Database Development Approach and Survey Design for Travel Mode Shift Behavior Study with Respect to Mass Transit in a Metropolitan Context

Nipa A. Desai^{1,2} · G. J. Joshi²

Received: 3 December 2014 / Accepted: 11 August 2016 / Published online: 29 August 2016
© Springer International Publishing Switzerland 2016

Abstract Travel behavior study is one of the most complex studies as it governs multi-parameter evaluation at a time and involves collection of huge data pertaining to human behavior in variety of situations (Rastogi and Krishna Rao in Segmentation analysis of commuters accessing transit: Mumbai study. *J Transp Eng* 135(8):506–515, 2009). Success of travel behavior studies is mostly dependent on healthy data collection. Aim of this paper, is to evaluate and suggest the most suitable database development approach and research strategies which can be adopted for travel behavior studies especially in context of Indian metro cities. It also presents entire survey design procedure adopted for travel mode shift study along with researchers' experience in execution of survey and difficulties faced by field teams. Three level filtration processes are adopted for selection of variables for development of mode shift models. Initially the least important variables were omitted based on variable ranking results. Top ranked seven variables are identified for final survey design, out of which travel time (including transit access time, waiting time and in-vehicle time), travel cost and number of transfers are directly measurable variables while comfort, convenience, safety and security are latent variables. The latent variables are measured using psychometric scale constructed from various latent variable indicators and then

converted into latent classes using principal component analysis. Research instrument developed and adopted for the present study was tested through pilot survey and corrected as per data compatibility requirement. Finally 1017 valid responses out of 1500 distributed questionnaires are analyzed to evaluate the effects of enhanced level of hypothetical public transport services on travel mode shift behavior through development of binary logistic regression models. Changes in the model performance with inclusion of latent classes are also evaluated using various statistical tools. This paper includes original work based on primary data from a field survey, findings of which are expected to provide a better understanding of entire database development process and effect of different variables on mode shift behavior in context of Indian metropolitan area.

Keywords Travel mode shift · Survey design · Sampling · Pilot survey · Psychometric scale · Principal component analysis (PCA) · Logit model

Introduction

Asian economy is growing at very rapid pace and India is one of the fastest growing economies in the Asian continent. Public transportation plays an important role in achieving high quality of urban life and sustainability. However, there are many valid reasons due to which, commuters' are attracted towards personalized modes. In the absence of an adequate and efficient bus transit system, potential bus users are currently using private and para-transit transport modes [2]. To analyze the negative shift towards public to private transportation various studies have been carried out in India and worldwide but these comparative studies of travel behavior differ greatly in

✉ Nipa A. Desai
nipa_adesai@yahoo.co.in
G. J. Joshi
gjsvnt92@gmail.com

¹ Civil Engineering Department, Babaria Institute of Technology, BITS edu Campus, Vadodara, Gujarat, India

² Sardar Vallabhbhai National Institute of Technology, Surat, India

methodologies, data, units of analysis, and explanatory variables. Vedagiri and Arasan [2, 3], Krishna Rao et al. [4], Rastogi [5], Rastogi and Krishna Rao [1, 6–8], and Tiwari [9] studied travelers' behavior for cities like Mumbai, Chennai, Delhi, Calcutta and Bangalore. Joshi and Katti [10], Joshi and Mokal [11], Desai and Vashi [12] have evaluated performance of existing public bus transportation services in Surat and Vadodara respectively. Desai and Joshi [13] have also thrown some light on geographical and cultural variations in factors affecting travel mode choice in Indian metropolitan contexts. However travel mode choice and evaluation of modal shift for sustainable transport planning, considering user preferences are still awaited especially with respect to Indian metro cities.

Travel behavior study is one of the most complex studies in transportation engineering as it governs multi-parameter evaluation at a time [1]. These studies involve collection of huge data pertaining to human behavior in variety of situations and its success is mostly depending on the judgment regarding sampling techniques, data collection methodologies and research instrument adopted. The experiences with different data collection techniques and variants of survey instrument format in developing country are different from the reported experiences of industrialized and developed countries [7]. The motivation for this article is to share experience in design, execution and analysis of survey for database development related to travel mode choice and mode shift study.

Methodology adopted for this study is based on evaluation of commuters' behavior and preferences in variety of situations using various discrete choice and multi-criteria decision modeling tools. The study involves a reasonable number of house hold interviews to evaluate commuters' behavior, their choices and preferences. Sample size for the study was calculated based on available statistical theories and budget criteria. Overall information about study area such as current population, population growth trend, vehicular growth rate, topography, road network, socio-economic profile of residents and municipal ward details were collected from various local authorities and past records. Initially designed research instrument was tested through two level of pilot survey with more than 150 and more than 100 samples respectively to formulate research instrument for execution of the main survey.

After finalizing research instrument two main surveys were executed. Survey-1 was related to development of income group categories and psychometric construct for latent variables (Executed during November' 13 to January' 14) and survey-2 for travel mode shift behavior study (Executed during January' 15 to March' 15). In the first survey, total 278 responses were received back out of 320 distributed questionnaires which were used for

psychometric construct of latent variables and income group categorization. The latent variables measured during survey-1 are reduced into three factors using principal component analysis based on which different attraction levels of comfort & convenience and safety & security were designed. Finally 1500 survey forms were distributed for main survey which was exclusively conducted for mode shift behavioral analysis.

Database Development

Studies related to transportation planning require huge data pertaining to travel behavior. Following sections describe entire process of database development adopted for travel mode shift behavior study.

Study Area Profile

The study is carried out for Vadodara metropolitan area; the third largest city in the State of Gujarat, India. As per census 2011, population of Vadodara urban area is 2.1 million. The public transport system of Vadodara city, which was earlier operated by the state government transport authority, is presently run by VTCOS private operators on private contract basis. Like other metro cities of India, remarkable imbalance is prevailing in demand and supply of public transport facilities in Vadodara also. In recent years, a huge reduction in the use of public transport; in other words a negative shift from public transport to personalized travel modes is observed [12]. Findings from the past studies also shown that more than 85 % of Vadodara city commuters are diverted to personalized vehicles or other intermediate transportation modes [12]. This study is taken up to help transit service providers in estimating potential transit users, who have propensity to shift from their personalized travel mode to mass transit for work purpose travel with implementation of set of transit service quality improvement measures.

Income Group Categorization for Travel Mode Shift Behavior Analysis

Household income is one of the most important factors influencing choice of mode for travel. Different income group people behave differently while choosing their travel mode [13]. For most of the transport related studies, due to lack of proper guidelines income group categories are either adopted from previous studies or from government minimum wage rates. Alpizar and Carlsson [14] adopted average monthly income and 10th deciles of average

Table 1 Survey results of average monthly expenditure based on different living standards and different category of family members

Family composition	Monthly personal average expenditure in INR (Indian rupee)	INR Just for survival	INR For normal living	INR For satisfactory living
Earning adult	Stay/house rent	500	1000	2000
	Household	2000	4500	8000
	Transport	1000	1500	2500
	Medical	300	500	700
	Miscellaneous	700	1000	3500
	Total expenditure in Rs.	4500	8500	16700
Non earning adult	Stay/house rent	500	1000	2000
	Household	2000	4000	5000
	Transport	300	800	1500
	Medical	300	500	700
	Miscellaneous	700	800	1500
	Total expenditure in Rs.	3800	7100	10,700
School/college going children	Stay/HR	500	1000	2000
	Household	2000	4000	5000
	Transport	500	1200	2000
	Medical	300	500	500
	Miscellaneous	700	500	1500
	Education	500	1300	1000
	Total expenditure in Rs.	4500	8500	12,000
Elderly person	Stay/HR	500	1000	2000
	Household	2000	3000	4000
	Transport	200	500	500
	Medical	1000	2500	3500
	Miscellaneous	500	1000	1000
	Total expenditure in Rs.	4200	8000	11,000

INR Indian rupee

monthly income in colon (currency of Costa Rica) as a reference to classify income groups in their studies. Rastogi and Krishna Rao [7] adopted income ranges as 1501–5000, 5001–10,000, 10,001–15,000 and 10,001–20,000 in INR (Indian rupee) for their study without specifying criteria about formation of these categories. Income group categories adopted in latest studies are cross validated through minimum wages prescribed by the state government but those income group categories are not matching with the existing rate of inflation and hence new income group categories are formed for the study based on income and expenditure pattern of a household. A preliminary level household survey was carried out to know minimum family expenditure and per person essential expenditure just for survival (Category-1), normal living (Category-2) and satisfactory living (Category-3). First category survey was carried out amongst bus drivers, housekeepers, gardeners, tempo drivers, auto rickshaw drivers and peons to know

their own and other family members' monthly personal expenditure. The second category survey was carried out amongst grocery shop owners, working class people, school teachers and rickshaw users who are staying in one room kitchen unit. The third category survey is carried out amongst service and business class people owning more than two 2-wheelers or one car costing <0.4 million INR and who are living in semi detached or detached houses but not in individual bungalows with large plot area, Tables 1 and 2 show summary of survey results.

As shown in Table 2 income group categories were formed to be used exclusively for present study. These categories are purely with respect to the ratio of family monthly income and household size rather than only family monthly income because irrespective of individual personal income, travel mode choice decisions are more dependent on family income and expenditure pattern [13].

Table 2 Income group categorization based on the ratio of family income/number of family members

Most common household categories	Description	Average expenditure in INR		
		Survival	Normal living	Satisfactory living
2 adult + 2 children	Earning adult	4500	8500	16,700
	Non earning adult	3800	7100	10,700
	Child-1	4500	8500	12,000
	Child-2	4500	8500	12,000
	Total family exp	17,300	32,600	51,400
	Per person family expenditure	4325	8150	12,850
2 adult + 1 children + 1 elder	Earning adult	4500	8500	16,700
	Non earning adult	4200	8000	11,000
	Child-1	4500	8500	12,000
	Elder	4200	8000	11,000
	Total family exp	17,400	33,000	50,700
	Per person family expenditure	4350	8250	12675
2 adult + 2 children + 1 elder	Earning adult	4500	8500	16,700
	Non earning adult	3800	7100	10,700
	Child-1	4500	8500	12,000
	Child-2	4500	8500	12,000
	Elder	4200	8000	11,000
	Total family exp	21,500	40,600	62,400
Average per person family expenditure irrespective of family type and size	Per person family expenditure	4300	8120	12,480
		4325	8173	12,668
Income group category based on the ratio of family income/number of family members		<5000	5000–10,000	10,000–15,000
Category names		LIG	MIG	UMIG

Design of Survey and Research Instrument

A huge data pertaining to commuters' information is required to fulfill study objectives. Exhaustive literature review related to survey design, sampling methodologies, variable selection and variable reduction techniques is found very much helpful for further development of research. Alpizar and Carlsson [14] studied a group of policies aimed at discouraging the use of private transportation, especially cars during peak hours at metropolitan Costa Rica. Their sampling strategy excludes the low-income segments of society from study, since they cannot afford to own a car. Study carried out by Ho et al. [15] for Klang Valley involved personal interview survey, randomly executed at the designated petrol stations and hypermarket parking lots. Out of total, about 10 % of samples were excluded in their study due to incompleteness. The major reason identified by researchers behind incomplete questionnaire was reluctance of respondents to answer several personal questions and the limited time they could spend to answer the survey. Adjemian and Williams [16] investigated the potential usefulness of a proxy approach for modeling discrete choice vehicle ownership

using data from the 2000 Bay Area Travel Survey (BATS) and the U.S. Census. Dillman [17] identified that with the development of internet services, a new mode of data collection has turned out as one of the most popular survey methodology. Olsen [18] provided evidence that internet surveys can be a viable alternative to traditional mail surveys. Israel [19], Bartlete et al. [20], Venigalla [21], Taylor-Powell [22], Yoshida et al. [23] have thrown some light on selection of sample frame and sample size.

Kipping in view the past research findings about survey execution and city municipal corporation zoning pattern the city is divided into ten administrative wards forming four zones. Residents residing and working within the limit of Vadodara Municipal Corporation are included in the sample frame. Stratified random sampling technique is used for data collection as this method gives every individual in the defined population an equal and independent chance of being selected and allows identified subgroups in the defined population to be represented in the same proportion that they exist in the population.

For healthier data collection, utmost care has been taken while deciding sample size. There are several approaches

Table 3 Summary on sample size based on different precision level

Sr. no	Response rate in percentage (%)	Calculated samples in numbers			
		95 % CI, 5 % ME	95 % CI, 3 % ME	99 % CI, 5 % ME	99 % CI, 3 % ME
1	50	768	2133	1327	3683
2	60	640	1778	1105	3069
3	70	549	1524	948	3069
4	80	480	1333	829	2302

Sample size calculation summary for 50 % variability

CI confidence interval, *ME* margin of error

Bold signifies calculated samples for present study based on 95 % confidence interval, 5 % Margin of error and 50 % response rate

mentioned in literature by Israel [19], Watson [24], Olsen [18], Morikawa et al. [25], Bartlete et al. [20] to determine sample size. Most of the popular approaches are depending either on population mean or standard deviation (S.D) but due to unavailability of past records methods based on population mean or S.D were not found suitable for present study and hence the method suggested by Watson [24] is adopted. The method is based on precision level of the study, acceptable margin of error (ME), confidence interval (CI) and degree of freedom (DF) was adopted for sample size calculations. Equation 1 shows sample size calculation formulae and Table 3 shows calculated samples for different categories of CI and ME on the basis of which final sample size is decided for the present study.

$$n = \frac{\left[\frac{p(1-p)}{\frac{A^2}{Z^2} + \frac{p(1-p)}{N}} \right]}{R} \quad (1)$$

where, n sample size required, N number of people in the population p : estimated variance in population, as a decimal: (0.5 for 50–50) A : precision desired, expressed as a decimal (i.e., 0.03, 0.05, for 3 %, 5 %) Z : based on confidence level (1.96 for 95 %, 1.65 for 90 % and 2.58 for 99 %) R : estimated response rate, as a decimal.

First of all time frame for execution of survey was fixed and than as per the availability of manpower and funds, the criteria for sample size calculation was fixed as 95 % confidence interval and 5 % margin of error with 50 % response rate. Total number of samples required as per these criteria are 762 and the collected valid samples are 1017 with multiple number of data points which are higher than requirement.

Pilot Survey Data Analysis and Refinement of Research Problem

Pilot survey was carried out to test survey instrument and survey methodology before execution of final survey. First level stated preference pilot survey with more than 150

samples and second level pilot survey with more than 100 samples were carried out as per the zoning framework prepared for the main survey execution. Compatibility of collected data for travel mode choice and mode shift related studies were checked through initial modeling using GRETAL and SPSS software. It was observed from pilot survey results that most of the respondents found questionnaire lengthy, in order to reduce length of questionnaire, less influential variables are identified from pilot survey results and related questions are removed. Questions related to respondents' occupation is removed as no specific relationship is observed between Vadodara city commuters' mode choice and occupation during pilot survey. Literature findings and pilot survey analysis also show that irrespective of personal income, commuters' choice of mode is found more dependent on family's income Desai and Joshi [13] and hence question related to personal income was also omitted.

Initially the survey was designed for all categories of vehicles and all purpose trips but based on the results of pilot survey, the research problem was narrowed down by restricting the area to the categories which are having highest travel mode shift potential. Some interesting findings of pilot survey are presented below which justifies inclusion or exclusion of specific category of people or travel attribute for further study. The respondents who have shown unwillingness towards desired mode shift are studied separately. It was found that amongst all unwilling commuters, 83.6 % are from HIG who are owning one or more cars and their most preferred mode choice determinants are personal status, luxury and quality travel. Due to their inclination towards quality travel they are found less attracted towards public transport services and hence excluded from the study. Similar to the common research findings, share of work trips in Vadodara is also observed three times than any other purpose trips if the school going children are excluded and hence, as an initial focus of mode shift, only work related trips of LIG, MIG and UMIG group commuters' are considered looking to its highest potential in achieving targeted mode shift.

Variable Selection for Travel Mode Shift Study

Three level filtration is carried out for selection of variables. Step by step process adopted for inclusion or rejection of variables is presented below.

Selection of Variables Based on Variable Ranking Results

Studies carried out by Mohamad and Kiggundu [26], Abdul Kadir et al. [27], Kamba et al. [28] in Klang Valley, Ras-togi [5] and Tiwari [9] investigated factors affecting mode choice preferences. Geographical and cultural variations in factors affecting travel mode choice are studied by Desai and Joshi [13]. In order to improve effectiveness of ranking scale, it was decided to include carefully selected variables while designing mode shift study questionnaire. Pilot survey data analysis and literature review findings are considered as a base for selection of variables for mode shift problem.

Based on literature findings, ten important variables responsible for mode choice and mode shift behavior are identified and respondents have been asked to rank those variables on a scale of 1–10. (1—most important and 10—least important). Out of selected ten variables, travel time (TT), travel cost (TC), convenience (CN), comfort (CM), safety and security (SS) and number of transfers (NT) got maximum rankings in top five, while other variables viz., sending kids to school (SK), irregular working hours (IRW), stress (STR), personal status (PS) and privacy (PR) were given less priority by most of the commuters. Figures 1, 2, 3, 4, 5, 6, 7, 8, 9 and 10 graphically represents summary of importance rankings for all the ten selected variables based on which the research instrument is refined for second level of survey.

It was observed during pilot survey execution that respondents' were found attentive while ranking up to top 5–6 variables only. To come out of this issue, it was decided to retain only top five to six higher ranked variables for final survey. Other variables were omitted from final questionnaire as including too many variables in ranking scale may result into false ranking. Out of six selected variables based on the results of pilot survey travel time (TT), travel cost (TC) and number of transfers (NTR) are directly measurable variables but comfort (CM), convenience (CN) and safety and security (SS) are latent variables which need to be treated differently. Principal component analysis (PCA) is performed to reduce large number of indicator statements representing latent variables. Finally, the categories of potential mode shift variables are formed for final survey and further analysis.

Fig. 1 Importance ranking for travel cost (TC)

Fig. 2 Importance ranking for travel time (TT)

Fig. 3 Importance ranking for convenience (CN)

Fig. 4 Importance ranking for comfort (CM)

Fig. 5 Importance ranking for safety and security (SS)

Fig. 7 Importance ranking for privacy (PR)

Fig. 6 Importance ranking for number of transfers (NTR)

Measurement of Latent Variables

Recent findings in the field of transport modeling Johansson et al. [29] suggests that only socio-economic attributes and trip related attributes are not sufficient to characterize travelers and make forecasts about their travel behavior. These studies have shown integration of latent variables representing attitudes, perceptions, behavior and preferences into choice models and revealed suitability of psychometric scales for measurement of latent variables [30]. As summarized in another literature, mode choice decisions might not only depend on objective criteria like time, cost, and income but also on other criteria such as comfort,

Fig. 8 Importance ranking for stress (STR)

Fig. 9 Importance ranking for sending kids to school (SK)

Fig. 10 Importance ranking for irregular working hours (IRW)

convenience, safety and flexibility [30]. Ben-Akiva et al. [31] presented a behavioral framework and related logical formulations for modeling these latent variables. Scheiner and Holz-Rau [32] studied travelers' behavior based on five key elements viz. life situation, lifestyle, location attitudes, and urban form. Yang et al. [33] formulated a two-stage choice model which accommodated taste heterogeneity of an individual. In their study, authors have computed service attributes such as reliability, convenience and comfort by the use of incomplete pair wise comparisons. Johansson et al. [30] studied the problem of unobservable latent preferences in mode choice models and identified that latent variables enriched discrete choice models. They have also found that the construct reliability of the "attitudinal" latent variables is higher than that of the "behavioral"

Table 4 Indicators contributing in construction of latent variables

Latent variable indicator code	Variable indicators
C1	Seat availability during travel
C2	Type of seats
C3	Reserved seats for females
C4	Cleanliness of buses and bus stops
C5	A.C vehicles
C6	Frequent cleaning of vehicle
C7	Vehicle with low bus platform heights for boarding
C8	Politeness and friendliness of the vehicle crew
C9	Vehicle with curtains/sun film
C10	Sufficient standing space
C11	Space to keep luggage
C12	Express bus routes
C13	Parking facility near bus stop
C14	Ease of boarding with luggage
C15	Availability of washrooms at interchange junctions
C16	Stress-free and relaxed travel
C17	Sheltered bus stops
C18	Route coverage as per requirement
C19	Travel in noise free environment
C20	Monthly passes and smart card payment facilities
C21	Published Timings and schedules at every bus stop
C22	Spaciousness and enough leg space
C23	Vehicles with music facility
C24	Vehicle with automatic door operations
C25	Riding quality
C26	Prior seat reservation facility
C27	Automatic ticketing machine with card swipe
C28	Load factor/overcrowding in buses
S1	Security of luggage and valuables during journey
S2	Safely driven buses
S3	Travelling with strangers
S4	Fire safety measures in vehicle
S5	Queue formation for safe boarding and alighting
S6	CCTV cameras in vehicles and stops for safety
S7	Lighting and other facilities at bus stops
S8	Personal safety at bus stop
S9	Appointment of security officer at important terminals
S10	Safe location of bus stops

latent variables. Morikawa et al. [25] included modal comfort and convenience in their analyses of mode choice and modeled latent variables through attitudinal indicators. Temme et al. [34] also highlighted valuable inputs in summarizing integrated choice and latent variable models.

Based on the findings from above stated literature, a separate research instrument is developed for measurement of latent variables using explanatory indicators of variables like comfort, cleanliness, convenience, security and safety. Total 40 indicator statements presented in Table 4 are identified to measure commuters' importance ranking for

indicators of selected latent variables. To increase dataset accuracy, experts and travelers of different category were asked to rank only 9 randomly assigned indicator statements on seven point Likert scale (1—least important and 7—most important) out of 40. Finally responses from all the respondents are combined and average importance rating of all 40 latent variable indicators are calculated.

Average importance rating of 40 indicator variables shows that few of them were found with very less average ranking. As suggested by Yang et al. [33], attributes that respondents deemed unimportant with average importance

Table 5 Extraction of communalities for selected latent variables related to comfort and convenience

Latent variable code	Latent variables	Initial communality	Extracted communality
C1	Seat availability during travel	1	0.715
C2	Type of seats	1	0.861
C4	Cleanliness of buses and bus stops	1	0.761
C5	A.C vehicles	1	0.923
C7	Vehicle with low bus platform heights for boarding.	1	0.912
C9	Vehicle with Curtains/sun film.	1	0.596
C10	Sufficient standing space	1	0.533
C11	Space to keep luggage	1	0.940
C15	Availability of washrooms at interchange junctions	1	0.745
C16	Stress free and relaxed driving	1	0.829
C17	Sheltered bus stops	1	0.853
C18	Route coverage as per requirement	1	0.674
C20	Monthly passes and smart card payment facilities	1	0.818
C21	Published timings and schedules at every bus stop	1	0.693
C25	Riding quality	1	0.914
C27	Automatic ticketing machine with card swipe	1	0.543
C28	Load factor/overcrowding in buses	1	0.703

Fig. 11 Scree plot of extracted components for comfort and convenience

rating <2.5 are eliminated and finally 17 higher ranked indicator statements related to latent variable comfort & convenience and 8 higher ranked indicator statements related to latent variable safety & security are selected for construction of different classes of these latent variables.

Construction of Latent Classes from Latent Variables

The findings from literature review advocate principal component analysis (PCA) and a Varimax rotation as one of the best techniques for latent variable data structuring and data reduction [35]. The major objective behind

carrying out principal component analysis (PCA) in this study is to reduce number of questions from main questionnaires and to develop latent classes out of too many latent variables which can directly be adopted for data collection pertaining to mode shift study. Here, the area of interest is to evaluate number of best suitable components and what these components represent rather than to examine interrelationship between latent variables. The PCA method differs from correlation analysis because correlation analysis only gives interrelationship between two different variables at a time, however when a set of variables are correlated with each other, there may be a possibility that there is some other more general variable with which they all are correlated. This variable is known as factor or component and principal component analysis is a statistical technique that attempts to uncover such factors. As suggested by Schussler and Axhausen [35], the initial step for the measurement of latent variable is adopted as principal component analysis (PCA) carried out using SPSS with the aim to reduce representative indicators to meaningful factors based on which different levels of comfort & convenience and safety & security can be constructed.

Component Extraction for Comfort and Convenience

Responses related to comfort and convenience are separated and analyzed on SPSS platform. Extracted communalities for all the 17 indicator variables are presented in

Table 6 Results of PCA for comfort and convenience in a form of rotated component matrix

Code	Latent variable	Component		
		1	2	3
C1	Seat availability during travel	–	0.629	–
C2	Type of seats	–	–	0.916
C4	Cleanliness of buses and bus stops	–	–	0.860
C5	A.C vehicles	0.952	–	–
C7	Vehicle with low bus platform heights for boarding	–	0.946	–
C9	Vehicle with curtains/sun film	0.645	–	–
C10	Sufficient standing space	–	–	0.989
C11	Space to keep luggage	0.948	–	–
C15	Availability of washrooms at interchange junctions	0.844	–	–
C16	Stress free and relaxed driving	–	–	0.899
C17	Sheltered bus stops	–	0.810	–
C18	Prior seat reservation facility	0.807	–	–
C20	Monthly passes and smart card payment facilities	0.868	–	–
C21	Published timings and schedules at every bus stop	–	0.807	–
C25	Riding quality	–	0.941	–
C27	Spaciousness and enough leg space	0.720	–	–
C28	Load factor/overcrowding in buses	–	–	0.666

Extraction method: principal component analysis. Rotation method: Varimax with Kaiser normalization

Table 5. The communalities are computations of the extent to which a variable is explained by the components. If extracted communalities are higher than PCA method is suitable for data structuring.

Most of the calculated communalities in Table 5 are high, which indicates that the extracted components represent variables well. Here, ‘sufficient standing space’ has the lowest communality, which indicates that this variable is less explained by the analysis than any other variable. The first three extracted components with eigen value greater than one explain nearly 75 % of variability in the original seventeen variables, and hence complexity of the data set by using these components can be considerably reduced, albeit with 25 % information loss. The eigen value of each component in the initial solution is plotted (Refer Fig. 11) to verify whether the extracted four components out of sixteen variables are proper or not. Generally, the components on the steep slope contribute more to the solution while the components on the mild slope contribute little to the solution. The scree plot of the data justifies three extracted components as only first three components are on steep slope with the eigen value more than the cut off value one.

Next, rotated component matrix is calculated to estimate loadings of extracted components. Rotated component matrix is the key output of principal component analysis. It contains estimates of the correlations between each of the variables and the estimated components. When multiple variables have very high positive or negative correlations

on a component in the rotated component matrix, the implication is that they are redundant and some of them can be removed or grouped together. Rotated component matrix presented below in Table 6 shows the most prominent indicators with component score greater than 0.5. The third component in Table 6 is highly correlated with sufficient standing space followed by type of seats, stress free travel and cleanliness. Component three is named as CC_3. All the variables associated with CC_3 are pertaining to basic level of comfort and convenience desired by any traveler. Component two named as CC_2 is found highly correlated with seat availability, low platform buses, sheltered bus stops, published schedules and good riding quality which is one level higher than the previous one. Component one named as CC_1 is related to air-conditioned vehicle with additional bus stop facilities and technology driven services. Out of all the five significant variables which are explaining component one can be summarized as high level of services opted by respondents. The positive value of variables indicate that all seven

Table 7 KMO and Bartlett’s test for observed data

Kaiser–Meyer–Olkin measure of sampling adequacy	0.778
Bartlett’s test of sphericity	
Approx. Chi square	912.220
Df	136
Sig.	0.00

Table 8 Extraction of communalities for selected latent variables related to safety and security

		Initial	Extraction
S1	Security of luggage and valuables during journey	1	0.847
S2	Safely driven buses	1	0.834
S4	Fire safety measures in vehicle	1	0.859
S5	Queue formation for safe boarding and alighting	1	0.911
S6	CCTV cameras in vehicles and stops for safety	1	0.883
S7	Lighting and other facilities at bus stops	1	0.735
S9	Appointment of security officer at important terminals	1	0.664
S10	Safe location of bus stops	1	0.854

Fig. 12 Scree plot of extracted components for safety and security

variables are having greater positive impact on each other in other words commuters' with positive score of **CC_1** prefer technology driven specious A.C vehicles with curtain/sun film and good bus stop facilities.

Kaiser–Meyer–Olkin (KMO) and Bartlett's test are performed on the data to validate suitability of data for structure detection and factorization. The KMO measure of sampling adequacy is a statistic that indicates the proportion of variance in our variables that might be caused by underlying factors. Generally, values of the Kaiser–Meyer–

Olkin measure of sampling adequacy higher than 0.5 and significance value less than 0.05 indicates that PCA is useful for analyzing the data. Table 7 shows KMO test results for study data. The value 0.778 of KMO measure; which is greater than 0.5; shows suitability of PCA for our data structuring.

Component Extraction for Safety and Security

Extracted communalities for all the selected eight indicator variables representing safety and security are presented in Table 8. Most of the calculated communalities in Table 8 are high, which indicates that the extracted components represent variables well.

The first three extracted components with eigen value greater than one explain nearly 82 % of variability in the original eight variables, and hence complexity of the data set can be reduced considerably by using these components, with 18 % loss of information. The scree plot of study data shown in Fig. 12 also justifies high contribution of three extracted components.

Rotated component matrix presented below in Table 9 shows the most prominent indicators with component score greater than 0.5. Here the third component named as **SS_3** is highly correlated with safely driven buses and security of belongings which shows fundamental level of safety and

Table 9 Results of PCA for safety and security in a form of rotated component matrix

		Component		
		1	2	3
S1	Security of luggage and valuables during journey	–	–	0.833
S2	Safely driven buses	–	–	0.869
S4	Fire safety measures in vehicle	–	0.879	–
S5	Queue formation for safe boarding and alighting	0.937	–	–
S6	CCTV cameras in vehicles and stops for safety	0.925	–	–
S7	Lighting and other facilities at bus stops	–	0.852	–
S9	Appointment of security officer at important terminals	0.786	–	–
S10	Safe location of bus stops	–	0.822	–

Extraction method: principal component analysis. Rotation method: Varimax with Kaiser normalization

Table 10 KMO and Bartlett's test for observed data

Kaiser–Meyer–Olkin measure of sampling adequacy	0.669
Bartlett's test of sphericity	
Approx. Chi square	255.323
Df	28
Sig.	0.00

security desire by any traveler. Component two named as SS_2 is correlated with installation of fire safety alarm in buses and safe location of bus stops with good lighting facility which is one level higher than the fundamental requirements. Component one named as SS_1 is related to installation of CCTV cameras, appointment of security officers at terminals and provision for safe boarding and alighting which is again one level higher than the previous one.

The value of Kaiser–Meyer–Olkin measure of sampling adequacy shown in Table 10 is obtained as 0.669 which is greater than 0.5 and significance value less than 0.05 shows suitability of PCA method for structuring of latent variables related to safety & security.

Considering results of PCA three different levels of extracted components of comfort & convenience and safety & security are constructed and further utilized for mode shift final questionnaire design. Level-I (basic level) of comfort & convenience comprises of CC_3, level-II (advanced level) of CC_3 + CC_2 and level-III (high level) is of CC_3 + CC_2 + CC_1 while the first (basic), second (advanced) and third (high) level of safety and security comprises of SS_3, SS_3 + SS_2 and SS_3 + SS_2 + SS_1 respectively. Table 11 shows three different levels of

comfort & convenience and safety & security developed for mode shift questionnaire.

Commuters were proposed with three different levels of comfort & convenient and safety & security to choose from. Level-I with very basic facilities, level-II with little improved facilities and level-III with modern facilities. As it is required to restrict length of research instrument, individual variables were not proposed to commuters for giving their preference but simultaneously low, simultaneously medium and simultaneously high levels of all the factors were proposed which may be one of the limitation of this study as analysis of the trade-offs between individual variables may not be possible.

Development of Travel Mode Shift Model

The final survey was executed after refinement of research instrument with two levels of pilot survey during January' 15 to March' 15. Total 1500 survey forms were distributed and 1263 were received back. The incomplete forms or forms observed with illogical response were identified and rejected during data validation. Finally, 1017 valid responses combined from personalized two wheeler, car and auto users are used for further data analysis and model development using GRETEL software. Two models are developed—one without inclusion of latent classes and the other with inclusion of the same to check probable changes in model explanatory power due to inclusion of latent categories. Table 12 shows input data format used for development of model.

Table 11 Latent classes of comfort and convenient and safety and security constructed from latent variables

CC_Level-I, Normal & clean buses with good support for standing and stress free riding.	CC_Level-II - Level I with low boarding platform less crowded buses with more seating space, curtain/sun film+ Improved riding quality, more routes, sheltered bus stops and published schedule on stops.	CC_Level-III - Level II with A.C buses, specious and comfortable seats with more sitting space. +smart ticketing & prior reservation system + wash room availability at main terminals.

SS_Level- I safely driven buses with normal bus stops for safety of belongings.	SS_Level II - Level I with fire safety alarm, safely located bus stops with lighting facility.	SS_Level III - Level II with CCTV in buses + queue formation mechanism for safe boarding and alighting + appointment of security officers at terminals & strict checking for drunken driving

Table 12 Input data format used for GRETTEL software

Variable code	Type	Variable description	Variable coding
GEN	Nominal	Gender	{1—Male, 2—female}
AGE	Scale	Age of respondents in years	None
HHSIZE	Scale	Household size	None
HHWORK	Scale	Number of working persons per house hold	None
INGCAT	Ordinal	Income group categories	{1—LIG, 2—MIG, 3—UMIG, 4—HIG}
TYPDU	Ordinal	Type of dwelling unit	{1—kaccha house, 2—apartment, 3—row house, 4—tenement, 5—others}
NO2W	Scale	Number of 2-wheelers per household	None
NOCAR	Scale	Number of cars per household	None
MLTDES	Nominal	Visiting multiple destination during work trip	{1—yes, 2—no}
DSTWRK	Scale	Distance of work place from home	None
TTONWY	Scale	Travel time for one way work trip by existing mode	None
DITC	Scale	Daily travel cost for work trip by existing mode	None
WAITT	Ordinal	Willingness to shift to mass transit w.r.to reduction in bus stop waiting time	{1—not more than 15 min, 2—not more than 10 min, 3—not more than 5 min, 4—case do not attract}
ACCTIME	Ordinal	Willingness to shift to mass transit w.r.to transit access time	{1—9 to 6 min, 2—6 to 9 min, 3—<3 min, 4—it does not attract me at all}
RTT	Ordinal	Willingness to shift to mass transit w.r.to reduction in travel time	{1—up to 10 min, 2—10 to 15 min, 3—>15 min, 4—case do not attract}
RTC	Ordinal	willingness to shift to mass transit w.r.to reduction in travel cost	{1—up to 7 INR/day, 2—7 to 15 INR/day, 3—>15 INR/day, 4—case do not attract} (Note: INR—Indian rupee)
INTCHG	Ordinal	willingness to shift to mass transit w.r.to number of bus interchanges	{1—Not more than 2 interchanges, 2—not more than 1 interchange, 3—prefer only direct bus, 4—it does not attract me at all}
CCLVL	Ordinal	willingness to shift to mass transit w.r.to enhanced level of comfort and convenience	{1—Level-I (basic), 2—level-II (average), 3—level-III (highest), 4—does not attract me at all}
SSLVL	Ordinal	willingness to shift to mass transit w.r.to enhanced level of safety and security	{1—Level-I (basic), 2—level-II (average), 3—level-III (highest), 4—does not attract me at all}
WILLINGSHIFT (dependent variable)	Nominal	Willingness to shift from personalized mode to mass transit with respect to enhanced level of services	{1—Yes, 2—no}

Mode-Shift Model-I with Respect to Personalized Vehicles and Mass Transit (Excluding latent classes)

For the development of first model all the variables mentioned in Table 12 are considered except the latent classes of comfort & convenience (CC) and safety & security (SS). The model is developed using GRETTEL software. Table 13 shows results of mode shift model with WILLINGSHIFT as dependent variable and all others as independent variables. The variables presented in Table 13 such as GEN, HHWORK, WAITT_2, WAITT_3, RTC_3, INTCG_3, ACCTIME_2 are found significant with significance value less than or very near to 0.05. Rests of the variables are identified as the least contributing variables for the present case study. A variety of informal tests can be applied to an estimated model to assess the reasonableness of the implications of estimated parameters [36]. The variables

included in Table 13 are tested for sign of parameters to check whether the associated variables have a positive or negative effect on the alternatives and comparable with theory, intuitions and judgment regarding the expected impact. The estimated coefficient of number of working people in household is negative, implying that utility of willingness to shift decreases with increase in number of working persons per household. This negative sign is justifiable as with increase in number of working person per family household income also increases and there are evidences in literature that more is the household income; less is the probability to shift to public transport. While in case of WAITT_2, WAITT_3, RTC_3, INTCG_3 and ACCTIME_2, sign of significant variables is observed positive. The positive sign of parameter in case of waiting time implies that commuters are attracted towards shifting to mass transit with the minimum waiting time up to 5 min and the probability of shift decreases with increase in wait

time. Similarly the probability of shift is higher for highest reduction in travel cost up to 15 INR/day or more, minimum number of bus interchanges, preferably direct buses and transit access time between 3 and 6 min. The value of McFadden R-square is 0.236 which is reasonably good. If the value is near to zero it reflects that the model is not fitting the data and if it is near to 0.4 than it shows that it perfectly fits the data [36]. The value of likelihood ratio test is also significant and number of cases correctly predicted are 82.5 % which shows good fit of model to the data.

Mode-Shift Model-II with Respect to Personalized Vehicles and Mass Transit (Including latent classes)

For development of second model, all the variables mentioned in Table 12 are considered including latent categories of comfort & convenience and safety & security. The variables marked with “*” in Table 14 are HHWORK, ICGCAT, WAITT_2, ACCTIME_2, CCLVL_III and SSLVL_III are found significant with significance value less than or very near to 0.05. Rest of the variables are identified as the least contributing variables after inclusion of latent classes though it has been identified significant in model-I, which shows that CC and SS are one of the most important parameters to attract modal shift. The estimated coefficient of comfort and safety categories are positive as expected, implying that probability of willingness to shift increases with increase in comfort and safety of journey.

The coefficient value of ICGCAT is negative as expected which shows that with increase in family monthly income willingness to shift to public transport decrease. The value of McFadden R-square is 0.269, which is found improved in comparison with model-I. The value of likelihood ratio test is also significant which shows good fit of model to the data. Improvement in the value of likelihood ratio and increased prediction success rate of 83.8 % show that the model performs slightly better with inclusion of latent classes in comparison with Model-I. Even lower value of Akaike criterion and Hannan-Quinn in model-II in comparison with model-I also justify little improvement in model performance with the inclusion of latent class variables.

Utility functions can be developed using significant variables of Model-I and Model-II and the output of utility value calculated from both the equations can be substituted in logistic regression equation to predict probability or proportion of personalized vehicle users shifting to mass transit. Further analysis to study the behavior of personalized two wheeler users, car users and auto users separately can also be carried out for more specific output.

Mode shift models of different income group people can also be developed separately using similar approach to evaluate differences in behavior of LIG, MIG and UMIG group of commuters. Based on different income group categories of people, their desire for comfort and convenience and safety and security levels differ, however no major difference is observed in preferences of any other parameter with reference to commuters' income group.

Table 13 Model I: Logit, using observations 1-1017

Variable	Dependent variable: WILLINGSHIFT			p-value
	Coefficient	Std. error	Z	
Const	−0.424196	0.776504	−0.5463	0.58487
GEN	0.468546	0.248303	1.8870	0.05916*
HHWORK	−0.229256	0.119538	−1.9178	0.05513*
WAITT_2	0.664531	0.283661	2.3427	0.01915**
WAITT_3	0.713479	0.258915	2.7557	0.00586***
RTC_3	0.461175	0.261781	1.7617	0.07812*
INTCHG_3	0.917294	0.245237	3.7404	0.00018***
ACCTIME_2	0.764892	0.28661	2.6688	0.00761***
Mean dependent var	0.753937		S.D. dependent var	0.430928
McFadden R-squared	0.236073		Adjusted R-squared	0.183154
Log-likelihood	−433.0669		Akaike criterion	926.1338
Schwarz criterion	1073.843		Hannan-Quinn	

Number of cases' correctly predicted' = 841 (82.8 %)

Likelihood ratio test: Chi square(29) = 267.658 [0.0000]

* means variables found significant with p-value less than or very near to 0.05

** means variables found with p-value less than or very near to 0.05 with little more impact on model explanatory power in comparison with variables with only one *

*** are having highest significance in the model

Table 14 Model II: Logit, using observations 1-1017

Variable	Dependent variable: WILLINGSHIFT			p-value
	Coefficient	Std. error	z	
const	−1.17131	0.804431	−1.4561	0.14537
HHWORK	−0.228282	0.122614	−1.8618	0.06263*
ICGCAT	−0.256615	0.133787	−1.9181	0.05510*
WAITT_2	0.736668	0.276098	2.6681	0.00763***
ACCTIME_2	0.581271	0.28839	2.0156	0.04384**
CC_LVL_III	1.0605	0.288896	3.6709	0.00024***
SS_LVL_III	0.928142	0.278913	3.3277	0.00088***
Mean dependent var	0.753937		S.D. dependent var	0.430928
McFadden R-squared	0.269376		Adjusted R-squared	0.205872
Log-likelihood	−414.1879		Akaike criterion	900.3757
Schwarz criterion	1077.626		Hannan-Quinn	967.6914

Number of cases' correctly predicted' = 851 (83.8 %)

Likelihood ratio test: Chi square (35) = 305.416 [0.0000]

* means variables found significant with p-value less than or very near to 0.05

** means variables found with p-value less than or very near to 0.05 with little more impact on model explanatory power in comparison with variables with only one *

*** are having highest significance in the model

Conclusion

This paper suggests precise and tested approach for development of database and design of survey for travel mode choice modeling. Refinement of research instrument using series of pilot surveys executed for this study is found as one of the ideal approaches for collection of healthier data. Reasonably good response rate is observed by researchers during execution of survey through personal interviews with appointment of skilled staff in comparison with internet survey. It was concluded from the experience of survey execution team that executing survey in local language facilitates local people to understand questionnaire properly which results into quick and correct responses. The method suggested by Watson [24] is found most suitable for sample size determination in the present study as due to unavailability of past records, most of the other popular methods which are dependent on population mean, standard deviation or variance are not found useful. Fixation of income group class boundaries considering ratio of family income and household size as base shows better representativeness of data in comparison with conventional methods used for income group categorization. The process shown in this paper for identification of potential riders for a particular mode and elimination of commuters with lesser mode shift potential could be helpful in narrowing the research problem and enhancing the specific target group of travelers. The psychometric construct developed using principal component analysis (PCA) methodology for measurement of latent variables may be readily adopted for related studies and can be

explored further. It was also observed that the significance level of variables in Model-II (with inclusion of latent categories) improved model performance compared to travel mode shift Model-I (without inclusion of latent categories). Waiting time, number of interchanges and access time are found as most significant variables in Model-I, while highest level of safety and highest level of comfort are found the most significant in Model-II. Though both the models are found significant, only minor improvements in model fit tests is observed with inclusion of latent classes. The approach suggested in this paper will be useful to the research scholars and scientists who are working in the similar area and rather new to this field. The logical inputs based on pilot survey results about sampling distribution and income group classification is likely to add some knowledge to the area of surveys related to transportation planning. It is finally suggested that the data collection technique and survey instrument formats should be thoroughly verified with the help of pilot survey. A well designed survey and research instrument definitely brings richness to the data and ultimately to the research outcomes.

References

1. Rastogi R, Krishna Rao KV (2009) Segmentation analysis of commuters accessing transit: Mumbai study. *J Transp Eng* 135(8):506–515
2. Vedagiri P, Arasan VT (2011) Modeling modal shift from personal vehicles to bus on introduction of bus priority measure. *Asian Transp Stud* 1(3):288–302

3. Vedagiri P, Arasan VT (2009) Estimating modal shift of car travelers to bus on introduction of bus priority system. *Transp Syst Eng IT* 9(6):120–129
4. Krishna Rao KV, Arasan VT, Rengaraju VR (1994) Mode choice analysis: a review. *Indian J Transp Manag* 18(11):701–712
5. Rastogi R (2010) Willingness to shift to walking or bicycling to access suburban rail: case study of Mumbai, India. *J Urban Plan Dev* 136(1):3–10
6. Rastogi R, Krishna Rao KV (2000) A policy framework for a sustainable multi-mode urban transportation planning. *Trasporti Europei* 5(13):65–75
7. Rastogi R, Krishna Rao KV (2002) Survey design for studying transit access behavior in Mumbai city, India. *J Transp Eng ASCE* 128(1):68–79
8. Rastogi R, Krishna Rao KV (2003) Travel characteristics of commuters accessing transit: case study. *J Transp Eng ASCE* 129(6):684–694
9. Tiwari G (2008) Public transport research challenges in India. Department of Civil Engineering & Transportation Research and Injury Prevention Program, Indian Institute of Technology, Delhi
10. Joshi GJ, Katti BK (2009) Dynamic trip distribution of workers for a metropolitan city through implicit gravity modeling. *Urban Transp J Inst Urban Transp (India)* 9:2
11. Joshi GJ, Mokul D (2010) Mode shift behavior study of low income group travelers' with respect to Patna bus rapid transit system. 9th International Workshop on Transportation Planning and Implementation Methodologies for Developing Countries held at IIT Mumbai
12. Desai N, Vashi BD (2007) Optimization & privatization of city bus network using GIS: a case study of Vadodara city in Gujarat state. *Indian Road Congress (IRC)*, vol 69, Part-2, Paper No: 543
13. Desai N, Joshi GJ (2013) Geographical and cultural variations in factors affecting travel mode choice: an analysis of travelers' behavior. *Indian J Transp Manag* 37:1–20
14. Alpizar F, Carlsson F (2001) Policy implications and analysis of the determinants of travel mode choice: an application of choice experiments to metropolitan Costa Rica. Working Papers in Economics no. 56, Department of Economics
15. Ho JS, Yahaya AS, Sadullah AF, Leong LV (2009) Segmentation of factors influencing car and motorcycle users in Klang Valley. In: *Proceedings of the Eastern Asia Society for Transportation Studies*, vol 7
16. Adjemian M, Williams J (2009) Using census aggregates to proxy for household characteristics: an application to vehicle ownership. *Transportation* 36:223–241
17. Dillman DA (2007) Mail and internet surveys: the tailored design method—update with new internet, visual, and mixed-mode guide. Wiley, New York
18. Olsen SB (2009) Choosing between internet and mail survey modes for choice experiment surveys considering non-market goods. *Environ Resour Econ* 44:591–610 (**Springer Science + Business Media B.V.**)
19. Israel GD (1992) Determining sample size. Program Evaluation and Organizational Development, IFAS, University of Florida. PEOD-6. National Science Foundation, Research and Development in Industry: 1992, NSF 95-324
20. Bartlett JE, Kotlik JW, Higgins CC (2001) Organizational research: determining appropriate sample size in survey research. *Inf Technol Learn Perform J* 19(1):43–50
21. Venigalla M (2001) Household travel survey data fusion issues. National Household Travel Survey Conference: Understanding Our Nation's Travel, Transportation Research Board
22. Taylor-Powell E (1998) Sampling program development and evaluation. University of Wisconsin Extension
23. Yoshida T, Shibasaki R, Abe K, Onodera H (2009) Future scenario of international economics based on questionnaire survey by the Delphi method. In: *Proceedings of the Eastern Asia Society for Transportation Studies*, vol 7
24. Watson J (2001) How to determine a sample size: Tipsheet-60. Penn State Cooperative Extension, University Park, PA
25. Morikawa T, Ben-Akiva M, McFadden DL (2002) Discrete choice models incorporating revealed preferences and psychometric data. *Econometric models in marketing*. Elsevier, Amsterdam, pp 29–55
26. Mohamad J, Kiggundu AT (2007) The rise of the private car in Kuala Lumpur, Malaysia—assessing the policy options. *IATSS Res* 31(1):69–77
27. Abdul Kadir IS, Radin RS, Haron K, Hariza MSA (2006) Mode choice model for vulnerable motorcyclist in Malaysia. *J Traffic Inj Prev* 7:1–5 (**Taylor & Francis**)
28. Kamba AN, Raok Rahmat, Ismail A (2007) Why do people use their cars: a case study in Malaysia. *J Soc Sci* 3(3):117–122
29. Johansson MV, Heldt T, Johansson P (2006) The effects of attitudes and personality traits on mode choice. *Transp Res Part A* 40:507–525
30. Johansson MV, Heldt T, Johansson P (2005) Latent variables in a travel mode choice model: attitudinal and behavioral indicator variables. Working paper, Department of Economics, Uppsala University
31. Ben-Akiva M, Walker J, Bernardino AT, Gopinath DA, Morikawa T, Polydoropoulou A (2002) Integration of choice and latent variable models. In: Mahmassani HS (ed) *In perpetual motion: travel behavior research opportunities and application challenges*. Elsevier, Amsterdam, pp 431–470
32. Scheiner J, Holz-Rau C (2010) Travel mode choice: affected by objective or subjective determinants. *Transportation* 34(4):487–511
33. Yang C-W, Duann L-S, Fang W-P, Chen Y-C (2003) Intercity bus choice models with choice set generation and heterogeneity. *J East Asia Soc Transp Stud* 5:1518–1533
34. Temme D, Paulssen M, Dannewald T (2008) Incorporating latent variables into discrete choice models—a simultaneous estimation approach using SEM software. *Bus Res* 1(2):220–237
35. Schussler N, Axhausen KW (2011) Investigating the influence of the environmentalism and variety seeking on mode choice. Working Paper in *Journal of Transport and Spatial Planning*, paper no 706, pp 1–28
36. Koppelman FS, Bhat C (2006) A self instructing course in mode choice modeling: multinomial and nested logit models. U.S. Department of Transportation Federal Transit Administration, Washington DC