

REVEALED PREFERENCE ANALYSIS FOR MODE CHOICE IN GRADUATE LEVEL ENGINEERING STUDENTS OF KATHMANDU VALLEY

¹Piyush Chataut, ²Dr. Pradeep Kumar Shrestha

¹Graduate Student, Department of Civil Engineering, Pulchowk Campus, Institute of Engineering, Tribhuvan University, Nepal

²Lecturer, Department of Civil Engineering, Pulchowk Campus, Institute of Engineering, Tribhuvan University, Nepal

Corresponding Email: pradeep.shrestha@pcampus.edu.np

Abstract

Proper planning which is the key element in ensuring infrastructure efficiency, relies on demand analysis. Among the various trips under the domain of demand analysis educational trips occupy a significant part and hence the knowledge about patterns and attitudes of these trips is important to policymakers and infrastructure planners. The current study analyzes the mode choice of graduate-level engineering students in Kathmandu valley where the current transportation system is facing multiple problems thus requiring a proper planning intervention. This study reveals the educational mode preference among the students of engineering colleges. The reveal preference survey was conducted at the various engineering colleges in Kathmandu valley. The study concludes that the travel distance, number of siblings, and vehicle ownership effects the selection of personal modes of transport and distance effects the selection of public transportation, walking option being base criteria for both the cases. It is recommended that existing walking conditions should be improved within the educational zones while public transport which are preferred options for long journeys be designed considering movement between zones.

Keywords: Educational mode choice -- Discrete choice model -- Mixed logit Model

1. Introduction

Transportation infrastructure has been found to impact positively on economic development [1] which is even more crucial in urban areas [2] The degree of this impact however depends upon the efficiency of the transportation system, which in turn extensively relies on proper planning [1].

About Kathmandu valley which the capital city and also the largest economic hub in the country, the existing transportation system seems to be facing multiple issues and thus seems unlikely to provide the above-stated benefits fully. Among the three components of transport public transportation system consisting of Bus Tempo Micro and Taxi is facing problems like congestion in the city center [3], overcrowding, insecurity [4], unreliability, and higher commute time [2]. Similarly, private modes are facing problems like higher travel costs, congestions, safety issues, etc. The users, however are reluctant to shift to public transportation due to the above-mentioned problems [3] in the public transportation system. Finally walking conditions too, with a walkability index of 557, are not user friendly either [5].

To ensure optimum contribution, proper planning through Travel Demand Analysis (TDA) is needed. TDA must be able to predict the ever-changing mobility demand for the present and future accurately through mathematical modeling. Four Step Transportation Model (FSTM) is the most frequently used model. It splits the TDA into 4 parts: trip generation, trip assignment, modal split & traffic assignment. Modal split analysis is an important step in TDA which describes trip preferences of users. Among the various trips to be considered under the TDA process, this study analyzed the modal split of educational trips for the population of graduate-level engineering level students of Kathmandu valley

Educational trips contribute to around 20% of travel demand in Kathmandu thus influencing a large portion of the demand, not only this these trips are supposed to have long term influence on the overall health [6,7] and future travel demands of the students [8].

A proper study into the factors affecting educational mode choice of the selected group will provide the planners with insights on the factors influencing the choice of mode in this group, shade some light on the nature of factors affecting the overall educational trips and aid the overall transportation planning process.

2. Literature Review

2.1 Utility based choice theory

Modal split involves studying the current mode shares and the factors behind it. Utility choice theory is the underlying principle for modal split analysis. Utility is a function of socio-economic, trip, and infrastructure characteristics. Utility based theory states that “an individual will select the alternative from their set of choices for which their utility is maximized.

2.2 Logistic Regression

Different models are used to form the utility functions, the selection of a particular model in general depends upon the nature of the trip. The models available can be divided into two types, aggregate choice models and discrete choice models. Discrete choice models which are developed based on the observed mode choice and socio-economic/trip characters of the individual are found to be better performing regarding transferability [9], predictability [10] and sensitivity [11] The most popularly used discrete choice models are the logit model family, they assume the CDF of the function to follow the logistic distribution. The different models in this family are logit, conditional logit, mixed logit, nested logit, etc.

The process of fitting data into the given model to determine the relationship between a set of dependent and independent variables is called regression. Logistic regression tries to fit the data into the logit model of choice and thus gives the coefficients of the model. The current performs multinomial regression using the mixed logit model for educational trips by graduate-level engineering students of Kathmandu valley. The mathematical formulation of the logit model is as follows

$$U_{ij} = V_{ij} + e_{ij} \quad \text{Equation 1}$$

$$V_{ij} = \beta_i X_{ij} + \beta_j X_i \quad \text{Equation 2}$$

Where

U_{ij} = Utility of the individual i for the j^{th} alternative

V_{ij} = Systematic part of the utility U_{ij}

e_{ij} = Random part of the utility U_{ij}

β_i = Coefficient for alternative variant regressors modeled as a random parameter with mean β and standard deviation σ_β

X_{ij} = Alternative variant regressors e.g. Price

β_j = Coefficient for alternative invariant regressors

X_i = Alternative invariant regressors e.g. income, age

Since only the systematic part of the utility function is known the mode choice is predicted based on the probability of the particular mode i.e. probability that an individual chooses a mode i equal the probability that the conditional utility of mode i is largest of the conditional utility of all other modes. Occurring. Probability is given by the CDF of the logit model as:

$$P_{ij} = \frac{e^{V_{ij}}}{\sum e^{V_{ij}}}$$

Equation 3

2.3 Previous works

Works conducted using different logit models were studied to determine the applicability of these models as well as the nature of factors affecting the different trips including educational trips.

Multinomial logit was found to be extensively used in choice modeling for a wide range of fields ranging from social to medical sciences, concerning mode choice it has been used in analysis of different types of trips. A primary school mode choice in Germany concluded that parental origin (German-born or not), parental employment status, age, and distance affected the mode choice for commutes [12]. Another mode choice study of school-going students in china [13] pointed the differences between the school education scenario in china and rest of the world like provision of one child only, vehicle ownership compulsion for admissions, etc. and found out that car ownership, departure time, parental employment and distance affected the mode choice, another similar study in China [14] analyzed the effect of weather condition on the educational commutes in China and concluded that along with other conventional factors sky condition temperature air quality and humidity effects the mode choice of students. Another educational mode choice study in India [15] analyzed student's mode choice in the Indian city of Kanpur and concluded that family income, vehicle ownership, and gender affects the mode choice. A mode choice study in 2012 [16] Investigated the mode choice of work trips in Kathmandu and found that the factors affecting the choice are waiting and travel time, total income, cost of commute, and age. Another similar study of work trips in Kathmandu in 2019 [17] concluded that mode choice is affected by travel cost, travel time, gender, and monthly income. A study in 2019 [18] Investigated intercity travels between Kathmandu and other cities using RP and SP survey methods with hypothetical alternatives of train and concluded that cost bearer, frequency, vehicle ownership, and family size affects the choice between the current modes of transportation while travel time, travel cost, and distance affect the choice for alternative transportation mode.

Mixed logit and conditional logit models been used in many fields, concerning mode choice analysis it has been used extensively in work and recreational trips. A study of educational mode choice in Vietnam [19] used the conditional logit model and investigated factors influencing mode choice of university students and also provided hypothetical conditions of efficient and reliable public transport, it concluded that age, gender, income, and travel time have a strong impact, and also that the students are likely to switch to public transport if effective and reliable transport is developed. A study in Korea [20] used the mixed logit to investigate the factors which will enhance or degrade the choice between the KTX (A new high-speed rail) and the existing aircraft and concluded that cost, journey time, safety, frequency and availability of duty-free shops affects the mode choice. A study in 2016 [21] compared the multinomial, mixed, and nested logit models to analyze the effect of various factors on the mode choice of general consumers with the main task of identifying the effect of cost adjustment, it concluded that all the models give almost the same result based on their fitting qualities, and also that the mode choice depends upon distance, gender and household size.

3. Study Area

This study analyzes educational trips in the Kathmandu valley which is the largest urban center and contributor in the country's economy. Demographically the valley contains a population of permanent inhabitants of 11,95,285 persons [22] in 3 Rural municipalities, 16 municipalities, 1 Sub Metropolitan Municipality, and 1 Metropolitan Municipality. The total residents on the other hand are as high as 4 million.

Travel patterns wise a survey of 18,100 households [3] conducted to understand the nature of trips made by them, estimated the number of trips to be 3,483,349 in number. The maximum number of trips was for education followed by that for work and business. Mode wise the maximum number of trips was done by walking followed by motorcycle and then by public transportation. The study also concluded that if present trends continue the number of trips will increase by 1.56 times and the modal share of motorcycle and car users will increase while that of walking and public transport using commuters will decrease

4. Methodology

4.1 Data collection

A questionnaire was prepared based on analysis of previous works, they contain data for socio-economic characteristics like sex, age, family income, family education status, family size, sibling number, and Revealed preference (RP) part with characteristics of trip like mode selected, route length, cost and journey time. Students were briefed about the questionnaire and then allowed to fill in data, any queries or confusion during the filling process were discussed and solved.

The minimum sample size to be collected is 384 at a 95% level of confidence. The total number of samples collected for the study is 631 of which 440 (Around two third) were used for validation and 191 (Around one-third) for calibration purpose. The calibration data set and validation data set were selected randomly from the whole data set.

4.2 Model estimation

The coefficients for various attributes in the model can be evaluated by the maximum likelihood method. In this method the coefficients are evaluated such that the value of the likelihood function is maximum i.e. the value of parameters is set such that the observed sample is most likely to occur. This is the most common method of evaluating the parameters of the logit model. The coefficients will be evaluated using the *mlogit* package in R- Studio statistical package

4.3 Model validation

A confusion matrix is used for model validation the rows of this matrix denote the actual mode used while the columns denote the mode predicted by the model. The prediction is based on the maximum probability among the probabilities developed by the CDF of the model. Accuracy is the ratio of correct predictions to total sample size i.e. the ratio of the sum of diagonal to the sum of matrix

5. Data analysis

The total number of data separated for model formation was 440, the nature of data based on various parameters can be seen in the table 1

Table 1: Category wise mode choice analysis

Category	Selected Mode			Total	
	Walk	Public Transport	Personal Modes	Count	Percentage
Gender					
Male	116	161	70	347	78.86
Female	33	49	11	93	21.14
Age Group					
18-22	127	190	64	381	86.59
22-26	22	19	14	55	12.50
26-30	0	1	3	4	0.91
Family Income Level					
< 15000	15	17	3	35	7.95
15000-25000	25	34	8	67	15.23
25000-35000	45	41	28	114	25.91
35000-45000	37	36	13	86	19.55
> 45000	27	82	29	138	31.36
Family Education level					
Both High School	49	81	30	160	36.36
One only High School	62	83	32	177	40.23
None High School	38	46	19	103	23.41
Family Size					
Upto 4	73	118	49	240	54.55
5 to 8	67	86	30	183	41.59
9 to 12	7	4	2	13	2.95
More than 12	2	2	0	4	0.91
Sibling Number					
One	21	31	18	70	15.91
Two	53	81	40	174	39.55
Three	35	46	15	96	21.82
Four or more	40	52	8	100	22.73
Trip Distance					
Upto 2 Km	93	14	5	112	25.45
2 to 4 Km	39	28	14	81	18.41
4 to 6 Km	10	50	14	74	16.82
6 to 8 Km	4	33	14	51	11.59
Greater than 8 Km	3	85	34	122	27.73
Friend Availability					
Available	94	95	41	230	52.27
Not Available	55	115	40	210	47.73
Footpath Availability					
Available	120	145	51	316	71.82

Not Available	29	65	30	124	28.18
Vehicle ownership					
Yes	84	115	77	276	62.73
No	65	95	4	164	37.27

Concerning gender the majority (78.86%) of students are male while age wise the majority (86.59%) of students are of age 18-22 years. The family income of the majority (31.36%) of the students is high followed by medium (25.91%), while only one parent of majority (40.23%) of students have a high school or higher degree. Size wise majority (54.55%) of students belong to families having four or fewer members and siblings wise most (40.55%) of the students have two siblings indicating a high frequency of nuclear families in the population. Concerning trip character, majority of students are either living in the vicinity (25.45%) or are living in greater distances (27.73%) this is because students who live in rented accommodations live near the college while those living with their families are living in residential areas most of which are far from the areas where the colleges are located, in terms of a friend sharing the same or majority of the route students having such friends (52.27%) and that not doing so (47.73%) are same, similarly in terms of footpaths majority of students have footpaths in their routes of commutes, while footpath availability of Kathmandu has been reported poor [2,3] the high amount of footpath availability seen here may be due to perception differences of these studies and that of students towards a standard footpath for footpath availability to be sanctioned. Lastly, vehicle ownership wise majority (62.73%) of students have vehicle ownership which they can bring to the college or can drop/pick up them.

Mode wise maximum number of students i.e. 47.73% were using public transport, 33.86% of students walked to campus, while the least number of students i.e. 18.41% used personal modes of transport. Table 1 presents the category wise mode choice of students

On analyzing the mode choices according to the categories of different attributes, gender wise both the male and female students have preferred public transport more while the percentage of those doing so is more for female (52%) and less for male students (46%). Age wise younger aged students have preferred public transportation mode with 50% in this category, medium aged students have slightly preferred the walking mode with 40% in this category, despite low in number more aged students have preferred personal mode of transport with 75% in this category, students in fresher years living far from college while those in middleman years living in the vicinity and final year students being involved in extra activities can be attributed to this effect. In terms of income students from low and high income families have preferred public transportation more, while those from medium income families prefer walking more, preference towards public transport is more clearly seen in the high income group while other groups show a slight difference only, the fact that the socio economic characteristics of medium income families allow them to live closer to educational districts is the main cause for this, similarly educations wise each family education level group have highest preference to public mode maximum preference for public transportation can be seen in families with both parents having higher education (50%). In terms of family size, students from smaller families prefer public transportation more while those from larger families despite being low in number prefer walking while in terms of sibling number each category based on siblings has the highest preference for public transport mode, therefore the increased walking tendency in larger families cannot be related to walking habits associated with siblings, and can be due to conservative views in larger families towards personal and public modes of transportation which are not a part of the study. In terms of friend availability students with friends on the route are seen to prefer walking and public transport mode equally while students

who do not have a friend on the route prefer public mode more, similarly in terms of footpath availability both the groups prefer public modes of transportation, while walking and personal modes are found to be equally preferred after public transportation mode in scenarios with non-availability of footpath, walking is much more preferred to personal modes of transport where there is footpath availability. Mode wise both the groups based on vehicle ownerships prefer public transport options, however, the proportion of students doing so is higher in the students lacking vehicle ownership, similarly Walking and personal are found to be equally preferred in groups with vehicle ownership, while walking is much more preferred to personal modes of transport where there is lack of vehicle ownership.

6. Model Specification and analysis

The mixed logit model was developed in R-program using walking as the base option for the mode choice. An initial model was formed considering all the factors whose predictability was very low, the model was then refined in multiple steps using different approaches like retaining only one attribute among the group of attributes which have a high degree of correlation, removing variables where model choice is equally distributed by observing the data analysis while doing so it was made sure that the Mc Fadden R squared value of the model was retained as much as possible while the accuracy of the model increases. The final model was formed with model attributes journey time, gender, trip distance, number of siblings, family income, friends availability, price as seen in table 2. In terms of goodness of fit the Mc Fadden R^2 value of 0.340 which is within the range specified of 0.2 to 0.4 for a good model [23], similarly the P-value of likelihood is significant at almost all levels of significance. Thus it can be concluded that the model is better than the base model or intercept only model in terms of fitting.

7. Model Parameter Estimates

For personal modes of transport, the trip distance, number of siblings, and vehicle ownership are significant at a 1% level of significance whereas journey time and gender are significant at 10% level of significance. Family income and friend availability are not significant but kept in the model as excluding these variables do not improve the overall model fitness. Both journey time and trip distance with a coefficient of 0.007 and 0.680 indicate that with an increase in trip distance and (or) journey time the odds of selecting personal modes of transport over walking increase. Gender with a coefficient of -0.901 indicates that the odds of using personal transport modes over walking decrease if the student is female provided all other factors are same. The greater preference for walking in females may be due to higher health concerns [24] or due to general liking towards walking with friends [15]. The number of sibling with a coefficient of -0.413 indicates that the odds of selecting personal modes over walking decreases as the number of siblings increase factors like personal mode of transport not able to accommodate many people and walking with siblings being less tiresome as compared to walking alone etc. are the causes behind it. In terms of vehicle ownership the coefficient of -3.127 indicates that lack of vehicle ownership decreases drastically the odds of selecting personal modes over walking. Except a narrow possibility of sharing someone else's mode of transport if a route or a major part of it matches, using a personal mode of transport requires that there is ownership of a vehicle, therefore, the nature of this coefficient seems logical.

For public modes of transport, only trip distance is significant at a 1% level of significance as shown in table 2. Trip distance with a coefficient of 0.656 indicates that with an increase in trip distance selecting public modes of transport over walking increases. The nature of the effects of other factors that are not statistically significant can be seen from table 2.

Table 2: Model information

Variable	Personal Modes of Transport		Public Transport	
	Coefficient	Std. Error	Coefficient	Std. Error
Journey Time	0.007	0.004*	0.007	0.004*
Gender (Reference = Male)	-0.901	0.493*	-0.0901	0.390
Trip Distance	0.680	0.0649***	0.656	0.048***
Number of Siblings	-0.413	0.108***	-0.094	0.106
Family Income (Reference = Less than 15000)	-0.025	0.177	0.174	0.142
Vehicle ownership (Reference = Vehicle ownership yes)	-3.127	0.602***	-0.050	0.335
Friend availability (Reference = Friend availability yes)	0.04	0.38	0.234	0.326
Price	0.0066	0.0066	0.0066	0.0066
Sd. of Price	3.53X10 ⁻⁵		3.53X10 ⁻⁵	
Model fitting information				
Mc Fadden R ²	0.340			
Likelihood ratio test	301.62			
P-value of likelihood	2.22X10 ⁻¹⁶ ***			

(Level of significance * 10% ** 5% *** 1%)

Confusion matrix developed from the validation and used for the validation of this model is shown in table 3. the rows of this matrix denote the actual mode used while the columns denote the mode predicted by the model.

Table 3: Confusion Matrix

	Public	Walk	Personal
Public	56	0	0
Walk	0	0	0
Personal	28	0	0
Accuracy	74.34%		

8. Conclusion and Recommendation

This study provides an insight into what factors affect the mode choice and thereby guides the policymakers, planners, and infrastructure-related stakeholders about the conditions that demand motorized or non-motorized commutes.

Since motorized commutation is more preferred in longer distances and walking in shorter ones, it is recommended to develop proper mass transit facilities between the residential areas and the educational hubs while developing proper walking conditions within the educational hub.

Similarly, females prefer walking more than motorized, based on which it is recommended that more gender-friendly walking conditions be promoted. Also, the probability of choosing public transport is more as compared to private modes for longer distances. Based on this, it is recommended to make the public transport more women-friendly.

Lastly, vehicle ownership is found to decrease the odds of walking or public transportation being used therefore planning and policy interventions are recommended in this field.

References

1. Sir Rod Eddington, The Eddington transportation study, Vol 1 – “*Transport’s role in sustaining the UK’s productivity and competitiveness*”, UK Government, UK, 2006.
2. Manav Kendrit Yatayat Abhiyan (MaYA), MaYA Fact sheet Number 4 – “*Public transport in Kathmandu valley restructuring and reforming*”, Clean Air Network Nepal (CAAN)/Clean Energy Nepal (CEN), Kathmandu, 2014 March.
3. JICA, “*Data collection on traffic improvement in Kathmandu valley*”, Japanese International Cooperation Agency(JICA), Kathmandu, 2012.
4. The World Bank, “*Gender and Public transport*”, World Bank, Kathmandu, 2013.
5. CANN, CANN Fact sheet no 1: “*Walkability in Kathmandu*”, Clean Air Network Nepal, Kathmandu, 2010
6. Noreen C. McDonald, Austin L. Brown, Lauren M. Marchetti, Margo S. Pedroso, “*US school travel*”, American Journal of Preventive Medicine Volume 41, Issue 2, August 2011, Pages 146-151, 2009,
7. Feng, J.-T., Lin, J.-J., “*School travel modes and children’s spatial cognition*”, Urban Study. 54 (7), pp 1578–1600, 2017.
8. Carlos J.L. Balsas. , “*Sustainable transportation planning on college campuses*”, Transport policy pp 35-39, 2003.
9. Baghestani, Amirhossein&Mamdoohi, Amir Reza &Abrishami, Seyed, “*A Comparison of Aggregate and Disaggregate Mode Choice Models Transferability.*” , International Journal of Transportation Engineering, Vol.2/No.2, pp 146 – 154, 2014.
10. Chansung Kim, Chang Gyu Choi, Seongkil Cho &DaehyonKimx, “*A comparative study of aggregate and disaggregate gravity models using Seoul metropolitan subway trip data.*” , Transportation Planning and Technology, 32:1, pp 59-70, 2014.
11. Robin Miles-Mclean, Michael Shelby, Clarisse Lula, Michael Sagan, Kenneth E. Train, “*Comparison of forecasts from aggregate and disaggregate models for personal vehicle energy consumption and CO2 emissions*”, Energy, Volume 17, Issue 4, April 1992, Pages 321-329, 1992.
12. Joachim Scheiner, Oliver Hubert, Stefan Lohmüllerb, “*Children’s mode choice for trips to primary school: a case study in German suburbia.*”, Travel Behaviour and Society 15 pp15–27

13. Rui Zhang, Enjian Yao, Zhili Liu. 2017, "School travel mode choice in Beijing, China.", *Journal of Transport Geography* 62 pp 98–110, 2019.
14. Lu Maa, Hui Xiongb, ZhongWangc, KaiqiangXieb. "Impact of weather conditions on middle school students' commute mode choices: Empirical findings from Beijing, China." ,*Transportation Research, Part D : Transportation and Environment*, Volume 68 , pp 39-51, March 2019.
15. Nishant Singha,VinodVasudevan, "Understanding school trip mode choice – The case of Kanpur (India).", *Journal of Transport Geography* 66 pp 283–290, 2018.
16. Rai Binaya, "Estimation of work trip mode choice model for public transit Riders in Kathmandu Valley", M.Sc Thesis, Department of Civil Engineering, Tribhuvan University – Institute of Engineering, Pulchowk, Lalitpur, Nepal, 2012.
17. Ghimire Anukshya, "Mode choice modeling for work trips in Kathmandu Valley", M.Sc Thesis, Department of Civil Engineering, Tribhuvan University – Institute of Engineering, Pulchowk, Lalitpur, Nepal, 2019.
18. Joshi Manmohan, "Mode choice modeling for Intercity travel in Nepal", M.Sc Thesis, Department of Civil Engineering, Tribhuvan University – Institute of Engineering, Pulchowk, Lalitpur, Nepal, 2019.
19. DuyQuy Nguyen-Phuoc, Richard Amoh-Gyimahb, Anh Thi Phuong Trana, Cao ThoPhana, "Mode choice among university students to school in Danang, Vietnam.", *Travel Behaviour and Society* 13 pp 1–10, 2018.
20. JoonKyu Le, Kwang-EuiYoo Ki-Han Song, "A study on travelers' transport mode choice behavior using the mixed logit model: A case study of the Seoul-Jeju route." , *Journal of Air Transport Management* pp 1-7, 2016.
21. Tongzhou Bai, Xuemei Lia, ZhiyuanSunc, "Effects of cost adjustment on travel mode choice: analysis and comparison of different logit models." , *Transportation Research Procedia* 25C pp 2653–2663, 2017.
22. CBS, "Population of 753 Local Unit", Central Bureau of Statistics (CBS), Kathmandu, 2074
23. Daniel McFadden, Antti P. Talvitie, and associates., *Urban Travel Demand Forecasting Project Phase 1 Final Report, Vol.V*, The institute of transportation studies, University of California, Berkley, California, USA, 1997
24. Kelly J Clifton and Andrea D. Livi., "Gender difference in walking behavior, attitudes about walking behavior and perceptions of the environment in three Maryland communities", *Research on women's issues in transportation- conference report, Volume 2 technical papers*, pp 79-88, 2004.