MODELING COMMUTER PREFERENCES FOR A BUS

Eric Nyarko, M. Phil
Caleb Owusu-Ansah, M. Phil
Charles Kwofie, M. Phil
Department of Statistics, University of Ghana, Legon, Ghana

Abstract

The research deals with the use of the discrete choice experiment technique and Random Utility Theory (RUT) to analyze bus choice behaviour for commuters in Asafo-market in Kumasi, Ghana. The principal aim of the study was the calibration of a logit model to forecast consideration sets. In order to estimate the model parameters, a specific survey was carried out inside the urban area of Kumasi. The survey focused on passengers and involved mainly employees, self employed workers and students (120 respondents). The data collected from the experimental survey was further calibrated and segmented according to gender. The magnitude of estimates generally indicates that commuters highly value buses with fixed departure time, spacious seats and Air Condition (AC). However, an increase in fare level will result in a disutility of bus choice.

Keywords: Bus, discrete choice experiment, discrete choice modeling, Random Utility Theory

Introduction

Transport is an indispensable element of development and socio-economic growth. As engine of economic integration, transport infrastructure and services facilities constitute a precondition for facilitating trade and the movement of goods and person. Long perceived as a tool for accessing national and regional trade in a radically changing global environment, transport infrastructure remains a pillar of development with a view to accelerating growth and reducing poverty (Okoko, 2006). Road transport is the predominant means of travelling in Ghana, which enhances high passenger travels and carting of goods and services. It provides essential role by linking the country to others in the entire West African sub-region. Transportation has developed rapidly in Ghanaian societies, but there is competition between privately owned cars and commercial vehicles. Despite the high growth rate in urban centers like Accra and Kumasi, there have been some shortfalls in public policy. This has contributed to longer shuttling period and journey delays, lengthy waiting times for commercial vehicles both at and between terminals, high accident rates, and localized poor air quality (Afful, 2011). However, as a result of the poor quality of travel in bus transportation systems in Ghana with a declining trend in commuters' choice of buses, policy-makers and transport operators are constantly in search of solutions for improving bus choice, especially in urban areas. In fact an increase in Bus use, with a concurrent reduction in the use of private cars, could help to reduce many problems like traffic congestion, air and noise pollution, and energy consumption. For these reasons, several works have been made by various studies on urban public transport; for example, Baidoo, Nyarko and Mettle (2015) in their study of modeling mode choice in passenger transport with discrete choice experiment revealed that in choosing a commercial vehicle, commuters generally took into consideration their safety, travel distance comfort and waiting time before making their choices. Travel safety is highly valued by commuters. Pavlyuk and Gromule (2010) in their study considered three transport options; car, coach, and train. A nested discrete choice model was used to analyze factors that influence passengers' choice. The authors concluded that departure time had a significant influence on bus/train choice. Passengers who choose price as a key factor in their selection prefer to use the train. The terminal point as a destination predictably increases the probability of train selection. Baidoo and Nyarko (2015) examined mode choice between bus and private car, with the habit of using a bus being one of the attributes. Binary logit model and its marginal effects were employed to assess commuters' behaviour with regards to their choice between different transportation modes in traveling to Accra central. The authors concluded that the level of noise, comfort, and time (morning trips) will result in a disutility of public transport choice. Van der Waerden, Borgers, Timmermans, and Berenos (2007) used MNL models to examine the choice between car, bus and bicycle for different journey purposes. They argued that the cost and time attributes dominate, obtaining a seat is significant across journey purposes. Baidoo and Nyarko (2015) employed Discrete Choice Experiment (DCE) and Random Utility Theory (RUT) to measure service quality in public transport. Probit model was calibrated and segmented based on gender. They concluded that an increase in the walking distance to bus stop and transport fare will result in a disutility of service quality attributes. Catalano, Lo Casto and Migliore (2008) employed random utility model to analyze travel mode choice behaviour for commuting urban trips in Palermo, Italy. The authors found out that, for the specific case of Palermo, the multinomial logit proved to be the best urban transport demand model, even if the choice set contained three car alternatives.

However, as far as the authors are concern, most of the studies that make use of DCEs are carried out in the Western world with paucity of information on users' attitude when they have a mode of choice between buses.

In this study, a Discrete Choice Experiment (DCE) and Discrete Choice Modeling (DCM) which is rooted in Random Utility Theory (RUT) are used to estimate bus passengers' attitude towards endogenous consideration sets. This will help to propose policy intervention issues in urban areas in developing countries.

Methods and Materials

Sample and Data Collection Procedure

The data were collected using Computer Assisted Personal Interview (CAPI). This explains a large response rate for a Discrete Choice Experiment (DCE). Commuters were sampled using simple random sampling techniques for the study. Commuters who board buses from the Kumasi-Asafo bus station to various destinations (Cape Coast and Accra, for example) and owned or have access to a private car were targeted since the study sought to analyze the hypothetical choice of a bus by these people. A sample size of 120 respondents was chosen for the study. A total sample of fifty (50) individuals each with 16 choice sets and fully generic parameter specification for design attributes and covariate effects might just be acceptable for choice experiment (Hensher et al., 2005).

Revealed/Stated Choice Design

The basic shortcomings of SP surveys are not present in RP surveys as they deal with existing actual situations being experienced by the user. But their general suitability is restricted (Kroes and Sheldon, 1988), the reasons being;

1. Observations of actual choices may not provide sufficient variability in the revealed data for constructing good models for evaluation and forecasting.
2. The observed behaviour may be dominated by a few factors making it very difficult to detect the relative importance of other variables or understanding trade-off between them.
3. RP data cannot be used in direct way to evaluate demand under conditions, which do not yet exist or in collecting responses for policies, which are entirely new.
4. RP data require that the explanatory variables can be expressed in objective or engineering units. Hence, they are normally used for primary

service variables and are rarely used to evaluate the effect of changes in secondary variables.

Stated choice experiment has been widely used in transportation. In this study, commuters were tasked to choose between hypothetical buses from a binary choice set. This approach requires commuters to trade-off the different aspects of the bus attributes/levels in a choice task. SP exercises provide an opportunity of trade-offs between the options available and increase the number of responses, as at each trade-off the commuter indicates his/her preference. However, the alternatives in the choice task were considered to be the main factors influencing bus choice in the Asafo-market bus station in Kumasi.

Determining Attributes and Associated Levels

The authors identified five bus attributes and accompanying levels to be the most important on the basis of extensive preparatory qualitative research. The initial list of potential attributes and their levels were established through in-depth interviews with local experts in bus transport and bus passengers at Kejetia bus station in Kumasi (the second largest city in Ghana). Adamowvic et al. (1998) opined that attributes are commonly identified from prior experience, primary or secondary research. However, focus group discussion was conducted to reduce the initial list of potential attributes to five (i.e., fare level, departure time, cleanliness of bus inside, arrangements of seats and air conditioning). The bus-choice attributes and their levels are defined in Table 1.

Table 1: Attributes and corresponding levels

Attributes	**Corresponding levels**
Departure time	Fixed
	Varies
Fare level	0-5% more
	10-15% more
	More than 30%
Cleanliness of bus inside	Clean enough
	Not clean enough
Air conditioning	Available
	None
Arrangements of seats	Spacious
	Congested

The next stage in DCEs is the experimental design process to elicit the choice sets to be presented to the commuters. The choice sets for the DCE questionnaire were generated using well-established statistical methods. We used DCE macros in the statistical programme SPSS to generate optimal orthogonal design with eight profiles. This method takes account of orthogonality (attribute levels are independent of each other), level balance (attribute levels appear with the same frequency), and minimal overlap (attributes do not take the same level within a choice set) (Kuhfeld, 2010). The profiles were combined to generate 28 choice sets, which according to literature, is within the acceptable range for DCE studies. Prior to the field data collection, the survey questionnaire was pretested to evaluate the reactions of the respondents, the appropriateness of the questions, and the suitability of format and wording of questions. During the main survey, commuters were tasked to evaluate the choice sets and indicate the kind of bus they would board or prefer when they have access in choosing between buses that are loading in a bus station (Asafo market in Kumasi). Table 2 portrays a choice set presented in the stated preference survey.

Table 2: Choice set submitted to commuters

Which of these two buses do you prefer to board?

Attribute	Bus 1	Bus 2
Departure time	Fixed	Varies
Fare level	0-5% more	0-5% more
Cleanliness of bus inside	Clean enough	Not clean enough
Air conditioning	Available	None
Arrangement of seats	Congested	Spacious
Which bus would you choose?	**Bus 1 [ ]**	**Bus 2 [ ]**

Econometric Specification

Discrete choice experiment modeling is rooted in Random Utility Theory (RUT). However, the utility can be modeled as;

$$U_{ci} = V_{ci} + \varepsilon_{ci} \tag{1}$$

Where V_{ci} is the deterministic term of the utility and ε_{ci} is the random term, taking care of the uncertainty. The deterministic term V_{ci} of each alternative is a function of the attributes of the alternative itself and the characteristics of the commuter.

McFadden (1974) opined that a utility can be characterized by a function;

$$U_{ci} = \alpha + \sum_{k-1}^{K} \beta_k X_{cki} + \sum_{m-1}^{M} \gamma_m Z_{mi} + \sum_{k-1}^{K} \sum_{m-1}^{M} \delta_{km} X_{cki} Z_{mi} + u_{ci} \tag{2}$$

Where vehicle choice $c = \{\text{Bus 1, Bus 2}\}$ and $i = 1...N$ refers to commuters, X is a vector of K attribute levels, and Z is a vector of M personal characteristics. The parameter β_k refers to the utility associated with bus attribute k and the parameter δ_{km} measures how this utility varies by a specific characteristic of the commuter. The term u_{ci} is random and represents unobservable influences on commuter choice. The framework assumes that the commuter chooses the bus which generates more utility. The utility gain from choosing bus 1 over bus 2 for commuter i is:

$$U_{Bi} - U_{Ai} = \sum_{k-1}^{K} \beta_k \left(X_{Bki} - X_{Aki} \right) + \sum_{k-1}^{K} \sum_{m-1}^{M} \delta_{km} \left(X_{Bki} - X_{Aki} \right) Z_{mi} + \left(u_{Bi} - u_{Ai} \right) \qquad (3)$$

The random component u_{ci} may be hypothesized to consist of three additive components; commuter specific component v_i, bus choice specific component e_c and a true iid random term. Of these, the commuter specific term cancels out. The bus choice specific component can be assumed to be zero, unless the respondents have a consistent tendency to be more or less likely to respond to bus 1 instead of bus 2. Suppose the commuter chooses bus 2 if $U_{Bi} - U_{Ai} > 0$. This takes place with the probability:

$$P\left[U_{Bi} - U_{Ai} > 0\right] = P\left[\sum_{k-1}^{K} \beta_k \left(X_{Bki} - X_{Aki} \right) + \sum_{k-1}^{K} \sum_{m-1}^{M} \delta_{km} \left(X_{Bki} - X_{Aki} \right) Z_{mi} + \left(u_{Bi} - u_{Ai} \right) \ > 0 \right]$$

$$= P\left[\left(u_{Ai} - u_{Bi} \right) < \sum_{k-1}^{K} \beta_k \left(X_{Bki} - X_{Aki} \right) + \sum_{k-1}^{K} \sum_{m-1}^{M} \delta_{km} \left(X_{Bki} - X_{Aki} \right) Z_{mi} \right] \qquad (4)$$

Assuming a distribution for $\left(U_{Ai} - U_{Bi} \right)$, for instance a logistic distribution, the probability in (4) can be expressed in terms of a logistic cumulative distribution and modeled accordingly with logit:

$$P\left[U_{Bi} - U_{Ai} > 0\right] = F\left[\sum_{k-1}^{K} \beta_k \left(X_{Bki} - X_{Aki} \right) + \sum_{k-1}^{K} \sum_{m-1}^{M} \delta_{km} \left(X_{Bki} - X_{Aki} \right) Z_{mi} \right] \qquad (5)$$

Where $F\left(x\right) = \dfrac{e^x}{1+e^x}$.

The logit model employed for the study based on RUT was therefore stated as;

$$Logit\left(Y = \tfrac{1}{X} \right) = Logit\left(U_{Bus1} > U_{Bus2} \right) \qquad (6)$$

This paper estimates equation (6) with a binary logit model where the levels of bus choice attributes are treated as separate dummy variables in the regression analysis. The response variable (bus choice) is assigned 1 if bus 1 is chosen and 0 if bus 2 is chosen.

Model Results and Discussion

The results reveal that there is goodness-of-fit of the model from the data. The likelihood ratio chi-square of 425.12 with a p-value of 0.000 tells us that the model as a whole is statistically significant, that is, it fits significantly better than a model with no predictors. All estimated coefficients have the expected sign and are significant at the 95% confidence level. However, buses that are clean enough and transport fare over 10% to

15% of normal fare are not significant. Bundle of buses such as fixed departure time, present of air condition and spacious seats increase the utility and thereby increase the probability of choosing a bus to those without. Buses with fare level over 0% to 5% of normal fare decreases the utility associated with bus choice, though it is significant. This attribute level will be traded-off for buses with fixed departure time, air condition and spacious seats.

Table 3: Logit estimates of the DCE model

Attributes	**Coefficient**	**Z Value**	**P>|Z|**	**[95% Conf. Interval]**	
Departure time (Fixed)	0.55960	11.08	0.000	0.46060	0.65861
Fare level (0-5% more)	-0.42478	-7.17	0.000	-0.54092	-0.30864
Fare level (10-15% more)	-0.03996	-0.70	0.486	-0.15237	0.07245
Cleanliness of bus inside (Clean enough)	-0.06891	-1.37	0.172	-0.16785	0.03002
Air conditioning (Available)	0.19829	3.93	0.000	0.09935	0.29723
Arrangement of seats (Spacious)	0.52305	8.81	0.000	0.40673	0.63938
Constant	-0.38494	-5.08	0.000	-0.53350	-0.23637
Number of observations	6716				
Prob> χ^2	0.000				
Likelihood χ^2	425.12				
Rho-square	0.0457				

Table 4: Marginal effects after logit model

Attributes	**dy/dx**	**Std. Error**	**Z Value**	**P>|Z|**	**[95% Conf. Int.]**	
Departure time (Fixed)	0.13899	0.01238	11.22	0.000	0.11472	0.1632
Fare level (0-5% more)	-0.10579	0.01465	-7.22	0.000	-0.13450	-0.0770
Fare level (10-15% more)	-0.00999	0.01434	-0.70	0.486	-0.03808	0.0181
Cleanliness of bus inside	-0.01722	0.01262	-1.37	0.172	-0.04195	0.0075
Air condition (Available)	0.04953	0.01259	3.93	0.000	0.02485	0.0742
Arrangement of seats	0.13002	0.01459	8.91	0.000	0.10143	0.1586

The result of the marginal effects from Table 4 indicates that for bus choice, attributes/levels such as fixed departure time, availability of air condition and spacious seats increase the utility as well as the change in the probability of bus choice. The level of fare decreases the change in the probability of bus choice. Cleanliness of bus is insignificant even though it decreases the change in the probability of bus choice.

Restricted model by gender

The results in Tables 5 and Table 6 show that buses with fixed departure time and spacious seats are all significant and increase the utility associated with the choice of buses to those without. These attributes also have the same effect as those estimated in the unrestricted model. However, there is difference in the choice of buses by gender; spacious seats increases the utility associated with male passengers' choice of buses, air condition in buses increases the utility associated with female passengers' choice. Generally, the level of bus fare will result in disutility of bus preference. However, this attribute level will be traded off for other factors.

Table 5: Restricted model by female passengers

Attributes	**Coefficient**	**Z Value**	**P>|Z|**	**[95% Conf. Interval]**	
Departure time (Fixed)	0.57307	9.09	0.000	0.44952	0.69661
Fare level (0-5% more)	-0.36516	-4.75	0.000	-0.51598	-0.21433
Fare level (10-15% more)	-0.01241	-0.17	0.862	-0.15252	0.12769
Cleanliness of bus inside (Clean enough)	0.02156	0.34	0.732	-0.10189	0.14501
Air conditioning (Available)	0.30002	4.76	0.000	0.17656	0.42347
Arrangements of seats (Spacious)	0.51950	6.74	0.000	0.36840	0.67060
Constant	-0.52109	-5.33	0.000	-0.71281	-0.32936
Number of observations	4312				
Prob> χ^2	0.000				
Likelihood χ^2	268.27				
Rho-square	0.0449				

Table 6: Restricted model by male passengers

Attributes	**Coefficient**	**Z Value**	**P>|Z|**	**[95% Conf. Interval]**	
Departure time (Fixed)	0.54490	6.42	0.000	0.37851	0.71129
Fare level (0-5% more)	-0.53295	-5.67	0.000	-0.71718	-0.34872
Fare level (10-15% more)	-0.08645	-0.90	0.370	-0.27529	0.10238
Cleanliness of bus inside (Clean enough)	-0.23865	-2.81	0.005	-0.40503	-0.07226
Air conditioning (Available)	0.01198	0.14	0.888	-0.15434	0.17831
Arrangements of seats (Spacious)	0.55419	5.89	0.000	0.36973	0.73865
Constant	-0.15256	-1.26	0.207	-0.38977	0.08464
Number of observations	2404				
Prob> χ^2	0.000				
Likelihood χ^2	173.02				
Rho-square	0.0519				

Conclusion

This study sought to explore passengers' attitude in choosing a bus from a bus station. Discrete choice experiment modeling which is rooted in Random Utility Theory was used to estimate endogenous consideration sets of Asafo-market bus users. The effects of certain attributes based on the findings from the study revealed that in choosing a bus, passengers generally took into consideration buses with fixed departure time, air condition and spacious seats before making their choices. Generally, buses with fixed departure time are highly valued by passengers Pavlyuk & Gromule, 2010), followed by spacious seats, and vehicles with air condition. Passengers' choice of buses generally decrease with an increase of fare level. A similar observation is reported by (Eboli & Mazzulla, 2008; Baidoo, Nyarko, & Mettle, 2015). However, generally, there is difference in the choice of buses by gender. This is consistent with the observation by Baidoo and Nyarko (2015).

The findings of this study may be used by transport operators and policy-makers to formulate strategies for the improvement of public transport in developing countries to help reduce traffic situation, air and noise pollution, and energy consumption.

Further developments of this study may be identified by considering D-efficient designs, a complex choice task; and also employ more complex logit models (i.e., the Hierarchical-logit or Mixed logit models).

References:

Adamowicz, W., Louviere, J., & Swait, J. (1998). *Introduction to Attribute-Based Stated Choice Methods.* Washinton, U.S.A: Resource Valuation Branch Damage Assessment Center, NOAA, U.S Department of Commerce.

Afful, D. (2011). *The Challenges Confronting Private Bus Operating in Ghana.* Netherland: Unpublished MBA Thesis.

Baidoo, I. K., & Nyarko, E. (2015). A Discrete Choice Modeling of Service Quality Attributes in Public Transport. *Research Journal of Mathematics and Statistics 7(1)* , 6-10.

Baidoo, I. K., & Nyarko, E. (2015). Stated Preference Modeling for a Preferred Transportation Mode 5(1). *Mathematical Theory and Modeling* , 134-139.

Baidoo, I. K., Nyarko, E., & Mettle, F. O. (2015). Modeling Mode Choice in Passenger Transport with Discrete Choice Experiment. *Mathematical Theory and Modeling* , 159-165.

Catalano, M., Lo Casto, B., & Migliore, M. (2008). Car sharing demand estimation and urban transport demand modelling using stated preference techniques. *European Transport* , 33-50.

Eboli, L., & Mazzulla, G. (2008). A stated preference experiment for measuring service quality in public transport. *Transportation Planning and Technology 31:5* , 509–523.

Hensher, D. A., Rose, J. M., & Greene, W. H. (2005). Applied Choice Analysis, A Primer. *Cambridge University Press, Cambridge* .

Kroes, E. P., & Sheldon, R. J. (1988). “Stated preference methods-An introduction”. *Journal of Transportation Economics and Policy 22(1)* , pp. 10-25.

Kuhfeld, W. F. (2010). Construction of efficient designs for discrete choice experiment. *Journal of Market Research 31* , 375-383.

Okoko, E. (2006). Quantitative techniques in urban analysis, Ibadan. *Kraft Books Limited* .

Pavlyuk, D., & Gromule, V. (2010). *Discrete Choice Model for a Preferred Transportation Mode.* Riga, Latvia: Transport and Telecommunication Institute, Lomonosova 1, LV-1019.

Van der Waerden, P., Borgers, A., Timmermans, H., & Berenos, M. (2007). *Users Evaluation of Transport Mode Characteristics with Special Attention to Public Transport.* Berkeley, USA: 11th World Conference on Transport Research.