RESEARCH ARTICLE

MODELING TRANSPORTATION PREFERENCES OF UNIVERSITY STUDENTS OF KARACHI USING PROMINENT FACTORS: A BINARY RESPONSE MODEL COMPARISON APPROACH.

Muhammad Ahsanuddin[1] and Zohaib Aziz[2].
1. Assistant Professor, Department of Economics, University of Karachi, Karachi, Pakistan.
2. Lecturer, Department of Statistics, Federal Urdu University of Arts, Science and Technology, Karachi, Pakistan.

……………………………………………………………………………………………………....

Manuscript Info	*Abstract*
……………………..	………………………………………………………………

This study examines the perceptions of students regarding the factors affecting their choice of the public or private transportation mode to come to their educational institutes. The data was collected through filling the questionnaires from 169 undergraduate and graduate students. Binary logit, probit and complementary-log-log model are used for the modeling of transportation mode variable through four explanatory variables namely gender, conservatives, distance and concerns. Binary logit model is found to be better than others. Gender and concern variable are found more influencing on the transportation medium than distance and conservative variables. The results in favor of high probability of private transport mode are alarming with respect to environmental concerns, maintenance of law and order as well as some ethical and moral values.

…………………………………………………………………………………………………....

Introduction:-
In modern times transportation is one of the basic human needs. It is important for manifold reasons that include travel, economic activity and mobility. Transport acts as a catalyst for migration which spurs up economic development and social transformation of many nations (Nistor & Popa, 2014). Transportation provides an effective way to transport people and commodities from one place to another and the common methods of transportation include buses, planes, trains, boats, cars and other vehicles that can be characterized in public and private transportation mode.

Students are one of the key users of transport. The objective of this study is to analyze the factors that affects the choice of transport mode used by the university students. In this context, the survey based analysis regarding the choices that affect the modes of transportation of students is needed. Therefore, we make the questionnaire to identify what students prefer and what are the main reasons behind their preference. We have considered various factors such as gender, concerns, family background, financial constraints, the distance to be covered and many other factors that affect one's decision and select the appropriate variables from it for the analysis.

Literature Review
Mohammed and Shakir (2013) models the modes of transportation chosen by the graduate students of university of Malaysia. They use the binary logit model to explore the effects of different variables on the transportation medium.

Corresponding Author:-Muhammad Ahsanuddin.
Address:-Assistant Professor, Department of Economics, University of Karachi, Karachi, Pakistan.

Survey based analysis is performed by taking 456 graduate student of university. The result suggests that reduction in traveling time and cost can be helpful to motivate the students to shift their transportation medium from private to public. This may be helpful to reduce the environmental population. Urban Institute Student Transportation Working Group (2017) investigated whether in urban areas student transportation is conducive or a barrier in accessing quality education.

Miskeen et al. (2013) investigates the intercity traveling medium (private car, bus or airplane) in Libya by using multinomial logit model. Survey based analysis is performed by taking the views of 1300 respondents. Two-third of data is used for modeling and remaining for its validation. The model indicates that gender, age, traveler nationality, monthly income, car availability, purpose of travel, duration of stay at destination, egress distance to airport/bus terminal, total travel cost and mode characteristics (privacy and convenience) are the impacted factors for intercity traveling medium in Libya.

Muller et al. (2008) models the travel modes of students in Germany by multinomial logit model to check the shifting of low transportation medium. To achieve this objective, comprehensive survey is conducted and data disaggregation and geocoding method is used to select the variables. They find that distance, weather and own transport availability are most influencing factors to shift the medium of transport.
Ewing et al. (2004) test the effect of income, time, license holding and other variables on the traveling mode (Car, Bus, Walk, Bike) for the K-12 students of Gainesville high schools, Florida to reach their schools. A multinomial logit model is fitted on 819 k-12 school trips. Findings reveal that time, distance and safe routes for walking are the significant factors on the travelling mode.

Data
The data is taken from the survey conducted on different public and private university students. A total of 169 respondents were involved in questionnaire filling process. Convenience sampling is performed to select the samples. Although several variables are collected through this survey, we have focused on few variables and just for convenience made some amendments. Response variable i.e. medium of transportation (MoT) is classified in two categories, viz public and private. Gender, Conservative, Distance and Concerns are considered the potential variables that may affect the transportation medium of students. Table 1 explains the description and scale of response and explanatory variables.

Table 1:-Description of response and explanatory variables

Variable	Description	Measurement
Medium of Transportation (MoT)	Binary variable for the students use the public and private transport	Public=0, Private=1
Gender	Binary variable for the Sex of the student	Male=0 or Female=1
Conservative	Binary variable for the student labels yourself and/or your family as conservative	Yes=0 or No=1
Distance	Binary variable for the distance $\leq$ 10 km versus > 10 km	Distance $\leq$ 10 km=0 or Distance > 10 km=1
Concerns	Binary variable for concerns of the student about the selection of mode of transportation.	Cost=0 or Convenience & Security=1

Methodology:-
Logistic Model
Let $X_i = [1, X_{i1}, X_{i2}, \dots, X_{ik}]$, for $i = 1, \dots, N$ be a row vector of explanatory variable. Let Y_i be a binary response coded as a 1 to denote a success and a 0 to denote a failure. The Y_i are independent Bernoulli random variables with success probabilities $P(Y_i|X_i) = \pi_i$ for $0 \leq \pi_i \leq 1$. The logistic regression response function is

$$\pi_i = E(Y_i|X_i) = \frac{\exp[\beta'X_i]}{1+\exp[\beta'X_i]} \tag{1}$$

Where β is a vector of unknown parameters. The logit transformation is

$$\text{logit}(\pi_i) = \log\left(\frac{\pi_i}{1-\pi_i}\right) = \beta'X_i \tag{2}$$

So,

$$\text{logit}(\pi_i) = \log\left(\frac{\pi_i}{1-\pi_i}\right) = \beta_0 + \beta_1 \text{Gender}_i + \beta_2 \text{Conservative}_i + \beta_3 \text{Distance}_i + \beta_4 \text{Concerns}_i \quad (3)$$

Probit Model
A probit model whose tolerance distribution is Normal distribution is described as for the given data

$$\Phi^{-1}(\pi_i) = \beta_0 + \beta_1 \text{Gender}_i + \beta_2 \text{Conservative}_i + \beta_3 \text{Distance}_i + \beta_4 \text{Concerns}_i \quad (4)$$

Where Φ denotes the cumulative probability function for the standard Normal distribution N (0,1)

Complementary-Log-Log Model
A complementary-log-log model whose tolerance distribution is the extreme value distribution is explained as

$$\log[-\log(1-\pi_i)] = \beta_0 + \beta_1 \text{Gender}_i + \beta_2 \text{Conservative}_i + \beta_3 \text{Distance}_i + \beta_4 \text{Concerns}_i \quad (5)$$

Diagnostics of Binary Response Model
1 Likelihood Ratio Test
Likelihood ratio test is appropriately used to check the diagnostics of the nested model. It has the statistic

$$\text{LR} = 2(L_u - L_r) \sim \chi^2(p - q) \quad (6)$$

Where L_u and L_r are the log-likelihood of model of interest with parameters p and minimal model with q respectively.

McFadden R^2
Pseudo R^2 represents the proportional improvement in the log-likelihood function due to terms in the model of interest, compared with the minimal model. It is defined as

$$\text{Pseudo } R^2 = 1 - \frac{L_u}{L_r} \quad (7)$$

Where L_u and L_r are the log-likelihood of model of interest with parameters p and minimal model with q respectively.

Results and Discussion:-
Table 2 shows the estimated coefficients and AIC and BIC of logit, probit and complementary-log-log models. If we compare the AIC and BIC of all three models, there is found no significant difference. Perhaps the logit model has the minimum value of information criteria. Moreover, magnitude of the coefficient estimates of the logit model are higher than the other models. It implies the greater probability of response variable for private transportation medium than other model. The interpretation of the coefficients for the qualitative response model is not significant in simple manners, thus, we resort to the other methods like odds ratio to interpret the effect of explanatory variables on response variable.

Table 2:-Coefficient estimates of different binary response models

Coefficients	Logit	Probit	Complementary-log-log
β_0	0.1318	0.0882	-0.2402
β_1	0.7521	0.4250	0.3680
β_2	0.0825	0.0686	0.0893
β_3	0.4668	0.2476	0.1864
β_4	0.7307	0.4438	0.4338
AIC	174.58	174.64	174.69
BIC	190.23	190.29	190.33

Table 3 shows the odds ratios and their 95 % Wald confidence interval for the estimated coefficients of logit, probit and complementary-log-log model. These odds ratios measure the effect a variable has on the probability of success while holding the other variables constant.

Interpretation of Logit Model

As we find logit model is better than others so we interpret the intervals of odds ratio of logit model. From Table 3, the estimated odds of private transportation medium are between 0.2 and 1.1 (1/0.90 and (1/4.98) times higher for female students than male. This result is reasonable to believe that there are some society and religious constraints for the female to avoid the public transports.

The estimated odds of private transportation medium are between 0.37 and 2.32 (1/0.43, 1/2.71) times higher for non-conservative students than conservatives. This result suggests that usually above medium or high-income class students can afford the private transport mode and they normally are considered non-conservatives or liberals in our society.

The estimated odds of private transportation medium are between 0.28 and 1.408 (1/0.71, 1/3.57) times higher for the students whose distance from house to university is greater than 10 km than those who have less than 10 km. This finding supports that students coming from longer distances usually prefer private transport over public.

The estimated odds of private transportation medium are between 0.18 and 1.28 (1/0.78, 1/5.56) times higher for the students whose concerns for the usage of private transportation medium are convenience and security than cost factor. According to this result, students are in the favor of private transportation medium due to comfort and security issues.

The Wald confidence interval of odds ratios for remaining two models namely Probit and Complementary-log-log models have no as much difference as logit has.

Table 3:-Odds ratios of estimated coefficients and their Wald confidence interval

Coefficients	Logit		Probit		Complementary-log-log	
	Odds ratio	95 % Wald C.I[*]	Odds ratio	95 % Wald C.I[*]	Odds ratio	95 % Wald C.I[*]
β_0	1.14	(0.34,3.88)	1.09	(0.53,2.25)	0.77	(0.39,1.58)
β_1	2.12	(0.90,4.98)	1.53	(0.95,2.47)	1.44	(0.95,2.19)
β_2	1.09	(0.43,2.71)	1.07	(0.64,1.80)	1.09	(0.69,1.72)
β_3	1.59	(0.71,3.57)	1.28	(0.81,2.03)	1.20	(0.81,1.79)
β_4	2.07	(0.78,5.56)	1.56	(0.86,2.83)	1.54	(0.85,2.79)

*95% Wald CI for odds ratio

Table 3 shows the likelihood ratio statistics and pseudo R^2 for three models as we can see our logit model is significant at 10 % and also has greater pseudo R^2 in three decimal places. Although logistic regression R^2type measures often appear alarmingly small even other measures suggest that the model fits the data well. Mittlbock and Heinzel (2001) give the reason that pseudo $R^2$is a measure of predictability of individual outcomes Y_i rather than the predictability of all the event rates. Accordingly, logit model is appropriate for this case.

Table 3:-Goodness of measures of fitted models

	Logit	Probit	Complementary-log-log
Likelihood Ratio Statistic (P-Value)	7.83 (0.09)	7.77 (0.10)	7.72 (0.10)
McFadden R^2	0.0453	0.0450	0.044

Conclusion:-

This research concludes that logit model is suitable for the modeling of transportation medium for the given case compared to probit and complementary-log-log model. Second, factors that have a high impact on the transportation medium are gender and concerns of students about their security and convenience. Other factors like distance and conservation also affect but are not found as much weighted according to odds ratios. Almost all factors are going in favor of private transport. But, the increasing tendency of private transport mode of students is alarming for us. Therefore, we recommend that public transport should be given preference over private transport to decrease the roads traffic, pollution etc. Nonetheless, the government should provide standard public transport, safety and respect for female students in the public transport by giving a separate safe compartment. Also, the government should

make strong laws for the security of students in public transport and for the distance to be covered in given suitable time so that the students can be motivated to prefer the public transport mode over the private.

References:-

1. Ewing, R., Schroeer, W., & Greene, W. (2004). School location and student travel analysis of factors affecting mode choice. Transportation Research Record: Journal of the Transportation Research Board, (1895), 55-63.
2. Miskeen, M. A. A. B., & Rahmat, R. (2012). Influencing Intercity Mode Choice Behavior in Libya. International Journal of Information and Electronics Engineering, 2(4), 547-551.
3. Mittlböck, M., & Heinzl, H. (2001). A note on R 2 measures for Poisson and logistic regression models when both models are applicable. Journal of clinical epidemiology, 54(1), 99-103.
4. Mohammed, A. A., & Shakir, A. A. (2013). Factors that affect transport mode preference for graduate students in the national university of Malaysia by logit method. Journal of Engineering Science and Technology, 8(3), 351-363.
5. Müller, S., Tscharaktschiew, S., & Haase, K. (2008). Travel-to-school mode choice modelling and patterns of school choice in urban areas. Journal of Transport Geography, 16(5), 342-357.
6. Nistor, F., & Popa, C. C. (2014). The Role of Transport in Economic Development.
7. Urban Institute Student Transportation Working Group. (2017). Student Transportation and Educational Access: How Students Get to School in Denver, Detroit, New Orleans, New York City, and Washington, DC. Washington, DC: Urban Institute.