Effect of captive riders in intercity mode-choice modeling

Syed Mohiuddin Afroz

Civil Engineering

April 1995

Abstract

The primary objective of this study is to investigate the effect of including captive riders in mode-choice modelling for intercity travel in Saudi Arabia. To achieve this objective, mode-choice models were calibrated and validated using the intercity travel data of Saudi Arabia. The calibration process was carried out in three stages. In the first stage, mode-choice models were calibrated using the data of choice riders only, the choices being those alternatives as stated by the travellers. In the second stage, mode-choice models were calibrated assuming full choices over all the alternatives available by the choice riders. Finally, in the last stage of model calibration, captive riders were included in the travel dataset. The calibrated models were compared with respect to their coefficient estimates and other goodness-of-fit statistics.

It was found from the study that the inclusion of captive riders in mode-choice models resulted in increased significance of the coefficient estimates which is mainly due to the increased sample size but decreased the overall fit of the models. When used for prediction purpose, only that model which was calibrated with choice riders data was capable of predicting the mode-choice of the travellers correctly whereas, the model calibrated with choice riders plus captive riders was unable to predict the mode-choice of the intercity travellers correctly.

Effect of Captive Riders in Intercity Mode-Choice Modelling

by

Syed Mohiuddin Afroz

A Thesis Presented to the

FACULTY OF THE COLLEGE OF GRADUATE STUDIES

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DHAHRAN, SAUDI ARABIA

In Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

In

CIVIL ENGINEERING

April, 1995

EFFECT OF CAPTIVE RIDERS IN INTERCITY MODE-CHOICE MODELLING

BY

SYED MOHIUDDIN AFROZ

A Thesis Presented to the

FACULTY OF THE COLLEGE OF GRADUATE STUDIES

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DHAHRAN, SAUDI ARABIA

In Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

In

CIVIL ENGINEERING

APRIL 1995

UMI Number: 1375582

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS
DHAHRAN, SAUDI ARABIA.
COLLEGE OF GRADUATE STUDIES

This thesis, written by

Syed Mohiuddin Afroz

under the direction of his Thesis committee, and approved by all its members, has been presented to and accepted by the Dean, College of Graduate Studies, in partial fulfillment of the requirements for the Degree of

Master of Sciences in Civil Engineering (Transportation)

Thesis Committee :

Dr. Hasan Musaed Al-Ahmadi (Chairman)

Dr. Gokmen Ergun (Co-Chairman)

Dr. Nedal Taisir Ratrout (Member)

Dr. Shukri H. Al-Senan (Member)

Dr. Hamad I. Al-Abdul Wahhab (Member)

Dr. AlFarabi Sharif
Department Chairman

Dr. Ala H. Al-Rabeh
Dean, College of Graduate Studies

Date: 20-8-95

Acknowledgements

Acknowledgement due to the University for providing me the financial support during my stay here.

Sincere gratitude and appreciation to Dr. Al-Ahmadi, H.M., who served as major advisor in my thesis committee. I would like to thank specially Dr. Gokmen Ergun for his help and suggestions at difficult times during the course of this work. Thanks are due to other committee members for their comments.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

TABLE NO. .... PAGE NO.

ABSTRACT

Name of the Student : Syed Mohiuddin Afroz

Title of the Study : Effect of Captive Riders in Intercity Mode-Choice Modelling

Major Field : Transportation Engineering

Degree Date : June 1995

The primary objective of this study is to investigate the effect of including captive riders in mode-choice modelling for intercity travel in Saudi Arabia. To achieve this objective, mode-choice models were calibrated and validated using the intercity travel data of Saudi Arabia. The calibration process was carried out in three stages. In the first stage, mode-choice models were calibrated using the data of choice riders only, the choices being those alternatives as stated by the travellers. In the second stage, mode-choice models were calibrated assuming full choices over all the alternatives available by the choice riders. Finally, in the last stage of model calibration, captive riders were included in the travel dataset. The calibrated models were compared with respect to their coefficient estimates and other goodness-of-fit statistics.

It was found from the study that the inclusion of captive riders in mode-choice models resulted in increased significance of the coefficient estimates which is mainly due to the increased sample size but decreased the overall fit of the models. When used for prediction purpose, only that model which was calibrated with choice riders data was capable of predicting the mode-choice of the intercity travellers correctly whereas, the model calibrated with choice riders plus captive riders was unable to predict the mode-choice of the intercity travellers correctly.

MASTER OF SCIENCE DEGREE

KING FAHD UNIVERSITY OF PETROLEUM AND MINERALS

DHAHRAN, SAUDI ARABIA.

APRIL-1995

خلاصة الرسالة

اسم الطالب الكامل : سيد محي الدين افروز
عنــوان الرسالـــة : تأثير الركاب ذوي الخيار الواحد على نماذج اختيار وسائط النقل بين المدن
التخصـــــــص : هندسة النقل .
تاريـخ الشهــادة : يونيو ١٩٩٥م .

إن التنبؤ باختيار المسافرين بين المدن لنوع معين من وسائط النقل المتوفرة (المركبة الخاصة ، والحافلات العامة، والقطارات ، والطائرات) يعتبر من المهام الصعبة . وفي هذا المجال فإن النماذج التي تبنى لمحاكاة عملية الاختيار هذه تعتبر من أهم الوسائل التي يمكن استعمالها للتنبؤ بتوزيع المسافرين بين المدن على وسائط النقل المتوفرة اعتماداً على خصائص هؤلاء المسافرين في جهة، وخصائص وسائط النقل المتوفرة من جهة أخرى . وعندما تبنى هذه النماذج على المعلومات التي يتم جمعها لعينة من المسافرين بما فيهم المسافرون ذوي الخيار الواحد (المسافرون الذين يمكنهم استعمال نوع واحد من وسائط النقل فقط لأي سبب كان)، فإن هذه النماذج والتنبؤات التي تبنى عليها لاحقاً قد لاتكون صحيحة .

وفي هذا الإطار فإن هذه الدراسة تهدف بالدرجة الأولى إلى دراسة مدى صحة هذه الفرضيه . وبعبارة أخرى فإن الهدف الأساسي هو استقصاء التأثير الناجم عن بناء نماذج محاكاة اختيار المسافرين بين المدن لوسائط سفرهم بالاعتماد على قاعدة معلومات تحوي مسافرين من ذوي الخيار الواحد على صحة هذه النماذج . ولتحقيق هذا الغرض فقد تم استعمال معلومات تم جمعها عن طريق اجراء مقابلات ميدانية لمسافرين بين مختلف المدن السعودية ، وتم بناء ثلاثة أنواع من النماذج هي :

- نماذج مبنية على الخيارات لوسائط النقل التي يمكن للمسافرين الذين تم مقابلتهم استعمالها فعلاً .
- نماذج مبنية على الافتراض بأن كل مسافر قادر على استعمال كافة وسائط النقل المتوفرة حتى تلك التي أخبر بأنه لايستعملها .
- نماذج مبنية على المعلومات الخاصة بالنوع الثاني بالإضافة إلى تضمينها المعلومات الخاصة بالمسافرين ذوي الخيار الواحد .

وقد تم مقارنة الأنواع الثلاثة من النماذج وإجراء الاختبارات الإحصائية لها وتم مطابقة نتائجها مع الخيارات الفعلية التي تم رصدها لعينة من المسافرين . وقد تم الاستنتاج بأن احتواء قاعدة المعلومات للمسافرين ذوي الخيارَ الواحد يؤدي إلى الحصول على نماذج ضعيفة وغير قادرة على محاكاة الواقع والتنبؤ به . وتبين أيضاً أن النموذج الأول هو الأكثر قدرة على التنبؤ الصحيح باختيارات المسافرين بين المدن لوسائط سفرهم ، وبناء عليه فإنه يوصي باستثناء المسافرين ذوي الخيار الواحد من قاعدة المعلومات المستعملة في بناء النماذج الخاصة بمحاكاة وتفسير اختيار المسافرين بين المدن لوسائط سفرهم وبناء نماذج خاصة بهذا النوع من المسافرين لأن احتواءهم في قاعدة المعلومات يؤدي إلى الحصول على نماذج ضعيفة يصعب الاعتماد عليها إن كان لتفسير هذه الظاهرة أو التنبؤ بها .

درجة الماجستير في العلوم
جامعة الملك فهد للبترول والمعادن
الظهران ، المملكة العربية السعودية

التاريخ : إبريل ١٩٩٥

CHAPTER 1

INTRODUCTION

1.1 GENERAL

Accurate forecasting of intercity demand by each mode is a difficult task and modal choice is crucial to plan for future modes as well as for the existing modes (1, 2). The mathematical relationships (termed as models), developed to estimate travel demand on each mode, assumes that each individual has a choice for travel decisions. But in some situations, travellers do not want to, or cannot choose certain modes. Travellers who do not have any choice other than the mode selected are termed as captive riders with respect to that travel mode.

Models built on travellers which include those riders who do not have any choice on mode are likely to give errors when used for predicting modal demand (3). The primary objective of this study is to see the effect of captive riders on model calibration (i.e., model parameter estimates etc,.) and model predictions in Saudi Arabian intercity travel environment.

1.2 PROBLEM DEFINITION

Mode choice is an essential step in transportation planning. Without it the prediction of future traffic growth in each mode cannot be determined. Recent work done by Al-Ahmadi et. al. (4) revealed that intercity travel in Saudi Arabia possesses some special characteristics. One of the characteristics is the high number of captive riders who are not willing to use any other mode than the one they use. It was found from the study that among the factors which contribute to this high number of captive riders are, distance of travel , travel cost and perception of privacy.

Mode-choice models calibrated on choice riders are considered to be the theoretically correct models (3). If captives are included in the model calibration stage then such a model is not likely to predict the exact travel demand. This research is aimed at investigating the effect of the inclusion of captive riders in model calibration and model predictions.

1.3 RESEARCH HYPOTHESIS

The mode-choice models assume that each individual has and exercises a choice for modes (3). The expected results when models are calibrated with datasets containing only choice riders and captive plus choice riders is explained by means of a flow chart shown in Figure 1.1.

The mode-choice models developed on choice data riders only is likely to represent the actual choice behavior and these models are likely to result in true predictions. Models calibrated on dataset containing both captive and choice riders may yield incorrect results during prediction. This is because, during calibration stages, the model may assume that captives also make choices and build models as if the whole dataset is of choice riders. Depending upon which mode has high number of captives in it, its utility is expected to be higher.

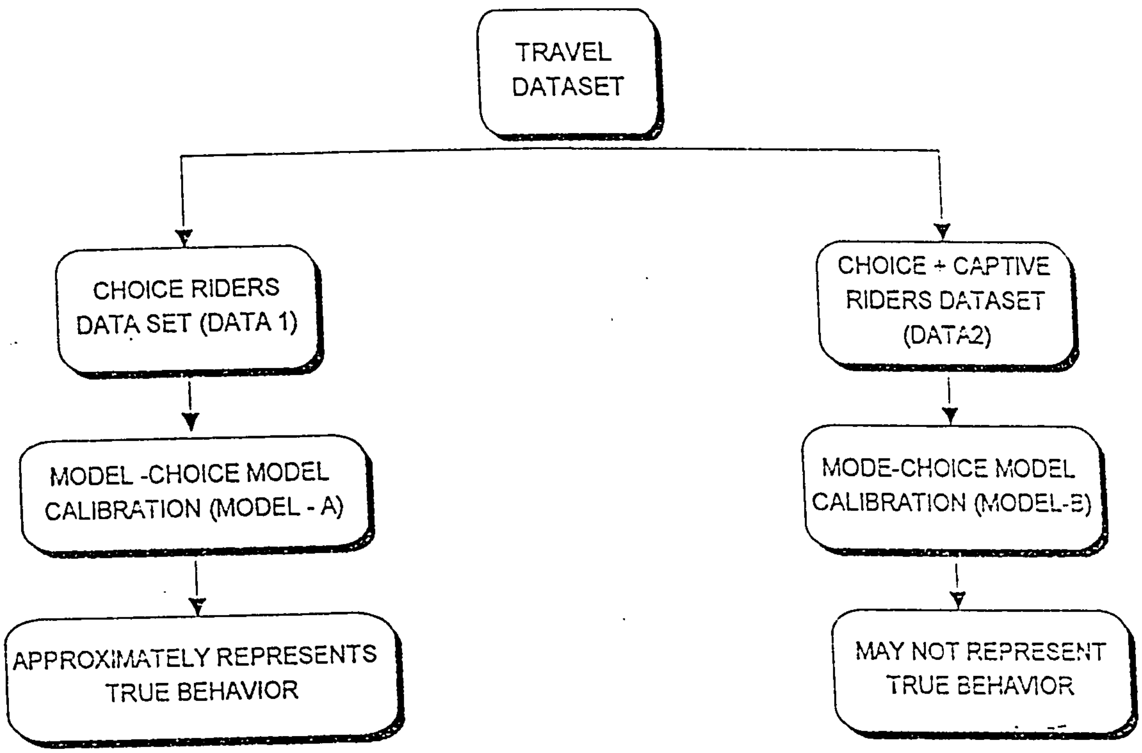


FIGURE 1.1: HYPOTHETICAL FRAMEWORK

The main goal of this study is to investigate the effects of including the captives in mode-choice model calibration and prediction stages. Hence, for this purpose mode-choice models are built, first with choice riders data only, then captive riders are included in the data of choice riders. The resulting models are used for validation purpose in later stages.

1.4 SPECIFIC OBJECTIVES OF THE PRESENT STUDY

Keeping in view the main goal of this study the objectives are formulated as follows:

1. to review the past work done by researchers in defining choice and captivity;
2. to determine the travel and socioeconomic characteristics of the intercity travellers of Saudi Arabia;
3. to build mode-choice models: first, using the data of choice riders with valid alternatives i.e., those alternatives which are acceptable to travellers (Model 1). Second, using the data of choice riders with all alternatives (Model 2). Third, using the data of choice riders with all alternatives and captive riders (Model 3);
4. compare the above developed models viz., Model 1, 2 and 3 with respect to their coefficient estimates, overall goodness-of-fit statistics, etc., to see the effect of the inclusion of captives in model calibration;
5. see the effects of using the above models viz.,Model 1, 2 and 3 separately on prediction datasets of choice riders with valid alternatives, choice riders with all alternatives and choice riders with all alternatives plus the captive riders.

CHAPTER 2

LITERATURE REVIEW

2.1 GENERAL

Modal split continue to receive primary attention from transportation planners because completely acceptable models have yet to be developed. Also, it is of most concern when deciding among alternative transportation proposals. In the initial part of this chapter the structure and features of travel demand models developed earlier will be discussed. The work on captive ridership determination and models developed in this area will be highlighted. The chapter is concluded with a brief comment on the information lacking in this subject.

2.2 OVERVIEW OF INTERCITY MODE-CHOICE MODELS

The model developed by Warren T. Adams (5) is one of the first modal split models to be devised. Since mid-1960's many mode choice models for intercity travellers were calibrated and used for prediction in various environments. Initial models developed were aggregate in nature wherein, the approximate travel demand emanating from a particular area was estimated.

However, studies revealed that models based on aggregate data are less likely to explain the behavior of an individual and is a poor prediction model. The inability of the models based on aggregate data to explain traveller's behavior led the researchers to propose another group of models i.e., the disaggregate models. The development of disaggregate models require the data that describes the behavior of an individual's characteristics and attitude towards the travel services provided by each mode.

In the following sections the historical development in this area will be discussed.

2.3 DEVELOPMENTS IN MODE-CHOICE MODELLING

Disaggregate travel demand models represent a recent innovation in travel forecasting procedures. The earliest research into disaggregate mode-choice models was done by Warner in 1962 (6).

Following his work, pioneering efforts were made by researchers during the period 1967-1969. Data was exclusively collected for research in this area and majority of them used binary mode-choice modelling with automobile as the base mode. Pioneers in this early age of disaggregate modelling include, David A. Quarmby (used discriminant analysis) (7), Thomas E. Lisco (used probit analysis) (8) and Peter R. Stopher (used regression and subsequently used logit analysis) (9). Besides, a wealth of literature is available on the development of these models (10-14).

However, it was found from the market research that, besides the socioeconomic and mode related characteristics, an individual evaluates his choice depending upon the Level Of Service (LOS) provided by the alternatives based on his socioeconomic characteristics. This led to the incorporation of factors such as, comfort, convenience, privacy and other mode related attitudinal indicators in the models. One of the earliest attempts to include traveller's attitude in mode-choice was by Ackoff in 1965 (15). His efforts were pioneering, in considering psychological factors in mode-choice. In pace with the development in various parts of the world, attempts were made by prominent researchers in transportation planning to incorporate the attitude of travellers in mode-choice models. This resulted in a series of models for different environments (16-19).

A number of studies have been conducted to build mode-choice models in Saudi Arabia. Al-Ahmadi (20) developed mode-choice model for two major corridors in the Kingdom, namely, Dhahran-Riyadh and Riyadh-Jeddah. Al-Ahmadi et. al. (4) developed a national intercity mode-choice model for Saudi Arabia for various trip purposes.

Recently, in 1994 AL-Sughaiyer (21) carried out a study investigating the effect of incorporating perceptual concepts of privacy, comfort, convenience and reliability in intercity mode-choice models for Saudi Arabia. He used these variables along with the traditional travel and socioeconomic variables. He excluded captive riders while calibrating the models. He recommended the inclusion of captive riders with choice riders during calibration stages and study their effect.

2.4 *EXTENSIONS TO MODE-CHOICE MODELLING*

With the growing popularity of mode-choice models, the existence of choice was reviewed keenly. Travellers who have choice of selecting alternatives were used for model building and travellers for whom data is available for only one alternative were eliminated. Such travellers were termed as captive with respect to that mode. The data for such travellers was kept apart and analyzed separately during model calibration and prediction for future estimation of their percentage.

Captivity to a particular travel decision was first studied in-depth during the Pittsburg Area Transportation Study (PATS), in 1958 (22). It was found from the study that in Pittsburg area, a large proportion of the transit trips were captive. Apart from work trips, probably 90% of the transit trips were carried out by people who had no alternative means of transportation. This was used as a justification for not incorporating the characteristics of the system in model calibration and prediction.

The following paragraphs bring out the details of the work done by various researchers on the captivity aspect in various places, both in the intercity and intracity travel context.

In 1966 planners in the Wilbur Smith and Associates (23) developed a model in which the proportion of captive and choice riders was related to the population of the area under study. The study was restricted to urban locations only. It was found that approximately 85-90 percent of the transit riders were captive to transit system in cities where population ranges one to two million people.

In 1973 Van der Tol et.al (24) made an attempt to relate the occupation of the intercity rider to his selection of a particular mode. From this study they found that almost twice as many persons in clerical and sales occupation were captive to public transport in intercity travel compared to the other occupation categories (labor, service, managerial and professional).

Morall and Morash (25) in 1973 developed two separate relationships, for choice riders alone, and for choice and captive riders together. The authors related transport system variable (transit/car travel time ratio) to percent trips by transit. Based on these relationships they concluded that the two groups possess different sensitivities to the used system variable.

M.G. Ferreri and Cherwony (26) in 1971 tried to determine the desirability and feasibility of developing two sets of modal-split models: direct generation for captive transit riders and trip interchange for choice transit riders. Captive transit trips produced and attracted were related only to the demographic and economic characteristics of the zone under study because captive riders are inelastic to the system variables. Whereas, the choice transit riders were related to some measures of the relative performance of the respective transportation systems (e.g., automobile time versus travel time) with diversion curves. The authors utilized linear regression technique to build models. The results of the analyses indicated that both the captive and choice modal-split models were reasonable and capable of accurately simulating existing transit travel patterns. Later, they also commented that the standard method of treating these groups as one usually produces models that are insensitive to system variables because the characteristics of the captives tend to dominate the models.

In 1972 Michael J. Demetsky (17) developed a model in which, he tried to identify the transit captive riders using only the socioeconomic characteristics of the trip makers. This model was first used to eliminate the transit captive riders from the population. The modal-split model was then applied to predict the split between the modes for the remaining part of the travellers who were assumed to be choice riders.

The effect of mode-choice model's prediction with and without captives in the calibration data was studied by P.R.Stopher (3) in 1980. He recognized that captivity to a mode exists in virtually all study situations. Through a theoretical discussion he stated that the inclusion of captives in the data for choice models, whether for calibration, prediction, or both, would have serious effects on the prediction results of the model. Furthermore, he suggested that captives have to be excluded from the travel data set during calibration of mode-choice models.

In 1981, Ogunjumo (27) developed a model for estimating the proportion of captive riders at Ile-Ife in Nigeria. In his work he calibrated a model for transit captive riders in that city by considering the socioeconomic characteristics of the travellers. Discriminant analysis was used in his study.

Morikawa, Ben-Akiva and Yamada (28) (1991), developed a model for forecasting intercity rail ridership by segregating the data into captive and choice riders for the alternatives available. It was found that, out of 274, 89 and 82 usable responses for rail, bus and car alternatives 133, 17 and 40 were of captive riders, which represents approximately 50% of the data under study. These figures indicate the seriousness of the captivity issue in the intercity context. However, they removed the captive riders from the travel dataset and calibrated models using Logit analysis.

Talvitie and Koskenoja (29) in 1991 calibrated a model for the availability of car mode using binomial logit model building technique. The alternatives were 'no cars available' (non-driver) and 'one or more cars available' (driver). If the person shared the use of atleast one car, he or she was classified as a driver. Being male, head of household, working away from home, having employment and high incomes, working long hours, living in a rural area in a one-family detached house or town house increased the probability of being classified as a driver. Although this was not an effort to model captive riders, the model being built for a different purpose, it does reveal the characteristics of potential captive car users (or non-drivers).

Ergun and Al-Ahmadi (30) calibrated a multinomial choice model with alternatives being choice riders, car captives and air captives. Their objective was to find out whether the intercity captive riders can be identified by their socioeconomic characteristics in Saudi Arabia. The model calibrated by them is shown in Table 2.1. The fit was marginal. They concluded from their study that captives can be identified by their socioeconomic characteristics and some general trip characteristics (such as trip

Table 2.1: Logit Model Calibrated by Ergun and Al-Ahmadi (30)

Variable Code	Explanation	Coefficient estimate	t stat
CHOICE	Constant for choice riders	4.134	14.28
CAR	Constant for car captives	1.453	3.58
DIST	Distance (Kms.) for car captives	-.0015	-4.01
DFMLYC	= 1, if No. of family members travelling together >2 for car captives = 0, otherwise	1.585	7.74
PRINA	Personal income for air captives	0.208	4.48
DAGE50	= 1, if age > 50 years for air captives = 0, otherwise	0.845	3.18
DMSTAT	= 1, if married for air captives = 0, otherwise	0.449	2.03
DNATIONA	= 1, if traveller is European or Australian = 0, otherwise	0.755	1.95
DEDUCPR	= 1, if only primary education for choice travellers = 0, otherwise	-.574	-2.81
DPAID	= 1, if the travel cost is paid by company for choice riders = 0, otherwise	-.288	-1.68
DSOCIAL	= 1, if purpose of the trip is social for car captives = 0, otherwise	0.654	3.26

Summary Statistics:

Number of Observations: 1632
Log-Likelihood at zero: -979
Log-Likelihood at Max: -862
LRTS Value: 233
Rho-Squared Value: 0.1191
Rho bar Squared Value: 0.1158

length, purpose, etc,.). They recommended that the inclusion of these riders in choice models is likely to result in biased estimates of the coefficients and should not be used in calibrating a choice model.

Reviewing the work done in the past by various researchers, it can be summarized that captivity to a particular travel decision virtually exists under all study situations. After the fundamentals of captivity has been identified i.e., understanding why people are captive and what underlying features make them captive they were separated out and then mode-choice models were built on travellers who exercise choice in their selection of mode.

The present work is not concerned with the fundamentals of captivity, or with the issues of modelling who is captive and who is not captive rather, it is aimed at identifying the effects captive riders will have on the end results of mode-choice models. This issue will be analyzed in-depth and the effect the captive riders is likely to have in model calibration and prediction will be investigated.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 INTRODUCTION

In this chapter, the various stages involved in the study are explained. The software that will be utilized for this purpose is also mentioned.

3.2 STAGES IN THE STUDY

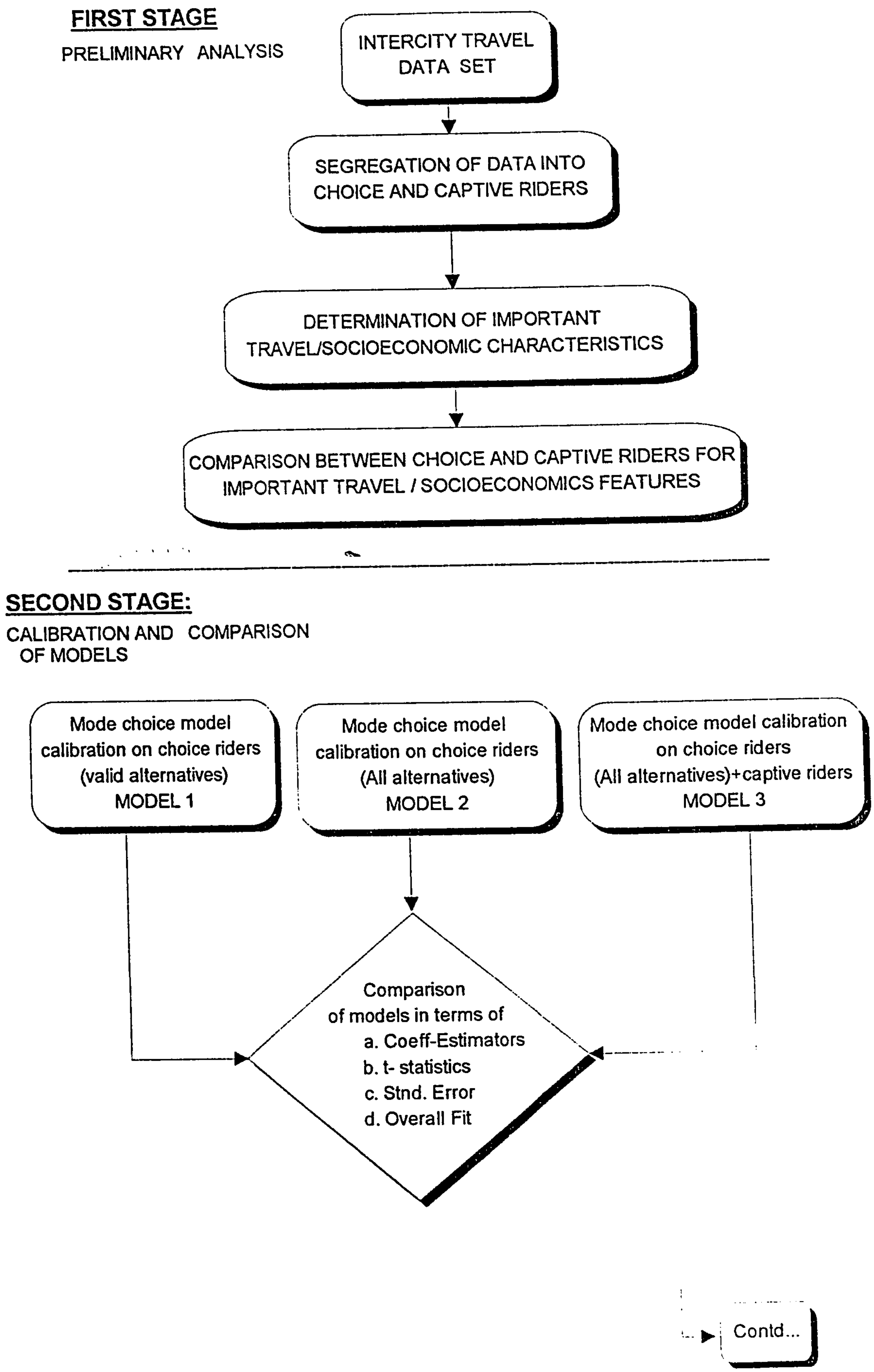
To achieve the goals and objectives of this research, the work is divided into the following three stages as can be seen in Figure 3.1.

First Stage: This part deals with the preliminary analysis of the intercity riders of Saudi Arabia. Percentage of choice and captive riders on each mode for various travel and socioeconomic features is determined.

Second Stage: In this stage, intercity mode-choice models are calibrated using data of:

1. choice riders with valid alternatives (Model 1),
2. choice riders with all alternatives (Model 2), and
3. choice riders with all alternatives and captive riders (Model 3).

The developed models are compared with regard to the coefficient estimates of the variables and also on their overall fit.


FIRST STAGE
PRELIMINARY ANALYSIS
INTERCITY TRAVEL DATA SET
SEGREGATION OF DATA INTO CHOICE AND CAPTIVE RIDERS
DETERMINATION OF IMPORTANT TRAVEL/SOCIOECONOMIC CHARACTERISTICS
COMPARISON BETWEEN CHOICE AND CAPTIVE RIDERS FOR IMPORTANT TRAVEL / SOCIOECONOMICS FEATURES
SECOND STAGE:
CALIBRATION AND COMPARISON OF MODELS
Mode choice model calibration on choice riders (valid alternatives) MODEL 1
Mode choice model calibration on choice riders (All alternatives) MODEL 2
Mode choice model calibration on choice riders (All alternatives)+captive riders MODEL 3
Comparison of models in terms of
a. Coeff-Estimators
b. t- statistics
c. Stnd. Error
d. Overall Fit
Contd...

THIRD STAGE:

VALIDATION OF MODELS

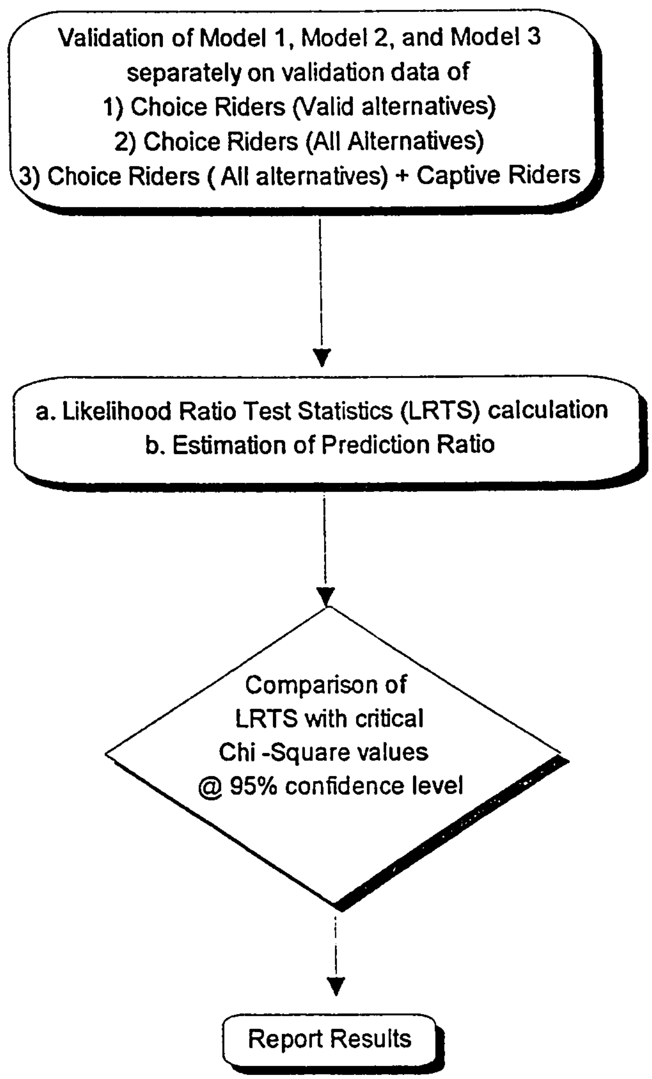


Fig 3.1 : RESEARCH METHODOLOGY

Third stage: This stage is primarily concerned with the model validation and predictions. Models developed in the previous stage viz., (Model 1, 2 and 3) are utilized separately for prediction on validation datasets of choice riders with valid alternatives, choice riders with all alternatives and choice riders with all alternatives plus the captive riders.

3.3 PRELIMINARY ANALYSES OF INTERCITY RIDERS

This is the first stage in the present study and in this stage, preliminary analysis of intercity travellers of Saudi Arabia is performed. Focus is on the determination of choice and captive riders for various travel and socioeconomic characteristics.

Statistical Analyses Software (SAS) available on the mainframe of King Fahd University of Petroleum and Minerals (KFUPM), Dhahran, is used to perform these analyses. The important travel and socioeconomic characteristics of the intercity travellers affecting mode choice are also determined with the help of Chi-Square and Cramer's V test.

3.4 MODE-CHOICE MODELS CALIBRATION AND COMPARISON

This is the second stage in the present study. To calibrate models, the concept of utility theory is used. The basic approach in this theory is that an individual selects an alternative out of the available ones such that his utility is maximized (5, 31). The mathematical details of this theory as well as the available procedures for model calibration is explained in Appendix

A. Multinomial Logit Analyses which relates the utility of the alternative to the probability of choice is used to calibrate models. Various computer packages for building logit models are available and examples of such packages includes BLOGIT (32), ULOGIT (33), SLOGIT (33), etc..

In this study, BLOGIT (Basic LOGIT) package is used to calibrate the desired models. This package uses the maximum likelihood technique to calibrate logit models (26). The output of this package includes various statistical performance indicators, such as;

1. t-test and the associated significance for parameter estimates ;
2. log of Likelihood function value $LL(\beta)$ at its maximum ;
3. log of Likelihood function value $LL(0)$ when all parameters are zero, in other words when all alternatives are assumed to have equal probability of being chosen ;
4. goodness-of-fit index rho-square (ρ^2) that measures the fraction of an initial likelihood value explained by the model, which is computed as:

$$\rho^2 = 1 - \frac{LL(\beta)}{LL(0)}$$

5. corrected goodness-of-fit index rho-bar square ($\bar{\rho}^2$) similar to rho square but corrected for the number of parameters estimated which is calculated as follows:

$$\bar{\rho}^2 = \frac{1-(LL(\beta)-K)}{LL(0)}$$

where, K is the number of parameters estimated in the model :

6. prediction success table that shows the number of individuals pre-

dicted to use each mode ;

7. various other optional outputs such as elasticities, etc,.

It is also desired that the signs of parameter estimates be logical. For instance, if travel time and travel cost coefficients have positive signs then a decrease in these variable values will decrease the demand and vice versa which is illogical.

3.4.1 Model Calibration

Mode-choice models are calibrated on 2/3rds of the data and the remaining 1/3rd is reserved for validation. First, mode-choice model is calibrated for choice riders with valid alternatives data only (Model 1). In the next phase of this stage assuming that the choice riders exercise full choice over all the available alternatives and keeping the same specification of Model 1, Model 2 is obtained. Finally captive riders were included in the dataset of choice riders with all alternatives, and Model 3 is calibrated.

3.4.2 Comparison of the Models

In this part, Models 1, 2 and 3 are compared with respect to the statistical performance indicators. Among the most important ones, that are studied, are the coefficient estimates of all the variables of the models, the goodness-of-fit statistics, rho square of the models and the rho-bar square values.

3.5 VALIDATIONS OF THE MODE-CHOICE MODELS

After the calibration process is completed and the models have been compared, validity of the mode-choice models is checked. Approximately 1/3rd of the reserved dataset of each category is used for this purpose. The three situations that is likely to result are tested:

1. prediction of models (viz., Model 1, 2 and 3) on data of choosers with valid alternatives data,
2. prediction of models (viz., Model 1, 2 and 3) on data of choosers with all alternatives as choices and,
3. prediction of models (viz., Model 1, 2 and 3) on data of captive riders plus choosers with all alternatives.

The validity of the models are tested by the Likelihood Ratio Test Statistics (LRTS). The null hypothesis formulated for this purpose is as follows;

H_0: There is no difference between the observed and the predicted behavior i.e., there is no difference between the parameter vectors obtained from calibration data and the validation data.

$$H_0 : \beta_i = \beta_j$$

where β_i and β_j are the estimated parameter vectors of the models obtained from the calibration and validation data (Note: same specification is needed for this test).

To obtain the LRTS value, the coefficients of the variables of a particular model will be restricted and the BLOGIT program is executed with the validation data. The program outputs two Log-Likelihood values. First value is the one which is computed by restricting the coefficients of the calibrated model while the second value is the one when the parameters are unrestricted for the validation data.

The LRTS value is obtained by the following relationship;

$$LRTS_i(\beta_j) = -2\{LL_i(\beta_j) - LL_i(\beta_i)\}$$

where,

$LRTS_i(\beta_j)$ – represents the Likelihood Ratio Test Statistics which restricts the parameters estimated from data j to be used to predict mode share in data i for same specification

$LL_i(\beta_j)$ – – Log-Likelihood Ratio value when the parameters are restricted in data j

$LL_i(\beta_i)$ – – Log-Likelihood value when the parameters are unrestricted in data j

The LRTS test discussed above is distributed as Chi-Square with K degrees of freedom , where K is the number of model parameters. If the LRTS values are less than the critical Chi-Square value (at 95% confidence level and degrees of freedom equal to K) then for that particular case the null hypothesis cannot be rejected otherwise, it is rejected.

CHAPTER 4

PRELIMINARY ANALYSES OF INTERCITY TRAVELLERS

4.1 INTRODUCTION

This chapter is concerned with the preliminary analyses of the intercity travellers. Initially, frequency analyses performed on the intercity travel dataset are highlighted. Analyses performed on choice and captive travellers to determine their percent share for various travel and socioeconomic characteristics is presented next. Later, the important travel and socioeconomic characteristics of the travellers that influence the decision for mode-choice is investigated. The chapter is concluded by comparing the choice and captive travellers with respect to the various travel and socioeconomic features.

4.2 DESCRIPTION OF THE DATA

The data utilized in this study was collected as part of a research project by Al-Ahmadi et. al. in February 1993 (4). A self administered questionnaire was developed for collecting intercity travel data for all modes of transport on all the major corridors of the Kingdom. Choice-based sampling technique was adopted to collect the intercity travel data i.e., sampling was done for each mode separately. The sampling strategy for air, bus, and train was done first by selecting a scheduled flight, bus or train randomly from the total weekly services and then systematically selecting every n^{th} incoming passenger at the airport gates, at bus stops, and trains. The questionnaires were distributed to the selected passengers and their responses collected. Interviewer's assistance was provided to those who faced difficulty in answering the questions. For latecomers, a self addressed envelope was given and they were asked to send their replies through mail. Interviews for auto mode were conducted at selected gas stations around major cities of the Kingdom (namely, Dammam, Riyadh, Madinah, Jeddah). It is observed in Saudi Arabia that a majority of the intercity trips are carried out either by air or by private cars. Hence, it was decided by Ahmadi et. al. (4) to collect more data about the air and the auto mode to improve the reliability. The English version of the questionnaire used for collecting the intercity air travel data is enclosed in Appendix B for reference. The actual sample sizes that were obtained through field interviews for the four modes are given in Table 4.1 (4).

The questionnaire contained the level-of-service variables of the modes, socioeconomic data of the traveller and data regarding the trip. Apart from various other questions, the questionnaire also contained the following question:

" Will you consider using the following means of travel for a trip similar to the one you are making "

The answer to the above question was indicated for modes other than the one he is using on the following scale:

1. **Will Never use it**
2. **May consider Using it**

Table 4.1: Actual Sample Sizes for the Field Interviews

Mode	Total sample size
Air	1436
Auto	1392
Bus	787
Train	264
TOTAL	4079

3. Definitely Consider Using it

If the traveller answered the first option for all modes then, he was considered to be a CAPTIVE RIDER to the mode he is currently using.

Some of the responses collected from the travellers were incomplete. The incomplete responses were not included in the data keeping in mind that this is likely to produce a bias in the collected sample. For some cases the traveller furnished information about his socioeconomic characteristics but didn't give response to the travel features i.e., travel time, travel cost, etc., engineering values were supplied to those missing components of travel time and travel cost. Altogether about 16% of the total data that contained missing information and those (Origin-Destination) O-D pairs for which all the three alternatives were not available i.e., air, auto, bus were deleted from the travel data. The data which was used for analyses purpose contains 3337 responses. This data is utilized for the present study.

4.3 GENERAL ANALYSES OF THE DATA

Initially frequency tables were obtained on the whole dataset to determine the distribution of travellers for various travel and socioeconomic characteristics. The results of this analyses are summarized in the form of frequency tables few of which are enclosed in Appendix C.

Table 4.2 shows the distribution of the intercity travellers for various trip purposes. As can be seen from the table, majority of the trips, around 40.7% were made for social/recreational purposes. Trips made for personal work and business were very few as compared to the social trips.

Table 4.2: Frequency Table for Various Trip Purposes

Trip Purpose	Frequency	Percent
Work	532	15.9
Personal Business	419	12.6
Social/Recreation	1359	40.7
Educational/Study	143	4.3
Umra	660	19.8
Others	224	6.7

Education/study trips constituted the least percent (4.3%). This might be because of the holiday break for the educational establishments when the survey was conducted.

Trips made for business purpose such as work, personal business possess different travel and socioeconomic characteristics compared to trips made for non-business purpose such as social and recreational trips. Furthermore, the purpose of the trip plays a prominent role in an individual's spending on the trip, selection of a particular destination, staying period at the destination, etc,. Thus, the purpose of the trip will have different sensitivities to the variables affecting mode-choice, such as travel time and travel cost.

It is observed in the collected data that the majority of the trips were made for non-business purpose. Hence, trips made for non-business purpose which constitute a large sample size, are used for model building and validation. This will enhance the reliability of the model constructed. The data pertaining to the trips made for non-business purpose is segmented from the collected dataset. The travel and socioeconomic characteristics of these trips are incorporated in mode-choice modelling. The following sections highlights the analyses performed on the data of these trips.

The data pertaining to non-business trips can further be classified into two more sub categories, namely, captive riders and non-captive riders. The distinction between captive and non-captive riders is made based on the response of the traveller to the question " Will you consider using . . . ." (Section 4.2). The response of the travellers to the above question is tabulated in Table 4.3. It can be observed from the table that 17.0 percent of the travellers indicated that they will never use air mode, 12.0 percent of the total travellers indicated that they will never use car, 50.7 percent of the total travellers indicated that they will never use bus and 29.5 percent of the travellers indicated that they will never consider using train for their future intercity trips. These figures shows the preference of the travellers to the modes other than the one they are currently using. A traveller is captive to a particular mode if he responded that he will never use the remaining modes. Separate analyses is performed on captive and non-captive riders i.e., choice riders.

4.4 DISTRIBUTION OF INTERCITY TRAVELLERS FOR VARIOUS TRAVEL AND SOCIOECONOMIC CHARACTERISTICS

After having separated captive and choice riders from the intercity travel dataset, statistical analyses were performed to determine the percentage-wise distribution of captive riders and choice riders on various modes for different trip characteristics such as trip purpose, trip length, trip duration and socioeconomic characteristics. Initially analyses performed on choice riders to determine their distribution for various demographic features is presented. Later, the procedure is repeated for captive riders.

Table 4.3: Consideration of Various Modes

Consideration	Percentage of Responses for			
	Air	Car	Bus	Train
Will Never use	17.00	12.00	50.70	29.46
May Consider using	27.90	22.80	26.80	30.75
Definitely consider	55.90	64.30	19.40	39.78

4.4.1 *Analyses on Choice Travellers*

The main objective of conducting frequency and cross-frequency analyses on choice riders is to determine the distribution of choice travellers on different modes for various trip characteristics, such as, trip purpose, trip length, duration of stay at destination and socioeconomic characteristics.

In order to study the effect of various trip and socioeconomic characteristics on the selected mode by a choice rider making a non-business trip, cross frequency analyses were carried out between the mode selected versus the various travel and socioeconomic characteristics. A sample cross frequency table between mode and duration of stay at destination is presented in Table 4.4. As can be seen from the table, the duration of stay for about 50% of the travellers by auto mode is greater than 3 days. It is evident from the dataset that if the trip involved short duration of stay, travellers preferred air mode. On the other hand, car mode is preferred for trips involving long duration of stay because cars can be used for intra-urban trips at the place of stay.

The decision of a choice rider to select a particular mode of travel depends to a large extent on the duration of stay at the destination. Cross frequency analyses is performed on the remaining trip and socioeconomic characteristics of choice riders. Few of those tables are included in Appendix C.

Table 4.4: Cross-Frequency Table for Mode by Duration of Stay

Mode	Duration				
	1-3 Days	4-7 Days	8-30 Days	>30 Days	Total
Air	293.00°	190.00	57.00	14.00	554.00
	21.88°°	14.19	4.26	1.05	41.37
	52.89±	34.30	10.29	2.53	
	41.21±±	45.24	40.74	20.59	
Car	255.00°	141.00	41.00	38.00	475.00
	19.04°°	10.53	3.06	2.84	35.47
	53.68±	29.68	8.63	8.00	
	35.86±±	33.57	29.29	55.88	
Bus	75.00°	77.00	34.00	11.00	197.00
	5.60°°	5.75	2.54	0.82	14.71
	38.07±	39.09	17.26	5.58	
	10.55±±	18.33	24.29	16.18	
Train	88.00°	12.00	8.00	5.00	113.00
	6.57°°	0.90	0.60	0.37	8.44
	77.88±	10.62	7.08	4.42	
	12.38±±	2.86	5.71	7.35	
Total	711.00	420.00	140.00	68.00	1339.00
	53.10	31.37	10.46	5.08	100.00

Frequency°
Percent°°
Row Percent±
Col. Percent±±

4.4.2 *Travel and Socioeconomic Characteristics of Captive Travellers*

In this section the distribution of captive travellers for different travel and socioeconomic characteristics is presented.

This analyses is performed in the following steps:

a. frequency analyses is performed mode-wise for each trip and socioeconomic characteristic,

b. percentage-wise distribution of captive riders with respect to choice riders is estimated for all modes combined.

Table 4.5 shows the distribution of captive riders for different durations of stay at the destinations. The entries along the row indicate the percent captivity to each mode for different durations. The last row represents the captivity combined over all modes for different durations. The last column in the table is the overall captivity to a particular mode summed over different durations. In the survey captivity to train mode is not reported therefore it is not considered for further analyses. It can be seen from Table 4.5 that the captivity to air mode decreases considerably when the duration of stay exceeds 30 days. In the case of car mode the captivity is little higher for the duration between 1-3 days. The duration of stay at the destination does not affect the captivity to bus mode.

The results of this analyses are later utilized for comparison with the choice travellers percentage in the form of tables.

Table 4.5: Percent Captivity on Different Modes for Various Durations of Stay

Mode	Duration				Total
	1-3 Days	4-7 Days	8-30 Days	>30 Days	
Air	23.50	24.00	16.18	26.32	23.06
Car	17.74	15.57	18.00	9.52	16.52
Bus	7.41	8.33	8.11	0.00	7.51
Total	19.51	18.56	14.84	12.50	17.08

4.5 RELATIONSHIP BETWEEN TRAVEL AND SOCIOECONOMIC CHARACTERISTICS AND INTERCITY RIDERSHIP

It is assumed that all the trip and socioeconomic characteristics play an equally important role in the decision of the traveller to select a particular mode of travel. In order to study the strength of relationship between the trip and socioeconomic characteristics and the captive ridership variable CAPTIVE (which contains the categories of CAPTIVE and NOT-CAPTIVE variables), Chi-Square test is performed. This test is performed on the cross frequency between the variable CAPTIVE and different trip and socioeconomic characteristics. The null hypothesis for this test is as follows:

H_0 : There exists 'No Relationship' between the CAPTIVE variable and the corresponding trip and socioeconomic variable under consideration

In order to test this hypothesis the Chi-Square test is applied to air, car , bus mode users separately taking one at a time. Train mode is not used because of lack of data.

For each combination of the captive ridership variable, CAPTIVE, and the trip and socioeconomic characteristics, the Chi-square value, its significance and Cramer's V statistics are computed. Cramer's V statistics which vary from 0-1 and it gives an indication of the strength of the relationship among the variables. The higher the Cramer's V value, the greater is the relationship with the CAPTIVE variable.

The null hypothesis is rejected if the significance level of the Chi-Square statistics is greater than 0.10 (i.e., confidence level is less than 90%). The results of this test is summarized in Table 4.6. The cases where the null hypothesis has been rejected at 90% confidence is indicated by asterisks.

From Table 4.6 it can be observed that the characteristics which affect the ridership variable at least for one mode of travel are duration of stay at destination, distance of travel, size of accompanied family, age, marital status, number of cars, driving license possession, number of licenses in a family, occupation, nationality and education. Among the above characteristics age, nationality, education are found to have the strongest relationship with the variable CAPTIVE because they have the highest Cramer's V statistics.

From the above statistical analyses it is found that the decision to select a particular mode of travel by a traveller depends on his socioeconomic background and to some extent on the trip characteristics i.e., trip purpose, trip length and trip duration.

4.6 *COMPARISON BETWEEN CHOICE AND CAPTIVE RIDERS FOR TRAVEL AND SOCIOECONOMIC CHARACTERISTICS*

In this section the trip and socioeconomic characteristics that influence the intercity ridership for captive riders are compared to non captive riders with the help of tables. These tables summarize the percent travellers on various modes for different travel and socioeconomic features.

Table 4.6: Test of Relationship between Intercity Ridership and Travel/Socioeconomic Variable

Variable Code Explanation	Chi-Square/Significance/Cramer's V		
	Air	Car	Bus
DURT (Duration)	2.09/0.55/0.05@	2.02/0.57/0.06@	0.99/0.80/0.07@*
DIST (Trip Length)	14.62/0.02/0.14*	9.39/0.15/0.13@*	2.31/0.88/0.02@*
PURP (Trip Purpose)	5.15/0.02/0.09	2.59/0.11/0.07@	0.08/0.77/-0.02@*
FMLY (Family size)	8.77/0.12/0.11@	15.12/0.20/0.10@*	16.81/0.00/0.28*
AGE (Age)	7.14/0.13/0.10@	5.99/0.20/0.10@	16.81/0.00/0.28*
MSTAT (Marital Status)	1.16/0.28/0.04@	3.32/0.07/0.08	1.08/0.29/0.07@
NUMCAR (Ownership car)	1.01/0.79/0.04@*	0.53/0.91/0.03@*	1.09/0.78/0.07@*
DLICE (License Possession)	3.48/0.06/0.07*	0.26/0.61/-0.02@	0.25/0.62/-0.03@
NUMDI (Licenses in a family)	11.61/0.04/0.13	6.62/0.25/0.11@	7.66/0.18/0.19@*
OCCUP (Occupation)	7.38/0.19/0.10@	8.39/0.14/0.12@*	15.73/0.01/0.27
NATION (Nationality)	5.49/0.48/0.09@*	11.99/0.06/0.15*	28.75/0.00/0.22
EDUC (Education)	8.44/0.02/0.11	9.16/0.01/0.13	2.57/0.28/0.11@
PERINC (Personal Income)	5.16/0.64/0.09@	15.81/0.03/0.17	14.95/0.04/0.16
HHINC (Household Income)	15.57/0.03/0.15/	@12.69/0.08/0.15	8.17/0.32/0.19@*

@ - Cases where null hypothesis is rejected
* - Some cells have counts less than 5. Chi-Square may not be a valid test

Train mode is not considered since it lacks data for captive riders.

a. **Duration of Stay:** Table 4.7 shows the percentage of captive and choice travellers on various modes for different durations of stay at destination. For durations less than a week choice and captive riders seem to prefer air mode. As the duration increases the preference of travellers seems to shift towards the car mode. The behavior of captive and choice travellers is similar for the selection of a particular mode for various durations.

b. **Distance of Travel:** Table 4.8 shows the distribution of captive and choice riders for various trip lengths. It seems that for distances less than or equal to 600 kilometers the preference of choice and captive travellers is car mode. As the distance increases, choice and captive riders seem to prefer air mode.

c. **Family Size:** The distribution of travellers for different sizes of accompanied family is shown in Table 4.9. It can be observed that in the case of air mode as the size of the family increases, captivity to air also increases. For small families (less than 3 members) air seems to be the preferable mode for both choice and captive riders. It can be inferred that when travellers (captive or choice) move with their spouses they prefer air and as the size increases they shift to car mode. Reason to this might be the cost of travel.

d. **Age:** Table 4.10 shows the distribution of choice and captive riders for different age groups. As the age increases choice riders seem to prefer air. This preference could be because of its comfort and convenience. Sim

Table 4.7: Comparison between Choice and Captive Travellers on Various Modes for Different Durations of Stay at Destination

Duration of Stay	Percent Travellers on Various Modes					
	Air		Bus		Car	
	Choice	Captive	Choice	Captive	Choice	Captive
1-3 Days	76.50	23.50	92.59	7.41	82.26	17.74
4-7 Days	76.00	24.00	91.67	8.33	84.43	15.57
8-30 Days	83.82	16.18	91.89	8.11	82.00	18.00
> 30 Days	73.68	26.32	99.99	0.01	90.48	9.52

Table 4.8: Comparison between Choice and Captive Travellers on Various Modes for Different Trip Lengths

Trip Length (Kms.)	Percent Travellers on Various Modes					
	Air		Bus		Car	
	Choice	Captive	Choice	Captive	Choice	Captive
< 400	80.23	19.77	94.51	5.49	89.56	10.44
400-600	72.18	27.82	89.19	10.81	87.50	12.50
600-800	66.67	33.33	89.47	10.53	82.22	17.78
800-1000	81.97	18.03	93.02	6.98	80.70	19.30
1000-1200	64.29	35.71	99.99	0.01	79.17	20.83
1200-1400	66.67	33.33	88.24	11.76	77.65	22.35
> 1400	99.99	0.01	99.99	0.01	77.27	22.73

Table 4.9: Comparison between Choice and Captive Travellers on Various Modes for Different Sizes of Accompanied Family

Family Size	Percent Travellers on Various Modes					
	Air		Bus		Car	
	Choice	Captive	Choice	Captive	Choice	Captive
1	80.12	19.88	93.57	6.43	79.89	20.11
2	82.03	17.97	85.29	14.71	91.00	9.00
3	79.03	20.97	99.99	0.01	79.41	20.59
4	71.43	28.57	87.50	12.50	96.77	3.23
5	67.53	32.47	87.50	12.50	99.99	0.01
>5	75.19	24.81	87.50	12.50	99.99	0.01

Table 4.10: Comparison between Choice and Captive Travellers on Various Modes for Different Age Groups

Age (Years)	Percent Travellers on Various Modes					
	Air		Bus		Car	
	Choice	Captive	Choice	Captive	Choice	Captive
< 20	69.44	30.56	97.06	2.94	83.61	16.39
20-29	80.91	19.09	93.75	6.25	85.92	14.08
30-39	74.77	25.23	92.16	7.84	83.46	16.54
40-49	76.85	23.15	99.99	0.01	[illegible]	20.59
50 & over	66.67	33.33	71.43	28.57	[illegible]	[illegible]

ilar liking towards the air mode seems to exist among captive riders also. As the age increases captivity to air and car modes seems to increase whereas, the percentage of choice riders seem to decrease.

e. **Marital Status:** Table 4.11 shows the percent travellers on various modes for marital status. As can be seen from the table, the captivity to air and car decreases if the traveller is not married, but for choice riders the phenomenon is opposite. Unmarried choice travellers prefer to move by air or car mode. Thus, the behavior of captive and choice riders seem to differ for this feature.

f. **Car Ownership:** Table 4.12 shows the distribution of choice and captive riders for car ownership. Captivity to car seems to decrease as the number of cars owned increases by the captive rider. On the other hand, in case of choice riders, as the number of cars possessed increases their percentage also increases. This feature seems to have very little effect on the distribution of air travellers (choice and captive).

g. **License Possession:** Table 4.13 shows the distribution of choice and captive riders for license posession. It can be observed from the numerical values in the table that choice travellers who have no driving license prefer to travel by car. In the case of bus mode, ridership seem to be least affected and its patronage is same whether the traveller (choice/captive) possesses a valid driving license or not.

h. **Number of licenses in a family:** Table 4.14 shows the distribution of captive and choice riders for number of licenses in a family. As the license

Table 4.11: Comparison between Choice and Captive Travellers on Various Modes for Marital Status

Status	Percent Travellers on Various Modes					
	Air		Bus		Car	
	Choice	Captive	Choice	Captive	Choice	Captive
Married	75.84	24.16	90.83	9.17	81.04	18.96
Unmarried	79.53	20.47	94.62	5.38	86.78	13.22

Table 4.12: Comparison between Choice and Captive Travellers on Various Modes for Car Ownership

Number of Cars	Percent Travellers on Various Modes					
	Air		Bus		Car	
	Choice	Captive	Choice	Captive	Choice	Captive
0	76.47	23.53	92.05	7.95	80.65	19.35
1	77.78	22.22	93.07	6.93	83.51	16.49
2	73.50	26.50	88.24	11.76	84.81	15.19
> 2	76.09	23.91	99.99	0.01	85.00	15.00

Table 4.13: Comparison between Choice and Captive Travellers on Various Modes for License Possession

Status	Percent Travellers on Various Modes					
	Air		Bus		Car	
	Choice	Captive	Choice	Captive	Choice	Captive
Have	76.46	23.54	93.10	6.90	83.76	16.24
Don't Have	94.74	5.26	91.18	8.82	81.25	18.75

Table 4.14: Comparison between Choice and Captive Travellers on Various Modes for Number of Licenses in a Family

No. of Licenses	Percent Travellers on Various Modes					
	Air		Bus		Car	
	Choice	Captive	Choice	Captive	Choice	Captive
0	77.07	22.93	94.52	5.48	80.98	19.02
1	76.33	23.67	96.36	3.64	79.76	20.24
2	78.00	22.00	91.18	8.82	87.80	12.20
3	78.26	21.74	88.89	11.11	79.35	20.65
4	61.82	38.18	93.33	6.67	87.50	12.50
5	87.32	12.68	77.78	22.22	89.13	10.67

holding members in a family increases captivity to air seems to decrease, whereas, choice rider's percentage increases. Similar phenomenon is observed in the case of car mode also. Bus mode seems to attract travellers from those households which have no license holding member.

Thus, in summary, it can be concluded that for trip features viz., duration of stay at destination and trip length, the behavior of captive and choice riders seem to be more or less the same, whereas, for socioeconomic features like car ownership, license possession, number of licenses in a family, age, marital status the behavior of captive riders seem to be different to that of the choice riders.

CHAPTER 5

MODEL CALIBRATION, COMPARISON AND VALIDATION

5.1 INTRODUCTION

In this chapter development of mode-choice models is presented in detail. The results obtained from the preliminary analyses of the intercity travellers are utilized for segmenting the travellers data and as well as for selecting the best potential variables that affect the mode-choice. Engineering values have been used for level-of-service variables viz., travel time components and travel cost components. The calibration process is carried out in three stages. In the first stage, mode-choice model is built on the data of choice riders with valid alternatives only (those modes which are perceived as valid by the choice riders, this is termed as Model 1). In the second stage, keeping the same specifications of the previous model mode-choice model is built on choice riders data with all alternatives, (this is termed as Model 2). The reason for keeping the same model specification is to facilitate the comparison of sensitivities to variables across different modes. Finally, captive riders are included in the data used for calibrating Model 2 and mode-choice model is built with the same model specification used in Models 1 and 2 (this is termed as Model 3).

The models built are compared to each other with regard to their coefficient estimates and other goodness-of-fit statistics. Later, the elasticities of the models (Model 1, Model 2, Model 3) for travel variables are compared. Finally, the chapter is concluded by validating the calibrated models and the prediction results obtained are reported.

5.2 SEGMENTATION OF TRAVELLERS DATASET

From the preliminary analyses of the intercity travellers data it is noticed that the mode shares are affected by travel and socioeconomic characteristics. As stated in chapter four, non-business trips are selected for this study.

Non-business trips are those which are made for purposes other than business. Social, recreational and trips classified in the category of other purposes are included in this segment. The dataset of the travellers making trips for non-business purposes is divided into three sets. Set I contains only the data of choice riders with valid alternatives. Set II contains data of choice riders with all alternatives. Set III contains data of choice riders with all alternatives and captives. Each category of the data is further divided into two parts. 2/3rd of the data is reserved for calibration and 1/3rd for validation. The selection of the data for calibration and validation is randomized by systematic sampling. Every third case is used for validation and the others are used for calibration. The size of each type of data that is used for calibration and validation is shown in Table 5.1.

5.3 MODEL VARIABLES

The list of variables used in model calibration with their abbreviations and explanations are presented in Table 5.2.

Table 5.1: Data Segmentation For Non-Business Trip Purpose

Data Type	Calibration Sample Size	Validation Sample Size
Choice Riders (Valid)	813	386
Choice Riders (All)	813	386
Captive+Choice (All)	995	487

Table 5.2: Variable Codes and Descriptions

Variable Name	Description of the Variable
TKCOST	Ticket Cost
TOTCOST	Total Travel Cost
TOTTT	Total Travel Time
INVTT	In-Vehicle Travel Time
OUTVTT	Out-of-Vehicle travel Time
ACSSEGST	Access and Egress Time
WAITT	Waiting Time
PERINCA	Personal Income (Specific to Air Mode)
PERINCB	Personal Income (Specific to Bus Mode)
HHINCA	Household Income (Specific to Air Mode)
HHINCB	Household Income (Specific to Bus Mode)
DHHINCB	Dummy for Household Income (Specific to Bus Mode) = 1 if household income is 2500 SR's p.m. or less = 0 otherwise
DISTC	Distance specific to Car Mode = 1 if distance is 900 Kms. or less = 0 otherwise
DFMLYC	Dummy for family (Specific to Car Mode) = 1 if family by Air and Car Modes > 1 = 0 otherwise
DNCARB	Dmmy for Number of Cars (Specific to Bus Mode) = 1 if Number of Cars owned is zero = 0 otherwise
DDURTA	Dummy for Duration of Stay = 1 if duration is 1-3 days = 0 otherwise
DNATIONA	Dummy for Nationality (Specific to Air) = 1 if nationality is American/European/Australian = 0 otherwise
DNATIONB	Dummy for Nationality (Specific to Bus) = 1 if nationality is Asian/Africans/Other Arabs = 0 otherwise

The specifications of the model are formulated using the following criteria:

1) coefficient estimates not significant at 95% confidence level are dropped,
2) variables which give counter-intuitive signs are not considered,
3) variables related to the Level-of-Service (i.e., supply variables) are used either in ratio forms (cost/income) or simpl form as cost.

5.4 NON-BUSINESS TRIP MODELS --- CALIBRATION, COMPARISON AND VALIDATION

In this section, mode choice models calibrated for non-business trips are presented. The process involved in selecting the best model specifications is also highlighted. The models calibrated are compared to each other with respect to their coefficient estimates and other goodness-of-fit statistics. The validation process is carried out later.

5.4.1 Development of mode-choice models for choice riders with valid alternatives (Model 1)

For calibrating mode-choice models for non-business trips, a systematic procedure is adopted. First of all, the best attribute of travel time and travel cost are determined by introducing into the model different components and formulations of these variables in a sequential manner. These attributes result in the best specification of the model that can replicate the behavior of the travellers.

Initially ticket cost (TKCOST) and in-vehicle travel time (INVTT) variables are included as generic variables with other specifications in the model. Other specifications included the household income specific to air mode (HHINA), distance specific to air (DISTC), dummy variable for number of cars specific to bus mode (DNCARB) and dummy variable for nationality specific to the bus mode (DNATIONB). The incorporation of these variables resulted in a coefficient estimates significant at 95 percent confidence interval and yielded a rho square value of 0.364 and rho bar-square value of .361 .The parameter estimates and other statistics of the model with the above mentioned specifications is shown in column 1 of Table 5.3 as Model 1.A. In order to further improve the statistics of the model, another variable, namely, the dummy variable for family size specific to car mode (DFMLYC) is included in the developed model. The variable is found to be statistically significant and furthermore, its inclusion resulted in an increase in the rho-square value from 0.364 to 0.389 as can be seen in Table 5.3. This improved model is Model 1.B.

To further improve the fit of the model, in Model 1.C the variable ticket cost (TKTCOST) is replaced by the total cost variable (TOTCOST) and another dummy variable for nationality specific to air mode (DNATIONA) is introduced alongwith the variables of Model 1.B. The coefficient of DNATIONA is found to be insignificant at 95 percent confidence level. However the rho-square value increased to 0.395. In Model 1.D, DNATIONA is removed because of its insignificance and BLOGIT program is run. This resulted in a good fit with all the variables being significant icant at 95 percent confidence level with the rho-square value of 0.393. In

Table 5.3: Intercity Mode Choice Models for Non-Business Trips using data of choice riders with valid alternatives

Variable Name	Estimated Variable Coefficients (t-stat)				
	Model 1.A	Model 1.B	Model 1.C	Model 1.D	Model 1.E
ASC AIR	1.031(2.26)	1.452(-3.17)	1.124(1.95)	1.674(3.56)	1.532(3.15)
ASC-BUS	1.634(4.89)	1.997(6.02)	2.066(6.09)	2.095(6.17)	2.118(6.19)
INVTT	-.149(-4.69)	-.138(-4.25)	-.148(-4.52)	-.147(-4.53)	-.153(-4.63)
TOTCOST			-.002(-5.12)	-.002(5.14)	-.002(-5.05)
TKCOST	-.004(-8.60)	-.002(-4.67)			
DNCARB	2.101(6.76)	1.983(6.33)	1.987(6.28)	2.001(6.34)	1.985(6.29)
DFMLYC		1.397(6.32)	1.310(5.93)	1.327(5.98)	1.337(6.02)
DNATIONA			0.585(1.65)		
DNATIONB	0.938(3.27)	1.132(3.82)	0.866(2.59)	1.112(3.74)	1.127(3.78)
DDURTA					.226(1.21)
HHINCA	0.215(4.52)	0.184(3.78)	0.171(3.48)	0.182(3.75)	0.182(3.74)
DISTC	2.392(7.61)	2.139(7.04)	2.114(6.86)	2.176(7.08)	2.186(7.08)
LL(β)	-521.54	-501.24	-496.46	-497.88	-497.15
LL(O)	-821.12	-821.12	-821.12	-821.12	-821.12
Rho sq. ρ^2	.364	.389	.395	.393	.394
Rho-bar sq.	.361	.386	.391	.390	.390

Model 1.E to further increase the fit of the model, the dummy variable duration of stay at destination specific to air mode i.e., DDURTA is added in Model 1.D. The inclusion of this variable increased the rho-square value to 0.396 but the coefficient of the variable is insignificant at 95 percent confidence level.

From the above analyses, it can be concluded that Model 1.D which contains DNCARB, DFMLYC, DNATIONA, HHINCA, DISTC, besides TOTCOST and INVTT represents the best model specification and can be used in explaining the behavior of the intercity rider for non-business trips. All the variable coefficients are found to be statistically significant at 95 percent confidence level and it had a good rho-square value.

The utility function for each mode for the best model (i.e., Model 1) is given below,

For air mode(A):

$$U(A) = 1.674 - 0.1476(INVTTA) - 0.0022(TOTCOSTA) + 0.1828(HHINA)$$

For bus mode(B):

$$U(B) = 2.095 - 0.1476(INVTTB) - 0.0022(TOTCOSTB) + 2.001(DNCARB) + 1.112(DNATIONB)$$

For auto mode(C):

$$U(C) = -0.1476(INVTTC) - 0.0022(TOTCOSTC) + 1.327(DFMLYC) + 2.176(DISTC)$$

5.4.2 *Calibration of mode-choice model for choice riders with all alternative choices (Model 2)*

In this section the procedure adopted for calibration of Model 2 is presented. It involves taking the specification of Model 1and running BLOGIT program for the data of choice riders with all alternatives as choices. The statistical results of this model are presented in Table 5.4. It can be observed from the table that the coefficient estimates of all the variables are significant at 95 percent confidence level but the rho-square value of this model is (0.391) which is slightly lower than that of model 1 (0.393). The utility functions for each mode under the assumption of full choice by the choice riders are as follows:

For air mode (A):

$$U(A) = 1.67 - 0.14(INVTTA) - 0.0022(TOTCOSTA) + 0.21(HHINA)$$

For bus mode (B):

$$U(B) = 1.56 - 0.14(INVTTB) - 0.0022(TOTCOSTB) + 2.13(DNCARB) + 1.48(DNATIONB)$$

For auto mode (C):

$$U(C) = -0.14(INVTTC) - 0.0022(TOTCOSTC) + 1.36(DFMLYC) + 2.30(DISTC)$$

Table 5.4: Mode-Choice Model for Non-Business Trips for Choice Riders with All Alternatives

Variable Name	Coefficient Estimates	t Statistics	Standard Error
ASC-Air	1.67	3.68	0.45
ASC-BUS	1.56	4.71	0.33
INVTT	-.14	-4.60	0.03
TOTCOST	-.0022	-5.31	0.0003
DNCARB	2.13	7.92	0.27
DFMLYC	1.36	6.43	0.21
DNATIONB	1.48	5.52	0.27
HHINCA	0.21	4.39	0.047
DISTC	2.30	7.68	0.30

Summary Statistics:

Number of Observations : 817
LL(B) = -581.672
LL(0) = -821.124
Rho Square = 0.391
Rho bar Square = 0.387

5.4.3 Calibration of mode-choice model with captive riders (Model 3)

In this section the procedure adopted for the calibration of Model 3 is presented. The objective of this is to see the effect of captive riders in intercity mode-choice modelling. In this procedure captive riders are included in the dataset of choice riders of all alternatives and the model is built using the same specifications of the previous model. Table 5.5 highlights the parameter estimates of the model along with their t-statistics and standard error. It can be seen from the table that the coefficients of all the variables are significant at 95 percent confidence level but the rho-square value dropped to 0.311 when captive riders are included in the model.

The utility functions for all the riders regardless of whether they are choice riders or captive riders are as follows:

For air mode (A):

$$U(A) = 1.61 - 0.16(INVTTA) - 0.0022(TOTCOSTA) + 0.211(HHINA)$$

For bus mode(B):

$$U(B) = 1.53 - 0.16(INVTTB) - 0.0022(TOTCOSTB) + 2.04(DNCARB) + 1.34(DNATIONB)$$

For auto mode (C):

$$U(C) = -0.16(INVTTC) - 0.0022(TOTCOSTC) + 1.34(DFMLYC) + 2.35(DISTC)$$

Table 5.5: Mode-Choice Model for Non-Business Trips for Choice Riders with All Alternatives and Captive Riders

Variable Name	Coefficient Estimates	t Statistics	Standard Error
ASC-Air	1.61	3.92	0.41
ASC-BUS	1.53	5.03	0.30
INVTT	-.16	-5.59	0.03
TOTCOST	-.0022	-6.21	0.0004
DNCARB	2.04	8.09	0.25
DFMLYC	1.34	6.78	0.19
DNATIONB	1.34	5.38	0.25
HHINCA	0.21	4.87	0.043
DISTC	2.35	8.67	0.27

Summary Statistics:

Number of Observations : 995
LL(B) = -974.099
LL(0) = -670.889
Rho Square = 0.311
Rho bar Square = 0.308

5.4.4 *Comparison between the calibrated mode-choice models*

In this section the calibrated models are compared to each other with regard to their coefficient estimates and goodness-of-fit statistics. The three models developed in the previous sections are presented in Table 5.6. It can be observed from the table that the coefficients of the variables of the three models are more or less the same but the t-statistics of all the variables increased from Model 1 to Model 3. This resulted in decrease variances of all the variables. This decrease in variances is also due to the increased number of observations.

Stopher in his work (3) made a comment that the inclusion of captive riders in mode-choice models would result in increased variances of the variables and drive the alternative specific constants to the market shares. In the present case, the alternative specific constants remained more or less the same and the significance of these constants increased. The coefficients of the other variables also showed a similar phenomenon wherein, the inclusion of captive riders yielded increased fit of the variables. This might be because of the increased sample size of the calibration data. The inclusion of the captive riders though increased the significance of the individual variables but decreased the overall fit of the models. The rho-square value decreased to 0.311 for Model 3 from 0.393 in case of Model 1. The drop is marginal from Model 1 to Model 2. The rho-bar square value has shown a drastic drop from 0.390 (for Model 1) to 0.308 (for Model 3).

In order to find whether coefficient estimates of each individual variable for the three models differ significantly or not, studentized t-test is

Table 5.6: Comparison of Mode-Choice Models

Variable Name	Estimated Coefficients (t-stat)		
	Model 1	Model 2	Model 3
ASC-AIR	1.67/3.56	1.67/3.68	1.61/3.92
ASC-BUS	2.09/6.17	1.56/4.71	1.53/5.03
INVTT	-.15/-4.63	-.14/-4.60	-.16/-5.59
TOTCOST	-.002/-5.14	-.002/-5.31	-.0022/-5.31
DNCARB	1.98/6.29	2.13/7.92	2.04/8.09
DFMLYC	1.33/5.98	1.36/6.43	1.34/6.78
DNATIONB	1.11/3.74	1.48/5.52	1.34/5.38
HHINCA	0.18/3.75	0.21/4.39	0.21/4.87
DISTC	2.18/7.08	2.30/7.68	2.35/8.67
No. of obs:	813	813	995
LL(β)	-497.464	-581.672	-974.099
LL(O)	-821.124	-821.124	-821.124
Rho sq. ρ^2	.393	.391	.311
Rho-bar sq.	.390	.387	.308

performed. The null hypothesis formulated for this purpose is stated as follows:

H_0: There is 'No Significant Difference' between the coefficients of a variable obtained in Model (*i*) and Model (*j*).

where,

i = 1, 2, and 3

j = i+1 but not greater than 3

In order to test the above hypothesis t-statistics is calculated by using the following relationship:

$$t_{calc} = \frac{X_i - X_j}{\sqrt{SE_i^2 + SE_j^2}}$$

where,

X_i = coefficient estimates of Model (*i*)

X_j = coefficient estimates of Model (*j*)

SE_i = Standard error of the estimate in Model (*i*)

SE_j = Standard error of the estimate in Model (*j*)

The t_{calc} value is compared to the table t-value at degrees of freedom equal to $(n_i + n_j - 2)$ at 95 percent confidence level, where, n_i, n_i are the number of observations used for calibrating Model (*i*) and Model (*j*) respectively. If $t_{calc} < t_{table}$ the null hypothesis cannot be rejected, otherwise it can be rejected.

The results of this test is reported in Table 5.7. This test is performed for all the three possible combinations of the models taking each

Table 5.7: Results of the t-test between the coeffcients of the calibrated Models

Variable	t-test results (t-calc.)		
	Model 1 and Model 2	Model 1 and Model 3	Model 2 and Model 3
ASC-AIR	0.002	0.105	0.107
ASC-BUS	1.115	1.234	0.074
INVTT	-.125	-.281	0.424
TOTCOST	0.060	-.434	0.509
DNCARB	-.320	-.117	0.231
DFMLYC	-.117	-.039	0.083
DNATIONB	-.915	-.569	0.369
HIINA	-.353	-.446	-.082
DISTC	-.295	-.439	-.132

Note: t-table value at 95% confidence level and $(n_i + n_j - 2)$ degrees of freedom = 1.645

variable separately. The cases where the null hypothesis cannot be rejected are indicated in the table by marking asterisks beside them. It can be seen from the table that the null hypothesis cannot be rejected for almost all the cases, which indicate that the coefficient estimates of the models are more or less the same. Thus, the inclusion of captive riders resulted in parameter estimates of the variables that are same but decreased the overall fit of the models.

5.4.5 *Validation of the calibrated models*

After the calibration process is completed and the models are compared to each other, validation of the models is carried out. The process of validation is performed by restricting the coefficients of the variables of the calibrated models and then running the BLOGIT program using the datasets reserved for the purpose of validation. There are three types of validation datasets:

1) Validation data for choice riders with valid alternatives (Type 1)
2) Validation data for choice riders with all alternatives (Type 2)
3) Validation data for choice riders with all alternatives and captive riders (Type 3)

Each of the three models calibrated earlier are used for validation purpose for each data set separately. This resulted in 9 test categories. The null hypothesis that is formulated states that there exists no significant differences between the observed behavior and the predicted behavior for mode-choice for validation. Likelihood Ratio Test Statistics (LRTS) is used in order to test this hypothesis. The results are reported in Table 5.8. The

Table 5.8: Validation Test Results for Calibrated Models

Model Type	Validation Data Type	Log-Likelihood Ratio Values UnRestricted	Restricted	LRTS Values	Number of Parameters	Critical Chi-Square	Remarks
Model 1	Type: 1	-251.243	-257.341	12.1959	9	16.92	Can't Reject
	Type: 2	-290.667	-299.905	18.4746	9	16.92	Reject Hyp
	Type: 3	-353.996	-366.014	24.0356	9	16.92	Reject Hyp
Model 2	Type: 1	-251.243	-264.779	27.0724	9	16.92	Reject Hyp
	Type: 2	-290.667	-297.807	14.2798	9	16.92	Can't Reject
	Type: 3	-353.996	-359.832	11.6719	9	16.92	Can't Reject
Model 3	Type: 1	-251.243	-270.763	39.0396	9	16.92	Reject Hyp
	Type: 2	-290.667	-302.247	23.1587	9	16.92	Reject Hyp
	Type: 3	-353.996	-363.522	19.0513	9	16.92	Reject Hyp

cases where the LRTS values are less than the critical Chi-square values at degrees of freedom equal to the number of model parameters at 95 percent confidence level indicate that there is no significant difference between the observed behavior and the predicted behavior i.e., the null hypothesis cannot be rejected for that particular model and vice-versa.

It can be seen from Table 5.8 that, when Model 1 is used for predicting the behavior of Type 1 data, the LRTS test revealed that there exists no significant difference between the observed and the predicted behavior for mode-choice. Therefore the model is valid when it is used with the Type 1 data. When this model is used for validation for datasets of Type 2 and Type 3 the LRTS test confirmed the rejection of the null hypothesis of "No difference" between the predicted and observed behavior. Hence, it can be inferred from this test that Model 2 best replicates the behavior of choice riders with valid alternatives, i.e., it is valid for the kind of data it was calibrated with. However, it is likely to result in inaccurate results when used for prediction on data of choice riders with all alternatives and captive riders.

When Model 2 is used for predicting the datasets of Type 2 and Type 3, the LRTS values obtained are smaller than the critical Chi-square values. This indicates that there exists no difference in the prediction of the observed and predicted behavior of the riders by the model. When this model is used for data of Type 1 i.e., choice riders, it rejected the null hypothesis of no significant difference and the LRTS value obtained is much higher than the critical Chi-Square value. Thus, Model 1 failed in predicting the behavior of the choice riders in their mode-choice decision taking stage. It can be inferred from this test that Model 2 is likely to result in true predictions when used for data of its own kind (Type 2) and captives (Type 3).

Finally, Model 3, i.e., model calibrated using the data comprising of captive riders and choice riders with all alternatives as choice is used for validation of all the three data types viz., Type 1, 2, and 3 by restricting the coefficients and calibrating the models with the respective datasets. It is found that the LRTS values obtained for all the cases were greater than the critical Chi-Square values, thus rejecting the null hypothesis of "No Difference". This model failed to validate its own data of choice riders plus captive riders. It can be concluded from this test that this model is not a valid model and it is likely to result in inaccurate predictions when used for forecasting.

5.4.6 Prediction Capability of the Models

A computer program, written in SAS software, is used to estimate the modal share for non-business trip purpose for the three models. The program calculates the utility of each tripmaker, then the probability of using each alternative is estimated. The alternative which has the highest probability is predicted to be the chosen mode for that particular individual. The number of tripmakers correctly predicted is summed up for each alternative and compared with the actual share to yield the prediction ratios.

First of all, Model 1, is used to compute the prediction ratios for calibration datasets of all the three types of datasets and then for validation datasets of three types. The procedure is repeated for Model 2 and Model 3. The results are shown in Tables 5.9 , 5.10 and 5.11. It can be noticed from the tables that the models yielded prediction ratios of around 0.66-0.72 for the three models which indicates that the models are capable of predicting 66-72% of the choices of the tripmakers correctly. The numerical figures do not indicate considerable changes in model's prediction capability, however this overall prediction capability is not a very good indicator of the models performance as explained elsewhere (25).

5.5 COMPARISON OF ELASTICITIES OF THE CALIBRATED MODELS

Elasticities are obtained as optional output from the BLOGIT software. In this section, we are interested in finding the change in elasticities of In-Vehicle Travel Time and Total Travel Cost variables when captive riders are included in the dataset.

The results are tabulated in Table 5.12. It can be observed from the table that in the case of In-Vehicle Travel Time variable, (INVTT) the model calibrated with choice riders data only has high direct elasticity values for all the three modes whereas, the inclusion of captive riders resulted in decreased values of the elasticities for all the three modes. This shows that when Model 3 is compared to Model 1 the percent decrease in share of that alternative is more when there is an increase of 1% travel time. Similarly there is a decrease in the direct elasticity values for all the three alternatives between Model 1 and Model 2. In the case of cross elasticities,

Table 5.9: Prediction Success table for Mode-Choice Model 1

Data Type	Criteria	Calibration Data			Validation Data		
		Alternatives			Alternatives		
		Air	Bus	Auto	Air	Bus	Auto
Type: 1	No. of Tripmakers choosing this alternative	323	122	368	146	63	177
	No. of Tripmakers correctly predicted by the model	242	57	286	96	28	140
	Prediction Ratio	0.75	0.46	0.78	0.66	0.44	0.79
	Overall Prediction Ratio		0.72			0.68	
Type: 2	No. of Tripmakers choosing this alternative	323	122	368	146	63	177
	No. of Tripmakers correctly predicted by the model	227	51	281	94	27	134
	Prediction Ratio	0.70	0.42	0.76	0.64	0.43	0.76
	Overall Prediction Ratio		0.69			0.66	
Type: 3	No. of Tripmakers choosing this alternative	384	129	476	176	69	232
	No. of Tripmakers correctly predicted by the model	277	54	367	114	30	175
	Prediction Ratio	0.72	0.42	0.77	0.65	0.43	0.75
	Overall Prediction Ratio		0.71			0.67	

Table 5.10: Prediction Success table for Mode-Choice Model 2

Data Type	Criteria	Calibration Data Alternatives			Validation Data Alternatives		
		Air	Bus	Auto	Air	Bus	Auto
Type: 1	No. of Tripmakers choosing this alternative	323	122	368	146	63	177
	No. of Tripmakers correctly predicted by the model	244	51	291	96	25	141
	Prediction Ratio	0.76	0.42	0.79	0.66	0.40	0.79
	Overall Prediction Ratio		0.72			0.68	
Type: 2	No. of Tripmakers choosing this alternative	323	122	368	146	63	177
	No. of Tripmakers correctly predicted by the model	234	45	287	95	23	136
	Prediction Ratio	0.72	0.36	0.78	0.65	0.36	0.77
	Overall Prediction Ratio		0.69			0.66	
Type: 3	No. of Tripmakers choosing this alternative	384	129	476	176	69	232
	No. of Tripmakers correctly predicted by the model	285	48	374	114	25	179
	Prediction Ratio	0.74	0.37	0.79	0.65	0.36	0.77
	Overall Prediction Ratio		0.71			0.67	

Table 5.11: Prediction Success table for Mode-Choice Model 3

Data Type	Criteria	Calibration Data			Validation Data		
		Alternatives			Alternatives		
		Air	Bus	Auto	Air	Bus	Auto
Type: 1	No. of Tripmakers choosing this alternative	323	122	368	146	63	177
	No. of Tripmakers correctly predicted by the model	242	47	302	94	22	145
	Prediction Ratio	0.75	0.39	0.82	0.64	0.35	0.82
	Overall Prediction Ratio		0.73			0.68	
Type: 2	No. of Tripmakers choosing this alternative	323	122	368	146	63	177
	No. of Tripmakers correctly predicted by the model	236	41	299	93	20	140
	Prediction Ratio	0.73	0.34	0.81	0.64	0.32	0.79
	Overall Prediction Ratio		0.71			0.66	
Type: 3	No. of Tripmakers choosing this alternative	384	129	476	176	69	232
	No. of Tripmakers correctly predicted by the model	288	43	390	112	22	183
	Prediction Ratio	0.75	0.33	0.82	0.64	0.32	0.79
	Overall Prediction Ratio		0.73			0.66	

Table 5.12: Comparison of the elasticities of the Travel Time and Travel Cost Variables of the calibrated Models

Variable	Elasticity	Alternatives (Model 1/Model 2/Model 3)		
		Air	Bus	Car
Travel-Time	Direct	-.069/-.071/-.078	-.611/-.704/-.845	-.244/-.251/-.259
	Cross	0.145/0.134/0.157	0.516/0.377/0.394	0.461/0.428/0.513
Travel-Cost	Direct	-.355/-.369/-.404	-.259/-.309/-.368	-.069/-.072/-.075
	Cross	0.531/0.499/0.569	0.172/0.125/0.131	0.180/0.171/0.211

there is an increase in the elasticity values for air and car alternatives when data containing captive riders is also used in calibrating the model as compared to the model calibrated by choice riders data only. This shows that the percent change in the share of air and car modes is more when Model 3 is used as compared to Model 1 when there is a one percent change in an attribute of another value. In case of bus mode there was a decrease in the cross elasticity value.

In the case of the variable Total Travel Cost (TOTCOST), a similar phenomenon is observed. The direct elasticities of all the three alternatives decreased from Model 1 to Model 3 but the cross elasticity values increased when they are computed using models calibrated with the data of choice and captive riders.

Thus in brief, the models calibrated with captive riders yielded elasticity values different than those that obtained from the models calibrated with choice riders data only. This indicates that, if captive riders are included during model calibration stages, the developed model is likely to give inaccurate results when policy changes are to be tested.

5.6 *IMPORTANT FINDINGS OF THIS RESEARCH*

Some of the noteworthy findings of this research can be summarized as follows:

1) Models calibrated with choice plus captive riders data although increased the fit of the individual parameter estimates, but decreased the overall fit of the models.

2) The parameter estimates of each individual variable from three different models were found to be statistically same.

3) The model calibrated on choice riders data only was used for validation of the data for choice riders only. The hypothesis of "No Difference" between the parameter vectors obtained from calibration data and the validation data could not be rejected. When the same model was used to validate the data of choice plus captive riders it rejected the hypothesis of "No Difference".

4) When the model calibrated on the data of choice plus captive riders was used for validation purposes, it failed to validate its own data type.

Thus, models calibrated on data of choice riders seem to be valid model whereas, models calibrated on data of choice plus captive riders is not a valid model.

CHAPTER 6

SUMMARY AND CONCLUSIONS

6.1 SUMMARY

The main objective of this research was to investigate the effect of including captive riders (i.e., those riders who do not have choice for travel mode) in intercity mode-choice models. Reviewing the past work done by various researchers in intercity mode-choice model building, it was found that the inclusion of these riders in mode-choice models results in incorrect predictions. In this study the effect of their inclusion was investigated.

To achieve the objective, mode choice models were calibrated first with choice riders and then captive riders were included in the models. Later these models were used for prediction. The data was collected in February 1993 as part of a research project by Al-Ahmadi et.al.(4) for the development of a National Intercity Modal Split model for Saudi Arabia. Analyses were performed using the Statistical Analyses Software (SAS) available at the King Fahd University's mainframe computing facility at Dhahran.

Initially, choice riders were separated from captive riders and a general analyses of the data were performed. First, various travel and socioeconomic characteristics of choice and captive riders were analyzed with the help of various tables. Second, a statistical analysis using Chi-Square test was employed for determining the socioeconomic and travel characteristics that significantly affect the mode-choice decision.

In the next stage of the present study, mode-choice models were calibrated for non-business trips. The calibration of the models was done using the BLOGIT software. The calibration process was carried out in three stages. In the first stage focus was on the development of mode-choice models for choice riders with valid alternatives as choice. The results obtained from the preliminary analyses of the intercity travellers were used for selecting the best potential travel and socioeconomic variables which explains the mode-choice behavior of the traveller. By using different combinations of variables, different models were obtained. Those model specification which yielded significant coefficient estimates (t-statistics value greater than 95 percent confidence level) for all the variables and high rho-square value was selected as the best model. This model was termed as Model 1. Utility functions for each mode were established for this combination of variables.

In the second phase, by assuming that the choice riders have all the available alternatives as choices, Model 2 was obtained. The specification that was obtained earlier was used and Model 2 was calibrated. Utility functions for each mode were then formulated based on this assumption.

In the final phase of model calibration, captive riders were also included in the data that was used in the second phase. Keeping the same model specification BLOGIT program was run to obtain Model 3. Table 6.1 shows the three models that were calibrated for non-business trips using the three different datasets.

The inclusion of captive riders in intercity mode-choice models yielded almost same coefficient estimates, it increased the significance of the variables but resulted in an overall decrease in the rho-square value of the model. The rho-square value dropped from 0.393 to 0.310 when captive riders were included in the model calibration.

The next stage in the present study involved validation of the models. One third of the dataset that was reserved for validation purposes was used. The three models that were calibrated were separately used for validating the following three types of datasets.

Type 1: Validation data for choice riders with valid alternatives

Type 2: Validation data for choice riders with all alternatives

Type 3: Validation dataset for captive riders and choice riders with all alternatives

The null hypothesis which states that there exists no difference between the observed and predicted behavior by the mode-choice model for validation data, was tested using the Likelihood Ratio Test Statistics (LRTS). The validation test results are summarized in Table 6.2. It was

Table 6.1: Comparative Parameter estimates of the Models

Variable Name	Estimated Coefficients (t-stat)		
	Model 1 Choice Riders with Valid Alternatives Data	Model 2 Choice Riders with All Alternatives Data	Model 3 Choice Riders with All Alternatives and Captive Riders Data
ASC-AIR	1.67/3.56	1.67/3.68	1.61/3.92
ASC-BUS	2.09/6.17	1.56/4.71	1.53/5.03
INVTT	-.15/-4.63	-.14/-4.60	-.16/-5.59
TOTCOST	-.002/-5.14	-.002/-5.31	-.0022/-5.31
DNCARB	1.98/6.29	2.13/7.92	2.04/8.09
DFMLYC	1.33/5.98	1.36/6.43	1.34/6.78
DNATIONB	1.11/3.74	1.48/5.52	1.34/5.38
HHINCA	0.18/3.75	0.21/4.39	0.21/4.87
DISTC	2.18/7.08	2.30/7.68	2.35/8.67
No. of obs:	813	813	995
LL(β)	-497.464	-581.672	-974.099
LL(O)	-821.124	-821.124	-821.124
Rho sq. ρ^2	.393	.391	.311
Rho-bar sq.	.390	.387	.308

Table 6.2: Validation Test Results

	Type 1	Type 2	Type 3
Model 1	Can't Reject	Reject Hyp	Reject Hyp
Model 2	Reject Hyp	Can't Reject	Can't Reject
Model 3	Reject Hyp	Reject Hyp	Reject Hyp

(Null Hypothesis tested is as follows:

Hyp: There is no difference between the observed and the predicted behavior by the model for that corresponding data Type) found that Model 1 (calibrated on data of choice riders with valid alternatives) gave satisfactory results and did not reject the null hypothesis for its same type of validation data but rejected the hypothesis for the validation data of choice plus captive riders. Model 2 (calibrated on choice riders with all alternatives) rejected the null hypothesis for Type 1 data but accepted it for Type 2 and Type 3 data. Model 3 when used for validation purposes resulted in rejection of the null hypothesis for all the three types of validation data, including its own type. Therefore, Model 3 was the worst as expected.

6.2 CONCLUSIONS

This research provided a better understanding of the effect of inclusion of captive riders with choice riders in mode choice modelling. Based on the findings of this study the following conclusions are drawn:

1) General analyses of the intercity traveller's data revealed that captivity to air and car modes is more pronounced as compared to bus mode.

2) Among the various travel and socioeconomic determinants, purpose of trip, trip length, age, occupation, nationality, accompanied family size and educational level plays a significant role in influencing traveller's decision for mode-choice.

3) Captive riders have to be removed from calibration data while building mode-choice models because their inclusion yielded models with lower overall fit, and with questionable validity.

4) The models calibrated with choice riders data only seem to explain the behavior of mode-choice of the travellers very well and they were shown to be valid.

5) Inclusion of captive riders yielded the worst model and its validity was rejected even with the type of data it was calibrated. So, this result indicates the need for the exclusion of captives from the model calibration.

6.3 *RECOMMENDATIONS FOR FURTHER RESEARCH*

Following recommendations are proposed for further research.

1) Investigations can be made by identifying the fundamentals of captivity i.e., as to why and how people become captive to a particular mode and how they can be modelled. This can be done by building a discriminant, probit or logit models.

2) The way the captive riders affect the property of Independence of Irrelevant Alternatives (IIA) can be investigated.

REFERENCES

1. Thamizarasan, V. and Rengaraju, V.R., "A Methodology of Approach for Intercity Travel Demand Modelling," Indian Highways, Dec. 1986, pp. 23-33.

2. Gupta, J., "Intercity Commuter Demand Model," Indian Highways, Dec. 1983, pp. 79-91.

3. Stopher, P.R., "Captivity and Choice in travel behavior models," Transportation Engineering Journal, ASCE, July 1980, pp. 427-435.

4. Al-Ahmadi, H.M., Ergun, G., Al-Senan, S.H. and Ratrout, N.T., "Development of National Intercity Modal Split Model for Saudi Arabia," Final Report to Research Financed By King Abdul Aziz City for Science and Technology," Department of Civil Engineering, Transportation Division, King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia, September 1993.

5. Adams, W.T., "Factors Influencing Mass Transit and Automobile Travel in Urban Areas," Public Roads, 1959, Vol. 30, pp. 256-260.

6. Warner, S.L., Stochastic Choice of Mode in Urban Travel: A Study in Binary Choice, North Western University Press, Evanston, Illinois, U.S.A. 1962.

7. Quarmby, D.A., "Choice of Travel Mode for Journey to Work: Some Findings," Journal of Transport Economics and Policy, March 1967,

Vol. 1, pp. 273-314.

8. Lisco, T.E., "The Value of Commuter's Travel Time: A Study in Urban Transportation," Ph.D. Dissertation, Department of Economics, University of Chicago, Chicago, U.S.A. 1967.

9. Stopher, P.R., "A Probability Model of Mode-Choice for the Work Journey," *Highway Research Record No. 283*, 1969, pp. 57-65.

10. Stopher, P.R. and Meyburg, A.H., *Urban Transportation Planning and Modelling*, Lexington Books, Lexington, U.S.A. 1975.

11. Koppelman, F.S., Kuah, G.H. and Hirsh, M., *"Review of Intercity Passenger Travel Demand Modelling: Mid-60's to Mid-80's,"* Draft Report, Department of Civil Engineering, North Western University, Evanston, Illinois, U.S.A., September 1984.

12. Leake, G.R. and Underwood, J.R., "Entropy Model and Intercity Modal Split in UK," *Transportation Engineering Journal, ASCE*, May 1976, Vol. 5, pp. 215-228.

13. Lyles, R.W., "Modal Split Model Considering Carpool Mode," *Transportation Engineering Journal, ASCE*, March 1979, Vol. 3, pp. 149-163.

14. Stopher, P.R., "Work-trip mode-choice models for South Africa," *Transportation Engineering Journal, ASCE*, November 1976, Vol. 11, pp. 595-608.

15. Ackoff, R.L., "Individual Preferences For Various Means Of Transpor-

tation," _Management Science Center_, Transportation Studies Center, University of Pennsylvania, Pennsylvania, U.S.A., May 1965.

16. Anas, A., "Discrete Choice Theory, Information Theory and The Multinomial Logit and Gravity Models," _Transportation Research_, January 1984, Vol. 17B, No. 1, pp. 13-23.

17. Demetsky, M.J., "Behavioral Approaches To Modal Demand Modelling," _Report No_: _CMUTRI-TP-72-70_, _Transportation Research Institute_, Carnegie Mellon University, Pittsburg, Pennsylvania, U.S.A., May 1972.

18. Parashker, J.N., "Mode Choice Models With Perceived Reliability Measures," _Transportation Engineering Journal_, _ASCE_, May 1979, Vol. 11, pp. 251-262.

19. Ortuzar, J.D. and Fernandez. J.E., "On the Stability of Discrete Choice Models in Different Environments," _Transportation Planning and Technology_, October 1985, Vol. 10, pp. 209-218.

20. Al-Ahmadi, H.M., "Transferablity of Mode-Choice Models Within Saudi Arabia," Ph.D. Dissertation, Department of Civil and Transportation Engineering. Michigan State University. Lansing, Michigan 1989.

21. Al-Sughaiyer, M.A., "Effect of Perceptual Concepts in Intercity Mode-Choice Models of Saudi Arabia," Ph.D. Dissertation, Department of Civil Engineering, Transportation Division, King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia, September 1994.

22. Basmaciyan, H. and Schmidth, J.W., "Developemnt of a Modal Split Model for the Puget Sound Region," PSRTS Staff Report: No. 12, May 1958.

23. Smith, W. and Associates, "Transportation and Parking for Tomorrow's Cities," Automobile Manufacturers Association, Detroit, Michigan 1966.

24. Tol, A.V.D., Hutchinson, B.G. and Shortreed, J.H., "Two Stage Modal Split Models," Paper Presented to American Society of Civil Engineers, New York, 1973.

25. Morall, J.F. and Morash, L.H., "Transit Planning Models for Calgary," Paper Presented at Canadian Transport Research Forum, 1973.

26. Ferreri, M.J. and Cherwony, W., "Choice and Captive Modal Split Models," Highway Research Record, No. 369, 1971, pp. 80-90.

27. Ogunjumo, E.A., "Captive Transit Ridership," Transportation Planning, May 1984, pp. 18-23.

28. Morikawa, Takayuki, Ben-Akiva, M. and Yamada, K., "Forecasting Intercity Rail Ridership Using Revealed Preference and Stated Preference," Transportation Research Record, No. 1328, 1991, pp. 30-35.

29. Talvitie, A.P. and Koskenoja, M., "Socioeconomics of the Individual and Costs of Driving: New Evidence for Travel Demand Modeler," Transportation Research Record, No. 1328, 1991, pp. 58-68.

30. Ergun, G. and Al-Ahmadi, H.M., "Characteristics of Intercity Captive Riders of Saudi Arabia," Unpublished Research Paper, Department of Civil Engineering, Transportation Division, King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia, January 1994.

31. Ben Akiva, M. and Lerman, S.R., *Discrete Choice Analysis*, MIT press, Cambridge, Massachusetts 1985.

32. Crittle, F.J. and Johnson, F.W., *Basic Logit(BLOGIT)*, Technical Manual, Australian Road Research Board, October 1989.

33. Transportation Research Board, "Application of Disaggregate Travel Demand Models," *TRB Special Report*, No. 253, 1982.

APPENDIX A
MODEL CALIBRATION PROCEDURES

In this appendix the concept of utility theory that was utilized for model calibration is explained in detail. Apart from this the various available model calibration procedures is also highlighted and the usage of MultiNomial Logit analyses for the present study is justified.

1. RUDIMENTS OF UTILITY THEORY

The basic approach in this theory is that an individual selects an alternative out of the available alternatives such that his utility is maximized (10,29,30). Since the behavior of the traveller cannot be predicted with 100% certainty , the probability of trip maker's choice is estimated. Utility function is partitioned into two components: a representative component V and a random component ε to take into account the utility of unobserved factors.

The utility function is given as

$$U_{ki} = V_{ki} + \varepsilon_{ki}$$

where,

U_{ki} = utility function of mode k to trip maker i

V_{ki} = systematic component of utility function of mode k to trip maker i

ε_{ki} = random component of utility function of mode k to trip maker i

The systematic component of utility function for mode choice can be stated mathematically as follows:

$$V_{ki} = \sum_{n=1}^{N} \beta_n X_{kin} + \sum_{l=1}^{L} \alpha_l S_{il}$$

Where,

V_{ki} = the utility of mode *k* to trip maker *i*

X_{kin} = value of attribute *n* of mode *k* for trip maker *i*

S_{il} = the socioeconomic characteristic *l* of trip maker *i*

β_n = coefficient of attribute *n* and

α_l = Coefficient of socioeconomic characteristic *l*;

Using this principle, individual will rank his alternative modes to their utility function value. The one with the highest value of utility will be chosen by him.

2. *MODEL CALIBRATION PROCEDURES*

The important mathematical calibration methods available to develop disaggregate mode-choice models are as follows (10):

a. Discriminant analysis

b. Probit analysis

c. Logit analysis

a. *Discriminant Analysis*

It is one of the earliest nonlinear techniques used for building behavioral travel demand models. It is based on the observation that their exists in a population two or more distinct subgroups that can be distinguished by means of a discriminating function (10,25):

$$D = d_1 z_1 + d_2 z_2 + \text{-----------} + d_p z_p + K$$

where,

D is a non dimensional 'discriminant score'

d^s are weighting coefficients

z^s are standardized values of the variables considered

K - being a constant

The discriminant function is not useful for the mode-choice model because the goal of fitting procedure in this is to maximize the separation between the subgroups and furthermore it is used for building mode-choice models for two alternatives at a time and in the present situation their exists more than two modes for all corridors in the Kingdom.

In the present study this technique was not used in building models.

b. Probit Analysis

In this type of analysis it is assumed that members of a population are subjected to a stimulus that generates a normally distributed response pattern over the range of values of the stimulus(10). If the stimulus is the related utility of one alternative over other, the choice response of one alternative is assumed to be normally distributed over the range of relative utility. The ordinate of the cumulative choice response distribution curve yields the probability of an individual having responded by the stimulus. It is given in the model form as

$$P_k^i = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{Y_k^i} \exp(-1/2t^2)$$

where,

P_k^i = the probability that a customer i chooses alternative k out of the available alternatives

Y^i_k = the stimulus expressed as a linear function of the relative attributes of alternative k

t = dummy parameter for normal distribution

In this type of analysis the individual makes up his mind and is not influenced by the decision of other members of the population. The probit analysis procedure is a detailed one requiring probabilities for population members and then calculate the utility value and then the integral of the stimulus curve.

In the present study probit analysis was not used for building models because of the complexity involved in it.

c. *Logit Analysis*

It is the most commonly used technique for building behavioral travel demand models. Logit analysis is considered to be a good representation of the traveller's behavior in making a mode selection for his trip. Based on the availability of alternatives and type of analysis performed they are segregated into following categories (10):

- Binary logit or binomial logit analysis
- Nested logit analysis
- Multinomial logit analysis

Binomial Logit Analysis

In this the individual has the choice between two available modes, so he will evaluate the attributes of these modes simultaneously and selects the one which has the maximum utility. Mathematically the model is represented as (5):

$$P_k^i = \frac{Exp(G^i X_{(jk)})}{1 + Exp(G^i(X_{jk}))}$$

where,

X_{jk} = Relative measures of attributes of alternative *k* against alternative *j*

Nested Logit Analysis

This type of analysis assumes that their exists a certain sequence of hierarchy in selecting a particular mode out of a number of available modes. It is used when individual evaluates one composite alternative consisting of two or more modes against another alternative. For ex: a traveller decides to travel he will evaluate ground modes as one composite alternative consisting of modes available to him on ground (viz., car, bus, train etc.) against air alternative. If he decides to choose the ground modes he then evaluates the attributes of ground modes to choose the one that maximizes his utility. Thus he hierarchically comes to a decision on which mode to use for his travel.

Multinomial Logit Analysis

When more than two alternatives are available to a trip maker then multinomial logit analysis is used and the models developed by this technique is referred to as MNL models. The general form used for model building is in a probabilistic form and it is written as

$$P_{ki} = \frac{Exp(V_{ki})}{\sum_{i=1}^{n} Exp(V_{ji})}$$

where,

P_k^i = the probability of choosing mode k out of the available alternatives and

V_{ki} = the systematic component of utility function of mode k to trip maker i

Binary logit model was excluded from the analysis since such models compare only two modes while their exists more than two modes in the intercity travel environment in the kingdom along all the major corridors. On the other hand nested logit models were also excluded from the analysis since these models complicate the purpose of this research. In addition, many researchers preferred the simultaneous models rather than sequenced or nested models since no unique natural sequence exists when trip maker takes his decision.

Hence, the most suitable type of analysis for evaluating the present conditions in the Kingdom is the MultiNomial Logit (MNL) analysis because, there exists more than two modes as choices for an intercity traveller. Thus, MNL analysis technique was made use of in building mode-choice models.

APPENDIX B
AIR TRAVELLER'S QUESTIONNAIRE
(ENGLISH VERSION)

بسم الله الرحمن الرحيم

Ministry of Higher Education
King Fahd University of Petroleum & Minerals
COLLEGE OF ENGINEERING SCIENCES
CIVIL ENGINEERING DEPARTMENT

وزارة التعليم العالي
جامعة الملك فهد للبترول و المعادن
كلية العلوم الهندسية
قسم الهندسة المدنية

AIR TRIPMAKER'S QUESTIONNAIRE

Dear Traveller,

The civil engineering department of the King Fahd University of Petroleum and Minerals is conducting a research to understand the travel needs between the cities of the Kingdom of Saudi Arabia. This research is funded by the *King Abdulaziz City for Science and Technology*.

You are asked not to put your name on the questionnaire and there is no way to link your answers to your name. It takes about ten minutes to complete.

We greatly appreciate your cooperation. If you have any questions, you may call Dr. Hasan M. Al-Ahmadi at 860-4269 during weekdays.

Sincerely,

Dr. Alfarabi Sharif
Chairman, Civil Engineering Department

Dr. Hasan M. Al-Ahmadi
Principal Investigator

PART I. **THE FOLLOWING QUESTIONS ARE ABOUT THIS TRIP.**

1. **Where did you begin this trip?** City ________________
2. **Where is your destination?** City ________________
3. **Where is your place of residence?** City ________________

4. **What is the purpose of your trip? Check only one answer:**

[ ] Work [ ] Educational/Study
[ ] Personal business [ ] Aumra
[ ] Social/Recreation [ ] Other (specify)__________

5. **How long did you stay or are you planning to stay away from your home? Check only one answer:**

[ ] one -3 days [ ] 4-7 days [ ] 8-30 days [ ] more than 30 days

6. **Please give your travel time in the following categories:**

Time to get to the airport ______ hours _____ minutes
Time spent waiting at the airport ______ hours _____ minutes
Travel time spent in plane ______ hours _____ minutes
Estimated time from airport to final destination ______ hours _____ minutes

7. **How many people from your family are travelling with you? (Please do not count yourself)**

______ Adults (full ticket) ______ Adults (half ticket) ______ Children

8. **Please give your travel costs in the following categories including family members if they are travelling with you (only fill in those that are appropriate):**

Please indicate if costs are for: [ ] round-trip [ ] one way

Ticket (fare) SR.__________
Limousine or Taxi (if used) SR.__________
Other (specify) __________ SR.__________

9. **Who paid for this trip?**

[ ] Yourself [ ] The Goverment or your company [ ] Others (specify) ----------

10. **Indicate if the plane was on time or late?**

[ ] on time [ ] late, How late? ______ hours ______ minutes

PART II. THE FOLLOWING QUESTIONS RELATE TO OTHER MEANS OF TRAVEL.

1. Will you consider using the following means of travel for a trip similar to the one you are making now?

	Will never use it	May consider using it	Definitely consider using it
BUS	[ ]	[ ]	[ ]
TRAIN	[ ]	[ ]	[ ]
AUTO	[ ]	[ ]	[ ]

2. If you are to choose a means of travel other than plane for this trip, what would be your estimates of time for the listed categories?

TIME CATEGORY	BUS	TRAIN	AUTO
Time to get to bus (or train) station in minutes.			
Time spent waiting at bus (or train) station or rest areas in minutes.			
Total travel time spent in the vehicle in hours.			
Time from bus (or train) station to the final destination (your estimate) in minutes.			

3. If you are to choose a means of travel other than plane for this trip, what would be the cost for each category in Saudi Riyals?

COST CATEGORY	BUS	TRAIN	AUTO
Ticket cost (one way).			
Taxi (or limousine) cost from origin to bus or train station.			
Taxi (or limousine) cost from bus or train station to your destination.			
Total cost (one way).			

PART III. THE QUESTIONS BELOW ARE FOR STATISTICAL PURPOSES ONLY. THEY WILL BE CONFIDENTIAL AND NO INDIVIDUAL WILL BE IDENTIFIED IN THE RESEARCH

1. What is your age? ___________

2. Are you married? [ ] Yes [ ] No

3. How many cars are there in your household? ________

4. Do you have a driver's license? [ ] Yes [ ] No

4. How many persons in your household have a driver's license? _____

5. What is your occupation?

   [ ] Student
   [ ] Employee
   [ ] Businessman
   [ ] Unemployed
   [ ] Retired
   [ ] Others (specify) ______________

6. What is your nationality? [ ] Saudi [ ] non Saudi, Pls. specify ______________

7. Indicate your level of education

   [ ] Elementary
   [ ] Intermediate or Secondary
   [ ] University degree

8. What is your personal and household monthly income? Tick one in each category.

Personal Income	Household Income
[ ] less than 1,000 SR.	[ ] less than 1000 SR.
[ ] 1,000-2,500 SR.	[ ] 1,000-2,500 SR.
[ ] 2,501-5,000 SR.	[ ] 2,501-5,000 SR.
[ ] 5,001-7,500 SR.	[ ] 5,001-7,500 SR.
[ ] 7,501-10,000 SR.	[ ] 7,501-10,000 SR.
[ ] 10,001-12,500 SR.	[ ] 10,001-12,500 SR.
[ ] 12,501- 15,000 SR.	[ ] 12,501 - 15,000 SR.
[ ] more than 15,000 SR.	[ ] more than 15,000 SR.

PART IV. EVALUATION OF VARIOUS MEANS OF TRAVEL WITH RESPECT TO THEIR CHARACTERISTICS.

In the following page you will be asked to evaluate alternative means of travel for the trip you are making now. This evaluation will be made by selecting the most appropriate answers for certain statements.

Let us say, for example, that you are evaluating different means of travel with respect to speed and safety. A possible response to this for bus, auto, air and train is given below. Note that the responses are indicated by circling numbers, where "1" indicates the worst and "5" indicates the best and other numbers indicate judgement in between.

Please evaluate the following means of travel (for this trip) with respect to...

MEANS OF TRAVEL	Speed	Safety
	WORST ... BEST	WORST ... BEST
Bus	1 (2) 3 4 5	1 2 (3) 4 5
Auto	1 2 (3) 4 5	1 (2) 3 4 5
Air	1 2 3 4 (5)	1 2 3 (4) 5

Now please go to the next page and evaluate the indicated means of travel with respect to ALL the listed characteristics.

PLEASE TRY TO RESPOND FOR EACH STATEMENT.

PART V. IN SELECTING THE MEANS OF TRAVEL FOR THIS TRIP, HOW IMPORTANT TO YOU WERE THE FOLLOWING CHARACTERISTICS. WE HAVE PROVIDED A SCALE WHICH RANGES FROM "EXTREMELY IMPORTANT" TO "OF NO IMPORTANCE". FOR EACH CHARACTERISTIC PLEASE ENCIRCLE THE NUMBER WHICH YOU FEEL BEST INDICATES THE IMPORTANCE.

	Of no importance				Extremely important
Having Some Privacy	1	2	3	4	5
Freedom in Choosing Locations and Time of Stops	1	2	3	4	5
The Effect of Weather Condition on the Trip	1	2	3	4	5
Cleanliness of Vehicle	1	2	3	4	5
Feeling of Independence	1	2	3	4	5
The Feeling that the Vehicle would not be Delayed for Repair	1	2	3	4	5
Comfort	1	2	3	4	5
The Effect of Weather Condition on Travel Time	1	2	3	4	5
The Need for Advance Planning	1	2	3	4	5
Talking Freely with Family Members or Friends	1	2	3	4	5
Accessibility from Home	1	2	3	4	5
Feeling Tired at the End of the Trip	1	2	3	4	5

THANK YOU FOR YOUR COOPERATION

Please return the questionnaire to the person who distributed it

Please evaluate the following means of travel (for this trip) with respect to ...

MEANS OF TRAVEL	1. Having Some Privacy.	2. Freedom in choosing location and time of stops.
	WORST — BEST	WORST — BEST
Bus	1 2 3 4 5	1 2 3 4 5
Auto	1 2 3 4 5	1 2 3 4 5
Air	1 2 3 4 5	1 2 3 4 5
	3. The effect of weather condition on the trip.	4. Cleanliness of vehicle.
	WORST — BEST	WORST — BEST
Bus	1 2 3 4 5	1 2 3 4 5
Auto	1 2 3 4 5	1 2 3 4 5
Air	1 2 3 4 5	1 2 3 4 5
	5. Feeling of independence.	6. Feeling that vehicle would not be delayed for repair.
	WORST — BEST	WORST — BEST
Bus	1 2 3 4 5	1 2 3 4 5
Auto	1 2 3 4 5	1 2 3 4 5
Air	1 2 3 4 5	1 2 3 4 5
	7. Comfort.	8. The effect of weather condition on travel time.
	WORST — BEST	WORST — BEST
Bus	1 2 3 4 5	1 2 3 4 5
Auto	1 2 3 4 5	1 2 3 4 5
Air	1 2 3 4 5	1 2 3 4 5
	9. The need for advance planning.	10. Talking freely with family members or friends.
	WORST — BEST	WORST — BEST
Bus	1 2 3 4 5	1 2 3 4 5
Auto	1 2 3 4 5	1 2 3 4 5
Air	1 2 3 4 5	1 2 3 4 5
	11. Accessibility from home.	12. Feeling tired at the end of trip.
	WORST — BEST	WORST — BEST
Bus	1 2 3 4 5	1 2 3 4 5
Auto	1 2 3 4 5	1 2 3 4 5
Air	1 2 3 4 5	1 2 3 4 5

APPENDIX C
TABLES
(FREQUENCY & CROSS FREQUENCY)

Frequency Table for Different Trip Lengths

Trip Length (km)	Frequency	Percent
< 400	1569	47.1
400 - 600	416	12.5
600 - 800	205	6.1
800 - 1000	677	20.3
1000 - 1200	81	2.4
1200 - 1400	346	10.4
> 1400	43	1.3

Frequency Table for Various Age Groups

Age (Years)	Frequency	Percent
< 20	197	5.9
20 - 29	1325	39.7
30 - 39	1034	31.0
40 - 49	528	15.8
> 50	253	7.6

Frequency Table for Payment of Trip

Who Paid for Trip	Frequency	Percent
Self Paid	2752	82.5
Govt/Co. Paid	448	13.4
Others	126	3.8
Missing	10	0.3

Frequency Table for Different Occupations

Occupation	Frequency	Percent
Student	561	16.8
Govt. Employee	1554	46.6
Businessman	452	13.5
Unemployed	73	2.2
Retired	58	1.7
Private Sector	639	19.1

Cross-Frequency Table for Mode by Trip Purpose

Mode	Trip Purpose		
	Social/ Recreation	Others	Total
Air	495.00° 36.97°° 89.35± 43.27±±	59.00 4.41 10.65 30.26	554.00 41.37
Car	379.00° 28.30°° 79.79± 33.13±±	96.00 7.17 20.21 49.23	475.00 35.47
Bus	177.00° 13.22°° 89.85± 15.47±±	20.00 1.49 10.15 10.26	197.00 14.71
Train	93.00° 6.95°° 82.30± 8.13±±	20.00 1.49 17.70 10.26	113.00 8.44
Total	1144.00 85.44	195.00 14.56	1339.00 100.00

Frequency°
Percent°°
Row Percent±
Col. Percent±±

Cross-Frequency Table for Mode by Nationality

Mode	Nationality							
	Saudi	Americans	European/ Austral.	Asian and Far East	Other Arabs	Africans	Unspecf. Non-Saudis	Total
Air	472.00° 35.28°° 85.20± 42.52±±	0.00 0.00 0.00 0.00	1.00 0.07 0.18 20.00	8.00 0.60 1.44 28.57	63.00 4.71 11.37 36.84	3.00 0.22 0.54 50.00	7.00 0.52 1.26 77.78	554.00 41.41
Car	439.00° 32.81°° 92.62± 39.55±±	6.00 0.45 1.27 66.67	3.00 0.22 0.63 60.00	3.00 0.22 0.63 10.71	20.00 1.49 4.22 11.70	1.00 0.07 0.21 16.67	2.00 0.15 0.42 22.22	474.00 35.43
Bus	109.00° 8.15°° 55.33± 9.82±±	2.00 0.15 1.02 22.22	1.00 0.07 0.51 20.00	13.00 0.97 6.60 46.43	70.00 5.23 35.53 40.94	2.00 0.15 1.02 33.33	0.00 0.00 0.00 0.00	197.00 14.72
Train	90.00° 6.73°° 79.65± 8.11±±	1.00 0.07 0.88 11.11	0.00 0.00 0.00 0.00	4.00 0.30 3.54 14.29	18.00 1.35 15.93 10.53	0.00 0.00 0.00 0.00	0.00 0.00 0.00 0.00	113.00 8.44
Total	1110.00 82.96	9.00 0.67	5.00 0.37	28.00 2.09	171.00 12.78	6.00 0.45	9.00 0.67	1339.00 100.00

Frequency°
Percent°°
Row Percent±
Col. Percent±±

Frequency Missing = 1

Cross-Frequency Table for Mode by Marital Status

Mode	Marital Status		
	Married	UnMarried	Total
Air	383.00° 28.60°° 69.13± 46.26±±	171.00 12.77 30.87 33.46	554.00 41.37
Car	265.00° 19.79°° 55.79± 32.00±±	210.00 15.68 44.21 41.10	475.00 35.47
Bus	109.00° 8.14°° 55.33± 13.16±±	88.00 6.57 44.67 17.22	197.00 14.71
Train	71.00° 5.30°° 62.83± 8.57±±	42.00 3.14 37.17 8.22	113.00 8.44
Total	828.00 61.84	511.00 38.16	1339.00 100.00

Frequency°
Percent°°
Row Percent±
Col. Percent±±

Cross-Frequency Table for Mode by Education

Mode	Level of Education			
	Primary level	Inter Mediate	Univ. Level	Total
Air	63.00° 4.71°° 11.37± 48.09±±	245.00 18.30 44.22 41.32	246.00 18.37 44.40 40.00	554.00 41.37
Car	27.00° 2.02°° 5.68± 20.61±±	209.00 15.61 44.00 35.24	239.00 17.85 50.32 38.86	475.00 35.47
Bus	31.00° 2.32°° 15.74± 23.066±	96.00 7.17 48.73 16.19	70.00 5.23 35.53 11.38	197.00 14.71
Train	10.00° 0.75°° 8.85± 7.63±±	43.00 3.21 38.05 7.25	60.00 4.48 53.10 9.76	113.00 8.44
Total	131.00 9.78	593.00 44.29	615.00 45.93	1339.00 100.00

Frequency°
Percent°°
Row Percent±
Col. Percent±±

Cross-Frequency Table for Mode by License Possession

Mode	License Possession		
	Have	Don't Have	Total
Air	536.00° 40.03°° 96.75± 45.27±±	18.00 1.34 3.25 11.61	554.00 41.37
Car	423.00° 31.59°° 89.05± 35.73±±	52.00 3.88 10.95 33.55	475.00 35.47
Bus	135.00° 10.08°° 68.53± 11.40±±	62.00 4.63 31.47 40.00	197.00 14.71
Train	90.00° 6.72°° 79.65± 7.60±±	23.00 1.72 20.35 14.84	113.00 8.44
Total	1184.00 88.42	155.00 11.58	1339.00 100.00

Frequency°
Percent°°
Row Percent±
Col. Percent±±

APPENDIX D
(CURRICULUM VITAE)

VITAE

* Born on 6th of October 1969 in Hyderabad, India.
* Obtained Bachelors Degree in Civil Engineering in 1990 from College of Engineering, Osmania University, Hyderabad, India.
* Obtained Master of Technology Degree in Transportation Engineering in 1992 from Regional Engineering College, Warangal, India.
* Obtained Master of Science Degree in Civil Engineering (Transportation Option) in 1995 from King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia.