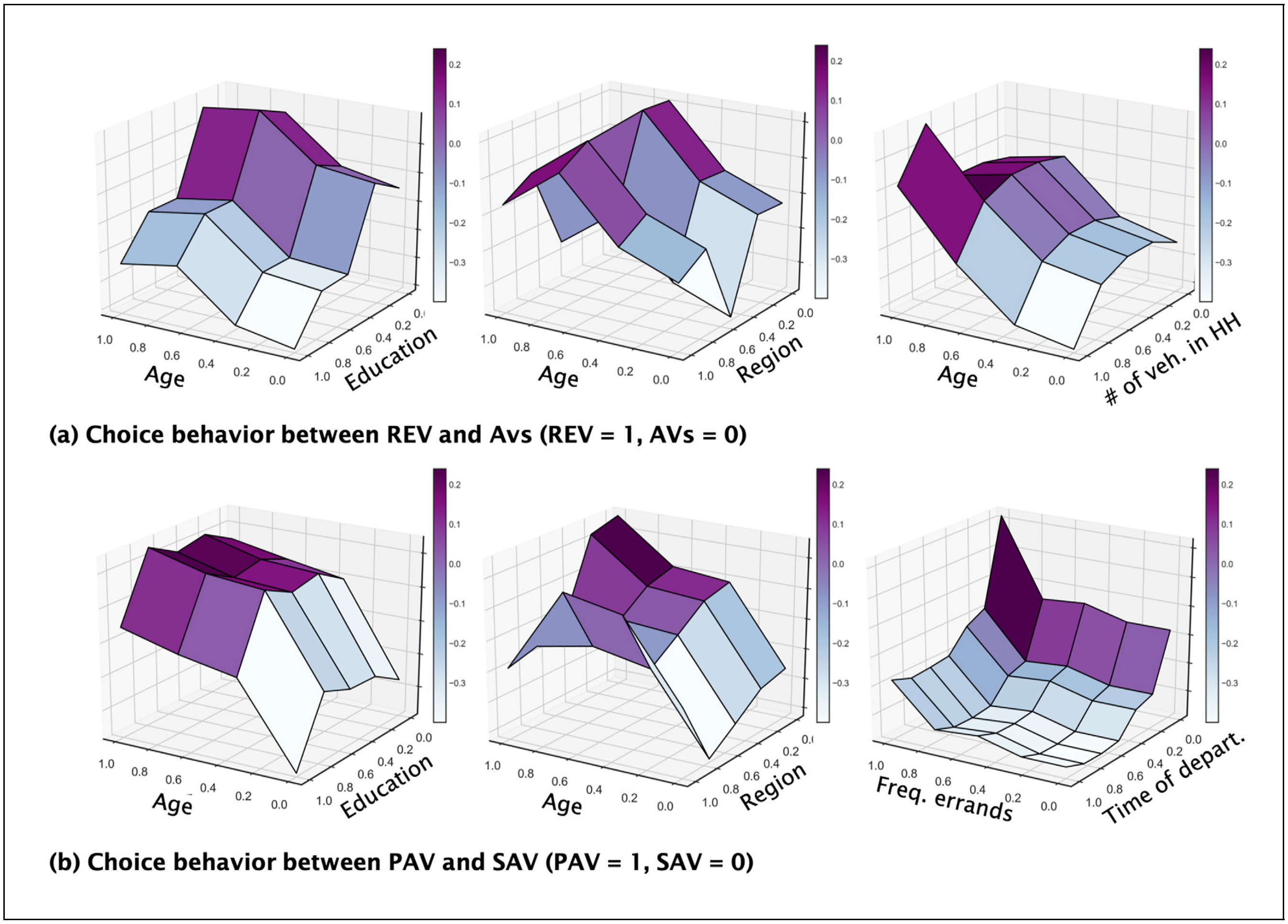

Figure 7. Compares the bivariate PD, especially for inter-relational effects of household demographic characteristics on making decisions.

model inherently includes uncertainty and it can cause biased estimation and misinterpretation of results when applying this model to the other instances. To increase the usability and generality of this model for future applications, it is important to measure the uncertainty of the prediction.

Acknowledgments

This research was supported, in part, by the National Science Foundation (NSF) CAREER Award 1551731 and the NSF CPS Award 1646395.

Author Contributions

The authors confirm contribution to the paper as follows—study conception and design: DL, SD; data collection: CJH, YS; analysis and interpretation of results: DL, JM, SD, YS; draft manuscript preparation: DL, JM, SD, YS. All authors reviewed the results and approved the final version of the manuscript.

References

1. Derrible, S. Urban Infrastructure is Not a Tree: Integrating and Decentralizing Urban Infrastructure Systems. *Environment and Planning B: Planning and Design*, Vol. 44, No. 3, 2016, pp. 553–569. https://doi.org/10.1177/0265813516647063.
2. Derrible, S. Complexity in Future Cities: The Rise of Networked Infrastructure. *International Journal of Urban Sciences*, Vol. 21, No. 1S, 2016, pp. 68–86. https://doi.org/10.1080/12265934.2016.1233075.
3. Fagnant, D. J., and K. Kockelman. Preparing a Nation for Autonomous Vehicles: Opportunities, Barriers and Policy Recommendations. *Transportation Research Part A: Policy and Practice*, Vol. 77, 2015, pp. 167–181. https://doi.org/10.1016/j.tra.2015.04.003.
4. Litman, T. *Autonomous Vehicle Implementation Predictions*. Victoria Transport Policy Institute Victoria, Canada, 2017.
5. Anderson, J. M., K. Nidhi, K. D. Stanley, P. Sorensen, C. Samaras, and O. A. Oluwatola. *Autonomous Vehicle Technology: A Guide for Policymakers*. Rand Corporation, Santa Monica, Calif., 2014.

6. Haboucha, C. J., R. Ishaq, and Y. Shiftan. User Preferences Regarding Autonomous Vehicles. *Transportation Research Part C: Emerging Technologies*, Vol. 78, 2017, pp. 37–49. https://doi.org/10.1016/j.trc.2017.01.010.
7. Golshani, N., R. Shabanpour, S. M. Mahmoudifard, S. Derrible, and A. Mohammadian. Modeling Travel Mode and Timing Decisions: Comparison of Artificial Neural Networks and Copula-Based Joint Model. *Travel Behaviour and Society*, Vol. 10, 2018, pp. 21–32. https://doi.org/10.1016/j.tbs.2017.09.003.
8. Robert, C. *Machine Learning, a Probabilistic Perspective*. Cambridge, MA, MIT Press 2014.
9. Lee, D., S. Derrible, and F. C. Pereira. Comparison of Four Types of Artificial Neural Network and a Multinomial Logit Model for Travel Mode Choice Modeling. *Transportation Research Record: Journal of the Transportation Research Board*, 2018. 2672: 101–112.
10. Wang, F., and C. L. Ross. Machine Learning Travel Mode Choices: Comparing the Performance of an Extreme Gradient Boosting Model with a Multinomial Logit Model. *Transportation Research Record: Journal of the Transportation Research Board*, 2018. 2672: 35–45.
11. Friedman, J., T. Hastie, and R. Tibshirani. *The Elements of Statistical Learning*. Springer Series in Statistics New York, 2001.
12. Zhou, Z. H. *Ensemble Methods: Foundations and Algorithms*. Boca Raton, FL, Chapman and Hall/CRC, 2012.
13. Friedman, J. H. Greedy Function Approximation: A Gradient Boosting Machine. *Annals of Statistics*, Vol. 29, No. 5, 2001, pp. 1189–1232.
14. Derrible, S., and N. Ahmad. Network-Based and Binless Frequency Analyses. *PLoS One*, Vol. 10, No. 11, 2015, p. e0142108.
15. Akbarzadeh, M., S. Memarmontazerin, S. Derrible, and S. F. Salehi Reihani. The Role of Travel Demand and Network Centrality on the Connectivity and Resilience of an Urban Street System. *Transportation*, 2017, pp. 1–15. https://doi.org/10.1007/s11116-017-9814-y.
16. Wisetjindawat, W., S. Derrible, and A. Kermanshah. Modeling the Effectiveness of Infrastructure and Travel Demand Management Measures to Improve Traffic Congestion During Typhoons. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. in press, 2018.
17. Ahmad, N., S. Derrible, T. Eason, and H. Cabezas. Using Fisher Information to Track Stability in Multivariate Systems. *Royal Society Open Science*, Vol. 3, No. 11, 2016, p. 160582. https://doi.org/10.1098/rsos.160582.
18. Wisetjindawat, W., A. Kermanshah, S. Derrible, and M. Fujita. Stochastic Modeling of Road System Performance during Multihazard Events: Flash Floods and Earthquakes. *Journal of Infrastructure Systems*, Vol. 23, No. 4, 2017, p. 04017031. https://doi.org/10.1061/(ASCE)IS.1943-555X.0000391.
19. Ahmad, N., S. Derrible, and H. Cabezas. Using Fisher Information to Assess Stability in the Performance of Public Transportation Systems. *Royal Society Open Science*, Vol. 4, No. 4, 2017, p. 160920. https://doi.org/10.1098/rsos.160920.
20. Golshani, N., R. Shabanpour, S. M. Mahmoudifard, S. Derrible, and A. Mohammadian. Modeling Travel Mode and Timing Decisions: Comparison of Artificial Neural Networks and Copula-Based Joint Model. *Travel Behaviour and Society*, Vol. 10, 2018, pp. 21–32.
21. Ghasri, M., T. Hossein Rashidi, and S. T. Waller. Developing a Disaggregate Travel Demand System of Models Using Data Mining Techniques. *Transportation Research Part A: Policy and Practice*, Vol. 105, 2017, pp. 138–153. https://doi.org/10.1016/j.tra.2017.08.020.
22. Nasrabadi, N. M. Pattern Recognition and Machine Learning. *Journal of Electronic Imaging*, Vol. 16, No. 4, 2007, p. 049901. https://doi.org/10.1117/1.2819119.
23. Lu, Y., and K. Kawamura. Data-Mining Approach to Work Trip Mode Choice Analysis in Chicago, Illinois, Area. *Transportation Research Record: Journal of the Transportation Research Board*, 2010. 2156: 73–80.
24. Paredes, M., E. Hemberg, U. M. O'Reilly, and C. Zegras. Machine Learning or Discrete Choice Models for Car Ownership Demand Estimation and Prediction? *Proc., 2017 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*, Naples, Italy, 2017.
25. Hagenauer, J., and M. Helbich. A Comparative Study of Machine Learning Classifiers for Modeling Travel Mode Choice. *Expert Systems with Applications*, Vol. 78, 2017, pp. 273–282. https://doi.org/10.1016/j.eswa.2017.01.057.
26. Rasouli, S., and H. J. P. Timmermans. Using Ensembles of Decision Trees to Predict Transport Mode Choice Decisions: Effects on Predictive Success and Uncertainty Estimates. *European Journal of Transport and Infrastructure Research*, Vol. 14, No. 4, 2014, pp. 412–424.
27. Hasegawa, H., T. Naito, M. Arimura, and T. Tamura. Modal Choice Analysis Using Ensemble Learning Methods. *Journal of Japan Society of Civil Engineering*, Vol. 68, No. 5, 2012, pp. 773–780.
28. Semanjski, I., and S. Gautama. Smart City Mobility Application—Gradient Boosting Trees for Mobility Prediction and Analysis Based on Crowdsourced Data. *Sensors*, Vol. 15, No. 7, 2015, pp. 15974–15987.
29. Zhang, Y., and A. Haghani. A Gradient Boosting Method to Improve Travel Time Prediction. *Transportation Research Part C: Emerging Technologies*, Vol. 58, 2015, pp. 308–324.
30. Lee, Y. J., and O. Min. Comparative Analysis of Machine Learning Algorithms to Urban Traffic Prediction. *Proc., 2017 International Conference on Information and Communication Technology Convergence (ICTC)*, Jeju, South Korea, 2017.
31. Mehmood, R., R. Meriton, G. Graham, P. Hennelly, and M. Kumar. Exploring the Influence of Big Data on City Transport Operations: A Markovian Approach. *International Journal of Operations & Production Management*, Vol. 37, No. 1, 2017, pp. 75–104. https://doi.org/10.1108/IJOPM-03-2015-0179.
32. Bansal, P., K. M. Kockelman, and A. Singh. Assessing Public Opinions of and Interest in New Vehicle Technologies: An Austin Perspective. *Transportation Research*

Part C: Emerging Technologies, Vol. 67, 2016, pp. 1–14. https://doi.org/10.1016/j.trc.2016.01.019.
33. Bansal, P., and K. M. Kockelman. Are We Ready to Embrace Connected and Self-Driving Vehicles? A Case Study of Texans. *Transportation*, Vol. 45, No. 2, 2018, pp. 641–675. https://doi.org/10.1007/s11116-016-9745-z.
34. Choi, J. K., and Y. G. Ji. Investigating the Importance of Trust on Adopting an Autonomous Vehicle. *International Journal of Human-Computer Interaction*, Vol. 31, No. 10, 2015, pp. 692–702. https://doi.org/10.1080/10447318.2015.1070549.
35. Xie, C., J. Lu, and E. Parkany. Work Travel Mode Choice Modeling with Data Mining: Decision Trees and Neural Networks. *Transportation Research Record: Journal of the Transportation Research Board*, 2003. 1854: 50–61.
36. Yu, Z., F. Haghighat, B. C. M. Fung, and H. Yoshino. A Decision Tree Method for Building Energy Demand Modeling. *Energy and Buildings*, Vol. 42, No. 10, 2010, pp. 1637–1646. https://doi.org/10.1016/j.enbuild.2010.04.006.
37. Tso, G. K. F., and K. K. W. Yau. Predicting Electricity Energy Consumption: A Comparison of Regression Analysis, Decision Tree and Neural Networks. *Energy*, Vol. 32, No. 9, 2007, pp. 1761–1768. https://doi.org/10.1016/j.energy.2006.11.010.
38. James, G., D. Witten, T. Hastie, and R. Tibshirani. *An Introduction to Statistical Learning*. New York, NY, Springer, 2013.
39. Kuhn, M., and K. Johnson. *Applied Predictive Modeling*. New York, NY, Springer, 2013.
40. Molnar, C. Interpretable Machine Learning: A Guide for Making Black Box Models Explainable. 2019.
41. Poursabzi-Sangdeh, F., D. G. Goldstein, J. M. Hofman, J. W. Vaughan, and H. Wallach. Manipulating and Measuring Model Interpretability. *arXiv preprint arXiv:1802.07810*, 2018.
42. Ribeiro, M. T., S. Singh, and C. Guestrin. Model-Agnostic Interpretability of Machine Learning. *arXiv preprint arXiv:1606.05386*, 2016.
43. Natekin, A., and A. Knoll. Gradient Boosting Machines, a Tutorial. *Frontiers in Neurorobotics*, Vol. 7, 2013. https://doi.org/10.3389/fnbot.2013.00021.
44. Doshi-Velez, F., and B. Kim. Towards a Rigorous Science of Interpretable Machine Learning. *arXiv preprint arXiv:1702.08608*, 2017.

The Standing Committee on Transportation Demand Forecasting (ADB40) peer-reviewed this paper (19-02893).