BUS AND RAIL CHOICE MODEL FOR THE COLOMBO METROPOLITAN REGION

Saman J. Widanapathiranage and Amal S. Kumarage
Department of Civil Engineering, University of Moratuwa, Sri Lanka.

Paper Presented at the Annual Sessions
Institution of Engineers Sri Lanka

October 2004
BUS AND RAIL CHOICE MODEL FOR THE COLOMBO METROPOLITAN REGION

Saman J. Widanapathiranage and Amal S. Kumarage
Department of Civil Engineering, University of Moratuwa, Sri Lanka.

Abstract- Mode choice models of the earliest stage were developed based on predictions. However, the recent models are generally by the hypothetical basis developed with considering traveler behavior. These models are categorized as behavioral choice models. This is a powerful decision making tool in transportation planning studies.

Bus and rail mode choice model in this paper intend to capture travel behavior of inter-zonal public transportation in the Colombo Metropolitan Region (CMR). The model contained three variables that described travel constraints in the travel markets such as, (a) passenger waiting time for the mode (b) in-vehicle travel time and (c) loading levels.

The paper attempt to explain a mode choice model that has been formulated with fewer travel attributes to capture the inter zonal travel behavior. The advantage of this model is that there are fewer transportation data requirement for its calibration and prediction.

Key Words: Choice Modelling, Passenger Transportation.

1 Background

1.1 General

The study concentrates on the Colombo Metropolitan Region (CMR) is identical to the Western provincial council administrative area. The land area extends about 3721 sq. kms.[1]. Population lives in Colombo, Gampaha and Kalutara districts was estimated at 5.4 million in 2001 [3].

In public transportation, passengers are equally consuming the service and represents about 70% of the total transportation market. There are normal and intercity bus services that operate from the Metro bus companies and private cluster companies in these markets. The bus service covers for all traffic zones within the region. The traffic zone is defined as the area of travel generates or destinates where the administrative boundary is similar to area of the Divisional secretary divisions.

The market is the inter-zonal origin– destination travels in which 528 markets available in the CMR. There are 171 travel markets operate both bus and railways. In the annexure 01, shows the four railway corridors that connected these traffic zones for inter –zonal travel.

In fact, there is considerably lower demand for railway inter-zonal travel when compared to more popular bus service. It is oblivious for indirect rail travel markets. That is mainly due to poor reliability of the service, frequent delays and inadequate rolling stocks that causes to increase the passenger waiting time for the journey and in-vehicle travel time.

With reference to travel observations in 1998, the bus service produced about 1.5 millions of round trips per day where as, the railway was about 0.1 millions travel per day. In each year, the passenger travel demand increases due to relocation of permanent houses in the remote travel zones of the CMR and the Government institutions and private companies locates in and around the Colombo district. In last years, bus and rail service have being shown a slow growth [2]. As a result, a large portion of travel standees are available for inter-zonal travel. These groups of standees generate an inconvenience journey
not only themselves but also the seated passenger group as well. As such, the public transportation service in the CMR is being deteriorated day by day. Therefore these travel constraints highly correlate to passenger decision making for mode choice process.

1.2 Choice Modelling in Literature

The use of choice models is established in all aspect of transport planning since their initial development [7],[10]. The root of these models based on the application of micro-economic theories. The econometric models are the idea of utility maximization. That is an assumption that a traveler will select the alternative of the choice situation and of the person which influence this utility. Based on the assumptions leading from these concepts it can be shown that the derivation of a utility function and its maximization leads us to well known choice model formulation such as logit modeling.

In the literature a choice model developed based on logit theory for inter city travel between bus and rail transportation in Sri Lanka [9]. It is a generalized choice model that can be applicable to the CMR. Five exogenous variables are found to explain the causes for significant variations between bus and rail services such as, the difference in fare, the difference in travel time, the difference in perceived waiting time, the average load factor on buses and the average travel time. The functional form of the calibrated model is,

\[
\begin{align*}
\text{P}(t) = & \frac{P(t)}{1 - P(t)} = -4.183 + 2.188 \times (IVIC) - 0.2995 \times (F_t - F_b) + 0.3144 \times \frac{V_b - S_b}{S_b} - 0.0218 \times (T_t - T_b) - 0.0625 \times (H_t - H_b) + 0.618 \times \ln \left( \frac{T_t + T_b}{2} \right) \\
& - 0.2183 \times (T_t - T_b) - 0.0625 \times (H_t - H_b) + 0.618 \times \ln \left( \frac{T_t + T_b}{2} \right)
\end{align*}
\]

Where, 4.183 is the stochastic element which derived from the unobservable parts of each variable that depends on data gathered from person to person. These stochastic error terms will correlate between alternatives capturing their similarity in the eyes of the traveler. The other factors are explained the scalar and vector parameters of each variables.

\[
P(t) = \text{Rail ordinary ticket passengers of inter-zonal travel.}
\]

\[
V_t = \text{Total inter-zonal passengers carried by railway during one time period.}
\]

\[
V_b = \text{Total inter-zonal passengers carried by bus in the same period.}
\]

\[
T_t = \text{The scheduled travel time for the train(s) leaving in the time period. In the case of more than one train, the mean was considered.}
\]

\[
T_b = \text{Mean travel time for buses in minutes, where } T_b = 0.5 \times (T_M + T_p)
\]

\[
T_M = \text{Mean travel time of Metro bus in minutes.}
\]

\[
T_p = \text{Mean travel time of Private bus in minutes.}
\]

\[
IVIC = 1 \text{ or } 0, \text{ set equal to one when the rail is an intercity express and equal to zero when it is not.}
\]

\[
F_t = \text{Fare for rail, Third class ordinary ticket one way cost in Rupees.}
\]

\[
F_b = \text{Bus fare in Rupees.}
\]

\[
S_b = \text{Total seating space supplies for buses dispatched in the time period.}
\]

\[
f_q = 0 \text{ or } 1, \text{ set to one when } V_b > S_b \text{ and set to zero at other times.}
\]

\[
H_t = \text{Perceived headway for buses in minutes, where perceived headway is computed as:}
\]

\[
H_{b,t} = \begin{cases} 
H_o & \text{if } H_a \leq H_o \\
H_o + (H_a - H_o)^{0.65} & \text{if } H_a > H_o 
\end{cases}
\]

Where, \( H_a \) is the actual headway in minutes.

As the model (1) explained that it requires a large number of data for the estimation of mode choice and therefore it would contained a high weightage of errors in predictors. As such a problem arises to functioning of the model due to complex model structure and the preparation of the travel attributes for their information’s.

2 Objective
The aim of this paper is to present a model that will estimate passenger choice between bus and rail service in the CMR. It is built according to the logit theory includes only a few representative attributes such as, (a) passenger waiting time, (b) in-vehicle travel time and (c) loading factors of each mode.

3 Methodology

Methodological approach of this study consists of two stages. First is the theoretical formulation of a logit model and Secondly the model calibration and validation.

3.1 Theoretical Modelling

In the logit modeling, theoretical concepts implied for their stability and make a simple model structure. Adding the assumption of perfect information of the traveler about choice of the alternatives, one arrives at the following two part model[8].

(a) The measurable system part $V_{ij}$, is the deterministic term is given by a linear combination of attributes. That is representing the value of objective utility of an alternative $j$ choosing a person $i$ [6]. Then,

$$V_{ij} = \beta_0 + \beta_1 x_{ij}^1 + \beta_2 x_{ij}^2 + .... + \beta_k x_{ij}^k$$

If we assume $T$ is a scalar or a vector parameter, then we can write

$$V_{ij} = \beta^T_k x_{ij}$$  \hspace{1cm} (3)

(b) The stochastic part representing error $\eta_{ij}$, considering unobserved evaluation by each user. In the analysis of error term $\eta_{ij}$, it used the Gumbel distribution method in statistics with parameters $(\eta, \mu)$ which is denoted by $G (\eta, \mu)$ are mutually exclusive and independently follow the distribution[5]. The probability density function of $G (\eta, \mu)$ as,

$$f(x) = \exp[-\exp(-\mu(\eta - x))]$$ \hspace{1cm} (4)

In according to the utility maximization theory that any decision maker chooses the attributes that maximized his or her utility of $i$ choosing the alternative $j$ is not deterministic but it is probabilistic [6]. Therefore, the alternative that maximizes traveler’s utility may change probabilistically. Then the probability $P$ is written as,

$$P_{ij} = \Pr[U_{ij} \geq U_{il}, l \in C, l \neq j]$$ \hspace{1cm} (5)

$$P_{ij} = \Pr[U_{ij} \geq \max_{l \in C, l \neq j} U_{il}]$$ \hspace{1cm} (6)

Assuming a $U_{ij}^*$ as,

$$[U_{ij}^* = \max_{l \in C, l \neq j} U_{il}]$$ \hspace{1cm} (7)

Then, it can be written as,

$$P_{ij} = \Pr[U_{ij} \geq U_{ij}^*]$$ \hspace{1cm} (8)

One of the interesting property of the Gumbel distribution is that if $n$ variables $x_1, x_2, ..., x_n$ independently follow the Gumbel distribution with parameters $(\eta_1, \mu_1)$, $(\eta_2, \mu_2)$, ..., $(\eta_n, \mu_n)$, then the maximum of $x_1, x_2, ..., x_n$ also follow the Gumbel distribution with parameters [5].

$$\left(\frac{1}{\mu} \log \sum_{j=1}^{n} \exp(\mu \eta_j), \mu \right)$$ \hspace{1cm} (9)

Consequently, $\eta_{ij}$ follow the Gumbel distribution with parameters,

$$\left(\frac{1}{\mu} \log \sum_{l \in C, l \neq j} \exp(\mu V_{ij}), \mu \right)$$ \hspace{1cm} (10)

The probability $P$ is,

$$P_{ij} = \Pr[U_{ij}^* - U_{ij} \leq 0]$$

Where $U_{ij}$ and $U_{ij}^*$ follow the Gumbel distribution with parameters,

$$\left(\frac{1}{\mu} \log \sum_{l \in C, l \neq j} \exp(\mu V_{ij}), \mu \right): (V_{ij}, \mu)$$ \hspace{1cm} (11)

Second property of the Gumbel distribution said that, if two variables $x_1$ and $x_2$ independently follow the Gumbel distribution with parameters $(\eta_1, \mu_1)$ and $(\eta_2, \mu_2)$, Then $x_1 - x_2$ follow the logistic distribution. Therefore probability distribution is given by,

$$f(x_1 - x_2) = \frac{1}{1 + \exp(\mu x_2 - \mu x_1)}$$ \hspace{1cm} (12)

Using this property it developed,

$$P_{ij} = \frac{1}{1 + \exp(\frac{1}{\mu} \log \sum_{l \in C, l \neq j} \exp(\mu V_{ij}) - V_{ij})}$$
Then, the logit model is formulated as,

$$P_{ij} = \frac{1}{1 + \exp(-\mu V_{ij})} \sum_{l \in C, l \neq j} \exp(\mu V_{il})$$  \hspace{1cm} (13)$$

Then, the logit model is formulated as,

$$P_{ij} = \frac{\exp(\mu V_{ij})}{\sum \exp(\mu V_{il})}$$ \hspace{1cm} (14)$$

It indicates that $P_{ij}$ is proportional to $V_{ij}$, the deterministic part of utility. In the traveler choice set, three variables included for this study that explained constraints in public transportation, such as a difference in waiting time represents an available constraint, travel time saving constraints and standees variable as a comfort constraints. These constraints are aggregated in to the deterministic component $V_{ij}$, using elastic demand theory in micro-economic. The elasticity as a proportional change in one variable over the proportional change in another variable[11]. In the application of this theoretical framework, the deterministic linear model form(3) is formulated as,

$$\ln \left( \frac{P(j)}{1-P(j)} \right) = \beta_1 + \beta_2 \ast \left( \frac{1}{WT_{ij}^k + \lambda_1} \right) + \beta_3 \ast \left( \frac{1}{TT_{ij}^k + \lambda_2} \right) + \beta_4 \ast \left( \frac{(LF_j - 1)}{(LF_i + LF_j - 2)(LF_i - 1)} \right)$$  \hspace{1cm} (15)$$

Where,

$P(j)$ = The ratio of rail ordinary ticket passengers to total passenger inter-zonal travel.

$WT_{ij}^k$ = Perceived reliable waiting time difference between mode i and j is observed by $k^{th}$ traveler.

$TT_{ij}^k$ = Observed reliable travel time difference between mode i and j is observed by $k^{th}$ traveler.

$LF_i$ = Load factor of $i^{th}$ mode is express as the ratio of total number of passengers to total number of seating spaces on $i^{th}$ mode.

$LF_j$ = Load factor of $j^{th}$ mode is express as the ratio of total number of passengers to total number of seating spaces on $j^{th}$ mode.

$\beta_1$ is the stochastic constant and $\beta_2, \beta_3, \beta_4$ are the scalar or vector parameters. $\lambda_1$ and $\lambda_2$ are the Box-Tuky constants.

### 3.2 Model Calibration and Validation.

The choice model calibration was through multinomial logistic regression method using the statistical software SPSS™+V10.0. The calibration investigated a weightage of the choice set due to aggregation of the variables. The data sets were prepared to year 1998, based on the collected data from various studies during the period 1996 to 1998. The recent data useful to calibrate a better result of the model and the lack of these data was the major constraint to validate the model.

The table 01, is shown the calibrated parameters of the model coefficients. The coefficients are significantly small and it indicates the model has developed a better casual relationship between independent variable to the dependent variable.

<table>
<thead>
<tr>
<th>$\ln \left( \frac{P(j)}{1-P(j)} \right)$</th>
<th>Stochastic Part ($\beta_1$)</th>
<th>$\beta_2$</th>
<th>$\beta_3$</th>
<th>$\beta_4$</th>
<th>$\lambda_1$</th>
<th>$\lambda_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_{iBus}$</td>
<td>0.680</td>
<td>2.097</td>
<td>0.439</td>
<td>1.034</td>
<td>0.740</td>
<td>2.730</td>
</tr>
<tr>
<td>$V_{iRail}$</td>
<td>0.079</td>
<td>1.386</td>
<td>0.461</td>
<td>0.081</td>
<td>0.740</td>
<td>2.730</td>
</tr>
</tbody>
</table>
The calibrated model gives the following observations for inter-zonal travel.

- The traveler preference on bus is significantly change when the load factor of bus change.
- The traveler preference on bus is significantly change when bus waiting time change.
- Bus traveler concern about the load factor and waiting time whereas, the rail traveler concern about travel time and waiting time.

The predictors is estimated by the application of deterministic components to model (15) where it assumed the $\mu=1$.

In the model validation, it gives the accuracy level of the model predictions with respect to the field observations. The table 02 shows the validation results of a randomly selected inter-zonal travel.

**Table 02: Comparison of Predicted and Observed Travel Choices for Selected Inter-zonal Travels**

<table>
<thead>
<tr>
<th>Inter-zonal Travel Name</th>
<th>Observed Share %</th>
<th>Predicted Share %</th>
<th>Observed Share %</th>
<th>Predicted Share %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colombo - Dehiwala</td>
<td>86</td>
<td>89</td>
<td>14</td>
<td>11</td>
</tr>
<tr>
<td>Gampha – Mirigama</td>
<td>76</td>
<td>86</td>
<td>24</td>
<td>14</td>
</tr>
<tr>
<td>Kelaniya – Mahara</td>
<td>85</td>
<td>87</td>
<td>15</td>
<td>13</td>
</tr>
<tr>
<td>Katana – Ja-ela</td>
<td>89</td>
<td>94</td>
<td>11</td>
<td>6</td>
</tr>
</tbody>
</table>

In most cases, rail share varies below 20 % of the total travel market. Bus share is significantly a high value, may be due to reliability and better accessibility for inter zonal travel. The predicted values of the rail share were found under predicted in most cases. It might be a comparism of a small ordinary ticket rail traveler with respect to the large portion of bus ticket traveler. To avoid weakness of this model, it is recommended to calibrate and validate with the recent travel data, thus the predictors will give a better travel scenarios of the travel markets.

**4 Conclusion**

The main conclusion that reached in this study is that a choice model has been calibrated to the acceptable levels of accuracy even though it has fewer variables describing the travel characteristics that differentiate travel modes. Furthermore, unimportant variable namely, the loading factor has been successfully included in the model. This is confirmed by the results of the validation process for a few selected zonal pairs.

**References**

Acknowledgements

The authors gratefully acknowledges the Asian Development Bank financial assistance to conduct this research and facilities provided by the Department of Civil Engineering, University of Moratuwa, Sri Lanka.

Biographical Sketch

SAMAN J. WIDANAPATHIRANAGE, B.Sc. Eng. (Moratuwa), M.Sc. (Netherlands), AMIE (SL), is a Civil Engineer at Road Development Authority Sri Lanka. He specialized in the area of Environmental and Transportation planning, is presently reading for Ph.D. degree from the Department of Civil Engineering University of Moratuwa, Sri Lanka. His thesis has to do a bus and rail travel modeling for Colombo Metropolitan Region: A theoretical approach to mode choice modelling.

AMAL S. KUMARAGE, B.Sc. Eng. Hons. (Moratuwa), Ph.D. (Calgary), FCIT (London), AMIE (SL), is professor of Civil Engineering at University of Moratuwa. He is a member of several policy formulation committees in area of road safety, public transport, highway planning and transport infrastructure where he pursues the application of engineering disciplines and system in different aspects of transportation planning and design in Sri Lanka.

Annexure

Annexure 1: Bus and Rail Transportation Corridors in the CMR

- **Main line:** Fort to Mirigama (50.27 km)
- **Puttalam line:**
  - Fort to Ragama (15.54 km)
  - Ragama to Kochchikade (29.50 km)
  - Fort to Kochchikade (45.04 km)
- **Kelani Valley line:**
  - Fort to Padukka (37.18 km)
  - Padukka to Avissawella (23.99 km)
  - Fort to Avissawella (61.17 km)
- **Coast line:**
  - Maradana to Panadura (28.15 km)
  - Panadura to Aluthgama (33.41 km)
  - Maradana to Aluthgama (61.51 km)