DEVELOPING MODE CHOICE MODEL FOR REGULAR TRIPS BASED ON PILOT STUDY

Ahmad Javeed Faizi,
M.Tech student, Department of Civil Engineering,
Parul University,
Vadodara, Gujarat, India.

Dr. Jayesh Juremalani,
Assistant Professor, Department of Civil Engineering,
Parul University,
Vadodara, Gujarat, India.

Abstract— **The lack of an appropriate transportation planning in non-metropolitan cities of India, such as Vadodara city of Gujarat state leads to deficiency of proper transportation planning policies. To mitigate the transportation problems causing by urbanization and rapid growth of population, the mode choice model is one of the best option to provide a convincible solution for mix traffic cities. Many researchers used different models namely: logit model, probit model, extreme value model and etc for different cities in the world to predict the future demand. Majority of them focused on generic parameters such as travel cost and travel time for theirs models. This paper is an attempt made to evaluate the utilizing modes comparison for work trips in Vadodara city to develop the model for their maximum utilization of trip makers for workplaces. The aim is to predict mode choice model for work trips based on initial questionnaire (pilot study) data. The data for this study was considered 45 sample sizes from entire 400 samples as already calculated based on the population of Vadodara city. The sample was collected through questionnaire by conducting fact to face interview with commuters in ward ninth Navjivan, Ajwa Road, Vadodara city, Gujarat state, India. The questionnaire was designed for variables of personal characteristics such as age, gender, occupation, car ownership, household size, driving license, and trip characteristics such as travel time, travel cost, travel distance. Private car, two wheelers, shared auto, public bus, bicycle and walking modes were considered for the model. Commuters were asked for the preference mode of travel that 57.8% respondents prefer to two wheelers, 20% respondents to private car the remaining percentage was for other vehicles. In this paper an attempt is made to predict mode choice model using multinomial logit model. Statistical package for social science (SPSS) software was used for data analysis model. The estimation of utility function coefficients were calculated through regression by SPSS software. Model variables were evaluated using T-statistics at the confidence level of 95% and the variables which T-value are larger than critical t-value (1.96) was excluded from model and null hypothesis was rejected.**

I.   INTRODUCTION

The world has rapidly led to higher growth rates and income, this can also increase the demand of vehicle and mobility. Increasing the number of vehicles in cities will definitely cause some transportation problems such as congestion, accidents, air pollution, delay, economic loss and disruption of traffic. To alleviate these problems travel demand modeling is the fundamental solution. Mode choice modeling is probably one of the prime important models of travel demand modeling in the transportation planning [12].mode choice modeling is important for predicting the trip maker's mode choice behavior and also for the factors affect the traveler mode. Mode choice model itself divided into two approaches aggregate and disaggregates models. Aggregate mode choice model focuses on the average individuals of trips or the entire zone trips but disaggregate mode choice model considers the individual trip behavior and its functions characteristics of available alternatives [4].There are three types of disaggregate mode choice models namely: logit model, probit model, extreme value model. Among these types, logit model is widely used for calibrating the mode choice model .In addition researcher preferred for its simple mathematical calculations and accurate model formulations as compared to other two types. The mathematical framework of logit model is based on the theory of utility maximization hypothesis. Hypothesis means that individuals select their mode of travel at maximize utility [1].The choice of a mode depends on many factors such as: socio-economics characteristics like age ,gender, income, occupation, driving license car ownership, household size, residency and trip characteristics like travel time ,travel cost, distance etc. Moreover there are also factors that may affect the mode choice such as Reliability, comfort, safety, dust and noise which refer to the level of service factor [2]. Pilot study can be defined as small study to test research protocols, data collections instruments and other techniques in preparation for the large study. Pilot study is one of the important stages to identify the problems areas and deficiencies before the implementation of full study of research. The importance of

pilot study can be much more, some of them can be mentioned here briefly: the feasibility of study protocol and to identify the weakness of study, it can test that whether our intended questions are comprehensive and appropriate for the target population, to find out the fair technique for the data collection among the available techniques and also the processes of data analysis can all be cleared by performing the pilot study before taking decision for the main research work. The sample size for the pilot study should be minimum the 10% of the main study sample size (Zailinawati Abu Hassan, Peter Schattner, et al).

II. LITERATURE REVIEW

From the past few decades, many researchers have developed and analyzed of mode choice models with different methods, approaches and considered different factors as per their conditions .few of them were studied as background study of this paper.

Essam Almasri, Sadi Alraee. (2013) studied factors affecting mode choice model for work trips in developing cities, they have done their work in Gaza city which a developing city in Palestine country. They collected the data through initial (pilot) and final questionnaire, socio-economics attributes, trip characteristics and Likert scale data was collected. 552 questionnaires were randomly select from 700 and two by three was used for the model calibration and one by three was kept for the validation of model. They developed multinomial logit model based on utility function at confidence level of 95%.They evaluated the models by likelihood ratio test.

Elharoun M. et al. (2018) developed mode choice model for Mansoura city in Egypt .They collected the data online through Google form and considered some variables such as travel time, travel cost, driving license, residence status, occupation, income, gender. Microbus, taxi, private care, motorcycle, walking were their modes. Their model was 85% accurate. Their calculation shows that the probability of microbus mode using decrease by increasing its fare ,while the private and taxi modes using probability is more as compared to microbus. In addition, they found that the increase in microbus fare would increase the probability of people walking.

Ling Ding and Ning Zhang. (2006) purpose was to predict a mode choice model with individual grouping based on cluster analysis. They investigated for two travel mode transit and car. To collect data, they conducted preference survey in city of Nanjing city of China to collect traveler's socio-economics and trip information. The estimation of utility function coefficient were calculated through regression based on the SP/RP survey information. Finally they found that the accuracy rate of mode choice estimation using individual grouping is remarkable higher than without grouping.

Arnab Jana and Varun Varghese. (2016) analyzed mode choice model for inter-regional trips in India. They determined that residence is facilitated to travel by the public transport, primarily buses and trains, with air travel being mostly popular in the affluent class in India. According to their search, that long distance travel is affected by various personal information, trip attributes, purpose of trip and destination being traveled to .They created binary logistic mode choice model. Their model showcased that increase the companions will directly effect on the choice of private vehicle in urban area. However the probability of public transport using is higher in rural areas. Juremalani J. R. et al. (2017) developed mode choice model for work trips by using artificial intelligent techniques. Their attempt is made to develop mode choice behavior predicting model using MNL and Fuzzy logic model. They considered 252 sample size and four modes namely: two wheelers, car, shared auto, bus. Travel distance, travel cost, travel time are considered as trip characteristics variables and gender, income, occupation ,and car ownership were considered for personal information. The author used chi-square of independency to find out the dependency of categorical variables and Pearson test was applied on scale data to determine the correlation. Finally they observed that fuzzy logic model gives better prediction accuracy for mode choice.

III. STUDY AREA AND DATA COLLECTION

Vadodara is formerly known as Baroda and is the largest city of Gujarat state of India after Ahmedabad and Surat. As per data released by Govt. of India for Census 2011, Vadodara city has population of 1,670,806; according to this population the sample size would around 400 for work trips mode prediction at 95% confidence level and 5 % marginal error of data. As our study is concerned with pilot study, so we will take out 11 % of whole sample that will give us 45 sample sizes. Questionnaire was initially designed with trip and personal characteristics distributed in Navjivan, Ajwa Road which called the ward ninth of the city ,this ward is the populous ward in the city with 2.4lakh residents and 55% male and 11% female population are working population as per the information of Vadodara municipality corporation. Respondents were asked through face to face interview.

IV. GENERAL DATA ANALYSIS

The person Socio -economics characteristics such as age, gender, monthly personal income, vehicle ownership, driving license, occupation and preference mode were preliminary analyzed. The data set of preliminary descriptive statistics is given in the table 1 and table 2. Table 1 shows the descriptive individual profile of ward No .9 of Vadodara city, while Table2 shows the details of preference mode of individual trips as per the collected data.

Table-1 Descriptive statistics of respondents in ward No 9

Socio-characteristics	Category	Frequency	Percentage
Gender	Male	35	77.8
	female	10	22.2
Income	0	9	20.0
	0-10000	10	22.2
	10000-20000	22	48.9
	20000-30000	2	4.4
	40000-50000	1	2.2
	>60000	1	2.2
Family size	1-3	16	35.5
	4-6	23	51.2
	7-9	5	11.1
	10-12	1	2.2
Car ownership	Two wheelers	30	66.7
	car	2	4.4
	bicycle	3	6.7
	None	10	22.2
Driving license	yes	28	62.2
	No	17	37.8
occupation	Gov. employee	2	4.4
	Self-employee	18	40.0
	Student	5	11.1
	housewife	1	2.2
	Daily wage	1	2.2
	private employee	8	17.8
	None	10	22.2

Table-2 The preferred mode of travelers in Ward No.9

Preferred mode	Frequency	Percentage (%)
Private Car	9	20.0
Public Bus	3	6.7
Auto Rickshaw	5	11.1
Bicycle	1	2.2
Walking	1	2.2
Two Wheelers	26	57.8
Total	45	100.0

Chi–square test of independency was applied on socio – economics factors of categorical variable such as gender, occupation, driving license, and income and care ownership to find out the dependency of one variable on other. The result of test is chi square value, degree of freedom, and significance p value which is less than 0.05 at confidence level of 95%, this means that these variables are affecting on the mode choice behavior of travelers.

Table -3 chi square independency test of mode choice of respondents

Variable	Chi square test			Likelihood Ratio		
	Value	df	Assumption of significance(2-sided)	Value	df	Assumption of significance(2-sided)
Gender	5.679	1	0.017	5.571	1	0.018
Driving license	5.679	1	0.017	5.571	1	0.018
Occupation	17.357	6	0.005	18.48	6	0.008
Car ownership	22.311	3	0.000	26.04	3	0.000
Income	12.734	5	0.026	13.71	5	0.018

The correlations between trip characteristics variables such as travel cost (fare), travel time and distance were checked by Pearson test using SPSS software. This test is used for scale data and is one of the useful test to correlate two variables. The coefficients show that trip variable have strong relationship and affecting the mode choice behavior of commuters.

Table-4 The correlation coefficients of travel time with respect to other variables(Fare,Dist)

Mode	Variable	Travel time	Sig (2-sided)	Correlation
Car	Travel cost	0.972	0.000	Strong
	Distance	0.895	0.000	Strong
	Travel time	1	Nil	perfect
Two wheelers	Travel cost	0.749	0.000	Strong
	Distance	0.884	0.000	Strong
	Travel time	1	Nil	perfect
Auto rickshaw	Travel cost	0.920	0.000	Strong
	Distance	0.949	0.000	Strong
	Travel time	1	Nil	perfect
Public bus	Travel cost	1	0.000	perfect
	Distance	1	0.000	perfect
	Travel time	1	Nil	perfect
Bicycle	Travel cost	0.718	0.000	Strong
	Distance	0.297	0.01	Weak
	Travel time	1	Nil	perfect
Walking	Distance	0.693	0.000	Strong
	Travel time	1	Nil	Perfect

As the above table values show the significance level of data less than 0.05 at confidence interval of 95% so we can proceed for the model.

V. MODEL DEVELOPMENT

The first model (M1) in this paper is developed for the specific variables, these include travel time, travel cost and distance of each mode. By considering the components of utility function,

a specific constant are required that reflects the utility between other alternatives. The two wheelers (motorbike) is considered as reference mode when adding constant. The following is the empirical utility function form.

U auto = Constant + β_1(auto TT) + β_2 (auto COST) + β_3(auto DIST)

U bus = Constant + β_1(bus TT) + β_2 (bus COST) + β_3(bus DIST)

U car = Constant + β_1(Car TT) + β_2 (Car COST) + β_3(Car DIST)

U bike = Constant + β_1(bike TT) + β_2 (bike COST) + β_3(bike DIST)

Table-5 Estimation Result of model one (M1)

Model		M1		
Parameters	Mode	Estimated coefficients	T-statistics	Sig
Constant	Car	0.00	-	0.000
Constant	Motorbike	0.239	3.811	0.000
Constant	Auto rickshaw	0.026	0.599	0.553
Constant	Bus	0.00	0.000	1.00
Constant	Bicycle	-0.014	-0.366	0.716
Constant	Walking	0.121	0.859	0.395
Travel cost	Car	0	-	0.000
Travel cost	Motorbike	0.007	3.613	0.001
Travel cost	Auto rickshaw	0.080	9.861	0.000
Travel cost	Bus	0. 400	0.000	0.000
Travel cost	Bicycle	0.365	12.802	0.000
Travel time	Car	0.072	-	0.000
Travel time	Motorbike	0.029	3.200	0.003
Travel time	Auto rickshaw	0.280	11.240	0.000
Travel time	Bus	-	-	-
Travel time	Bicycle	0.077	13.057	0.000
Travel time	Walking	0.205	8.980	0.000
Distance	Car	-0.083	-	
Distance	Motorbike	-0.033	-3.047	-.055
Distance	Auto rickshaw	-0.669	-8.700	0.000
Distance	Bus	-	-	-
Distance	Bicycle	0.917	20.878	0.000
Distance	Walking	0.097	2.541	0.015

According to the values of above table 5 ,the T statistics values for each variable shows that most of the given result is larger than critical t- value at 95 % confidence level (1.96).Therefor, the null hypothesis is rejected ,it means that these variables have no effect on utility mode. To improve the model other variables were added to utility function of modes.

Table.6: Estimation Result of final model (M2)

Mode	Variable	Estimated coefficient	t-statistics
Two wheeler	(Constant)	.801	3.495
	Gender	-.030	-.314
	size of family	.015	.772
	monthly income	-.062	-2.087
	Car ownership	-.229	-4.624
	Driving licensing	.006	.058
	occupation	.004	.237
	Age	.002	.941
	Travel time	.012	1.442
	Distance	-.015	-1.557
	Travel cost	.004	2.721
Auto rickshaw	(Constant)	-.513	-2.117
	Gender	.288	2.474
	size of family	.030	1.154
	monthly income	-.004	-.108
	Car ownership	.042	.837
	Driving licensing	-.092	-.695
	occupation	.006	.271
	Age	.002	.804
	Travel time	.283	10.061
	Distance	-.680	-7.846
	Travel cost	.081	9.185
Car	(Constant)	-.185	-1.478
	Gender	.108	1.780
	size of family	-.011	-0.867
	monthly income	.106	5.228
	Car ownership	.033	1.324
	Driving licensing	-0.068	-.981
	occupation	0.001	0.053
	Age	0.00	-0.249

Mode	Variable	Coefficient	t-value
	Travel time	.018	14.853
	Distance	-	-
	Travel cost	-	-
Bus	(Constant)	-0.635	-1.193
	Gender	0.362	1.416
	size of family	0.034	.632
	monthly income	-0.019	-.234
	Car ownership	0.142	1.348
	occupation	0.038	.848
	Age	-0.002	-.302
	Travel time	-	-
	Distance	-	-
	Travel cost	-	-
Bicycle	(Constant)	-.268	-1.153
	Gender	.014	.127
	size of family	-.014	-.631
	monthly income	.019	.520
	Car ownership	-.021	-.425
	Driving licensing	.107	.845
	occupation	.012	.612
	Age	.003	1.119
	Travel time	.082	11.622
	Distance	.897	18.603
	Travel cost	.361	11.171
walking	(Constant)	.750	.918
	Gender	-.689	-1.736
	size of family	-.058	-.725
	monthly income	-.026	-.217
	Car ownership	.145	.861
	Driving licensing	.450	1.013
	occupation	-.083	-1.239
	Age	.000	.017
	Travel time	.179	6.809

Distance	.107	2.561

U bike = 0.801 – 0.030(Gender) – 0.015 (family size) – 0.062(income) – .229 (Car ownership) –.006 (Driving license) +.004 (occupation) +.002 (age) +.012 (TT) –.015 (TD) +.004 (TC)   (1)

U auto = – 0.513 +0.288 (Gender) + 0.030 (family size) – .004(income) +.042 (Car ownership) –.092 (Driving license) +.006 (occupation) +.002 (age) +.0238 (TT) –.680 (TD) +.081 (TC)   (2)

U car = – 0.185 + 0.108(Gender) – 0.011 (family size) + .106(income) +.033 (Car ownership) –.068 (Driving license) – .001 (occupation) +018 (TT)   (3)

U bus = – 0.635 + 0.362(Gender) +0.034 (family size) – .019(income) +.142 (Car ownership) +.038 (occupation) –.002 (age)   (4)

U bicycle = – 0.268 + 0.014 Gender) – 0.014 (family size) + .019(income) –.021 (Car ownership) +.107 (Driving license) +.012 (occupation) +.003 (age) +.082 (TT) –.897 (TD) –.361 (TC)   (5)

U walking = 0.750 – 0.689(Gender) – 0.058 (family size) – .026(income) –.145 (Car ownership) –.450 (Driving license) – .083 (occupation) +.179 (TT) –.107 (TD)   (6)

VI.   EVALUATION OF MODEL

According to the above table 6 estimated coefficients and t-statistics values, we can evaluate the models using two tests, the first test is logicality of model variables. Each variable appears with its expected sign indicating positive or negative impact of variable on the utility function model. We can evaluate the sign impact on each mode utility based on two wheeler. The second test is the T-value of each variable coefficient with the t –value critical at 95% confidence interval (1.96).The mode variable which has larger t-value than critical value. Therefore, the null hypothesis is rejected, which means these variables have no effect on modes utilities.

VII.   MODEL RESULT

Estimation result of model (M1) in the model development section for generic variables (travel time, travel cost) t –statistics shows that these variables are not only affecting the mode choice in the city .However there are some other variables such as socio economics factors that can also effect. on mode choice. According to the estimation result of model (M2) shows that majority of t-values are smaller than T critical (1.96) value, this means, these variables can effect on the mode choice model. In this case we can accept our null hypothesis and reject our alternative hypothesis the model. The mode preference result of respondents were 57.8% for two wheeler, 20% for private car, 6.7% for public bus,11.1% for auto rickshaw and 2.2% for walking and bicycle modes as shown in the figure 1.The probability of each mode on the basis of the above utility function estimated in the following table 7. If we see the estimation result of the each mode according to its

utility value that obtained from mean values of each variable by SPSS will give us satisfactory result. This means that the observed and the predicted result match as per the developed utility function. The developed model is valid with condition of the Vadodara city for the maximum utilization of commuters. impact of variable on the utility function model. We can evaluate the sign impact on each mode utility based on two wheeler. The second test is the T-value of each variable coefficient with the t –value critical at 95% confidence interval (1.96).The mode variable which has larger t-value than critical value. Therefore, the null hypothesis is rejected, which means these variables have no effect on modes utilities.

Table-7 prediction of modes

Mode	Utility Value	e^u	Probability	% Probability	Predicted	Observed
Bike	3.98	53.51	0.59	59.08	26.59	26
Auto	2.3	9.974	0.11	11.0	4.95	5
Bus	1.68	5.35	0.059	5.91	2.66	3
Car	2.88	17.81	0.196	19.66	8.85	9
bicycle	0.65	1.915	0.021	2.11	0.95	1
walking	0.69	1.99	0.022	2.20	0.99	1
Total		90.56	1	100	45	45

Fig.1:   Prediction of mode (%)
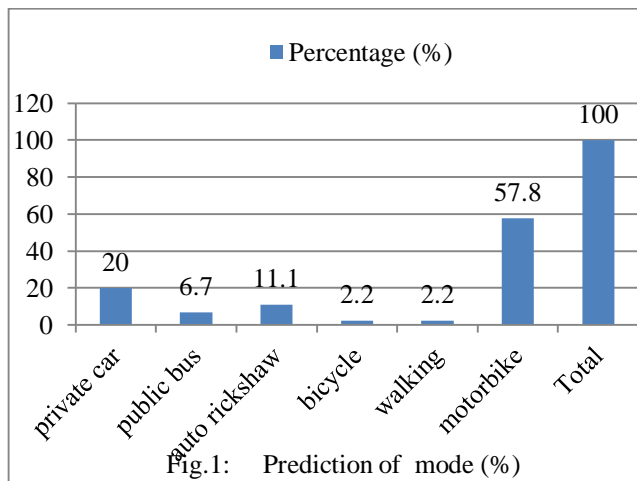
VIII.   CONCLUSION

The purpose of this paper is to predict the mode choice model for work trips in Vadodara city based on the pilot survey data. The 45 sample size collected through questionnaire by face to face interview. Respondent's personal and trip related data was analyzed through SPSS software. The utility function model was created for considered modes namely: private car, two wheelers, bus, shared auto, bicycle and waking through regression in SPSS. The model variables were checked statistically based on T-value statistics at confidence level of 95%. Some of variables such as age were not statistically important for car and walking modes because SPSS counted coefficient zero. The two wheelers mode was preferred by respondents 57.8%. As many researchers developed models for mode choice behavior using categorical and scale variables, it is better to consider some qualitative parameters such as comfort, safety in model for our main model development in future.

IX.   REFERENCES

[1] Essam Almarsri,Sadi Alraee.(2013).Factors affecting mode choice of work trips in developing cities –Gaza as a case study,(pp.247-259).

[2] Hasan M,Al-Ahmadi.(2006). Development of intercity mode choice models for Saudi Arabia, (pp.3-21).

[3] Ling Ding, Ning Zhang.(2006). A travel mode choice model using individual grouping based on cluster analysis, (pp786-795).

[4] Marwa Elharoun,Usama Elrawy Shahdah,Sherif M.El-Badawy. (2018). Developing a mode choice model for Mansoura city in Egypt,(pp.528-542).

[5] Nadezda Zenina,Arkady Borisov. (2011). Transportation mode choice analysis based on classification methods, (pp.49-56).

[6] Ortuzar.J,Willumsen.L.G.(2011). Modelling transport, ,(pp.607-611).

[7] Ram Chanda,Saparshi Sen, Sudip Kumar Roy. (2016) . Mode choice modeling of work trips: A case study of Kolkata, (DOI: 10.15224/ 978-1-63248-114-6-20).

[8] Richarson.A.J.(2003).Creative thinking about transportation planning,(pp.53-58).

[9] Sreerag SR,S.N.Sachdeva,Shri.S.Shameem.(2016). Mode choice modeling for work trips in Thiruvanthruram city,(pp.381-384).

[10] Sven Muller, Stefan Tscharaktschiew, Knut Haase. (2008).Travel to school mode choice modeling and patterns of school choice in urban areas, (pp.342-357).

[11] Jayesh Juremalani, Krupesh A. Chauhan.(2017). comparison of different mode choice models for work trips using data mining process,(pp.124-132).

[12] Minal,Ch.RaviSekhar. (2014). Mode choice analysis: the data,the modes and future ahead,(pp.269-285).

[13] Zailinawati.A.H,Peter.S,Danielle.M.(2006). Doing a pilot study: why is it essential?,(pp.70-74).

[14] B. Prasad,Kumar Molugaram.(2018). Development of mode choice model of a trip maker for Hyderabad metropolitan city, (pp.1-7).

[15] Arnab Janaa, Varun Varghesea. (2016). Analyzing mode choice for inter-regional travel in India ,(pp. 5220–5234).

[16] Faizi Ahmad Javeed,Jayesh Juremalani,Sanket Bagadia.(2019). Critical apprisal of mode mhoice for work trips   ,(pp.1766-   1770