Joint Discrete-Continuous Model of Travel Mode and Departure Time Choices

Ramin Shabanpour, Nima Golshani, Sybil Derrible, Abolfazl (Kouros) Mohammadian, and Mohammad Miralinaghi

This paper presents a cluster-based joint modeling approach to investigating heterogeneous travelers' behavior toward trip mode and departure time choices by considering those choices as a joint decision. First, a two-step clustering algorithm was applied to classify travelers into six distinct clusters to account for the heterogeneity in their decision-making behavior. Then, a joint discrete-continuous model was proposed for each cluster, in which the travel mode and departure time were estimated by a multinomial logit and a log-linear regression model, respectively. These two models were jointly estimated with a copula approach. For an investigation of the performance of the proposed approach, its results were compared with an aggregate joint model on all nonclustered observations to assess the potential benefits of population clustering. The goodness-of-fit measures and prediction accuracy results demonstrated that the proposed cluster-based joint model significantly outperformed the aggregate joint model. Further, the variations in the estimated parameters of different clusters indicated significant behavioral differences across clusters. Hence, the proposed cluster-based joint model, while offering higher accuracy, possesses a significant potential for transportation policy making because it has the capability to target different types of travelers on the basis of their decision-making behavior.

Travel mode and departure time choices are two key components of travel behavior that directly influence the spatial and temporal distribution of travel demand in a transportation system (*1*). Recognizing the key influencing factors in travelers' decision behavior, and especially in travel mode and departure time choices, is essential to devise effective transportation demand management policies (*2, 3*). These travel choices are closely intertwined, because although the choice of time of day substantially affects the attributes of travel modes (e.g., availability of travel mode, travel time, travel cost), the choice of departure time is also influenced by the expected travel time for each mode of travel. The first issue in modeling these decisions is the shared factors affecting them and the causal effects that they have on each other (*4, 5*). Hence, it is necessary to investigate these two travel decisions jointly to capture the unrestricted correlation between their unobserved influencing factors. Numerous studies have been conducted to consider multiple travel dimensions in a joint model structure (*6–8*).

Another critical issue in modeling these interrelated decisions is the heterogeneity of travelers in terms of decision-making criteria related to the various attributes of their available choices. Several studies show that heterogeneous decision makers naturally respond differently to travel demand management policies, which greatly affects their effectiveness (*9–11*). Therefore, capturing the heterogeneity of decision makers offers substantial benefits in developing a model that is both reliable and policy-sensitive in practice.

This study undertook to bring together these two lines of research by coupling clustering analysis approach with joint trip departure time and mode choice modeling. To this end, first, a clustering analysis using a two-step clustering algorithm was conducted to assign individuals to certain clusters, in which members of each cluster are relatively homogeneous in terms of their lifestyle specifications and decision behavior. Then, within each cluster a joint discrete-continuous model was proposed to investigate the joint travel decision behavior of departure time and mode choices. Travelers' mode choices and continuous departure time choices were estimated by a multinomial logit and a log-linear regression model, respectively. These two models were jointly estimated using a copula approach.

The copula-based approach introduced by Bhat and Eluru facilitates model estimation without imposing restrictive distribution assumptions on the dependency structures between the errors in the discrete and continuous model components (*12*). In this line of research, Eluru et al. used a copula-based approach to simultaneously estimate vehicle ownership, residential location, and vehicle miles traveled while capturing the correlation between the error terms (*13*). Born et al. estimated a copula-based joint continuous-discrete model of activity and accompanying type and activity duration (*14*). Results of these studies show significant improvement of copula-based models over the other models, which ignore the interrelated distribution of the variables. Other examples of applications of copula-based models in travel demand modeling include, but are not limited to, work by Karimi et al. (*15*), Sener et al. (*16*), Sener and Reeder (*17*), Rashidi and Mohammadian (*18*), and Golshani et al. (*19*). A rich set of introduced copula classes, including the Gaussian, Clayton, Gumbel, Frank, and Joe copulas, allows researchers to capture the most appropriate dependency structure between random variables (*15*).

The proposed model offers a powerful tool to assess the impact of transportation demand management policies on various population segments and presents substantial benefits for planning agencies in practice. Specifically, the proposed modeling approach is capable of forming the core of the trip planning module in the ADAPTS activity-based travel demand model (*20–25*), and so it can replace the current submodels with independent mode and departure time

R. Shabanpour, N. Golshani, S. Derrible, and A. Mohammadian, Department of Civil and Materials Engineering, College of Engineering, University of Illinois at Chicago, 842 West Taylor Street, Chicago, IL 60607-7023. M. Miralinaghi, Lyles School of Civil Engineering, Purdue University, West Lafayette, IN 47907. Corresponding Author: R. Shabanpour, rshaba4@uic.edu.

Transportation Research Record: Journal of the Transportation Research Board, No. 2669, 2017, pp. 41–51.
http://dx.doi.org/10.3141/2669-05

choices. Further, because trip purpose has an undeniable effect on travel mode and departure time decisions, the focus was on home-based nonmandatory trips in this study. Several studies have focused on mandatory work trips because of their direct impact on peak period congestion.

For instance, Nural Habib et al. estimated a joint model of mode choice and trip timing for commuting trips in the greater Toronto area (*26*). Nural Habib later added activity duration to the joint framework by incorporating it as an endogenous variable and concluded that the results were more accurate when the work duration was considered in the model (*8*). Paleti et al. presented a joint model of mode choice and time of day for work trips using both observed and stated preference data (*27*). They argued that using only observed data would result in less model sensitivity to travel demand policies. On the other hand, Kumar and Levinson stated that advances in telecommunication should enable more people to work at home, which means more flexibility for nonmandatory activities during the day (*28*). In addition, although the possibility of shifting the time of a trip or changing the travel mode in response to transportation demand management plans is much higher for nonmandatory trips (*1, 2*), limited attention has been devoted to modeling the mode and timing decisions for these trips.

The performance of the proposed cluster-based joint model was compared with an aggregate joint model on all nonclustered observations to assess the potential benefits of population clustering. The goodness-of-fit measures and prediction accuracy results demonstrated that the proposed cluster-based joint model significantly outperformed the aggregate joint model. Furthermore, the variations in the estimated parameters for the different clusters indicated significant behavioral differences across clusters. Among the copula functions explored in this study [i.e., the Frank, Gumbel, Clayton, and Joe copulas discussed by Bhat and Eluru (*12*)] to estimate the dependent structure of decisions, the model with Frank copulas provided the best statistical fit.

The next section describes the data preparation process. This is followed by a description of the modeling approach, followed by detailed model estimation results and behavioral implications. The study concludes with a summary of the major findings and recommendations for future studies.

DATA PREPARATION

The main data source used in this study was the Travel Tracker Survey that was conducted between January 2007 and February 2008 by the Chicago Metropolitan Agency for Planning. The data set was collected by surveying approximately 10,500 households, in which each member was asked to fill in a detailed travel diary for one or two assigned dates. The data set contains more than 210,000 trip observations and detailed information including trip purpose, origin and destination, mode, and departure time, along with demographic information for the travelers.

The survey data set was first analyzed and invalid records were eliminated. Home-based nonmandatory trips were formed and linked to the sociodemographic information of travelers. Only home-based trips, with the trip origin at home, were examined in this study, because the modes of non-home-based trips are highly associated with the mode of the home-based trip within the same tour. As for trip purposes, the nonmandatory trips were considered in this study and categorized into the following groups:

- Personal trips, such as religious, health care, and civic activities;
- Discretionary trips, such as dining out, visiting friends, and entertainment; and
- Shopping trips, such as grocery shopping.

There are in total about 31,000 nonmandatory trips in the Chicago Metropolitan Agency for Planning data set. A sample of 11,000 trips of 1,642 individuals that followed almost the same age distribution as that of the whole data set was selected for the purpose of this study.

This study employed the choice set generation methodology proposed by Javanmardi et al. to generate the personalized attributes (e.g., travel time, travel cost) of alternatives that are available for travelers at their departure times but not chosen (*29*). They developed software to query the attributes of trips from the Google Maps application programming interface and the Chicago Regional Transportation Authority trip planner on the same day of the week and at the exact time of day that the original trips were made. This information includes exact point-to-point travel times, available modes, disaggregate access and egress distances, the nearest available transit stations or stops, and transfer information. Relevant resource constraints, such as vehicle availability in the household or transit availability at the time of the observed trips, were also investigated to find out what modes were available to travelers. In addition, multiple variables, such as population density, transit density, and road density of travel analysis zones, were generated to be used as proxies for land use and built environment factors. Summary statistics of the key variables used in this study are presented in Table 1.

MODEL SPECIFICATION

This research aimed to develop a joint model of travel mode and departure time choices that takes into account the heterogeneity of travelers' decision-making behavior. To this end, individuals were assigned to clusters using a cluster analysis technique in such a way that members of each cluster are relatively homogeneous in terms of their lifestyle specifications. A joint copula-based discrete-continuous model was then estimated to investigate the interrelated decision mechanisms of these two travel dimensions within each cluster. This paper first presents the preparatory modeling steps, including principal component analysis (PCA) and cluster analysis; then the joint modeling approach applied in this study is detailed.

Principal Component Analysis

A descriptive analysis of the data revealed that several of the variables introduced above were highly correlated. For example, income level and education, or number of vehicles and number of workers in a household, are highly correlated. This possible multicollinearity between variables might cause misapprehension in identifying the influence of explanatory variables on travel behavior. To cope with the multicollinearity issue, this study applies the PCA technique, which reformulates a set of observed variables into a new set (usually fewer in number) of independent variables (*10*). Indeed, PCA attempts to determine the components of the data that reduce the dimensions of variations and may be given a possible meaning (*30*).

A total of 27 explanatory variables of individuals' and households' demographic and built-environment characteristics were chosen

TABLE 1 Description of Variables and Summary Statistics

Variable	Description	Mean	SD
Population_density	Population density of home TAZ (population per unit area)	0.01	0.01
Road_density	Road density of home TAZ (roads lengths per unit area)	0.16	0.11
Housing_density	Housing density of home TAZ (number of houses per unit area)	0.005	0.01
Transit_density	Transit density of home TAZ (number of stops per unit area) ($\times 10^6$)	1.68	3.58
CBD	1 if traveler lives in CBD; 0 otherwise	0.12	0.32
Walk_TT	Travel time for walk mode (h)	2.45	2.93
Bike_TT	Travel time for bike mode (h)	0.85	0.65
Drive_TT	Travel time for auto drive mode (h)	0.27	0.38
Transit_TT	Travel time for transit mode (h)	0.44	0.32
Drive_cost	Travel cost for auto drive mode ($)	1.16	1.39
Transit_cost	Travel cost for transit mode ($)	1.79	1.60
Walk_accessible	1 if walking distance to destination is less than 0.25 mi	0.09	0.29
Transit_egress	Egress distance to destination for transit mode (km)	0.81	1.12
Transit_access	Access distance from origin for transit mode (km)	1.26	2.09
Activity_dur	Duration of activity at trip destination (min)	210.11	216.93
Weekend	1 if the trip is made during weekend; 0 otherwise	0.11	0.31
Low_income	1 if household income is less than $50,000; 0 otherwise	0.39	0.49
Med_income	1 if income is between $50,000 and $100,000; 0 otherwise	0.46	0.50
High_income	1 if income is more than $100,000; 0 otherwise	0.15	0.36
HH_bikes	Number of bikes in household	1.38	1.67
HH_license	Number of licensed drivers in household	1.88	0.86
HH_size	Household size	2.69	1.36
HH_worker	Number of workers in the household	1.53	0.92
HH_student	Number of students in the household	0.76	1.08
HH_vehicle	Number of vehicles in the household	1.82	1.04
Part_work	1 if traveler works part time; 0 otherwise	0.14	0.34
Full_work	1 if traveler works full time; 0 otherwise	0.53	0.51
Age –16	1 if traveler's age is less than 16; 0 otherwise	0.06	0.22
Age_17–30	1 if traveler's age is between 17 and 30; 0 otherwise	0.12	0.29
Age_31–50	1 if traveler's age is between 31 and 50; 0 otherwise	0.38	0.46
Age_51–65	1 if traveler's age is between 51 and 65; 0 otherwise	0.29	0.44
Age +65	1 if traveler's age is greater than 65; 0 otherwise	0.15	0.36
Age –20	1 if traveler's age is less than 20; 0 otherwise	0.08	0.28
Age_40–65	1 if traveler's age is between 40 and 65; 0 otherwise	0.67	0.47
White_ethnicity	1 if traveler is of white origin; 0 otherwise	0.41	0.49
Black_ethnicity	1 if traveler is of black origin; 0 otherwise	0.36	0.48
Hisp/other_ethnicity	1 if traveler is of Hispanic or other origins; 0 otherwise	0.23	0.42
No_high_degree	1 if traveler is not a high school graduate; 0 otherwise	0.09	0.29
High_degree	1 if traveler has high school degree; 0 otherwise	0.16	0.37
College_degree	1 if traveler has college degree; 0 otherwise	0.21	0.41
Bachelor_degree	1 if traveler has a bachelor's degree; 0 otherwise	0.28	0.44
Grad_degree	1 if traveler has a graduate degree; 0 otherwise	0.26	0.44

NOTE: TAZ = travel analysis zone; CBD = central business district; TT = travel time; HH = household.

for the PCA. Models with five to 15 components were tested with different combinations of explanatory variables, and finally eight components were selected that can explain 76% of the variance in the data set. The transformation was conducted with the varimax method, a type of orthogonal rotation, with Kaiser normalization to derive uncorrelated factors. All of the eigenvalues for the calculated factors were greater than one. The factor loadings of each of the explanatory variables onto each of the factors are summarized in Table 2.

These eight factors, which provide an initial understanding of the interdependencies between each of the variables, can be identified as follows:

1. Household general information. This factor represents a household's six dominant variables, including the number of vehicles, workers, licensed drivers, bicycles, and members of the household. These variables are highly correlated in the data set.
2. Highly dense area. This factor represents dense urban areas with high population, housing, road, and transit densities. It defines dimensions relating to land use.
3. Lower income and education. This factor consists of variables representing individuals with low income (<$50,000) who are mostly uneducated or have a high school diploma or a college degree.
4. Higher income and education. This component is affected by three dominant variables that are representing individuals with medium or high income (>$50,000) who mostly have a graduate degree.
5. Middle-aged family. This factor refers to families comprising young to middle-aged parents (31 to 65 years old) with children.
6. Youngers. This factor represents young singles or couples with a bachelor's degree or degrees.
7. Black and Hispanic. The variables that indicate the minority groups are dominant parameters in this factor.
8. Seniors. This component defines the characteristics of senior persons (more than 65 years old).

Cluster Analysis

Many studies show that travelers' lifestyles, attitudes, and perceptions significantly affect their travel behavior, and heterogeneous travelers may response differently to transport policies [such studies include, e.g., Anable (*31*), McCarthy et al. (*32*), and Mahmoudifard et al. (*33*)].

TABLE 2 Principal Component Analysis Results

Variable	HH General Info	Highly Dense Area	Lower Income and Education	Higher Income and Education	Middle-Age Family	Youngers	Black and Hispanic	Seniors
HH_vehicle	**0.704**	–0.320	–0.160	–0.065	–0.095	0.212	0.062	0.027
HH_size	**0.806**	0.307	0.119	–0.144	0.132	0.053	0.189	–0.089
HH_worker	**0.661**	0.053	–0.346	–0.092	–0.076	0.237	0.235	–0.012
HH_student	**0.655**	0.424	0.178	–0.116	0.193	–0.069	–0.090	–0.104
HH_license	**0.729**	–0.160	–0.242	–0.132	–0.080	0.320	–0.398	–0.003
HH_bikes	**0.548**	0.240	–0.029	–0.002	0.090	–0.243	0.024	–0.083
CBD	–0.141	**0.184**	–0.173	0.105	–0.046	0.047	–0.093	0.163
Population_density	–0.357	**0.763**	–0.318	–0.166	–0.146	0.124	–0.088	–0.010
Housing_density	–0.410	**0.748**	–0.324	–0.145	–0.143	0.108	–0.002	–0.001
Transit_density	–0.397	**0.617**	–0.218	–0.001	–0.053	0.026	0.184	–0.014
Road_density	–0.077	**0.438**	–0.146	–0.196	–0.128	0.154	–0.005	–0.069
Low_income	–0.453	0.135	**0.510**	–0.032	0.389	0.270	0.236	–0.083
No_high_degree	0.373	0.474	**0.631**	–0.204	–0.220	–0.058	0.135	0.099
High_degree	–0.157	–0.200	**0.417**	–0.216	0.210	0.182	0.284	–0.035
College_degree	–0.118	–0.182	**0.496**	–0.323	–0.019	0.066	0.032	–0.038
White_ethnicity	0.434	–0.104	**–0.585**	0.086	0.503	0.377	–0.056	–0.045
Med_income	0.119	–0.211	–0.025	**0.596**	–0.565	–0.433	0.083	–0.001
High_income	0.363	0.091	–0.406	**0.649**	0.215	0.191	–0.024	0.080
Grad_degree	–0.067	–0.022	–0.444	**0.459**	–0.127	–0.169	–0.289	–0.443
Age –16	0.362	0.438	–0.237	0.262	**0.594**	–0.191	0.298	0.087
Age_31–50	0.135	0.097	–0.437	–0.239	**0.609**	–0.447	–0.159	–0.181
Age_51–65	–0.176	–0.306	–0.233	0.249	**0.556**	0.187	–0.248	–0.008
Age_17–30	0.142	0.107	0.034	–0.287	–0.003	**0.585**	0.328	0.120
Bachelor_degree	–0.025	–0.061	–0.329	–0.151	0.162	**0.838**	–0.021	0.202
Black_ethnicity	–0.043	0.149	0.454	–0.268	0.044	–0.029	**0.736**	–0.082
Hisp/other_ethnicity	0.049	0.094	0.289	–0.373	0.075	–0.045	**0.564**	–0.032
Age +65	–0.131	–0.280	0.254	0.006	0.126	0.020	0.035	**0.425**

NOTE: Dominant variables in boldface.

Cluster analysis is one of the most widely used approaches to deal with heterogeneity of individuals in terms of their decision-making criteria toward various choices, as shown in Everitt et al. (*34*) and Raihanian Mashhadi and Behdad (*35*). In this research, a distance-based clustering method was used that classified observations into relatively homogeneous collections by minimizing the variance across variables of interest within clusters and maximizing the variance between clusters (*10*).

Numerous clustering techniques have been proposed in the literature, among which K-means, hierarchical, and two-step clustering techniques are widely used in travel behavior studies, such as those of McCarthy et al. (*32*), Collum and Daigle (*36*), and Pronello and Camusso (*37*). This study applied the two-step method because of the accuracy of its reported results. Further, it is generally considered to be more efficient when one is dealing with large data sets that contain both continuous and categorical variables, as noted by Chiu et al. (*38*). In this method, the data were first divided into subclusters on the basis of log likelihood distance. All the records were checked one by one, and the current observation was either merged with the previously formed clusters or assigned to a new cluster on the basis of the distance criterion. In the second step, the resulting subclusters were further categorized into the desired number of clusters by comparing their distance measures with a specified threshold. If the decrease in log likelihood as a result of merging the two clusters was larger than the threshold, the two subclusters could be merged. The distance between two clusters can be calculated as follows (*38*):

$$d(i, j) = \eta_i + \eta_j - \eta_{\langle i,j \rangle} \tag{1}$$

where

$$\eta_v = -N\left(\sum_{k=1}^{K_{CO}} \frac{1}{2}\log\left(\hat{\sigma}_k^2 + \hat{\sigma}_{vk}^2\right) + \sum_{k=1}^{K_{CA}} \hat{E}_{vk}\right) \tag{2}$$

$$\hat{E}_{vk} = -\sum_{l=1}^{L_k} \frac{N_{vkl}}{N_v} \log \frac{N_{vkl}}{N_v} \tag{3}$$

and

$d(i, j)$ = distance between clusters i and j,
N_v = number of observations in cluster v,
K_{CO} = total number of continuous variables,
K_{CA} = total number of categorical variables,
L_k = number of categories in kth categorical variables,
$\hat{\sigma}_k^2$ = estimated variance of kth continuous variable in all data,
$\hat{\sigma}_{vk}^2$ = estimated variance of kth continuous variable in cluster v,
N_{vkl} = number of observations in cluster v whose kth categorical variable takes the lth category, and
$\langle i, j \rangle$ = index that represents the cluster formed by combining clusters i and j.

With the factors derived from the PCA that reflect socioeconomic and land use characteristics, the two-step clustering method was applied with 10% noise allowance to categorize the records into homogeneous clusters. This optimal number of clusters was determined with a Bayesian information criterion (*38*). At the end, the records were assigned to six clusters, each containing between 7.9% and 22.5% of the data set. These results can be informative in defining various lifestyles that are assumed to influence travelers' behavior. Table 3 presents a brief summary statistic and centroids of clusters, which represent the following lifestyles:

1. Forever worker. This cluster represents individuals with lower levels of income and education. They usually live in dense urban areas and are mostly full-time workers.
2. Affluent in suburbs. This group is composed of middle-aged persons with higher income. They usually live in nondense areas such as suburbs. They have bachelor's or graduate degrees and mostly have children.
3. Young achievers. The cluster is made up of young singles or couples, mostly with high education levels, that live in dense urban areas, such as downtowns.
4. Seniors. This cluster consists of older individuals who are mostly retired or work part time, with low income. They are mostly white, but some black or Hispanic people are included.
5. Mainstream families. This group represents white, middle-aged, working-class couples. They mostly have low income and young children. Most live in rural areas.
6. Minorities. This cluster represents black people, Hispanic people, or people from other minority groups. Most do not have graduate degrees. They mostly live in upper middle-aged families and have lower income levels.

TABLE 3 Cluster Statistics and Centers

Cluster	Cluster Name	Size	Percentage	General HH Info	Highly Dense Area	Lower Income and Education	Higher Income and Education	Middle-Aged Family	Youngers	Black and Hispanic	Seniors
1	Forever worker	1,557	14.2	0.24	1.07	2.35	–0.80	–0.66	–0.59	0.17	0.67
2	Affluent in suburbs	2,365	21.5	1.43	–0.29	–0.57	1.54	1.71	–0.12	–0.67	0.36
3	Young achievers	2,187	19.9	–0.36	1.81	–0.53	0.83	–0.23	2.08	0.08	–0.44
4	Seniors	869	7.9	–0.04	0.25	0.89	–0.30	0.12	–0.34	0.21	1.75
5	Mainstream families	1,550	14.1	0.84	–0.46	1.66	–0.11	1.72	0.18	–0.41	0.25
6	Minorities	2,472	22.5	0.62	–0.32	1.04	–0.82	0.99	–0.52	1.60	0.58
Total		11,000	100								

Joint Model of Travel Mode and Departure Time

As noted, a copula-based approach was applied in this study to jointly estimate trip departure time and mode choice decisions within each homogeneous cluster of travelers, identified in the last subsection. As the first component of this joint structure, a multinomial logit model was applied to predict the mode choice decision. The utility function of the choices can be written as Ben-Akiva and Lerman describe (*39*):

$$U_{ai} = V_{ai} + \varepsilon_{ai} = \beta_a x_{ai} + \varepsilon_{ai} \tag{4}$$

where

- U_{ai} = person-specific utility of mode a for individual i;
- V_{ai} = systematic utility, which is a function of a set of explanatory variables (x_{ai}) and corresponding parameters of weighting factors (β_a); and
- ε_{ai} = a random variable that is the error term of the utility corresponding to unobserved factors, which is assumed to have a standard Type 1 extreme value distribution.

On the basis of this assumption for the error terms, the closed form probability for selecting alternative a by person i would be expressed as given in Ben-Akiva and Lerman (*39*):

$$P_{ai} = \frac{\exp(\beta_a x_{ai})}{\sum_k \exp(\beta_k x_{ki})} \tag{5}$$

Departure time, as the second component of this model, was treated as a continuous variable in this study and was modeled with log-linear regression as

$$\ln(t_{ai}) = \alpha_a Z_{ai} + \zeta_{ai} \tag{6}$$

where

- $\ln(t_{ai})$ = natural logarithm of trip timing for person i and alternative a, only if choice a is selected as the trip mode;
- α = vector of estimable parameters;
- Z = vector of explanatory variables; and
- ζ = error term corresponding to unobserved factors.

The linkage between mode choice and trip timing decisions depends on the type and the extent of the dependency between the stochastic terms ε_{ai} and ζ_{ai}. This study applied the copula approach to capture this dependency between these two decisions. The copula presents the joint probability distribution of random variables with predefined marginal distributions, following Sklar (*40*):

$$F_{\varepsilon_{ai},\zeta_{ai}}(X_1, X_2) = C_\theta\left(u_1 = F_{\varepsilon_{ai}}(X_1), u_2 = F_{\zeta_{ai}}(X_2)\right) \tag{7}$$

where

- X_1 and X_2 = random variables,
- $C_\theta(. , .)$ = relevant copula function,
- $F_{\varepsilon_{ai},\zeta_{ai}}(. , .)$ = multivariate joint distribution,
- u_1 and u_2 = marginal distributions of X_1 and X_2,
- $F_{\varepsilon_{ai}}(.)$ and $F_{\zeta_{ai}}(.)$ = marginal distributions, and
- θ = dependence parameter.

In this study, the four different copula functions of Frank, Gumbel, Clayton, and Joe were explored to estimate the dependent structure, and the model with Frank copulas provided the best statistical fit. Although it is not possible to provide a detailed discussion of these models here because of space constraints, a comprehensive discussion of these copula structures can be found in Bhat and Eluru (*12*). The model specifications and estimation results of only the Frank copula are presented here. The copula function for the Frank copula with θ as the copula parameter is as follows (*12*):

$$C_\theta(u_1, u_2) = -\frac{1}{\theta}\ln\left(1 + \frac{\left(e^{-\theta u_1} - 1\right)\left(e^{-\theta u_2} - 1\right)}{e^{-\theta} - 1}\right) \tag{8}$$

This joint distribution is then used to form the likelihood function. Following Spissu et al. (*41*), the likelihood function would be formulated as follows:

$$L = \prod_{i=1}^{N}\left[\left\{\prod_{a=1}^{A}\frac{1}{\sigma_{\zeta_{ai}}} \times \frac{\partial C_{\theta a}\left(u_{i1}^a, u_{i2}^a\right)}{\partial u_{i2}^a} f_{\zeta_{ai}}\left(\frac{\ln(t_{ai}) - \alpha_a Z_{ai}}{\sigma_{\zeta_{ai}}}\right)\right\}^{R_{ai}}\right] \tag{9}$$

where

- R_{ai} = binary variable representing whether mode a is selected by person i,
- $f_{\zeta_{ai}}$ = probability density function of ζ,
- $\sigma_{\zeta_{ai}}$ = scale parameter of ζ, and
- $C_{\theta a}$ = copula corresponding to $F_{\varepsilon_{ai},\zeta_{ai}}(u_{i1}^a, u_{i2}^a)$, with $u_{i1}^a = F_{\varepsilon_{ai}}(\beta_a x_{ai})$ and $u_{i2}^a = F_{\zeta_{ai}}(\ln(t_{ai}) - \alpha_a Z_{ai}/\sigma_{\zeta_{ai}})$.

ESTIMATION RESULTS AND DISCUSSION

Following the PCA and clustering steps, six copula-based joint models were estimated to determine the significant factors in the joint trip mode and departure time decisions of the travelers within each homogeneous cluster. All variables and their relevant interactions were tested and the statistically significant coefficients (at confidence levels 90%, 95%, and 99%) are presented in Table 4.

The estimation results indicated that various sociodemographic, land use, and trip-related variables in both the discrete and continuous components remained statistically significant across all travelers' clusters. Further, the results showed that applying the clustering technique, which makes it possible to account for unobserved shared factors among members of each cluster, leads to variation in the magnitude of influence and significance level of variables across clusters. For example, travel time has a mixed effect on choosing walk as the mode of travel. The results demonstrated that, generally, people who face greater travel times are less inclined to walk to a destination. However, in the second cluster, travel time has no significant effect on selecting this mode. This is possibly because of the lifestyle specifications of people in this cluster. Referring to Table 3, this cluster mainly consists of people with higher income levels who live in nondense areas (e.g., suburbs), where generally shopping centers or other nonmandatory trip destinations are not as accessible as they are in dense urban areas (e.g., a central business district). Similarly, the effect of travel time on choosing bike varies significantly across clusters; it is even not significant in Cluster 2 (i.e., residents of nondense areas with high income) and Cluster 4, which mostly includes senior citizens. On the other hand, in the utility functions of the drive and transit modes, travel time plays a significant role in travel mode decisions of travelers in all user clusters.

Moving to transit-related variables, it is found that egress distance is a main factor in choosing transit as the travel mode across all clusters. The results indicate that people in all clusters are less likely to choose transit in trips with longer egress distances, whereas the access distance is significant in only half of the clusters. The partial significance of this variable is possibly because of the behavior of travelers who do not have access to other modes of transportation. Travel cost also has a variable negative effect on transit mode choice in all clusters except Cluster 2, where it is not significant.

With respect to trip departure time models, the results highlighted that travel time is one of the key contributors to trip departure time decision; that is, the increase in travel time in almost all modes and users' clusters results in earlier departure times, except for the group of seniors. Activity duration is also found to significantly affect trip timing; trips with longer activities at their destinations are generally taken at earlier times. The work status of the travelers is also an important factor in estimating the trip departure time, as part-time workers tend to perform nonmandatory activities sooner than individuals with full-time jobs.

Turning to household sociodemographic characteristics, the result showed that the number of vehicles in the household and household size significantly affect departure time; more vehicles in the household result in later departure times, possibly because of a reduction of shared trips in the household and subsequently more flexibility for travelers. The results also indicated that people start their trips sooner if they are traveling to high-density areas.

To investigate further the implications of the clustering approach in the modeling scheme, an aggregate joint model for all observations was also estimated; the results of this model are presented in the last column of Table 4. The estimated parameters showed a remarkable difference in magnitude of the estimated coefficients, their degree of significance, and, in some cases, their signs, between the aggregate model and the cluster-specific models. This fact indicates the potential behavioral differences across clusters and highlights the importance of segmentation analysis in devising transportation policies, which should account for diverse responses of all travelers in the network.

To assess the prediction capability of these models, they were used to simulate selected choices in test sets. Figure 1 presents the prediction accuracy (percentage of correctly predicted alternative) of mode choice components of joint models for both cluster-based models and the aggregate model. The analysis revealed a significant improvement in prediction accuracy in the proposed clustered-based joint models across all modes. The greatest improvement happened in the drive mode, with a 9.08% increase in prediction accuracy; the lowest improvement was in predicting bike, with a 1.14% increase. A smaller increase in the prediction accuracy for bike could be as a result of its few trip observations in the data set, which limits the model's performance in identifying bicycle-related travel attributes in each cluster; thus, the clustering step does not significantly improve the model prediction accuracy for this mode.

To analyze the potential effect of segmentation on departure time estimation, the mean absolute percentage error (MAPE) index was calculated for this component of copula models in both aggregate and clustered approaches. The MAPE index showed the average error in estimation of departure times. Following Washington et al. (*42*), the index is defined as

$$\text{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right| \tag{10}$$

where A_t is the actual value and F_t is the forecasted value for observation *t*. The MAPE measure for the aggregate model was 17.45% and for the cluster-based models was 10.51%, which shows a reduction of 6.94% in the average error. Figure 2 presents the predicted departure times versus the observed ones as visual support for the improved prediction accuracy in cluster-based models over the aggregate one. Overall, the results suggested that the segmentation technique significantly improved the prediction accuracy of mode choice and departure time models.

CONCLUSION

This study concerned an effort to analyze travelers' behavior regarding two of the most fundamental trip-related decisions: mode choice and departure time. Because of the interrelated decision mechanisms of these two travel dimensions, they should be considered jointly to allow for unrestricted correlation between their influencing factors. Moreover, because of heterogeneity in travelers' decision-making criteria toward these travel attributes, different travelers can respond differently to transport policies. Capturing this diversity among travelers is desirable to develop a reliable and policy sensitive model.

To achieve this goal, this study proposes a modeling approach that relies on cluster analysis and a joint copula-based modeling technique. First, principal component analysis was applied to remove the potential multicollinearity among explanatory variables and reduce the data dimension. Following that, a two-step clustering technique was adopted to assign travelers to homogeneous clusters. Then, a copula-based joint model of discrete mode selection and continuous departure time decision was estimated within each cluster. The mode choice model was estimated by a multinomial logit model and the departure time choice was modeled by means of a log-linear regression model.

To investigate the potential advantages of developing different models for heterogeneous clusters, an aggregate joint model on all observations was also established. The results showed that the overall clustered models outperformed the aggregate one in both decision components. The estimated prediction accuracy for mode choice component improved from 1.14% for bike mode to 9.08% for drive mode. In addition, applying the segmentation technique results in a 6.94% improvement in the prediction accuracy of trip departure time. Further, the variations in estimated parameters indicate the potential behavioral differences across clusters and highlight the importance of segmentation analysis in devising transportation policies, which should account for the diverse responses of all travelers in the network.

This study suggests several possibilities for future research directions. First, as an extension of this work, it is desirable to develop a joint model that considers the correlation of other activity attributes, such as activity duration, with these two dimensions to better simulate travelers' trip-making behavior. Applying other joint modeling techniques and comparing their results with the employed copula approach would be informative about their performance and coupling structures.

TABLE 4 Model Estimation Results for Each Cluster

Variable	Forever Worker		Affluent in Suburbs		Young Achievers	
	Parameter	*t*-Stat.	Parameter	*t*-Stat.	Parameter	*t*-Stat.
Mode Choice Model						
Walk_TT	–0.78***	–4.63	—	—	–2.29***	–7.08
Walk_accessible	2.1***	4.31	1.75***	8.55	0.89*	1.73
Walk_age –20	1.53***	4.43	—	—	0.83**	2.5
Walk_Low_income	–0.43*	–1.82	—	—	—	—
Bike_constant	–1.84***	–3.72	–3.2***	–6.42	–3.56***	–8.74
Bike_HH_bikes	0.42***	7.44	0.43***	6.81	0.41***	4.37
Bike_TT	–1.16***	–5.39	—	—	–0.83**	–2.19
Bike_age_40–65	—	—	–0.48*	–1.91	—	—
Drive_constant	–0.9**	–1.97	–1.85***	–4.62	–2.05***	–4.32
Drive_HH_vehicle	1.24***	15.48	1.15***	17	0.72*	1.83
Drive_TT	–0.11**	–2.07	–0.38*	–1.78	–0.18***	–4.02
Drive_age –20	–0.69***	–4.44	–0.06*	–1.69	—	—
Drive_cost	–0.12***	–4.14	–0.05**	–2.16	–0.09**	–2.49
Transit_constant	0.44*	1.77	–2.5***	–3.47	–1.88***	–5.32
Transit_TT	–0.26*	–1.93	–0.05***	–3.45	–0.09***	–2.68
Transit_cost	–0.16***	–2.93	—	—	–0.05**	–2.23
Transit_egress	–0.69**	–2.53	–0.21***	–3.68	–0.7***	–2.95
Transit_access	—	—	–0.52***	–3.46	–1.38***	–4
Transit_low_income	0.06**	2.11	—	—	—	—
Departure Time Model						
Walk_constant	6.91***	64.15	6.79***	45.91	6.98***	71.3
Walk_age –20	0.09*	1.68	–0.27***	–2.9	—	—
Walk_TT	–0.29***	–4.56	–0.18***	–3.74	–0.35**	–2.37
Walk_activity_dur	–0.12***	–5.48	—	—	–0.09**	–2.15
Bike_constant	6.88***	73.23	6.81***	44.11	6.95***	111.45
Bike_TT	–0.34***	–7.53	–0.39***	–4.78	–0.36**	–1.99
Bike_activity_dur	–0.10**	–1.99	—	—	–0.13*	–1.83
Bike_age –20	0.32**	2.42	–0.39***	–5.95	—	—
Bike_part_work	–0.06*	–1.88	—	—	–0.06**	–2
Drive_constant	6.93***	81.24	6.89***	91.43	7.02***	57.66
Drive_HH_vehicle	0.25***	3.91	0.23**	2.13	—	—
Drive_TT	–0.87***	–6.35	–0.81***	–2.79	–0.65**	–2.23
Drive_activity_dur	–0.06***	–5.04	–0.07*	–1.73	–0.01**	–2.22
Drive_full_work	0.29***	4.61	0.37***	6.02	0.17***	4.54
Drive_road_density	—	—	—	—	–1.64**	–2.29
Transit_constant	6.83***	120.91	6.75***	42.54	6.88***	55.19
Transit_HH_size	–0.26***	–9.25	—	—	—	—
Transit_age +65	—	—	—	—	—	—
Transit_TT	–0.61***	–3.06	–0.49***	–4.67	–0.54***	–8.57
Transit_activity_dur	–0.15***	–4.31	–0.11***	–9.57	–0.21***	–4.25
Transit_Weekend	–0.17*	–1.68	—	—	–0.26**	–2.38
Copula Parameters						
θ_{Walk}	–6.39***	–4.87	–5.55***	–6.76	–3.26**	–2.27
θ_{Bike}	–6.01***	–2.59	–7.62***	–5.37	–2.56***	–3.79
θ_{Drive}	–6.18***	–16.8	–9.89***	–4.14	–7.24***	–4.44
$\theta_{Transit}$	–8.11***	–9.34	–13.9***	–20.57	–4.65***	–5.56
Scale Parameters						
σ_{Walk}	1.03***	8.13	4.49***	15.34	1.02***	4.78
σ_{Bike}	2.94***	14.59	3.74***	14.31	1.00***	2.84
σ_{Drive}	1.78***	49.99	1.82***	54.35	1.00***	13.02
$\sigma_{Transit}$	1.00***	12.61	1.68***	14.17	1.00***	11.43

NOTE: — = not significant. Number of observations and log likelihood, respectively: forever worker = 1,557 and –2,683.71; affluent in suburbs = 2,365 and –4,183.31; young achievers = 2,187 and –3,976.58; seniors = 869 and –1,928.24; mainstream families = 1,550 and –2,590.64; minorities = 2,472 and –5,332.86; all clusters = 11,000 and –18,092.79.
*Significant at 90%; **significant at 95%; ***significant at 99%.

Seniors		Mainstream Families		Minorities		All Clusters	
Parameter	*t*-Stat.	Parameter	*t*-Stat.	Parameter	*t*-Stat.	Parameter	*t*-Stat.
−1.54***	−9.24	−0.52***	−3.64	−0.33***	−4.95	−1.65***	−16.76
2.63***	13.45	1.89***	8.11	1.59***	7.84	2.05***	17.17
—	—	1.76***	7.76	1.39***	8.62	0.73***	6.28
—	—	—	—	—	—	—	—
−0.67***	−2.97	−1.84***	−6.14	−1.4***	−5.86	−2.11***	−13.89
0.23***	4.89	0.44***	8.61	0.41***	10.4	0.47***	21.6
—	—	−1.9***	−5.78	−1.23***	−6.6	−2.11***	−14.1
−1.28***	−4.8	−0.96***	−3	−0.53**	−2.33	−0.56***	−5.2
0.58***	3.94	−0.71***	−4.01	−0.46***	−2.59	−0.09	−0.74
0.91***	12.66	1.09***	13.63	1.15***	17.47	1.16***	33.7
−0.23**	−2.32	−0.62***	−2.78	−0.29***	−4.35	−0.37**	−2.45
—	—	−0.89***	−2.9	−0.64**	−2.02	−1.62***	−14.56
−0.11***	−3.24	−0.11***	−3.42	−0.15***	−5.84	−0.16***	−11.1
—	—	—	—	0.47**	2.36	−0.23***	−27.15
−0.06***	−2.71	−0.42*	−1.77	−0.35**	−2.63	−0.61***	−4.84
−0.18***	−4.08	−0.11**	−2.14	−0.12***	−2.81	−0.66***	−10.28
−0.64***	−3.11	−0.37**	−2.73	−0.45**	−2.04	−0.06**	−2.54
−1.82***	−9.79	—	—	—	—	−1.18***	−17.01
—	—	—	—	0.09**	1.98	0.07***	1.72
6.56***	108	6.67***	101.95	6.58***	197.72	6.72***	229.12
—	—	−0.11**	−2.48	−0.08*	−1.71	—	—
−0.28*	−1.69	−0.37***	−3.82	−0.17**	−2.56	−0.28***	−3.38
−0.08**	−6.39	—	—	−0.11***	−4.69	−0.06***	−24.87
6.61***	46.91	6.71***	60.04	6.54***	93	6.53***	139.45
−0.46***	−2.68	−0.26*	−1.75	−0.24**	−2.06	−0.31***	−4.37
—	—	−0.12***	−3.81	−0.10**	−2.31	−0.06***	−11.44
—	—	−0.28*	−1.77	−0.16***	−3.69	—	—
—	—	—	—	−0.13***	−2.58	—	—
6.71***	61.52	6.75***	113.17	6.63***	116.37	6.25***	409.74
0.17*	1.82	0.16***	3.59	0.14**	2.43	0.15***	24.73
−0.62*	−1.71	−0.76*	−1.8	−0.59***	−3.88	−0.03**	−2.31
−0.08**	−2.37	−0.04***	−4.51	−0.07*	−1.86	−0.06***	−45.16
—	—	0.27**	2.09	0.13***	5.65	—	—
—	—	−0.98*	−1.79	—	—	−0.10***	−2.87
6.84***	120.47	6.75***	148.53	6.82***	258.25	6.36***	336.76
—	—	−0.07***	−3.49	−0.16***	−10.48	0.03***	4.64
−0.67***	−7.19	−0.49**	−2.02	—	—	—	—
−0.53***	−8.27	−0.51***	−5.38	−0.62***	−6.09	—	—
−0.18**	−2.17	−0.14***	−6.48	−0.13**	−2.49	−0.03***	−13.74
—	—	−0.24**	−2.68	−0.31***	−4.82	—	—
−0.48***	−3.15	−0.51***	−2.89	−0.67***	−5.38	−0.48***	−10.25
−0.49***	−3.49	−0.54***	−3.76	−0.69***	−5.14	−0.47***	−12.44
−0.42**	−2.26	−0.83***	−6.36	−1.00***	−7.56	−0.54***	−14.66
−0.55***	−8.16	−0.63***	−9.46	−0.77***	−8.9	−0.64***	−38.39
1.15***	11.43	1.00***	9.62	1.00***	11.7	1.00***	27.39
1.34***	6.47	1.19***	5.49	1.14***	7.02	1.18***	15.78
1.06***	25.2	1.00***	19.51	1.00***	25.15	1.00***	59.74
1.00***	16.34	1.00***	14.01	1.00***	16.53	1.00***	38.46

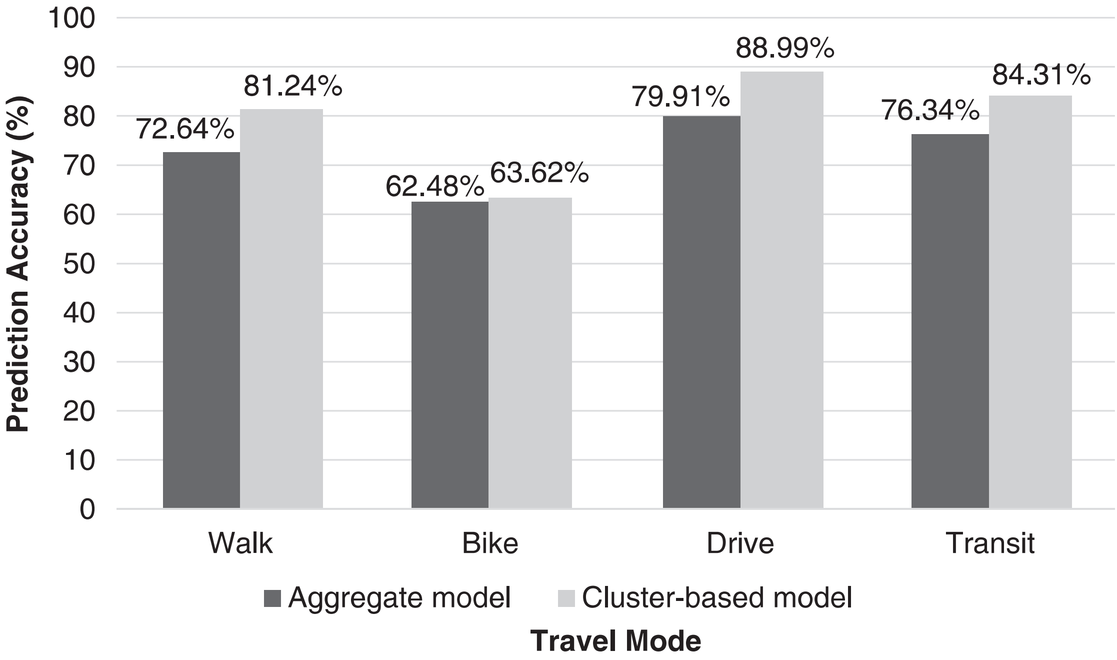

FIGURE 1 Travel mode prediction accuracy in clustered and aggregate models.

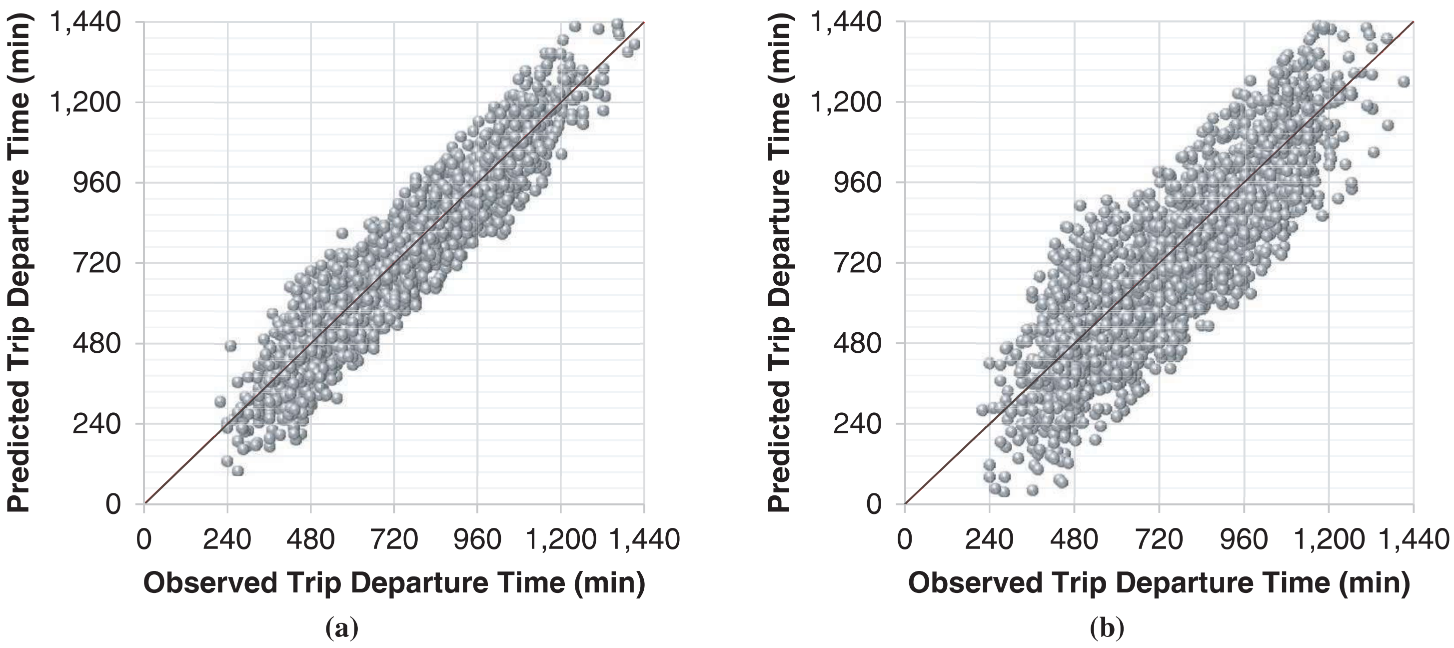

FIGURE 2 Departure time prediction accuracy for (*a*) cluster-based models and (*b*) aggregate model.

REFERENCES

1. Tringides, C.A., X. Ye, and R.M. Pendyala. Departure-Time Choice and Mode Choice for Nonwork Trips: Alternative Formulations of Joint Model Systems. *Transportation Research Record: Journal of the Transportation Research Board,* No. 1898, 2004, pp. 1–9. https://dx.doi.org/10.3141/1898-01.
2. Bhat, C.R. Analysis of Travel Mode and Departure Time Choice for Urban Shopping Trips. *Transportation Research Part B: Methodological,* Vol. 32, No. 6, 1998, pp. 361–371. https://doi.org/10.1016/S0191-2615(98)00004-6.
3. Amirgholy, M., and E.J. Gonzales. Demand Responsive Transit Systems with Time-Dependent Demand: User Equilibrium, System Optimum, and Management Strategy. *Transportation Research Part B: Methodological,* Vol. 92, 2016, pp. 234–252. https://doi.org/10.1016/j.trb.2015.11.006.
4. Bhat, C.R., and J.Y. Guo. A Comprehensive Analysis of Built Environment Characteristics on Household Residential Choice and Auto Ownership Levels. *Transportation Research Part B: Methodological,* Vol. 41, No. 5, 2007, pp. 506–526. https://doi.org/10.1016/j.trb.2005.12.005.
5. Rashidi, T.H., and T.T. Koo. An Analysis on Travel Party Composition and Expenditure: A Discrete-Continuous Model. *Annals of Tourism Research,* Vol. 56, 2016, pp. 48–64. https://doi.org/10.1016/j.annals.2015.10.003.
6. Heckman, J. Shadow Prices, Market Wages, and Labor Supply. *Econometrica,* Vol. 42, No. 4, 1974, pp. 679–694. https://doi.org/10.2307/1913937.
7. Lee, L.F. Some Approaches to the Correction of Selectivity Bias. *Review of Economic Studies,* Vol. 49, No. 3, 1982, pp. 355–372. https://doi.org/10.2307/2297361.
8. Nurul Habib, K.M. Modelling Commuting Mode Choice Jointly with Work Start Time and Duration. *Transportation Research Part A: Policy and Practice,* Vol. 46, No. 1, 2012, pp. 33–47. https://doi.org/10.1016/j.tra.2011.09.012.
9. Mohammadian, A., and Y. Zhang. Investigating Transferability of National Household Travel Survey Data. *Transportation Research Record: Journal of the Transportation Research Board,* No. 1993, 2007, pp. 67–79. https://dx.doi.org/10.3141/1993-10.
10. Raihanian Mashhadi, A., B. Esmaeilian, and S. Behdad. Simulation Modeling of Consumers' Participation in Product Take-Back Systems. *Journal of Mechanical Design,* Vol. 138, No. 5, 2016, p. 051403. https://doi.org/10.1115/1.4032773.
11. Miralinaghi, M., Y. Lou, Y.T. Hsu, R. Shabanpour, and Y. Shafahi. Multiclass Fuzzy User Equilibrium with Endogenous Membership Functions and Risk-Taking Behaviors. *Journal of Advanced Transportation,* 2016. https://doi.org/10.1002/atr.1425.
12. Bhat, C.R., and N. Eluru. A Copula-Based Approach to Accommodate Residential Self-Selection Effects in Travel Behavior Modeling. *Transportation Research Part B: Methodological,* Vol. 43, No. 7, 2009, pp. 749–765. https://doi.org/10.1016/j.trb.2009.02.001.

13. Eluru, N., R. Paleti, R.M. Pendyala, and C.R. Bhat. Modeling Injury Severity of Multiple Occupants of Vehicles: Copula-Based Multivariate Approach. *Transportation Research Record: Journal of the Transportation Research Board,* No. 2165, 2010, pp. 1–11. https://dx.doi.org/10.3141/2165-01.
14. Born, K., S. Yasmin, D. You, N. Eluru, C.R. Bhat, and R.M. Pendyala. Joint Model of Weekend Discretionary Activity Participation and Episode Duration. *Transportation Research Record: Journal of the Transportation Research Board,* No. 2413, 2014, pp. 34–44. https://dx.doi.org/10.3141/2413-04.
15. Karimi, B., Z. Pourabdollahi, R.S. Anbarani, and A.K. Mohammadian. A Mixed Joint Discrete-Continuous Model of Non-Mandatory Out-of-Home Activity Type and Activity Duration. Presented at 94th Annual Meeting of Transportation Research Board, Washington, D.C., 2015.
16. Sener, I.N., N. Eluru, and C.R. Bhat. On Jointly Analyzing the Physical Activity Participation Levels of Individuals in a Family Unit Using a Multivariate Copula Framework. *Journal of Choice Modeling,* Vol. 3, No. 3, 2010, pp. 1–38. https://doi.org/10.1016/S1755-5345(13)70012-5.
17. Sener, I.N., and P.R. Reeder. An Integrated Analysis of Workers' Physically Active Activity and Active Travel Choice Behavior. *Transportation Research Part B: Methodological,* Vol. 67, 2014, pp. 381–393.
18. Rashidi, T.H., and A. Mohammadian. Application of a Nested Trivariate Copula Structure in a Competing Duration Hazard-Based Vehicle Transaction Decision Model. *Transportmetrica A. Transportation Science,* Vol. 12, No. 6, 2016, pp. 1–18.
19. Golshani, N., R. Shabanpour, S.M. Mahmoudifard, S. Derrible, and A. Mohammadian. Comparison of Artificial Neural Networks and Statistical Copula-Based Joint Models. Presented at 96th Annual Meeting of the Transportation Research Board, Washington, D.C., 2017.
20. Auld, J.A., and A. Mohammadian. Activity Planning Processes in the Agent-Based Dynamic Activity Planning and Travel Scheduling (ADAPTS) Model. *Transportation Research Part A: Policy and Practice,* Vol. 46, No. 8, 2012, pp. 1386–1403. https://doi.org/10.1016/j.tra.2012.05.017.
21. Javanmardi, M., M.F. Langerudi, R. Shabanpour, and A. Mohammadian. An Optimization Approach to Resolve Activity Scheduling Conflicts in ADAPTS Activity-Based Model. *Transportation,* Vol. 43, No. 6, 2016, pp. 1023–1039. https://doi.org/10.1007/s11116-016-9721-7.
22. Langerudi, M.F., R.S. Anbarani, M. Javanmardi, and A.K. Mohammadian. Resolution of Activity Scheduling Conflicts: Reverse Pairwise Comparison of In-Home and Out-of-Home Activities. *Transportation Research Record: Journal of the Transportation Research Board,* No. 2566, 2016, pp. 41–54. https://dx.doi.org/10.3141/2566-05.
23. Shabanpour, R., M. Javanmardi, M. Fasihozaman, M. Miralinaghi, and A. Mohammadian. Investigating the Applicability of ADAPTS Activity-Based Model in Air Quality Analysis. *Travel Behaviour and Society,* 2017, https://doi.org/10.1016/j.tbs.2017.02.004.
24. Shabanpour, R., N. Golshani, J. Auld, and A.K. Mohammadian. Dynamics of Time-of-Day Choices in the Agent-Based Dynamic Activity Planning and Travel Simulation (ADAPTS) Framework. Presented at 96th Annual Meeting of the Transportation Research Board, Washington, D.C., 2017.
25. Shabanpour, R., N. Golshani, M.F. Langerudi, M. Javanmardi, and A.K. Mohammadian. Modeling Type and Duration of In-Home Activities in ADAPTS Activity-Based Framework. Presented at 96th Annual Meeting of the Transportation Research Board, Washington, D.C., 2017.
26. Nurul Habib, K.M., N. Day, and E.J. Miller. An Investigation of Commuting Trip Timing and Mode Choice in Greater Toronto Area: Application of a Joint Discrete-Continuous Model. *Transportation Research Part A: Policy and Practice,* Vol. 43, No. 7, 2009, pp. 639–653. https://doi.org/10.1016/j.tra.2009.05.001.
27. Paleti, R., P.S. Vovsha, D. Givon, and Y. Birotker. Joint Modeling of Trip Mode and Departure Time Choices Using Revealed and Stated Preference Data. *Transportation Research Record: Journal of the Transportation Research Board,* No. 2429, 2014, pp. 67–78. https://dx.doi.org/10.3141/2429-08.
28. Kumar, A., and D.M. Levinson. Temporal Variations on Allocation of Time. *Transportation Research Record,* No. 1493, 1995, pp. 118–127.
29. Javanmardi, M., M.F. Langerudi, R. Shabanpour, and K. Mohammadian. Mode Choice Modelling Using Personalized Travel Time and Cost Data. International Conference of the International Association for Travel Behavior Research (IATBR), Windsor, UK, 2015.
30. Stopher, P.R., and A.H. Meyburg. *Survey Sampling and Multivariate Analysis for Social Scientists and Engineers.* Lexington Books, Lexington, Mass., 1979.
31. Anable, J. "Complacent Car Addicts" or "Aspiring Environmentalists"? Identifying Travel Behavior Segments Using Attitude Theory. *Transport Policy,* Vol. 12, No. 1, 2005, pp. 65–78. https://doi.org/10.1016/j.tranpol.2004.11.004.
32. McCarthy, O.T., B. Caulfield, and M. O'Mahony. Technology Engagement and Privacy: A Cluster Analysis of Reported Social Network Use Among Transport Survey Respondents. *Transportation Research Part C: Emerging Technologies,* Vol. 63, 2016, pp. 195–206. https://doi.org/10.1016/j.trc.2015.12.015.
33. Mahmoudifard, S.M., A. Kermanshah, R. Shabanpour, and A. Mohammadian. Assessing Public Opinions on Uber as a Ridesharing Transportation System: Explanatory Analysis and Results of a Survey in Chicago Area. Presented at 96th Annual Meeting of the Transportation Research Board, Washington, D.C., 2017.
34. Everitt, B.S., S. Landau, and M. Leese. *Cluster Analysis,* 4th ed. Arnold Press, London, 2001.
35. Raihanian Mashhadi, A., and S. Behdad. Optimal Sorting Policies in Remanufacturing Systems: Application of Product Life-Cycle Data in Quality Grading and End-of-Use Recovery. *Journal of Manufacturing Systems,* Vol. 43, 2017, pp. 15–24. https://doi.org/10.1016/j.jmsy.2017.02.006.
36. Collum, K.K., and J.J. Daigle. Combining Attitude Theory and Segmentation Analysis to Understand Travel Mode Choice at a National Park. *Journal of Outdoor Recreation and Tourism,* Vol. 9, 2015, pp. 17–25. https://doi.org/10.1016/j.jort.2015.03.003.
37. Pronello, C., and C. Camusso. Travellers' Profiles Definition Using Statistical Multivariate Analysis of Attitudinal Variables. *Journal of Transport Geography,* Vol. 19, No. 6, 2011, pp. 1294–1308. https://doi.org/10.1016/j.jtrangeo.2011.06.009.
38. Chiu, T., D. Fang, J. Chen, Y. Wang, and C. Jeris. A Robust and Scalable Clustering Algorithm for Mixed Type Attributes in Large Database Environment. *Proceedings of the 7th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, San Francisco, Calif., 2001.
39. Ben-Akiva, M., and S. Lerman. *Discrete Choice Analysis: Theory and Application to Travel Demand*. MIT Press, Cambridge, Mass., 1985. https://doi.org/10.1145/502512.502549.
40. Sklar, A. Random Variables, Joint Distribution Functions, and Copulas. *Kybernetika,* Vol. 9, No. 6, 1973, pp. 449–460.
41. Spissu, E., A.R. Pinjari, R.M. Pendyala, and C.R. Bhat. A Copula-Based Multinomial Discrete-Continuous Model of Vehicle Type Choice and Miles of Travel. *Transportation,* Vol. 36, No. 4, 2009, pp. 403–422. https://doi.org/10.1007/s11116-009-9208-x.
42. Washington, S., M. Karlaftis, and F. Mannering. *Statistical and Econometric Methods for Transportation Data Analysis,* 2nd ed. Chapman & Hall/CRC, Boca Raton, Fla., 2010.

The Standing Committee on Transportation Demand Forecasting peer-reviewed this paper.