PUBLIC TRANSPORT CHOICE BEHAVIOUR IN THIMI BHAKTAPUR

Sarita Prajapati[1], Pradeep Kumar Shrestha[2]

[1]*Department of Civil Engineering, Khwopa Engineering College, Libali-8, Bhaktapur, Nepal*
[2]*Department of Civil Engineering, Institute of Engineering, Pulchowk Campus, Lalitpur, Nepal*

Abstract

The public transport users tend to shift public transport mode in search of a better one. The choice depends on the various attributes of public transport, such as travel time, waiting time, comfort, safety, etc. Since there were competition among different public modes, such as buses, minibuses, microbuses, and even three-wheeler (tempo), the sustainability of the public transport directly depends on the public or users. The users continue to use those public transport which meet their expected level of service. This study focuses on examining attributes of different public transport alternatives available, investigating factors affecting preference of public transport alternatives and developing choice model for public transport alternatives. Multinomial logistic model was developed from the SPSS to find the probability of selecting particular public transport. The public transport choice among the four alternatives, whose destination was same, was found to be influenced by the variables, such as monthly income, age, travel time, walking time, waiting time, fare and gender. The model helps to understand the choice behavior of traveler and can be used for transport planning in future.

Keywords: Public transport, Choice model, Travel behavior, Multinomial logistic model

1. Introduction

Road transport is one of the major land transportation systems that is used for the movement of people. Road transport system can be categorized as the public and private transport. Private transport system is the system which is not available for the general public. It is the system only used by the people for their personal use. The car, bus and motorcycle are motorized and bicycle, rickshaw and walking are non-motorized form of private transportation. Public transport system is a system that is shared by the public with specific fare. It generally operates on fixed routes and includes modes, such as three wheelers, mini/micro buses, buses, trains, trams taxi, and ferries. In the context of Nepal, the most commonly used public modes are mini/micro buses, large buses, Sajha bus, taxi, safa tempo, gas tempo, etc.

*Corresponding author: Sarita Prajapati
Department of Civil Engeneering, Khwopa Engineering College, Libali-8, Bhaktapur, Nepal
Email: prajapati.sarita@khec.edu.np
(Received: January 14,2020 Accepted: August 06, 2020)

Besides reducing congestion and air pollution by providing transportation services to a large number of people, high capacity public transport systems may also influence the urban form and quality of life in cities. A good public transport system makes efficient use of urban space, provides efficient and affordable mobility, and access to work, school/colleges, social, recreation and economic activities.

Large volume of daily trip in Kathmandu valley is still made by public modes. Out of approximately 3.4 million one-way person trips made each day, nearly 41% are made on foot, while almost 28% are made on public transport (MoPIT/JICA, 2012). Public transport vehicle represents less than 3% of total registered vehicle fleet in Kathmandu valley but their travel mode share is almost equal to that of private vehicles (cars and motorbikes) (DoTM, 2011). In the past 10 years, population in Kathmandu valley has increased by 4.32% per year and motorization has increased by 12% per year (CBS 2011; DoTM, 2011) while the modal share of public transport has remained stagnant (MoPIT/JICA, 2012).

Bus services are the only means of public transport in Kathmandu until the introduction of small vehicles such as 3 wheelers and microbuses as public transport mode. Now a days, the share of low occupancy vehicles, such as minibuses, microbuses and tempos operating within Kathmandu Valley accounts for 94% of total public transport vehicles; while share of large buses is only 6% (MoPIT/JICA, 2012). The popularity of microbuses, though they were introduced in 1998 as a replacement of outdated Vikram Tempos from Kathmandu, is surely due to its distinct pattern of service that are successful in attracting public transport users.

However, choice behavior of users among public transport modes is different. There is tendency of public transport users to switch between modes, which can be shown by recent increase in micro-buses in many Modes. Although they are public transport captive riders who do not have choice other than using public modes, they have numbers of the choices for among the public vehicles. That means people can select the public transport, such as mini bus, micro bus, bus running along the same mode. There are not many studies on factors affecting public satisfaction and preference among various pubic modes available for users in Kathmandu. Travel time factors, such as in-vehicle travel time, access time and egress time, waiting time, travel cost and number of transfers are considered to be important in mode choice. Similarly, Ben-Akiva and Morikawa (2002) found reliability, information availability, comfort, safety and security, availability are important factors. According to Gonzaga and Villoria (1999) stated that personal travel behavior is affected by income, age, number of workers, travel length and number of travel mode. Goulias (2003) stated that travel behavior research aims to understand how traveler values, norms, attitudes, and constraints lead to observed behavior.

This study mainly focused on studying mode choice behavior of public transport users among different public transport mode available to them. In the research, first of all the study of different attributes, such as size of vehicle, speed, fare etc. of different public transport mode available to users on the study area was done. Then the study was done for finding the different factors affecting mode preference of public mode. At the end, the development of mode choice model for public transport was done using the logit model.

2. Research Methodology

2.1. Study Area

The study area for the research was Naya-Thimi chowk, which lies in the Madhyapur Thimi municipality of the Bhaktapur district. There are 4 alternatives available which reach up to Kathmandu. The 4 alternatives were as follows: Alternative1: Large buses from Banepa to Kathmandu, Alternative2: Minibuses and buses from Kamalbinayak Bhaktapur to Bagbazar, Alternative3: Buses from Chyamhasingh Bhaktapur to Kathmandu and Alternative4: Micro-buses from Naya-Thimi to Kathmandu.

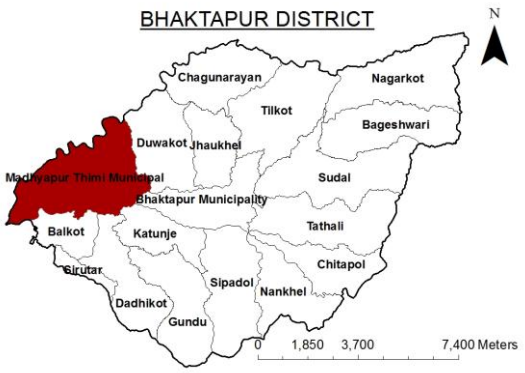

Fig. 1 Map of Bhaktapur District

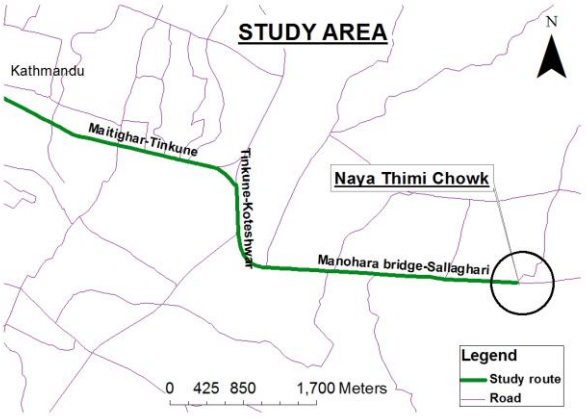

Fig. 2 Map of Study Area

2.2. Outline of Methodology

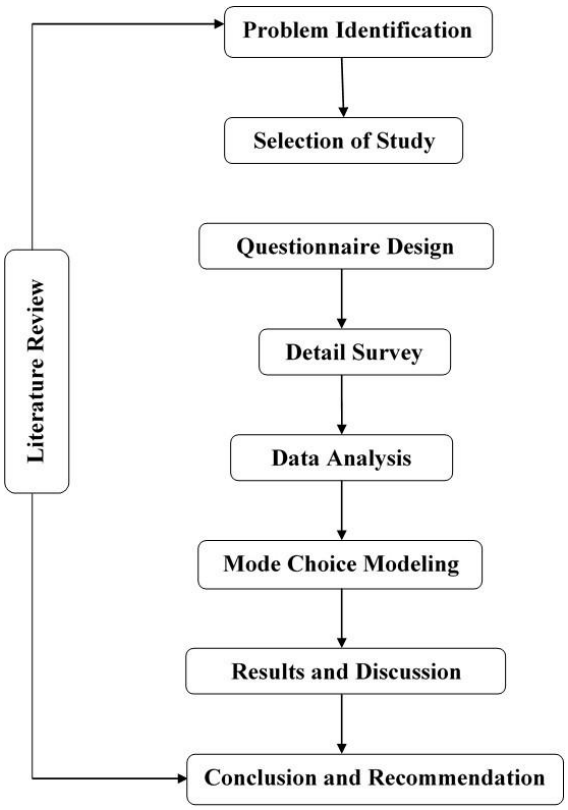

Fig. 3 Outline of methodology

2.3. Questionnaire Design

For the survey, questionnaire were prepared on the basis of the objectives of the research. In the questionnaire there were three parts.

Part I was socioeconomic information. In this section, socioeconomic characteristics such as gender, age group, number of family members, education level, occupation, monthly income(Rupees) were included.

Part II of the questionnaire was trip information. In this section trip characteristics, such as Trip purpose, trip origin and destination, facilities available at the origin and destination were included.

Part III of the questionnaire was public transport characteristics. In this section, attributes of the public transport were rated. The rating was done as 1 to 5(Not good to very good).

2.4 Sampling and Sample Size

Convenience sampling was used for the sampling. It (also known as availability sampling) is a specific type of non-probability sampling method that relies on data collection from population members who are conveniently available to participate in study. Convenience sampling is a type of sampling where the first available primary data source will be used for the research without additional requirements. In other words, this sampling method involves getting participants wherever you can find them and typically wherever is convenient.

The Slovin's formula is used for the calculation of the sample size. i.e:

$$n = \frac{N}{(1 + NE^2)}$$

where:

n = sample size

N = population size

E = Error tolerance

The total population for the study was calculated as the 40% population of the Madhyapur thimi municipality i.e.83,036 and 50% of the Dadhikot VDC population i.e. 11,629 (CBS,2011). Therefore, the total population was 39028.

Then the sample size was calculated as 203. But the questionnaire survey was conducted among 260 individuals. After filtering the incomplete or invalid questionnaires the final sample considered was 213.

2.5 Data Collection

Data collection were done by conducting roadside questionnaire survey which was done in peak hour (9:00 am to 10:00 am and 5:00 pm to 7:00 pm), off peak time, working days and holidays. Through the 3 observers the required data were collected though the questionnaire sheets.

2.6 Data Analysis

The surveyed data was managed and entered in excel and uncompleted answers were filtered from the data. The data analysis was done by using the SPSS and Excel. By using these applications the data was analyzed, and the results were presented from the graphs and tables. In SPSS Multi logistic regression was used for the modeling.

2.7 Choice Modeling

Choice modeling would do in the three steps, namely model formulation, model estimation and model validation.

2.7.1 Model Formulation

The formulation of the choice model needs extensive evaluation of observed data. The parameters such as gender (GEN), age (AGE), educational level (EDU), monthly income (INC), trip purpose (TRP), travel time (TRT), waiting time (WTT), walking time (WKT), fare (FAR), travel distance (TRD), safety (SAF), comfort (COM), behavior of driver and conductor (BDC), seat availability (SAV), reliability (REL), accessibility (ACE), vehicle condition (VCO), serviceability (SER) were considered for understanding the choice pattern of travelers.

To define the relations between characteristics and choice of the road users of, a multinomial logistic regression is applied to show the difference among the mode usage of road users. It is the simplest and most popular practical discrete choice model which is based on random utility theory. That is given as:

$$U_{in} = V_{in} + \epsilon_i$$

$$V_{in} = \sum_k \beta_{ik} x_{ikn}$$

where, ϵ_i = unobservable part.

x_{ikn} = the value of the characteristic k for individual n and the choice option i

β_{ik} = the coefficient associated with this characteristic. This coefficient can vary for each choice option (it varies with i) if it is specific. If it is generic, it does not vary with i and can simply be noted. These are the coefficients estimated by SPSS based on Multinomial logistic regression method.

Ben-Akiva and Lerman (1985) stated that the probability that the individual 'n' will choose 'i' as

$$P_{in} = \frac{e^{U_{in}}}{\sum_{j \in C_n} e^{U_{jn}}}$$

where,

$$0 \le P_{in} \le 1 \qquad for \quad all \quad i \in C_n$$

$$and \quad \sum_{i \in C_n} P_{in} = 1$$

2.9.2 Model Estimation

The coefficients were estimated by fitting the data to the model(s). The maximum likelihood estimation method was used. This method involves choosing values for the coefficients to maximize the likelihood (or probability) that the model will predict the same choices made by the observed individuals. Let N denote the sample size and define,

$$y_{in} = \begin{cases} 1 & if \ observation \ n \ chose \ alternative \ 'i' \\ 0 & otherwise \end{cases}$$

The likelihood function for a general multinomial choice model is

$$L^* = \prod_{n=1}^{N} \prod_{i \in C_n} P_{in}^{y_{in}}$$

where,

$$P_{in} = \frac{e^{U_{in}}}{\sum_{j \in C_n} e^{U_{jn}}}$$

The estimation of above function is done by Newton Rapshon Method. In this study SPSS software package will be used to solve it.

2.7.2 Model Validation

After the calibration process was completed, the developed models were validated. The models were validated using the data other than those were used for model calibration. The collected survey data was divided into two parts, the first part was used for model calibration whereas the second part was used for model validation. The validity of models was evaluated by comparing

similarities in the observed and predicted choices. Additionally, statistical validation will be done with following tests. i.e. Test of coefficient Estimates and The Goodness of Fit Measures.

3. Results and Discussions

3.1 Public transport Attributes

Table 1 Rating of public transport Attributes

Attributes of Public Transport	Alternative1	Alternative2	Alternative3	Alternative4
Safety	3.33	3.55	3.643	3.57
Comfort	3.41	3.18	3.79	3.33
Behaviour of drivers and conductors	2.81	2.88	3.28	3.34
Seat Availability	3.27	3.03	3.82	3.10
Reliability/ Punctuality	3.42	3.12	3.19	3.36
Accessibilty	3.23	3.27	3.16	3.49
Size of vehicle	4.62	3.73	4.01	3.08
Vehicle Condition	3.26	3.16	3.33	3.49
Service Quality	2.92	3.10	3.00	3.29

3.2 Analysis of Factor Affecting Preference among Public Transport Alternatives

3.2.1 Trip Purpose

The alternative2 and alternative3 were used by student at the same percentage, whereas the scenario of the alternative1 and alternative4 was also nearly equal.

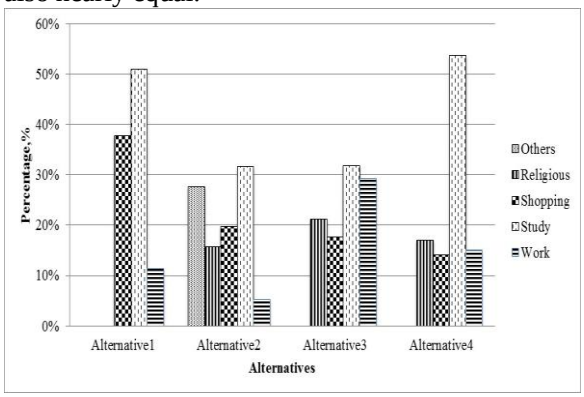

Fig. 4 Selection of Alternatives on the basis of Trip Purpose

3.2.2 Selection Reason

The alternatives were chosen on the basis of different reasons. Among them travel time was one of the main factors. More than 70% of the users used alternative4 due to the travel time. As the travel time of this was less than other 3 alternatives. Then after that the alternative3 was selected by user due to comfort of about more than 50%.

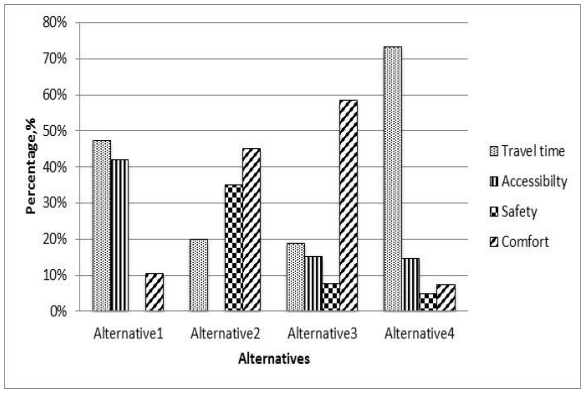

Fig. 15 Selection of Alternatives on the basis of Selection reason

3.2.3 Age Group

Among the 4 alternatives, 3 of them were mostly used by the age group of 15-20.

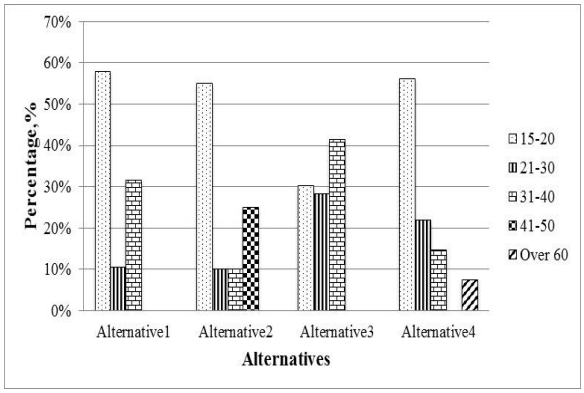

Fig. 16 Selection of Alternatives on the basis of Age group

3.2.4 Gender

The alternative2 and 3 were mostly preferred by the female and the alternative1 and 4 were preferred by male. The alternatives were selected by the users on the basis of the different factors.

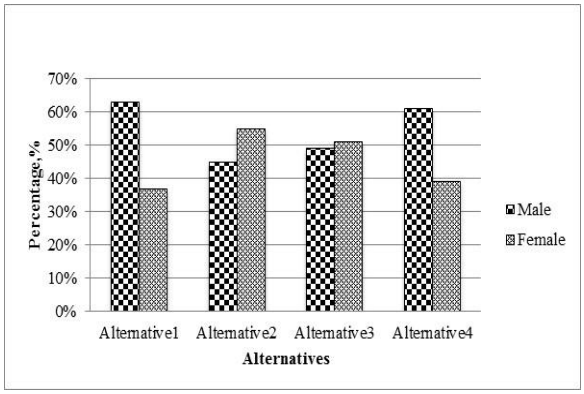

Fig. 17 Selection of Alternatives on the basis of Gender

Table 2 Parameter Estimates

Alternatives		B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp(B)	
								Lower Bound	Upper Bound
Alternative1	Intercept	12.086	2.774	18.978	1	0			
	Age	0.05	0.035	1.963	1	0.161	1.051	0.98	1.126
	Monthly Income	0	0	39.905	1	0	1	1	1
	Traveltime	0.102	0.017	35.486	1	0	1.107	1.071	1.145
	Waitingtime	0.206	0.15	1.896	1	0.168	1.229	0.916	1.648
	Walkingtime	-0.132	0.056	5.61	1	0.018	0.876	0.785	0.977
	FareRs	-1.122	0.189	35.389	1	0	0.326	0.225	0.471
	[Gender=1]	-1.261	0.435	8.401	1	0.004	0.283	0.121	0.665
	[Gender=2]	0[b]			0				
Alternative2	Intercept	1.381	3.022	0.209	1	0.648			
	Age	0.094	0.031	8.986	1	0.003	1.099	1.033	1.169
	MonthlyIncome	0	0	1.907	1	0.167	1	1	1
	Traveltime	0.112	0.019	35.933	1	0	1.118	1.078	1.16
	Waitingtime	0.842	0.186	20.588	1	0	2.322	1.614	3.34
	Walkingtime	0.281	0.054	26.839	1	0	1.324	1.191	1.472
	FareRs	-1.004	0.169	35.171	1	0	0.366	0.263	0.511
	[Gender=1]	-2.89	0.528	29.998	1	0	0.056	0.02	0.156
	[Gender=2]	0[b]			0				
Alternative3	Intercept	0.891	2.04	0.191	1	0.662			
	Age	-0.06	0.027	5.028	1	0.025	0.942	0.893	0.992
	MonthlyIncome	0	0	44.412	1	0	1	1	1
	Traveltime	0.053	0.014	13.159	1	0	1.054	1.024	1.084
	Waitingtime	0.263	0.11	5.688	1	0.017	1.301	1.048	1.615
	Walkingtime	-0.123	0.039	9.744	1	0.002	0.885	0.819	0.955
	FareRs	-0.182	0.109	2.809	1	0.094	0.833	0.674	1.031
	[Gender=1]	-1.214	0.358	11.476	1	0.001	0.297	0.147	0.6
	[Gender=2]	0[b]			0				
a. The reference category is: Alternative4.									
b. This parameter is set to zero because it is redundant.									

3.2.5 Occupation

Among the four alternatives, three of them were mostly used by the students, and one of them was used by the government officers.

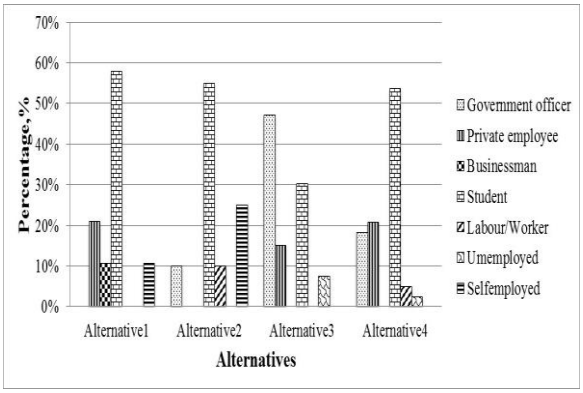

Fig. 17 Selection of Alternatives on the basis of Occupation

3.3 Choice Model

3.3.1 Model Development

The model for the mode choice was developed by using the survey data and the analysis tool, i.e. SPSS. For the modeling Multinomial logistic regression was used. In Multinomial logistic regression outcome also compared with reference category that was the model attempts to explain the relative effect of differing explanatory variables on the outcome. The tables show the result of the multinomial logistic regression. For the final analysis after a number of iterations the variables found significantly influencing the mode choice were only seven. They were monthly income(INC), travel time (TRT), waiting time (WAT), walking time (WKT), fare (FAR), age (AGE), gender male (GEN_M) and gender female (GEN_F).All the variables are statistically significant at the 0.05 probability level, and hence

all the independent variables have overall relationship to the dependent variable. Table 2 shows the parameter estimates.

3.3.2 Model Validation

a) Significance test of the model

The Initial Log Likelihood Function, (-2 Log Likelihood or -2LL) is a statistical measure like total sums of squares in regression. If independent variables have a relationship to the dependent variable, the accurate prediction of the dependent variable is done, and the log likelihood measure will decrease. The initial log likelihood value (944.661), as shown in table 3 was a measure of a model with no independent variables, i.e. only a constant or intercept.

The final log likelihood value (577.246), as shown in table 3 was the measure computed after all of the independent variables have been entered into the logistic regression. The probability of the model chi-square obtained was 0 as shown in table 3, less than or equal to the level of significance of 0.05, the existence of a relationship between independent variables and dependent variable was also supported.

Table 3 Model Fitting Information

Model	Model Fitting Criteria	Likelihood Ratio Tests		
	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	944.661			
Final	577.246	367.415	21	0

b) Measures of R^2

The next SPSS outputs indicate the strength of the relationship between the dependent variable and the independent variables, analogous to the R^2 measures in multiple regressions. The Cox and Snell R^2 measure operates like R^2, with higher values indicating greater model fit. From the table 4 the R2 value of the Cox and Snell and Nagelkerke was found as 0.578 and 0.621 respectively, which was greater than 0.5 so, the model was fit.

Table 4 R-square Value of model

Measures	R^2
Cox and Snell	0.578
Nagelkerke	0.621
McFadden	0.323

c) Significance of predictor variables

The Likelihood Ratio Tests indicate the contribution of the variable to the overall relationship between the dependent variable and the individual independent variables. In this model, the variables namely monthly income, travel time, waiting time, walking time fare, age, and gender were all significant contributors (significance less than 0.05 as shown in table 5) to explaining the public transport mode choice behavior of the users.

Table 5 Likelihood ratio test of predictor variables

Effect	Model Fitting Criteria	Likelihood Ratio Tests		
	-2 Log Likelihood of Reduced Model	Chi-Square	df	Sig.
Intercept	577.246[a]	0	0	
Age	596.817	19.571	3	0
MonthlyIncome	660.641	83.395	3	0
Traveltime	640.329	63.083	3	0
Waitingtime	605.942	28.696	3	0
Walkingtime	668.389	91.142	3	0
FareRs	667.064	89.818	3	0
Gender	616.088	38.842	3	0

d) Classification table showing the prediction

The classification matrix in logistic regression serves the same function as the classification matrix in Multinomial Logistic Regression, i.e. evaluating the accuracy of the model. If the predicted and actual group memberships are the same, then the prediction was accurate for that case. If predicted group membership and actual group membership were different, the model "misses" for that case. The overall percentage of accurate predictions (63.4% in this case, from the table 6) was the measure of a model that relies on most heavily for this analysis as well as for Multinomial Logistic Regression because it has a meaning that was readily communicated, i.e. the percentage of cases for which the model predicts accurately.

Table 6 Classification table showing the prediction of mode selection

Observed	Predicted				
	Alternative1	Alternative2	Alternative3	Alternative4	Percent Correct
Alternative1	52	0	8	16	68.40%
Alternative2	20	44	0	16	55.00%
Alternative3	12	8	50	36	47.20%
Alternative4	10	0	30	124	75.60%
Overall Percentage	22.10%	12.20%	20.70%	45.10%	63.40%

3.4 Model Interpretation

The logit of selecting was obtained by multiplying the coefficients from the table 2 of parameter estimates and the variables. They were as shown in equation below.

U_1 = 12.086 + 0.50AGE + 0.00011INC+ 0.102TRT + 0.206WTT - 0.132WKT – 1.122FAR - 10261GEN_M

U_2 = 1.813 + 0.94AGE + 0.000029INC + 0.112TRT + 0.842WTT + 0.281WKT - 1.001FAR - 2.890GEN_M

U_3 = 0.891 - 0.060AGE + 0.000090INC + 0.053TRT + 0.263WTT - 0.123WKT - 0.182FAR - 1.214GEN_M

$U_4 = 0$

Where,
U_1 = total utility for alternative1
U_2 = total utility for alternative2
U_3 = total utility for alternative3
U_4 = total utility for alternative4.

Probabilities of selecting each alternative are as follows:

$$P_1 = \frac{e^{U_1}}{e^{U_1} + e^{U_2} + e^{U_3} + e^{U_4}}$$

$$P_1 = \frac{e^{U_1}}{1 + e^{U_2} + e^{U_3} + e^{U_4}}$$

$$P_2 = \frac{e^{U_2}}{1 + e^{U_2} + e^{U_3} + e^{U_4}}$$

$$P_3 = \frac{e^{U_3}}{1 + e^{U_2} + e^{U_3} + e^{U_4}}$$

$$P_4 = \frac{1}{1 + e^{U_2} + e^{U_3} + e^{U_4}}$$

4. Conclusion

A multinomial logit model was developed to study the choice behavior of Bhaktapur district (Study area Naya-Thimi chowk). The study helped to draw the socio economic characteristics of the study area. The activity based approach really helped to study the travel pattern in a realistic way. For the empirical analysis, the data used was collected by means of questionnaire survey of sample population, and software tool SPSS was used to develop the model. For the public transportation, choice of the four alternatives monthly income, age, travel time, walking time, waiting time, fare and gender was found to be the most influencing factors for the choice of alternatives. The model developed could greatly help the policy makers and will be the area of interest for public transport service providers who are interested in attracting choice riders. From the analysis, alternative4 was the best alternative among the others. The analysis was totally based on the survey data and the data was collected with respect to the experiences of the users.

References

[1] Ben-Akiva, M. E. and Morikawa, 2002.Comparing Ridership Attraction of Rail and Bus. Transport Policy, Volume 9, Issue 2, p. 107-116.

[2] Ben-Akiva, M. E. and Lerman, S. R. (1985).Discrete Choice Analysis: Theory and Application to Travel Demand, MIT Press, Cambridge, Ma.

[3] CBS, 2011.National population and housing census 2011. Kathmandu, Nepal: Central Bureau of Statistics (CBS).

[4] DoR, 2013.Nepal Road Standard 2070. Kathmandu, Nepal: Department of Road.

[5] DoTM, 2011.Public Transport code of conduct, 2011. Kathmandu, Nepal: Department of Transportation Management(DoTM).

[6] Gonzaga J.T. and Villoria, 1999.An Analysis of Travel Activity Patterns in Metro Manila. Journal of The Eastern Asia Society for Transportation Studies.

[7] Goulias, K. G., 2003. Transportation Systems Planning Methods and Application.USA : CRC Press.

[8] MoPIT/JICA, 2012. Final Report: Data collection survey on Traffic improvement in Kathmandu Valley. Kathmandu, Nepal: Ministry of Physical Infrastructure and Transport.