Forecasting Paratransit Utility by Using Multinomial Logit Model: A Case Study

Waikhom Victory [#1], M. Ali Ahmed [*2]

[#] *Department of Civil Engineering, National Institute of Technology Manipur*
Manipur, India - 795004

[1] `waikhomvictory@gmail.com`

[*] *Department of Civil Engineering, National Institute of Technology Silchar*

Assam, India - 788010

[2] `ali.mokaddes@gmail.com`

Abstract— **Paratransit plays an important role in the urban passenger transportation system in the developing countries. Three cities viz. Imphal East, Imphal West and Silchar in India have been undertaken for the study. Household survey and traffic survey have been employed to collect data for the paratransit users. Modelling techniques and tools also have been used to forecast the utility of paratransit in the region. For this purpose, a Multinomial Logit Model (MNL) had been used. A total of seven variables were considered in the model estimation of which three are quantitative i.e. trip length (km), travel cost (rupees) and travel time (minutes) and four are qualitative variables i.e. reliability, comfort, road condition and convenience.**

Keyword- Multinomial logit model, forecasting, reliability, comfort, road condition, convenience, travel cost, travel time, trip length

I. INTRODUCTION

Paratransit can be defined as the intermediate mode between the privately owned automobiles and the conventional transit that has fixed routes and schedules [1]. The paratransit vehicles are operated on dual system either on shared ride or on hired basis. The service conveniently meets the needs and travel desire of the individuals. In some other cases, paratransit modes do not have fixed routes and schedules.

Paratransit system consists of less passenger capacity vehicles operated on flexible routing and scheduling or fixed routing and flexible scheduling according to the need of the trip makers. In Indian context, paratransit plays an important role for the urban and rural passenger transportation services, particularly in the small and medium size cities. The paratransit vehicles have been providing services as a replacement for buses in the urban hinterland since many years. The study was carried out in Imphal agglomeration in the state of Manipur and Silchar town in Cachar district of Assam.

The areas under study are part of the north-eastern India. The geographical structure of this region comprise of mostly hilly areas with few pocket of plains. So, the topography of the region is one of the major challenges which hinder development particularly in terms of infrastructure. In other words, the landlocked feature of the region is one major hindrance for development and also creates difficulty in road connectivity and transit system. However, there are many paratransit vehicles plying on urban roads to highways fulfilling the needs of the people in the region but mostly in plains and valleys. These paratransit constitute as a significant transport system of the region. The flexibility and its responsiveness towards sustain the demand and utility of such transportation system.

This paper is an outcome of the study which forecast the utility of paratransit mode in north-east part of India. The transport facility in this region is largely provided by privately owned paratransit vehicles such as auto-rickshaw, cycle-rickshaw and other paratransit [1](OP).

The utility of paratransit has been observed based on different variables which are categorized into two i.e. quantitative and qualitative. The travel cost, travel time and trip length are considered as quantitative variables and the latter consist of reliability, comfort, road condition, safety and convenience. The multinomial logit model is used for modelling and prediction of various possible outcomes based on the significance of different variables or parameters mentioned above. The model is estimated by using maximum likelihood method.

II. LITERATURE REVIEW

Paratransit comprises vans, jitneys, shuttles, microbuses, and minibuses [2]. Jitneys and privately operated minibuses are used extensively in developing countries where labor costs are low and transit services are less and inadequate [3]. Paratransit is served in many countries in different forms such as "Angkutan Kota" in

---

[1] Other paratransit is group of paratransit mode which comprised of Sumo, Magic, Max and Van etc. These vehicles are grouped into this category for the convenient of the study.

Indonesia, "Jeepney" in Philippines, "Tuk-Tuk and "Songtaew" in Thailand and "Mammy Wagons" (converted trucks and "Matatu" (converted vans) in Africa [4]. It also includes transport modes such as Taxis, jitneys (in developing countries), dial-a-ride, and other types of demand-responsive minibuses (in developed countries) usually provide public service [3]. The modal split of such modes comprised of approximately 40 - 50% in developing countries of which 70 - 80% are motorized transport. While in poorer cities of developing countries the modal share is 10 - 15% of which 20% are motorized [5], [6]. In most Indian cities, the share of private and paratransit transportation occupied optimal amount [7].

Paratransit has become important and efficient mobility in many developing countries which bridges the gap where the public transit systems are absent and it acts as an efficient feeder of a transport service [2]. In suburbs and areas where the conventional public transport is absent, paratransit plays important role by fulfilling the demand of mobility [8].

Paratransit efficiently fulfilled the requirement of the poor in terms of cost (fare) and flexibility is assumed [5]. As compare to other public transport mode, paratransit has more advantages such as easy mobility and high accessibility, optimal operating cost and low maintenance cost [9]. The service can be easily opened when needed; and at the same time it can be closed when number of user decreases [10]. It also provide job for low skilled workers and thus help society [8].

III. THEORETICAL APPROACH AND MODEL SPECIFICATION

It is observed that the most common theoretical framework for developing discrete choice model is the random utility theory which states that in order to maximize the net utility, individuals use their rational power depending on the legal, physical, social, time and money constraints [11]-[13]. Further, in the analysis of transportation demand and utility, the emergence of disaggregate travel demand model based on discrete choice analysis methods is quite substantial [14].

For discrete choice analysis and logistic regression, the logit function plays an important role [15]. Moreover, there is extensive and wide use of logit model in the field of transportation modelling and forecasting. Predominantly, since 1980, modelling based on discrete choice for disaggregate data were used whereas the aggregate based models were left behind [13]. The multinomial logit model can be used for finding the underlying choice probabilities without incorporating any multivariate analysis [16]. Earlier empirical work in the field suggest that under discrete choice model, multinomial logit model can be used quite suitably for transportation modelling [14], 17]. Further, discrete choice model suggests that the probability of individuals opting for a given mode is a function of their socioeconomic characteristics and the relative attractiveness of the mode [13].

MNL is most simplest and convenient modelling method. In the study, multinomial logit model (MNL) is used to predict the paratransit utility in the north-eastern region of India. The usage of MNL model is also well explained by [18] and it is given by

$$P_n(i) = \frac{e^{v_{in}}}{\sum_{j=1}^{k} e^{v_{jn}}} \qquad (1)$$

Where
$P_n(i)$ = Probability of individual n choosing mode i,
v_{jn} = Utility of mode j by n individual,
K = Number of different modes

The utility of a particular mode is directly proportional to the level of mode choice of that particular mode which is depend on the socio-economic characteristics of the individual. In other words, higher the level of choosing a particular mode, higher the utility of that mode. Therefore, probability of individuals choosing a given option is a function of their socio-economic characteristics and the relative attractiveness of the option [13].

The utility by an individual n from mode j, V_{jn}, is derived as a linear function of the explanatory variables as follows i.e. equation (2).

$$V_{jn} = a_{0j} + a_{1j} x_{1n} + a_{2j} x_{2n} + a_{3j} x_{3n} + a_{4j} x_{4n} + \cdots .. + a_{nj} x_{sn} \qquad (2)$$

Where V_{jn} = Utility of mode j by n individual,
a_{0j} = Alternative specific constant for mode j
$a_{1j}, a_{2j}, a_{3j}, a_{4j} \dots \dots \dots a_{nj}$ = Coefficient of the explanatory variables and
$x_{1n}, x_{2n}, x_{3n}, x_{4n} \dots \dots \dots x_{sn}$ = Explanatory variables for individual n
s = Number of explanatory variables included in the model

The designing of the mode choice model needs an extensive evaluation of observed data and the efficiency of the whole model system. In the present study, the paratransit utility is predicted based on specific parameters. Here, the variables comprise of randomly selected variables i.e. reliability (RY), comfort (CF), convenience (CV), travel cost (TC), travel time (TT) and trip length (TL) of the mode.Some of the parameters such as travel cost and travel time are considered to be substantial in literature while other variables are presented exclusively to deal with specific research problems.

IV. STUDY LOCATION

The data were collected through household survey and traffic survey. The household survey was aimed to collect primary data from the respondents in the study area. In the household survey, a total number of 970 respondents comprising of 300 from Imphal East, 320 from Imphal West and 350 from Silchar were involved.

In addition, the collection of household data was carried out through a questionnaire distributed to the respondents. The questions are set basically in English, however, depending on the requirement and situation, the questions are also being translated in preferred local language (Manipuri, Bengali or Hindi). The questionnaire was divided broadly into two parts in which one part aims to collect the daily trip information of the respondents while the other part was basically designed to collect the data pertaining to the socio-economic characteristics of the respondents. Further, the questionnaire is target to collect data pertaining to several parameters such as travel mode, travel behaviour, trip purpose, trip length and as well as the profile of the respondents which consist of age, gender, occupation, educational qualification, members of the family, vehicle ownership and monthly income.

V. MODEL ESTIMATION AND RESULT

A number of variables comprising of qualitative and quantitative were considered in developing the logit model. A few of the tested parameters have resulted into statistically less significant or inappropriate intuitive signs, and have been therefore invalidated. However, the model consider variable which shows fitness of good but had a counter-intuitive sign.

A multinomial logit model has been designed for auto-rickshaw and other paratransit to determine the aspects which might impact the utility of mode. The variables with insignificant coefficients were dropped from the model and hence the remaining explanatory variables were reliability (RY), comfort (CF), convenience (CV), travel cost (TC), travel time (TT) and trip length (TL). The trip length, travel cost and travel time are quantitative variables while comfort, road condition, convenience and reliability are qualitative variables.

The coefficients were estimated by fitting the data to the models. In the estimation, maximum likelihood 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. Usually this method produces highly accurate estimates.

The developed models were validated after the calibration process by using the data which were not used in model calibration. For the purpose of model validation, 320 observations have been used. In the process, the model was tested several times with various combinations of variables. Trip length, travel cost, travel time, comfort, reliability, convenience and road condition were found to be significant in determining the utility of paratransit modes. The estimated coefficients and standard errors of the explanatory variables for each mode considered are shown in Table I. The calibration of model was done in SPSS (a statistical software package) and estimated by using maximum likelihood (ML) method

Models
$$(P)_{Auto\ Rickshaw} = 2.2303 - 5.266\ (RY) + 2.463\ (CF) + 0.6408\ (CV) - 0.031\ (TC) + 0.0175\ (TT) - 0.0504\ (TL)$$

$$(P)_{Other\ Paratransit} = -6.627 + 5.419\ (RY) - 3.807\ (CF) - 1.961\ (CV) - 0.009\ (TC) - 0.0111\ (TT) - 0.041\ (TL)$$

In the above equation (3), Reliability (RY), Comfort (CF), Convenience (CV), Travel Cost (TC), Travel Time (TT) and Trip Length (TL) are the explanatory variables which represent an attribute of the travellers or of the mode considered. The impact of each attribute on the overall satisfaction provided by the mode or alternative is given by its coefficient [13]. However, in equation (3), 2.2303 is the alternative-specific constant which is generally represented the net influence of all unobserved attributes, or not explicitly included, characteristics of the individual and the option in its utility function [13], [18].

The basic test of the estimates are indicated by their signs (+ or -) and the impact of the corresponding variables [4]. The summary of estimations using the multinomial logit model is presented in Table 1. All the variables presented had significant estimates and logical signs. The level of service variables, such as total travel

cost (TC) and travel time (TT) is generally expected to have negative coefficients. As TC and TT increase, the probability of selecting the modes will decrease. A number of studies have considered travel cost to have a considerable adverse impact on the choice of travel mode.

TABLE I: THE MULTINOMIAL MODE CHOICE MODEL ESTIMATES

Modes	Independent variables	Co-efficient (B)	Standard Error
Auto-Rickshaw	Intercept	2.2303	0.854
	RY	-5.266	1.136
	CF	2.463	1.321
	CV	0.6408	0.8608
	TC	-0.031	0.0287
	TT	0.0175	0.0051
	TL	-0.0504	0.2700
Other Paratransit	Intercept	-6.627	0.727
	RY	5.419	1.743
	CF	-3.807	1.848
	CV	1.961	0.765
	TC	-0.009	0.024
	TT	-0.0111	0.0052
	TL	-0.041	0.232
Private	Intercept	7.084	0.9627
	RY	-0.201	1.123
	CF	-1.733	1.475
	CV	-2.498	0.9493
	TC	0.0065	0.0188
	TT	0.0131	0.0057
	TL	0.0286	0.181
Summary of Statistics			
Number of observations		651	
(-2) Initial LL		1058.071	
(-2) Final LL		857.19	
Cox& Snell's R^2		0.632	
Nagelkerke value		0.892	
McFadden's value		0.622	
Explanation of Variables Included in the Selected Model			
RY = Reliability		TC = Travel Cost	
CF = Comfort		TT = Travel Time	
CV = Convenience		TP = Trip Length	

Note: The cycle-rickshaw as the base case was set to zero.

Many recent studies on intercity and mode choice have revealed the significance of travel cost. Aljarad [19] developed intercity mode choice using disaggregate transport mode choice models and found that the travel cost was high significant and had negative coefficients. Kumar, et al. [20] have examined rural intercity bus services in India, to determine users' behaviours towards various features of service. As anticipated, the travel time and travel cost adversely impacted on paratransit utility. In the United States [21], [22] a national intercity travel demand model was designed and predicted that, as travel cost increase for a particular mode, the utility and choice is decreased for that particular mode. Mukala and Chunchu [23] predict that the income and total travel cost play an important role in the mode choice decision for the intercity transport. The travel cost variable is specified as generic in the model. This implies that an increase of travel cost has the same impact on modal utility for the two modes [21].

Some perceptual variables were also incorporated during the calibration to investigate the effect on paratransit utility. The reliability, comfort and convenience of mode significantly affected the mode choice of the traveller. However, the coefficients of reliability in case of auto-rickshaw and comfort in case of other paratransit were negative. This implies an increase in their absolute values will decrease the utility but in normal assumption, these two variables are expected to have a positive impact. The reason might be the influence of one mode over the other. The preference rankings of alternatives and perception indicators (such as comfort, convenience, and reliability, etc.) had been considered, and the preference index computed from the estimation using the multinomial logit model [21], [24]. The results have shown that most parameters had the correct signs.

VI. CONCLUSION

The fundamental goal of this research is to build a mode choice model, which would be substantial information to the policy-makers and transport planners. The proposed model might be beneficial to the planners as it is accessible to a number of parameters, and consequently, can be used to estimate the impact of changes in the several features or the particular policy changes of the demand for each of modes.

The requirements of model utility might support in further study to emphasise on which data is essentially required. The model indicated that travel cost, reliability, convenience, and comfort have impacted the utility of the mode. Almost the estimated coefficients possessed the expected signs.

The two R^2 values indicated the explanatory power of the model. The factors included in the model accounted for 89.2 % of the variation for the Negelkerke, while the Cox and Snell-t explained 63.2 % and 62.2 % of the McFadden's value. It is also notable that the "values of 0.2 to 0.4 for R^2 represent an excellent fit" [25].

In conclusion, the results of this study is distinctive because it is a primary endeavour to forecast the utility of paratransit in the region. This study will be extremely useful for travel demand analysis by transport department of the region. It will also help government and public transportation organizations as well as the private providers in making suitable choices and avoid under/over designing of required amenities and services.

The outcomes of this study can be applied to assess the public transportation, and predict future paratransit utility. Decision-makers can utilize the outcomes to enhance the planning of paratransit services and intercity transport service as well. The model developed in this study can also be utilized by the operators, to approximate the probable repercussions of specific policies and investment planning. Subsequently, the effect of policy changes can be replicated in mode choice models. Consequently, the probable reactions of travellers, both, in shifting the mode and stimulated need, are incorporated in the model. Therefore, this model enhances the accuracy of predicting utility of various paratransit modes and the same can be applied to study the current development of intercity carriers for instance buses.

ACKNOWLEDGMENT

REFERENCES

[1] V. Vuchic, *Urban Transit System and Technology*, New Jersey: John Wiley & Sons, 2007.
[2] R. Cervero, *The Transit Metropolis: A Global Inquiry*, Washington DC: Island Press, 1988.
[3] V. Vuchic, *Transportation for Liveable Cities,* New Jersey: The State University of New Jersey, 1999.
[4] T.B. Joewono, and H. Kubota, "Paratransit Services in Indonesia: User Satisfaction and Future Choice," *Transportation Planning and Technology*, vol. 31(3), pp. 325-345, 2008.
[5] R. Kaltheier, "Urban Transport and Poverty in Developing Countries: Analysis and Options for Transport Policy and Planning," *Eschborn: Divsion 44 Environmental Management, Water, Engery, Transport Deutesche GTZ,* 2002.
[6] S.B. Nugroho, A. Fujiwara, and J. Zhang, "Empirical Study on Fuel Consumption of Paratransit in Jakarta City," in *Proc. Infrastructure Planning Issue*, 2011.
[7] S. Singh, "Review of Urban Transportation in India," *Journal of Public Transportation*, vol. 8(1), pp. 79-97, 2005.
[8] A. Tarigan, Y. Susilo, and T. Joewono, "Negative Experiences and Willingness to Use Paratransit in Bandung, Indonesia: An Exploration with Ordered Probit Model," *Transportation Research Board*, 2010.
[9] DLLAJ, "The Guide of Passenger Public Transportation Price Calculation with Fixed Routes in Urban Areas," Bandung: DLLAJ, 2001.
[10] R. Cervero, and A. Golub, "Informal Transport: A Global Perspective," *Transport Policy*, vol. 14, pp. 445-457, 2007.
[11] T. Dominich, and D. McFadden, *Urban Travel Demand*, Amsterda: NHPC, 1975.
[12] H. Williams, "On the Formation of Travel Demand Models and Economic Evaluation Measures of User Benefit," *Environment Planning A*, vol. 9, pp. 285-344, 1977.
[13] J. Ortuzar, and L. Willumsen, *Modelling Transport*, 4th ed. New Delhi: Wiley and Sons, 2011.
[14] Ben-Akiva and S. Lerman, *Discrete Choice Analysis: Theory of Application To Travel Demand*, MIT Press Cambrigde, 1985.
[15] A.B.M. Manssour A., "Modeling a Multinomial Logit Model of Intercity Travel Mode Choice Behavior for All Trips in Libya," I*nternational Journal of Civil, Architectural, Structural and Construction Engineering WASET*, vol. 7(9), pp. 300-309, 2013.
[16] J. Hausman, and D. McFadden, "Specification Tests for the Multinomial Logit Model," *Econometrica*, vol. 52(5), pp. 1219-1240, 1984.
[17] T. Morikawa, "Incorporationg Stated Preference Data in Travel Demand Analysis," Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, 1989.
[18] D. Hensher, and L. Johnson, *Applied Discrete Choice Modelling*, London: Croom Helm, 1981.
[19] S. Aljarad, "Disaggregate Mode Choice Modelling for Intercity Non-Business Travelers in the Saudi Arabia-Bahrain Corridor," *Transportation Research Part A*, vol. 1(3), pp. 85, 1996.
[20] C. Kumar, D. Basu, and B. Maitra, "Modeling Generalized Cost of Travel for Rural Bus Users: A Case Study," *Journal of Public Transportation*, vol. 7, pp. 59-72, 2004.
[21] S. Ashiabor, H. Baik, and A. Trani, "Logit Models for Forecasting Nationwide Intercity Travel Demand in the United States," *Transportation Research Record*, pp. 1-12, 2007.
[22] S. Ashiabor, H. Baik, A. Trani, and N. Hinze, "Development of an Intercity Mode Choice Models for New Aviation Technologies," *Transportation Research Board*, pp. 61-77, 2007.
[23] P. Mukala, and M. Chunchu, "Mode Choice Modelling for Intercity Transportation in India: A Case of Guwahati to Five Metro Cities," *International Journal of Earth Sciences and Engineering*, vol. 4(6), pp. 364-374, 2011.

[24] D. Jung, T. Nishimura, and Y. Hino, "Analysis of Travel Mode Choice Models Considering the Subjective Indicators of Alternatives," *Memoirs of the Faculty of Engineering Osaka City University*, vol. 36, pp. 39-46, 1995.
[25] D. McFadden, "Quantitative Methods for Analyzing Travel Behavior of Individuals: Some Recent Developments" In D. A. Hensher & P. R. Stopher, eds. *Behavioural Travel Modelling*, London: Croom Helm, 1978.

AUTHOR PROFILE

Waikhom Victory is presently teaching as a lecturer at Department of Civil Engineering, National Institute of Technology Manipur. Her area of specialization includes transportation engineering and planning.

M. Ali Ahmed is currently teaching at National Institute of Technology Silchar in Assam. He is the professor of civil engineering at the institute. His area of specialization is transportation engineering and planning.