

# Mode choice Behaviour Analysis of Students in Thrissur city

Jerry Soman<sup>1</sup>, Vincy Verghese<sup>2</sup>

<sup>1</sup>PG scholar, Dept of Civil Engineering, JECC Cheruthuruthy, Kerala, India

<sup>2</sup>Assistant Professor, Dept of Civil Engineering, JECC Cheruthuruthy, Kerala, India

\*\*\*

**Abstract** – Travel behaviour plays an important role in transportation planning. Mode choice analysis is the process of arriving at a decision about which mode to use under a set of circumstances. Choice of a particular mode by a student affects the general efficiency of travel within the city. Main objective of the study is to find out the travel pattern, identify the factors to choose the mode and predict the model. Various techniques are available for mode choice modelling. Multinomial logit models are found to be efficient in estimating the different mode shares in a scenario where more than two choices of modes of travel are available for a student. Multinomial logistic regression is to analyse the mode choice behaviour of students in Thrissur city. SPSS is used to develop the model for different modes.

**Keywords:** Travel behaviour, SPSS, Multinomial logit model

## 1. INTRODUCTION

Transportation modellings plays an vital role in planning. Major roles of transportation modelling are forecasting travel demand based on variations in the transportation planning system. Different types of models are there that have been developed to create actual travel patterns of people and existing conditions. Models are used to predict changes in travel pattern and utility of the transportation system. Thrissur city is the third largest urban agglomeration in Kerala about 1,854,783 population and the 20th largest in India. There are several number of schools within the city and outside the city. Mode choice behaviour of students in urban and rural area is different. The sex ratio of male and female is 1092:1000. Development of models in rural and urban region helps the transportation planners to add in transport policies. And after work trips the second most occurrence of congestion are school trips at the peak time. Forecasting will help to design the transport planning systems and considering the travel behaviour of the study area. It also helps to develop a system that can accommodate the travel demand for the future. The problem faced is incomplete information that makes the uncertain conditions. If we need to reach on predictability; it can be achieved through the probability of individual decision and its leading characteristics. However, private automobiles remain the predominant school travel mode (45.3% mode share) followed by school buses (39.4 %), while walking or bicycling together account for only 12.7% of the total school transportation mode share in the USA as in 2000. In addition to personal and social factors such as parents' education,

income and attitudes, children's age, race, gender and attitudes, and peer influence studies have found many built environmental factors associated with school travel behaviours. Environmental approaches to promote walking to or from school have become increasingly popular with the recognition that a safe environment is a prerequisite to any promotional efforts. Further, environmental improvements are relatively permanent and can lead to population-level changes over time if successful. Among the environmental correlates, home-to-school distance has shown to be the most consistent and often the strongest predictor of school travel mode choice, followed by safety and weather. However, shortening the distance to school is not simple. It requires long-term, multilevel policy and environmental changes, such as school sitting, zoning, and land development policies.

## 1.1 Objectives

The main objectives of the study are:

- Study the mode choice behaviour of students
- Identifying the various factors that contribute to the selection of a particular mode in the city for educational trips.
- To develop a model to predict the mode choice behaviour of students Thrissur city for both rural and urban region

Before you begin to format your paper, first write and save the content as a separate text file. Keep your text and graphic files separate until after the text has been formatted and styled. Do not use hard tabs, and limit use of hard returns to only one return at the end of a paragraph. Do not add any kind of pagination anywhere in the paper. Do not number text heads-the template will do that for you.

Finally, complete content and organizational editing before formatting. Please take note of the following items when proofreading spelling and grammar:

## 2. LITERATURE SURVEY

The access travel characteristics are discussed with respect to mode use pattern, availability of access modes, satisfaction with the access environment and characteristics of the access leg of a trip in comparison to primary and egress legs. Acceptable trip lengths by walk and bicycle

modes in accessing transit stations are computed from their cumulative frequency distribution of trip lengths. The commuters are segmented into different socio economic subgroups using characteristics like occupation, type of house, vehicle availability and household income. The mean and acceptable walking and bicycling distances found for different subgroups. (Rajat Rastogi and K V Krishna Rao, 2003)

Matched home locations and different school travel modes were examined using conditional logistic regressions. Despite the same/similar objectively measured distance and home location, perceptions of distance, sidewalk and traffic conditions, park presence, and convenience of walking differed between walkers and automobile users. Parental attitudes and children's preferences were associated with the odds of walking. Safety concerns (traffic danger, stranger danger, and getting lost) were higher among drivers, but only significant in bivariate analyses. To promote walking to school, route/street improvements appear promising, but parallel educational and promotional efforts may be needed to address perceptual and attitudinal barriers. (Chanam lee and jeongjae yoon, 2013)

### 3. METHODOLOGY

The work intends to find a proper model for students trips to predict the future travel demand and finding important characteristics. It is aimed at identifying various factors that contribute to the selection of a particular mode in the city. MNL modelling was adopted in the study because of its capability in estimating the mode shares where more than two choices of modes of travel. The first phase is researching for topic and study literature, and then data was collected data of Thrissur city. Preliminary analysis is done to study the data. Student trips are distributed on different modes- Two wheeler, bus, car, three-wheeler, cycle and walk. Validation is the process of checking the model for its accuracy in prediction. The last phase is concluding the study. Numbers of studies were done under transportation planning and selectively on mode choice behaviour, integration of multi-modes, Intra-city level characteristics and attitudinal behaviour. To develop better model needs good data collection. As the data increases the model gives better output. The method used for data collection to development model for mode choice is questionnaire survey. Total number of student trips getting from the school is 1200. The available questionnaire survey data contains data of travel characteristics as well as traveller characteristics. From the data, the important variables were needed to identify. The considered travel characteristics are total travel time, travel cost, distance, vehicle ownership, gender, age and income. Model development is the most important step in this study and it is done by the help of Statistical package system software. The models used to study the mode choice behaviour are multinomial logit model. The common modelling method for mode choice are Multinomial Logit model (MNL) due to it leads to simpler, providing more

economic model and increasing the power to detect the relationships with other variables. The most preferred softwares used for multinomial logit models are SPSS (Statistical Package for the social sciences). The developed model should be validated.

### 4. DATA COLLECTION

Assuming the population to be normally distributed, empirical formulas given by Levy and Lemeshow (2008) [shown in Eqs. (2) and (3)] were used to determine the sample size.

$$n_0 = \frac{Z^2 pq}{e^2} \quad (2)$$

where  $n_0$  = sample size for infinite population;  $Z$  = statistical parameter corresponding to confidence level ( $Z$  is 1.96 for 95% confidence interval);  $e$  = desired margin of error (adopted as 5%);  $p$  = hypothesized true proportion for population (adopted as 0.5 to account for the worst case) and

$$q = 1 - p \quad n = \frac{n_0}{\left[1 + \frac{(n_0 - 1)}{N}\right]} \quad (3)$$

where  $n$  = sample size for finite population;  $N$  = population size.

For a very good model, study needs large quantity of data. In this a questionnaire form is prepared for students to find out their travel behaviour, factors which influence to select the mode. From the google map, names of the schools are got. From the number of schools, schools of mixed students are taken for the survey. Eliminate girls and boys only school from the list. Then select the schools which are in corporation ward. From that, schools are divided into urban schools and rural schools. After that from urban area, 4 schools are selected and from rural area, 4 schools are selected. From the selected schools 50 samples from each section (upper primary, high school and higher secondary) are collected. Total number of samples are 150 from each school. Urban area and rural area is determined by whether there is residential buildings, commercial buildings and highway. Four schools from urban area and four schools from rural area are randomly selected to collect the data. The questionnaire format is shown below. As the data increases model gives better output. Questionnaire format include travel characteristics, travellers characteristics. From the data important variables were needed to identify. Questionnaire surveys involve collecting the information regarding household characteristics, personal characteristics and the travel decisions made in the recent past.

### 5. DATA ANALYSIS

Stratification of the collected samples is discussed in the following sections. The collected data is entered in the Microsoft excel and the preliminary analysis is done. Data analyses are the process of identifying the right data for achieving the objective of the study. It is the process of obtaining raw data and converting into useful information and used for mode choice modelling which helps for decision making process. Data was collected through questionnaire survey and the objective is to find the most influencing variables on the mode choice behaviour of students and to develop a model of student mode choice in Thrissur City. The collected data is organised into a format given in the following tables to fit for analysis within a spread sheet for easy data processing. The data are divided into categories and coding is done for easy analysis. Table shows different categories of mode, gender, age, income, distance and vehicle ownership. The analysis was done by Microsoft excel using filtration tool. Preliminary analysis is done on the data. The data is presented graphically for gender wise mode selection, age wise mode selection, income wise mode selection, distance wise mode selection.

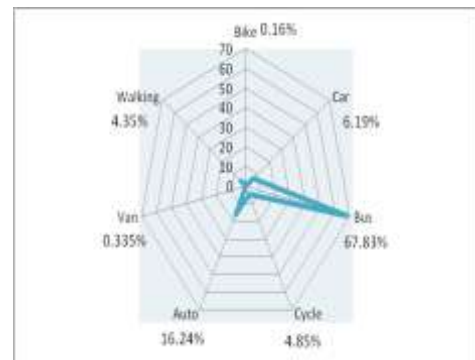


Fig -1: Mode split

Mode split data is the primary step of mode choice modelling. Most students prefers bus (67.83%) as their mode of choice. 4.85% of students are preferring cycle as their mode, three wheeler and van have 16.24% and 0.333%. walking and bike have 4.35% and 0.16%. from this graph we can clearly identified that more percentage of urban students are using bus as their mode and least percentage of mode choosing is for two wheeler.

Preliminary analysis of Rural region:

Table -1: Categories and coding pattern for the study

	Categories	Coding pattern
Mode	Two wheeler	1
	Car	2
	Bus	3
	Cycle	4
	Auto	5
	Van	6
	Walk	7
Gender (G)	Male	1
	Female	2
Age (A)	10-12	1
	13-15	2

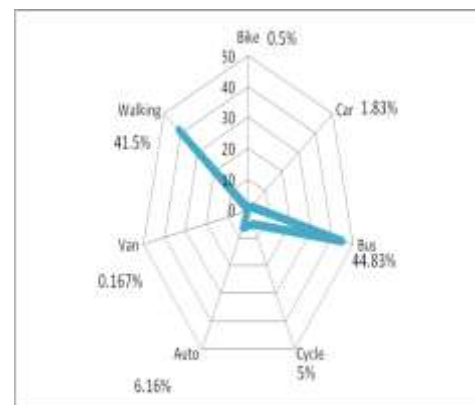


Fig -2: Mode split

Mode split data is the primary step of mode choice modelling. Most students prefers bus (44.83%) as their mode of choice. 5% of students are preferring cycle as their mode, three wheeler and van have 6.16% and 0.167%. walking and bike have 41.5% and 0.5%. from this graph we can clearly identified that more percentage of rural region students are using bus as their mode.

### 6. MODEL DEVELOPMENT

Mode choice modeling is the third important steps of the transportation planning after trip generation and trip distribution. Mode choice modelling helps to find the traveller's choice of mode. For modeling the mode choice, there is need of proper analysis of the data. Modelling predicts how road users will split themselves between transits and other vehicles. Mode choice modeling is to be done by identifying different modes, identifying the

attributes and characteristics and constructing an experimental design that is appropriate for those attributes and characteristics. Before doing the model correlation of variables have to be found out. When using MNL regression, one category of the dependent variable is chosen as the reference category. All parameters in the model are interpreted with reference to it. It is the standard/ reference category to which others (response category) would naturally be compared. The coefficients are estimated through an iterative maximum likelihood method. Odds ratios are determined for all independent variables for each category of the dependent variable except the reference category. Odds ratio represents the change in the odds of being in the dependent variable category versus the reference category associated with a one unit change in the independent variable. In analysis, the mode of conveyance is chosen as dependent variable, as the problem is concerned with choice of mode. In order to get the preference of bus relative to other modes, bus is chosen as the reference category. Bus was compared

It can be found that only travel time have significance less than 0.05. So we can conclude that the travel time used in the model have significant contribution toward predicting the mode choice behaviour of students in urban region.

From table variables of urban model MNL regression model of car in relation with bus, the factor travel time has a positive coefficient of 0.780, which means increase in travel time, students are choosing car over bus as their mode of travel. The odds ratio obtained in the output of travel time is 2.182, indicating that for each unit of increase in travel time, the mode choice of students increase by 2.182 times. In MNL regression model of auto in relation with bus, the factor travel time has a positive coefficient of 0.663, which means increase in travel time, students are choosing auto over bus as their mode of travel. The odds ratio obtained in the output of travel time is 1.940, indicating that for each unit of increase in travel time, the mode choice of students increase by 1.940 times. The factor reason=1 has a negative coefficient of 1.931, which means students are choosing auto over bus as their mode of travel because the reason comfort. The odds ratio obtained in the output of reason is 0.145, indicating that students are lesser comfort with the mode auto by 0.145 times over bus. In MNL regression model of walk in relation with bus the factor travel time has a negative coefficient of 0.723, which means decrease in travel time, students are choosing walk over bus as their mode of travel. The odds ratio obtained in the output of travel time is 2.062, indicating that for each unit of decrease in travel time, the mode choice of students decrease by 2.062 times.

**Table -2:** The likelihood ratio test for Urban region

Effect	Model Fitting Criteria	Likelihood Ratio Tests		
		Chi-square	df	Sig
Intercept	374.224 <sup>a</sup>	0.000	0	
Income	380.207 <sup>b</sup>	5.982	5	0.308
Distance	378.265	4.041	5	0.544
Travel Time	396.847	22.622	5	0.000
Vehicle ownership	375.159 <sup>b</sup>	0.934	5	0.968
Traffic	377.164	2.939	5	0.709
Comfort	378.376 <sup>b</sup>	4.151	5	0.528
Level Of Education	380.134	5.910	5	0.315
Reason	397.175	22.951	20	0.291

with car and two-wheeler separately to understand the factors The independent variables chosen are age group, gender, income, vehicle ownership, distance, travel time, travel cost, level of education, reason,

Table 3 shows a statistical comparison of intercept-only model and final model. Intercept only describes a model that does not include any predictor variables and simply fits an intercept to predict the output variable. Final describes a model that includes the specified predictor variables and has been arrived at through an iterative process that maximizes the log likelihood of the outcomes seen in the output variable. By including the predictor variables and maximizing the log likelihood of the outcomes seen in the data, the final model is an improvement on the intercept-only model. The likelihood ratio test shows the contribution of each variable to the model. This fact is revealed by the statistical tests on the two models as given in Table. The chi-square statistic is the difference between the -2 log-likelihoods of the null/intercept-only and final models. Because the significance level of the test is less than 0.05, it can be concluded that the final model is outperforming the null model. The pseudo R<sup>2</sup> value of the final model according to Nagelkerke for urban model are 0.500 and pseudo R<sup>2</sup> value for rural model are 0.400. Pseudo R<sup>2</sup> value indicates the proportion of variance of the response variable explained by the predictors and its maximum value is 1. Larger pseudo R<sup>2</sup> statistics indicate that larger percentage of variation can be explained by the model.

Comfort level and traffic. The independent variables are classified into various categories. The likelihood ratio tests that show the contribution of each variable to the model of urban area are shown below.

**Table -3:** Values of likelihood ratio tests for Urban area:

Model	Model fitting criteria	Likelihood ratio tests		
	-2 log likelihood	Chi square	df	Sig
Intercept only	475.860			
Final	374.224	101.636	55	0.000

**Utility equation for Urban model**

$$U_{(Car)} = 0.78 * TT$$

$$U_{(Auto)} = 0.663 * TT - 1.931 * R_1$$

$$U_{(Walk)} = -0.723 * TT$$

**Table -4:** Variables for urban region

MODE <sup>a</sup>		B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence interval for Exp(B)	
								Lower Bound	Upper Bound
								Car	TT
Auto	TT	0.663	0.223	8.845	1	0.003	1.940	1.254	3.003
	R=1	-	0.957	4.074	1	0.044	0.145	0.022	0.946
Walking	TT	0.723	0.306	5.583	1	0.018	2.062	1.131	3.757

**6.1 validation**

From the 1200 data collected analysis is performed by taking 800 samples and remaining 400 are taken for validation. The procedure followed by validation is given below:

- i. Score of the utility model is found out
- ii. Found out the mode choice of students using this equation.

$$P(Y=1) = \frac{e^{score(Xi,1)}}{1 + \sum_{k=1}^{k-1} e^{score(Xi,k)}}$$

- iii. Maximum value of each mode is taken as predicted output
- iv. Compared the value of observed and predicted
- v. Determine the prediction percentage
- vi. prediction percentage for rural model is 72.5% and for urban model it is 42.85%

**7. CONCLUSIONS**

The study Mode Choice behaviour of students in Thrissur City|| aims at developing a model for school in Thrissur city using SPSS software. The study considered some selected modes only which are dominant. The main conclusions drawn from the study are:

- Factors influencing the mode choice of students in rural region are age, gender, level of education, distance ,travel time, comfort and travel cost
- Factors influencing the mode choice of students in urban region are income, distance, travel time, vehicle ownership, traffic, reason for choosing mode and comfort level.
- Common variables coming under urban and rural model are distance and travel time.
- Both rural and urban region most of the students are using bus as their mode of travel.

**REFERENCES**

- [1] Guerraa, E.,(2018).”Urban form, transit supply, and travel behavior in Latin America: Evidencefrom Mexico’s 100 largest urban areas”Transport Policy 69 ,98–105
- [2] V Sohonia,A., Thomasa,M.,(2017)”Mode shift behavior of commuters due to the introduction of new railtransit mode”Transportation Research Procedia , 2603–2618
- [3] K.M. Keyes,A., Crawford-Brown,D.,(2018) “The changing influences on commuting mode choice in urban England under Peak Car: A discrete choice modelling approach”Transportation Research Part F ,167–176
- [4] Roider,O.,Rudloff,C.,(2012),“an individualized travel information systemfor optimizing mode and route choice behaviour of commuters” Procedia - Social and Behavioral Sciences ,1948 – 1957
- [5] Xianyua,J.,(2013).”An exploration of the interdependencies between trip chaining behavior and travel mode choice”Procedia - Social and Behavioral Sciences,1967 – 1975
- [6] Ding,L., Zhang,N.,“A Travel Mode Choice Model Using Individual Grouping Based on Cluster Analysis”Procedia Engineering ,786 – 795

- [7] Kaewkluengkloma,R.,Satiennamb,W.,(2017),” Influence of psychological factors on mode choice behaviour: Case study of BRT in Khon Kaen City, Thailand”Transportation Research Procedia ,5072-5082
- [8] Sreerag ,S., Sachdeva,S,”Mode Choice Modelling for Work Trips in Thiruvanthpuram City”journalof Basic and Applied Engineering Research ,Vol 3, Issue 4
- [9] Verma,M., Manoj,M.,(2016),“Analysis of the influences of attitudinal factors on car ownership decisions among urban young adults in a developing country like India” Transportation Research Part F , 90-103