INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

NOTE TO USERS

The original manuscript received by UMI contains pages with indistinct print. Pages were microfilmed as received.

This reproduction is the best copy available

UMI

ANALYSIS OF THE TEMPORAL TRANSFERABILITY OF WORK TRIP DISTRIBUTION “GRAVITY MODELS”

by:

ABDISHAKOR MOHAMOUD ELMI, B.Sc.Eng.

University of Garyounis
Benghazi, Libya, 1989

A thesis submitted to
the Faculty of Graduate Studies and Research
in partial fulfilment of
the requirements for the degree of

MASTER OF ENGINEERING

Department of Civil and Environmental Engineering
Carleton University, Ottawa, Ontario

© Copyright
1998, Abdishakor Mohamoud Elmi

The Master of Engineering in Civil and Environmental Engineering Program is a joint
program with the University of Ottawa, administered by the
Ottawa-Carleton Institute for Civil Engineering

National Library
of Canada

Acquisitions and
Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale
du Canada

Acquisitions et
services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-32396-X

The undersigned hereby recommend to
the Faculty of Graduate Studies and Research
acceptance of the thesis,

**“Analysis of the temporal transferability of
work trip distribution “gravity models”**

submitted by

Abdihakor Mohamoud Elmi, B.Sc.Eng.

in partial fulfilment of the requirements
for the degree of Master of Engineering

Eric J. Miller (University of Toronto)
Thesis Co-Supervisor

Ata. M. Khan (Carleton University)
Thesis Co-Supervisor

J.L. Humar
Chair, Department of Civil and Environmental Engineering

Carleton University
April, 1998

ABSTRACT

This thesis investigates the temporal transferability of trip distribution models in the Greater Toronto Area. Empirical studies of both strategic and tactical forecasts were carried out using travel behaviour databases collected in 1964, 1986, and 1996. Different model specifications, ranging from a proportional flow model [PFM] to various market segmentations of auto travel time models [ATTM], were tested to investigate the relationship between model specification and transfer effectiveness. Results indicate that both long-range and short-range transferred models provide very useful information about travel behaviour in the application contexts. Pragmatically, these transfer models exhibited comparable results with that of estimated models. At the disaggregate level, the relatively short-range [10-year] transferability test showed better results compared with longer-range tests. Aggregate prediction measures showed mixed results for both short and long term forecast applications. This might be attributed to aggregation errors. Consistently, the well specified models such as occupationally disaggregated models and driver's licence/household auto-ownership models transferred better.

ACKNOWLEDGMENT

During the course of this research I have benefited from the assistance and contributions of many individuals. My supervisors at the University of Toronto, Professor Eric Miller and Professor Daniel Badoe, contributed significantly to the development of this thesis. Their thought provoking suggestions and insights played a pivotal role in this research. I cannot thank them enough for their patience and guidance. I am also very thankful to my supervisor at Carleton University, Professor Atta Khan, for his support and advice during this research. The invaluable assistance and advice that I received from my colleagues at the University of Toronto is also gratefully appreciated. Asmus Georgi was really helpful with computers and data management. This thesis has also benefited from the valuable comments of Professor Bhagwant Persaud of Ryerson Polytechnical University. I am also indebted to Mrs. Martine Johnson of International Student Centre for editing the first draft. I am grateful to Mrs. Marian J. Khalaf for typing and revising the final manuscript. My first ever tutor and undoubtedly my best friend, my mother Halima A. Abdi and my niece Sadia Ali deserve the greatest appreciation for their support and love throughout my graduate studies. Finally, I would like to express my gratitude towards my late father Mr. Mohamoud Elmi, who taught me with discipline and perseverance all goals can be achieved. May God bless him.

TABLE OF CONTENTS

	Page
ABSTRACT.....	ii
ACKNOWLEDGMENTS.....	iii
TABLE OF CONTENTS.....	iv
LIST OF TABLES.....	vii
LIST OF FIGURES.....	ix
CHAPTER 1: INTRODUCTION.....	1
1.1 Study Approach.....	1
1.2 Thesis Organization.....	3
CHAPTER 2: THEORY [LITERATURE REVIEW].....	5
2.1 Entropy Maximization and Information Minimization.....	5
2.2 Information Theory.....	11
2.3 Transferability Issues.....	17
2.3.1 Testing of Aggregate Models for Forecasting Capability of the Gravity Models.....	17
2.3.2 Disaggregate Choice Model Transferability Studies.....	21
2.4 The Relationship between Entropy Maximization and Logit Model.....	29
CHAPTER 3: DATA ANALYSIS.....	30
3.1 Descriptive Analysis of 1964 MTARTS, 1986 TTS and 1996 TTS Databases.....	30
3.2 Descriptive Analysis of GTA Urban Structure in 1964, 1986 and 1996.	40

3.2.1 Demographic and Socio-Economic Changes.....	41
3.2.2 Socioeconomic Development of GTA.....	49
3.2.3 Analysis of Trip Start Times for Total and Home-based Work Trips in 1964, 1986 and 1996 Databases.....	56
3.2.4 The Importance of Planning District 1 to the GTA.....	61
3.2.5 Changes in Transportation Infrastructure.....	64
CHAPTER 4: ANALYSIS METHODS	69
4.1 Model Calibration Techniques.....	69
4.2 Outline of the Model Systems Under Investigation.....	71
4.2.1 Model 1: The Proportional Flow Model.....	77
4.2.2 Model 2: Auto Travel Time Model.....	80
4.2.3 Model 3: Generalized Cost 1986 Model.....	82
4.2.4 Travel Time Variable.....	87
4.3 Goodness of Fit [Model Performance Measures].....	89
4.4 Transferability Tests.....	90
4.4.1 Measures of Transferability.....	90
CHAPTER 5: ESTIMATION AND TRANSFERABILITY RESULTS.....	96
5.1 Estimation Results for Different Model Specifications.....	96
5.1.1 ATTM Results and its Comparison with PFM.....	97
5.1.2 Spatial Segmentation Results "Planning District 1 Vs. Other".....	107
5.1.3 Socioeconomic Segmentation Results.....	114
5.1.4 Results of Model 3 "Generalized Cost Model"	122

5.2 Transferability Results.....	124
5.2.1 Disaggregate Measures of Transferability of 1	
964 Model Transferred to 1986 Context.....	125
5.2.2 Aggregate Measures of Transferability Results [1964 Model].....	130
5.2.3 Results of 1964 Model Transferred to 1996 Model.....	131
5.2.4 Results of 1986 Model Transferred to 1996 Context.....	132
CHAPTER 6: SUMMARY, CONCLUSIONS AND FUTURE	
RESEARCH DIRECTION.....	135
6.1 Summary of Results.....	135
6.1.1 GTA Trends.....	136
6.1.2 Model Estimation and Specification Test Results.....	138
6.1.3 Transferability Test Results.....	139
6.2 Conclusions and Direction for Further Research.....	142
REFERENCES.....	145
APPENDIX A: Trip Length Distributions for All the Models.....	150
APPENDIX B: Prediction Errors.....	161
APPENDIX C: Goodness-Of-Fit Equations.....	169
APPENDIX D: Input Parameter Files.....	172

LIST OF TABLES

3.1	The Distribution of the Zones in the Six Regional Municipalities.....	33
3.2a	Distribution of Population of 1964, 1986 and 1996 by Region.....	43
3.2b	Household Size by Regional Municipality (1961-1996).....	46
3.3	Distribution of Population by Household Size.....	48
3.4	Distribution of Households by Number of Workers.....	51
3.5	Percentage of Households Segmented by Gender with Driver's Licence.....	52
3.6	1986 Database Car Ownership.....	53
3.7	Distribution of Households by Number of Vehicles Owned.....	54
3.8	1964 and 1986 Trip Rates by Municipality Morning Peak Period Trips.....	55
3.9	GO Transit System.....	65
3.10	Changes in 1964 to 1996 GTA Transportation System.....	68
4.1	PFM Trip Matrix Table.....	79
4.2	Proportion of Trips in PFM Matrix	80
5.1	1964 Estimation in Log Likelihood Values.....	98
5.2	1964 Model Parameter Estimates.....	99
5.3	1986 Estimation in Log Likelihood Values.....	99
5.4	1986 Model Parameter Estimates.....	100

5.5	1996 Estimation in Log Likelihood Values.....	100
5.6	1996 Model Parameter Estimates.....	101
5.7a	1964 TTS Occupation Data Base Segmented by Gender.....	120
5.7b	1996 TTS Occupation Data Base Segmented by Gender.....	121
5.8	Home-to-Work Trip Distribution Model Parameter Estimation.....	123
5.9	Results from Transferring 1964 to 1986 Application Context.....	126
5.10	Results from Transferring 1964 to 1996 Application Context.....	127
5.11	Results from Transferring 1986 to 1996 Application Context.....	128

LIST OF FIGURES

2.1	Information: The Difference between a Priori and Posteriori knowledge.....	12
2.2	Model Uncertainty.....	17
3.1	The Regional Municipality within the Greater Toronto Area.....	32
3.2	Household size for the GTA.....	47
3.3	Labour Force Participation Rate by Gender for Toronto CMA.....	50
3.4	1964, 1986 and 1996 Total Trip Start Times Distribution.....	57
3.5	1964, 1986 and 1996 Total Work Trip Start Times Distribution.....	58
3.6	1964, 1986 and 1996 Home-based Work Trip Start Times Distribution.....	58
3.7	1964, 1986 and 1996 Percent of Total Trip Start Times Distribution.....	59
3.8	Location of Planning District 1 within Metropolitan Toronto.....	63
3.9	The Six Rail Corridors Network and Stations.....	66
3.10	TTC Subway and RT System.....	67
4.1	1964 Model Specifications and its Transferability to 1996 Context.....	73
5.1	1964 Observed and Predicted Trip Length Distribution Using PFM.....	98
5.2	ATTM Parameter Values for 1964, 1986 and 1996 Data bases.....	102

5.3	Observed Average Travel Times for 1964, 1986 and 1996 Data bases.....	103
5.4	Observed and Predicted Trip Length Distribution for 1964 Workers.....	104
5.5	Observed and Predicted Trip Length Distribution for 1986 Workers.....	105
5.6	Observed and Predicted Trip Length Distribution for 1996 Workers.....	105
5.7	1964 Cumulative Trip Frequency Distribution for All Workers.....	106
5.8	1986 Cumulative Trip Frequency Distribution for All Workers.....	106
5.9	1996 Cumulative Trip Frequency Distribution for All Workers.....	107
5.10	Observed Average Travel Time for PD1 Workers.....	109
5.11a	Trip Length Distribution for PD1 Workers in 1964.....	110
5.11b	Trip Length Distribution for non-PD1 Workers in 1964.....	110
5.12	Cumulative Observed and Predicted Trips for PD1 workers in 1964.....	111
5.13	Cumulative Observed and Predicted Trips for	

	non-PD1 workers in 1964.....	111
5.14	1986 Trip Frequency Distribution for Workers that Destined in PD1.....	112
5.15	1986 Trip Frequency Distribution for Workers that Destined in non-PD1.....	112
5.16	1996 Trip Frequency Distribution for Workers Destined in PD1.....	113
5.17	1996 Trip Frequency Distribution for Workers Destined in non-PD1.....	113
5.18	1964 Model Travel Time Parameters Segmented by Occupation	117
5.19	1996 Model Travel Time Parameters Segmented by Occupation	118
5.20	1996 Female-Model Travel Time Parameters Segmented by Occupation	118
5.21	1996 Male-Model Travel Time Parameters Segmented by Occupation.....	119

CHAPTER 1

INTRODUCTION

1.1 Study Approach

Urban travel demand analysis is based on the Urban Transportation Modeling System [UTMS], also called the Urban Transportation Planning System [UTPS]. The UTMS involves the sequential application of the following four stages: trip generation, trip distribution, modal split, and trip assignment.

This thesis deals with the second stage, trip distribution, which estimates the flow of trips from origin to destination and predicts the future trip matrix. In this stage, the estimates of trips from origins to destinations are obtained from a number of trip distribution models developed for this purpose such as the gravity model. The gravity model which is applied in forecasting is estimated using data typically drawn from cross-sectional surveys.

The implicit assumption is that spatial structure and travel behavior are captured by the model, and that the determined behavioral relationship will remain reasonably stable over time.

The validity of these assumptions in long-range forecasting, in which significant travel-related factors may change, has not yet been fully explored. Changes in urban spatial structure, particularly from monocentric to polycentric structure and the significant changes in the socioeconomic base of an urban area motivate investigation into whether these cross-sectional models are stable over time.

Quite a few transferability studies are documented in the literature. Most of those studies dealt with disaggregate behavioral models, particularly the logit model of mode choice.

There are no studies done so far that explicitly address the transferability of trip distribution models in new situations. Obtaining a transferable trip distribution model from the estimation context that can describe or explain a significant component of information to the application context will be very cost effective to the planning process in terms of time and cost of data collection and model development. The objective of this study was to investigate the temporal transferability of gravity trip distribution model. This investigation addressed two major issues

1. Given that three different data sets are available for the Greater Toronto Area [i.e., the 1964 MTARTS, the 1986 TTS and the 1996 TTS travel surveys] which are suitable for the purpose of specification and transferability test, it was possible to explore the applicability of the gravity model in both strategic and tactical forecasting. The first strategic forecasting test was to transfer the 1964 model into the 1986 historical forecast context [22-year period]. The second strategic test was to transfer the same 1964 model into the 1996 context [32-year period]. The tactical forecasting test was to transfer the 1986 model to 1996, a relatively short-range transfer [10 year period].
2. It is believed that a well specified model transfers better than other models. To verify this, the relationship between model specification and transferability was investigated. Transferability is assessed by the ability of the model to predict both individual and aggregate trip patterns in the forecast context. It is hoped that the result of the study will add some significant insight into the transportation planning process.

1.2 Thesis Organization

In addition to this brief introductory chapter, the thesis contains six chapters. Chapter 2 provides a brief review of the literature dealing with background

information, including entropy models, information minimization models, the relationship between entropy and logit models, and, finally, some transferability issues.

Chapter 3 presents the data used in the study, and also describes the structure of the Greater Toronto Area [GTA] in 1964, 1986 and 1996. Chapter 4 discusses the analysis methods, that is, the calibration program used in this study; the model performance measures such as goodness-of-fit, the log-likelihood method, t-statistic and the prediction errors. This chapter also discusses the transferability measures that were used to judge the ability of a model to be transferable, and lastly the various models that were tested in this report are also outlined in this section. Chapter 5 presents the results of the exercise as well as a summary of these results. Chapter 6 presents the summary and conclusions that can be drawn from this study. Finally, a series of appendices will summarize the outputs from running different programs.

CHAPTER 2

THEORY/LITERATURE REVIEW

This chapter outlines the literature review of entropy maximization theory [information minimization method, see section 2.1 and 2.2], relationship between the entropy method of trip distribution process and logit method of mode choice process, and finally, the transferability issue which is the main interest of this thesis.

2.1 Entropy Maximization/Information Minimization Models

This section provides the derivation and evaluation of the entropy maximization model as an analytical tool for calculating present and forecasting future urban distribution patterns. Entropy maximization methods form the basis for the models derived and tested in this thesis.

The guiding principle used in this approach is that of entropy, a concept borrowed from physical science, such as thermodynamics, and statistical mechanics. In thermodynamics, entropy relates energy and absolute temperature. It measures the

disorder of particles in a closed system. The greater the disorder, the greater is the entropy of the system [Snickars and Weibull, 1977].

In statistical mechanics, the most probable permissible macro distribution is the one produced by the largest number of different micro states without regard to the missing information. In other words, the most likely state of a system is that particular state for which the number of combinatorial possibilities is the greatest. In this branch of physics, the log quantity of the following equation is referred to as entropy [Wilson 1967].

$$w\{T_{ij}\} = \frac{T!}{\prod_i T_{ij}!} \quad [2.1]$$

where

T = is the total number of trips for one purpose.

T_{ij} = probability of the distribution occurring or is the distribution

which gives rise to the most probable distribution.

$\prod_i$ the summation is over all the distribution which satisfy constraints

[2.3] - [2.5].

w = the number of distinct arrangements of individuals which give rise to the distribution $\{T_{ij}\}$.

Let us take an observed distribution of trips in an area as a state of the transportation system. When the system has the most probable disordered pattern, it is said to have the maximum entropy. The trip distribution model is designed to estimate flows on the transport system and predict the future trip matrix on the system. Since this model is the second step of all urban modeling transportation system, we assume here that the estimates of trip ends i.e., the total number of trips originating from a zone O_i and those destined for each zone D_j are estimated in the first step of the travel demand modeling process (trip generation model). Now, the second step of this process answers the question of how many of these trips leaving a given zone will terminate at each possible destination.

Wilson, in 1967, derived the first theoretically sound approach of trip distribution from information theory called the entropy maximization method. The number of micro states $w\{T_{ij}\}$ (number of distinct arrangements of individuals) associated with macro states T_{ij} is given by:

$$w(T_{ij}) = \frac{T!}{T_{11}!(T - T_{11})!} \cdot \frac{(T - T_{11})!}{T_{12}!(T - T_{11} - T_{12})!} \cdots = \frac{T!}{\prod_j T_{ij}!} \quad [2.2]$$

To be feasible (i.e., internally consistent), the estimated trip matrix must satisfy

$$\sum_j T_{ij} = O_i \quad \forall i \text{ (for all values of } i) \quad [2.3]$$

$$\sum_i T_{ij} = D_j \quad \forall j \text{ (for all values of } j) \quad [2.4]$$

To obtain the effect of the deterrence function, it is required to add C constraints which usually take the form:

$$\sum_i \sum_j T_{ij} c_{ij} = C$$

c_{ij} = the impedance function of travelling between i and j. [2.5]

This constraint implies that the total amount of time or cost spent on trips in a region at this point in time, is a fixed amount C or, equivalently, that the predicted average trip cost equals the observed average [obtained by dividing both sides of equation 2.5 by T].

One can maximize $\log w$ instead of w with respect to [2.1], [2.3], [2.4] and [2.5].

$$\log w = \log \frac{T!}{\prod_i T_i!} = \log T! - \sum_i \log T_i! \quad [2.6]$$

Stirling's (short) approximation for $\log X! = X \log X - X$, can be used to make it easier to optimize.

$$\log w = \log T! - \sum_i (T_i \log T_i - T_i) \quad [2.7]$$

The term $\log T!$ is usually constant and can be omitted from the optimization problem.

Now, to obtain the set of T_{ij} 's which maximizes $w(T_{ij})$ subject to the constraints [2.3], [2.4] and [2.5]

$$L = \log w + \sum_i \lambda^{(1)} (O_i - \sum_j T_{ij}) + \sum_j \lambda^{(2)} (D_j - \sum_i T_{ij}) + \beta (C - \sum_i \sum_j T_{ij} c_{ij}) \quad [2.8]$$

where $\lambda^{(1)}$, $\lambda^{(2)}$ and β are Lagrangian multipliers.

and taking its first partial derivatives with respect to T_{ij} and equating to zero gives

$$\frac{\partial w}{\partial T_{ij}} = -\log T_{ij} - \lambda_i^{(1)} - \lambda_j^{(2)} - \beta c_{ij} = 0 \quad [2.9]$$

and this reduces to

$$T_{ij} = \exp[-\lambda_i^{(1)} - \lambda_j^{(2)} - \beta c_{ij}] \quad [2.10]$$

Substituting in (2.3) and (2.4) to calculate $\lambda_i^{(1)}$ and $\lambda_j^{(2)}$:

$$\exp(-\lambda_i^{(1)}) = O_i / \left[\sum_j \exp(-\lambda_j^{(2)} - \beta c_{ij}) \right] \quad [2.11]$$

$$\exp(-\lambda_j^{(2)}) = D_j / \left[\sum_i \exp(-\lambda_i^{(1)} - \beta c_{ij}) \right] \quad [2.12]$$

dividing each side by O_i gives

$$\exp(-\lambda_i^{(1)}) / O_i = 1 / \left[\sum_j \exp(-\lambda_j^{(2)} - \beta c_{ij}) \right] \quad [2.13]$$

calling A_i = the right side part of [2.11] for convenience gives

$$A_i = \exp(-\lambda_i^{(1)}) / O_i \quad [2.14]$$

$$B_j = \exp(-\lambda_j^{(2)}) / D_j \quad [2.15]$$

and then

$$T_{ij} = A_i B_j O_i D_j \exp(-\beta c_{ij}) \quad [2.16]$$

substituting $\exp(-\lambda_i^{(1)})$ and $\exp(-\lambda_i^{(2)})$ of equations 2.14 and 2.15

by $A_i O_i$ and $B_i D_i$ respectively, the right hand sides of equations

2.11 and 2.12 will have the following forms:

$$A_i = \left[\sum_j B_j D_j \exp(-\beta c_{ij}) \right]^{-1} \quad [2.17]$$

$$B_j = \left[\sum_i A_i O_i \exp(-\beta c_{ij}) \right]^{-1} \quad [2.18]$$

equations [2.17] and [2.18] are equivalent to the balancing factors of the classic gravity model. As a result, the most probable trip distribution of the entropy model whose statistical derivation gives a new theoretical base for the trip distribution model is the same as that of the gravity model. It is not necessary that the cost constraint equation [2.5] is known. Parameter β can be statistically estimated, as discussed below.

Wilson's statistical theory of entropy maximization can be summarized as: "given total numbers of trip origins and destinations for each zone for a homogeneous person-trip purpose category, given the costs of traveling between each pair of zones, and given that there is some fixed total expenditure on transport in the region, then there is a most probable distribution of trips between zones, and this distribution is the same as the one normally described as the gravity model distribution" [Wilson, A. G, 1970].

2.2 Information Theory

Information theory was first established by Shannon and Weaver [1963], when they launched their statistical theory of communication. After that many other different science disciplines used this method in order to solve their problems.

From the transportation planning point of view, A. G. Wilson applied this theory in 1967 to trip distribution modeling process based on probability principles.

To explain the concept let us borrow the idea of Henri Theil [1967]. Suppose you know that an event E will occur [a priori knowledge] with a certain probability P , $0 \leq P \leq 1$, and some time later you have been notified that the event E has indeed occurred [a posterior knowledge]. There will be no surprise for you, $P=0.99$, because it was practically certain that this event will occur. In other words the message has very little "information content" when P is close to 1. That is, if the occurrence of the event will give you no information then $h(P) = 0$, and a priori knowledge is equal to a posterior knowledge. However, the situation will change when probability is smaller [close to zero] that this event will occur; you will be greatly surprised by the message when the event does occur. The message has very great "information content".

Hence, the only assumption made here is that the information content is a function of the probability P that the event will occur before realization.

To further clarify this point, let us see Figure 2.1, a priori knowledge is the total amount of information you have before the event E occurs.

Figure 2.1 Information : the Difference between a Priori Knowledge and Posterior Knowledge

A posterior knowledge is the total amount of information you get after the event does occur. In that case, information can be defined as the difference between a posterior knowledge and a priori knowledge of a phenomenon after realization [Pluym 1979].

It is clear from the above observation that information content M is a decreasing function of the probability P , [the more unlikely this event does occur before the

realization, the larger the information content of the message]. Now the decreasing function becomes the reciprocal of this probability.

$$M(P) = \ln \frac{1}{P} = -\ln P \quad [2.19]$$

This is the fundamental formula of statistical information theory.

The reason for this logarithmic function among all decreasing functions is the additivity in the case of independent events.

The general information theory-based derivation of trip distribution models can now be defined as follows:

$$M_{\{ij\}} = -\ln p_{\{ij\}} \quad [2.20]$$

For a set of events I [$i=1, \dots, I$] the mathematical expectation of the information content is:

$$h = - \sum_{j=1}^I P_i \ln P_i \quad [2.21]$$

$$\text{with } \sum_{i=1}^I P_i = 1 \quad [2.22]$$

so that h reflects the amount of uncertainty represented by a discrete probability distribution. In the framework of a spatial interpretation model where no prior

information is available about the trip distribution [except the marginal conditions and the cost budget], it seems reasonable to employ a probability distribution with a maximum uncertainty i.e.

$$\text{Max } v = - \sum_{i=1}^I \sum_{j=1}^J P_{ij} \ln P_{ij} \quad [2.23]$$

where

$\text{max } v$ - represents the maximum number of combinatorial assignments of trips over i and j .

P_{ij} - is the trip probability from i to j satisfying the additivity condition

$$\sum_{i=1}^I \sum_{j=1}^J P_{ij} = 1 \quad [2.24]$$

Furthermore, the marginal conditions and the travel budget may also be introduced

$$\sum_{j=1}^J P_{ij} = \frac{O_i}{t} \quad [2.25]$$

$$\sum_{i=1}^I P_{ij} = \frac{D_j}{t} \quad [2.26]$$

$$\sum_{i=1}^I \sum_{j=1}^J P_{ij} C_{ij} = \frac{c}{t} \quad [2.27]$$

maximizing [2.21] subject to the constraints [2.25], [2.26], and [2.27] implies

$$\begin{aligned} mv = & - \sum_{i=1}^I \sum_{j=1}^J P_{ij} \ln P_{ij} + \sum_{j=1}^J u_j [O_j / t - \sum_{i=1}^I P_{ij}] + \sum_{i=1}^I \lambda [D_i / t - \sum_{j=1}^J P_{ij}] \\ & + \beta [c / t - \sum_{i=1}^I \sum_{j=1}^J P_{ij} C_{ij}] \end{aligned} \quad [2.28]$$

t = the total number of trips.

mv = the lagrange function for a constrained maximum of v

Equation [2.28] is similar to equation [2.8] which was presented earlier.

It is clear that this specification gives rise to the same results as the entropy model discussed in the previous section. Thus the maximization of the number of micro-states of the system leads to the same results as the maximization of the expected information content of uncertain events.

This information measure quantifies the dispersion of a distribution and the distance and association between distributions that will be used to interpret and compare the results of these trip distribution models [see Pluym, 1978].

When dealing with the trip distribution process it is assumed that the only information known is the estimates of trip ends. The information available to a planner can come from different sources and may not directly be compatible with each other. In that case, the entropy model is based on uncertainty gained in this process. The aim of entropy maximizing methods is to maximize the amount of uncertainty in the model with the intention that if the uncertainty is maximum then the resulting model will have the most reliable information.

Examining Figure [2.2] there are three models constructed and calibrated in a single data set. The three 'bars' represent three models. The shaded areas in

Figure 2.2 [i] represent the amount of uncertainty each model entails based on real data. Since the models are applied to the same data set, they must have the same amount of uncertainty. Now it is obvious that the two outer models contain some bias. The shaded region in Figure 2.2[ii] represent the amount of bias between model 1 and 2 and between model 2 and 3.

There is a tendency for a planner to prefer model 3 under the impression that it is the best of the models. However, as unappealing and counter-intuitive as it may seem, model 2 is the most reliable and least deceptive because removing all the bias or uncertainty indicated in this model leaves the planner with pure information and therefore with the model with the highest certainty.

Figure 2.2 Model Uncertainty

2.3 Transferability Issues

2.3.1 Testing of Aggregate Models for Forecasting Capability of the Gravity Model

In general, forecasting constitutes an essential input into decision making. All predictions are not perfect since future events practically always entails some uncertainties.

The main purpose of this section is to evaluate the forecasting capability of the gravity model. Trip distribution models are designed to capture the overall changes in trips between some base year and the forecast year and also the major changes in commuting structure. Several previous studies regarding the reliability and usefulness of the gravity model as a forecasting tool are well documented, in particular the following two papers.

Volet and Hutchinson [1986] investigated the forecasting capabilities of the doubly- constrained gravity model and the Fratar model. They examined the ability of these models to reproduce interzonal trips of 1971 and 1981 in Toronto region and the ability of the 1971 model to forecast the 1981 flows. The data were collected from Toronto and Hamilton census metropolitan areas and Oshawa census area between 1971 and 1981.

The estimation and prediction of the gravity model are compared with that of Fratar. The census journey-to-work data bases of the study area in 1971 and 1981 were used. During that time [1971-1981], the spatial distribution of land development and the transport system experienced a major change that might influence the explanatory and forecasting capabilities of the trip distribution models. The authors' analysis showed that between 1971 and 1981 the total population in the studied area increased by 470,000, the resident labour force by about 500,000, and total employment by about 514,000.

The calibration criteria [mean trip-length calibration criterion, trip length-frequency distribution and mean absolute error between observed and simulated trip interchanges], were used to evaluate the relative accuracy of the models.

The goodness-of-fit measure used in this study is similar to that of the work done by Smith and Hutchinson [1981].

$$\frac{\Phi}{T} = \frac{1}{T} \sum_i \sum_j T_{ij} \left| \ln \frac{T_{ij}^*}{T_{ij}} \right| \quad [2.36]$$

where

T - is the total number of home-to-work trips

T_{ij}^* - is the observed trips from i to j and

T_{ij} - is the simulated trip interchange value.

Φ - modified entropy measure (the larger the value of Φ the poorer the model fit).

Calibration of both 1971 and 1981 databases indicated that the doubly constrained gravity model was unable to forecast well the longer trips from the fringe residential areas [such as Mississauga and Oakville] to the central business district [CBD], as well as from the newer residential areas to the northeast. The under-

prediction of longer distance trips from the above areas is related to the socioeconomic change in jobs and resident labour forces over the period between 1971 and 1981. The authors concluded that the overestimated shorter commuting trips from the residential areas close to the Toronto balanced the underestimation of longer commuting trips.

Duffus and Soliman [1987] used the peak hour trip data sets for 1962, 1971, 1976, and 1981 from the city of Winnipeg, Manitoba to investigate the forecast performance of a gravity model. The study area was divided into 139 traffic zones. The city of Winnipeg experienced a steady population growth rate between 1962-1981. This study investigated the most reliable F-factors and adjustment factors [K-factors] that can be used for the prediction purposes.

The calibration criterion is said to be reasonable when the residual of observed trip length frequency distribution for the trips and the calibrated frequency distribution is minimum. The K-factors are then introduced in order to account for the socioeconomic effects that are not incorporated in the model.

Like the models of Volet and Hutchinson [1986], the authors found that the model overestimates the short trips and underestimates the long trips.

The criteria used to measure the accuracy of the model prediction were the mean absolute error [MAE] and the root-mean square error [RMSE]. The mean

absolute error expresses the average absolute differences between corresponding observed and calibrated O-D trips, while the RMSE is a measure of the standard deviation of the differences between observed and estimated trips. The RMSE is more sensitive to large deviations than the absolute mean error.

The calibration analysis indicated that the estimated trip length frequency distributions were comparable with the observed distributions for each year.

Consequently, the F-factors of all the four study year periods were found to be stable. Nevertheless, the result showed that the values of K-factors were mixed and irregular; the authors concluded that K-factors were not necessary to incorporate in the model since they resulted in larger errors.

Finally, the authors noted that the model can be used with greater confidence to predict the future when F-factors alone are used.

2.3.2 Disaggregate Choice Model Transferability Studies

In recent years government budgets have become tighter while travel demand has increased; on the other hand developing adequate behavioral models is expensive due to the cost of data collection. In light of this constraint, quick and useful response of travel estimation is needed. The study of a model transferability,

which is transferring existing models to new situations and areas, became one of a possible option that gives a quick response to the travel demand forecasting procedures.

A demand model is considered to be transferable if the model's predictive ability is not affected by empirical and policy conditions. If the model cannot successfully predict travel behavior in the area for which it was estimated, there is no reason to expect it to function better in any other areas. Several studies have been conducted to evaluate the transferability of aggregate and disaggregate behavioral models. Transferability could be either spatial or temporal.

Disaggregate models are more likely to be transferable because they represent the average behavior of the individual traveler, and it is reasonable to expect individual travel behavior to be essentially the same in one area as in another (specially, the areas of similar characteristics such as North America). The analyst's options are either to improve the model's specification or to segment the travel choices to eliminate the sources of mis-specifications.

The results of these studies have been mixed in their assessment of transferability. In four studies, the model parameters estimated in one context were applied without any modification to a second context [Atherton and Ben Akiva, 1976; Watson and Westin, 1975; Talvitie and Kirshner, 1978; Stopher and Wilmot,

1979]. The other studies reviewed in this thesis were concerned with model updating. Each of these studies is discussed here in turn.

Atherton and Ben-Akiva [1976] carried out an intra-urban travel study for work trip mode choice model calibrated on Washington D.C. in 1968 to predict the mode share of car drivers, car passenger & transit travelers for work trips in New Bedford, Massachusetts [1963] and Los Angeles, California [1967]. The authors used the level-of- service variables [i.e. in-vehicle times, out-of-vehicle times, and out-of-pocket costs] to test if this model of Washington and some other socio-economic variables can be transferred to New Bedford & Los Angeles.

The test was to use the specification of the Washington model and then calibrate new models using New Bedford and Los Angeles data and compare the model coefficients of the new models with that of Washington; the criteria for transferability were judging the model by the equality of model coefficients, classifying model evaluations [i.e. goodness-of-fit], and model accuracy. The authors cited above stated, as Ben Akiva and Koppelman argued earlier, that no model specification or any cultural or socio-economic differences between the cities would certainly affect the comparison of any of these models.

The overall conclusion of this study was that the model was transferable in spite of the differences in size, urban characteristics, year and manner in which the data were collected.

Dehghani, et al. [1981] examined spatial transferability of disaggregate, mode choice models similar to that of Atherton and Ben Akiva, but developed from the Baltimore, Minneapolis, St. Paul and San Francisco [pre & post-BART] household surveys. The result indicated spatial transferability of models between cities with different transit service. They concluded that transferability is limited to areas that have similar travel and urban characteristics. The analysis of transferability took the form of a comparison of model coefficients calibrated for each city.

In 1978 Talvitie and Kirshner used the two data sets for the San Francisco Bay Area. These two data sets were collected in 1973 before Bay Area Rapid Transit was constructed and 1975 after the completion of this project. The authors stressed that if a model is not temporarily transferable, it has no use as a prediction tool. The authors chose the same model specification as that of Atherton & Ben-Akiva [1976].

Three modes were considered: drive alone, transit with walk access, and shared ride. The authors rejected the temporal transferability and concluded by saying

that if the transferability of models is judged using the statistical criteria, there are few grounds to claim that the extensive work trip mode choice models are transferable. The authors believed that the differences in coding network and the manner in which outlying data were handled led to the central conclusion of their research.

Parody [1977] used a temporal longitudinal or time series data base at MIT. The prediction study was made after some quantifiable changes in transportation system occurred in the University. To evaluate well the prediction and the accuracy of the model, a forecast of before the change and after the change was conducted. A before survey was conducted after the introduction of changes in parking prices and regulations. The logit model was used.

Parody compared the predictive performance of a model that contains only level of service variable with a model that included both level of service and socio-economic variables. The latter one was found to perform better than the former one. The author concluded that the parameter estimates were similar in the two contexts and that the model was transferable.

Watson and Westin [1975] present one of the first model transferability approaches to investigate the spatial transferability of work trip logit mode choice models (inter-city trip making) between Glasgow and Edinburgh. The modes

considered were automobile and train. Six identically specified binomial logit models were developed to predict the modal splits of six spatially defined market segments, [i.e. five geographic locations other than the area from which the original sample of travelers was drawn]. The results showed that the models that have one or both end in the central city forecast very well among themselves. The authors warned that disaggregate models should be refined before any transferability issues are considered.

The overall conclusion of Watson & Westin was that the models are capable of predicting the forecast context.

Train [1979] used models which were developed and estimated using data for workers in the San Francisco Bay Area in 1972. Workers were considered to have a choice which included four modes: auto alone; bus, with auto access; bus, with walk access; and car pool. The author first investigated the model specification and then examined the relationship between the specifications and the transferability. The author tested several model specifications till he obtained the specification which is consistent with a priori expectations and fits the data fairly well; he then evaluated the transferability by comparing the equality of the model coefficients in the local and the applications contexts. He found the model coefficients of the level-of-service variables to be comparable in the two contexts. However, the socio-economic coefficients were very different in the estimated and

application contexts. He finally concluded that the complex model specifications predict well but still rejected transferability.

McCarthy [1982] examined the temporal characteristics or stability of work trip behavior in the San Francisco Bay Area. This research utilized the same data that Talvitie and Kirshner used, that is, from two household surveys collected in the San Francisco Bay Area as part of the BART [Bay Area Rapid Transit] Impact Travel Project. The null hypothesis was that “the parameter estimates associated with the service related attributes and the automobile specific variables remain stable over the short-run period”. Finally McCarthy concluded that all pre-BART model coefficients [i.e. set of service related variables, alternative specific variables, and the alternative specific constant] were temporally transferable.

Hensher and Johnson in 1977 tested temporal transferability. The data used were collected before 1971 and after 1973 in the Sydney Metropolitan Area, Australia. The authors used a simple specification of level-of-service variables. Their sample consisted of 226 trips. The alternatives considered were auto driver and rail transit travelers between the financial district of Sydney [CBD] and the suburban areas. Between the above dates there was a remarkable change in transit fares and the transportation system performance. The coefficient estimates were found to be very similar between the estimated context and the application contexts for all the four logit models calibrated in 1997 data and used to test the model transferability

in 1973. Finally the authors concluded that the models developed were extremely transferable.

McCoomb [1986] tested inter-urban transferability. This study used Labour Force Survey - transportation supplement data. The author examined spatial transferability of ten identically specified models for 10 Canadian cities all calibrated from the same high-quality data source. The use of the above consistent and uniform data indicated that the models were reasonably comparable based on the similarities in the model coefficients. Similarity in structure, size and transportation system of the estimated and the application context maximizes the likelihood of transferability. McCoomb concluded that there is no single "universal" model that can be transferable from one city to another.

2.4 The Relationship Between Entropy Maximization and Logit Model

The gravity model for trip distribution was designed to balance origin-destination trip tables so that the predicted zone-to-zone trips are equal to the trips generated at each origin and the trips terminating at each destination. There has been a long argument about the validity of these models.

The development of Wilson's first theoretically valid derivation of the gravity model in 1967 from statistical information minimizing principles called the "entropy maximizing" method, brought a behaviourally sound basis to the gravity

model and correctly balanced the trip tables. The model developed based on this method uses aggregate data.

On the other hand, McFadden [1973], Ben-Akiva [1973] and others began formulating the problems of travel mode and location decisions as problems in microscopic economic consumer choice among discrete alternatives. Unfortunately, this approach, unlike that of Wilson, did not succeed in providing a theoretical grounding of empirically established concept until finally, McFadden's derivation of the Logit Model from the utility maximizing concept introduced the new area of research called "behavioural demand modelling". This method is based on disaggregate small-sample data in the estimation of models via maximum likelihood.

Annas [1983] viewed the strong similarity between these two approaches the "information minimizing" and "utility maximizing" models of spatial interaction. He proved that the two approaches are identical in that the multinomial logit model can be derived and identically estimated by either method. He also proved that the doubly-constrained entropy model derived by Wilson [1967] is identical to a multinomial logit of joint origin-destination choice, consistent with stochastic utility maximization [see also Williams 1977].

CHAPTER 3

DATA ANALYSIS OF 1964 MTARTS, 1986 TTS AND 1996 TTS DATABASES

3. Introduction

This chapter discusses the data sets used in this study and also the details of the structure of the study area, that is, the Greater Toronto Area (GTA). These data sets are, the 1964 Metropolitan Toronto and Region Transport Study (MTARTS), the 1986 Transportation Tomorrow Survey (TTS) and the 1996 TTS databases.

The GTA consists of the Municipality of Metropolitan Toronto (as of January 1, 1998, Metro Toronto now corresponds to the new amalgamated city of Toronto) and the five regional municipalities of Durham, York, Peel, Halton, and Hamilton-Wentworth [see figure 3.1 and 3.2]. The data extracted from the 1964, 1986 and 1996 surveys have the following traffic zone system: The 1964 model has 817 traffic zones, defined over the range of zone numbers 1-939. While 1986 and 1996

have the same traffic zones, that is, 1404 traffic zones, defined over the range of zone numbers 1-1770 for the six municipalities mentioned above.

To facilitate the analysis this zone system has been further aggregated into 46 zones or "Planning Districts". This zone system in the GTA is developed and maintained by the Data Management Group of the Joint Program in Transportation in the University of Toronto.

Figure 3.1: The Regional Municipalities within the Greater Toronto Area
Source: MTO [1987] The Transportation Tomorrow Survey: Design and Conduct of the Survey.

Table 3.1: The distribution of the zones among the six regional municipalities.

Municipality	Zones	Planning Districts
Metro	1 - 460	1 - 16
Durham	501- 710	17 - 24
York	801- 978	25 - 33
Peel	1101-1325	34 - 36
Halton	1401-1563	37 - 40
Hamilton	1601-1770	41 - 46

All three databases contain all trips collected in a 24-hour period in the GTA. Home-based work trips made in the morning peak period from 6:00 to 8:59 a.m. were then extracted from that database and used in the analysis for this study. This is because home-based all-purpose trips usually constitute about 80% of the total trips in an urban area and because work-related trips are far more peaked than the other non-work trips and constitute the majority of trips in urban areas [RTAC, 1974].

The analysis of home-based work trips may be called a long-term process in the trip distribution process [Kanifani, 1983]. This is because these trips are stable and changes only in the long run, either by the change of residential location or of employment. They have been considered to be more important than any other kind of trips since it is believed they dictate the size of transportation facilities [Meyer and Miller, 1984]. Person trips were used because it will be necessary to evaluate the modal shares [highway and public services].

Intrazonal trips were not considered in this analysis, because auto and transit times from their respective assignment procedures, and walk distance associated with walk trips, are dependent on the centroid coordinates of the trip origin and destination zones. These coordinates of intrazonal trips are the same, hence their computed travel times and distances have numerical values of zero. The measure of spatial separation between zones (t_{ij}) used in this study was calculated for the road network using the EMME/2 program. Time was used because it was recognized to be the most realistic measure of spatial separation. EMME/2 modeling software was also used to batch out the auto and transit travel times. Deterministic user equilibrium assignment was used for the road network and aggregate transit assignment option used for the transit network.

Seven modes of travel were used in the analysis of the 1986 and 1996 work trip models: Auto-passenger allway; Transit allway; Subway with auto access (drive or passenger); Go-rail with transit (or walk)access; Go-rail with auto access (drive or passenger); Walk allway; and Auto-drive allway. The definition and the assumptions of these modes are listed in section 4.2. In 1964 only four modes (Auto-passenger allway, Transit allway, Auto-driver allway, and Walk allway) were used due to the unavailability of Go-Rail and Subway trips in that year.

As a transportation system characteristic, travel time was first considered to be the only explanatory variable that had a great effect on the trip distribution

process, but the result indicated that this single variable is not enough to explain the difference of travel behavior. To further understand travel patterns and how workers behave in the space, we segmented the three data sets into different market segments. The full description of the models tested are discussed in Section 4.2 [see also Figures 4.1].

The 1964 MTARTS, 1986 TTS and 1996 TTS databases used in this study to analyze the travel behavior in the Greater Toronto Area will be discussed below.

1. The 1964 MTARTS database

The 1964 MTARTS database was a 24-hour weekday home interview survey conducted in Metropolitan Toronto and its surrounding regions. The study area bounds extended from Oshawa and Bowmanville in the east to Hamilton and Saltfleet Township in the west and from Guelph and Barrie in the north to Lake Ontario in the south. The objective of this survey was to collect the information on travel by people living and working in the study area.

Over 24,700 households in the survey area were interviewed about household and travel information, from which close to 24,000 usable questionnaires were obtained.

The above figure constituted a representative sample of 3.3 percent of all households in the study area. To obtain a more accurate representation of the

travel information of residents in all the study area, additional households in the Brampton city were interviewed. Detailed information of this survey can be found in the report by [Traffic Research Corporation Limited, 1965]. The sample of households selected for the survey was obtained from special listings of customers of municipal hydro commissions and of the Ontario Hydro Commission for rural districts. This source of sample households gave a representative coverage of all households, since over 99.5% of households were serviced by Hydro, and less than 7% of these were on bulk meters. The selected households were surveyed during the period beginning on April 28, 1964, and ending on June 30, 1964. Approximately equal numbers of households were interviewed on each weekday, Monday through Friday. This was necessary as the frequency of travel is known to vary throughout the week, i.e. Monday travel is approximately 10% below the average and Friday travel is 10% above the average [1965, Municipal Directory, Ontario Department of Municipal Affairs]. Initial contact with the householders selected was made by mailing post cards which informed them that they had been chosen to participate in the survey. A few days later, the households were telephoned and asked for their cooperation in the survey. On agreeing to participate in the survey, the householders were interviewed by face to face.

For each data zone the number of survey households was compared with total household counts received for the region, and the basic expansion factor was based on the ratio of actual households to the number of sample households.

The zone system for 1964 is different from that of 1986 and 1996. The 1964 zone system comprises 817 zones, while in 1986 and 1996 the zones comprise 1404 zones.

2. The Transportation Tomorrow Survey Program

Prior to the 1986 Transportation Tomorrow Survey, the last comprehensive travel survey conducted in the GTA was that discussed above, the 1964 MTARTS. In 1977, the six Regional Municipalities in the GTA, the Ministry of Transportation of Ontario; the Toronto Transit Commission; and the GO Transit commuter rail agency, established the Toronto Area Transportation Planning Data Collection Steering Committee [TATPDCSC] to coordinate data collection efforts. They agreed to perform a series of large telephone based travel surveys in the GTA at five year intervals. In 1985, the TATPDCSC began planning for a major survey to coincide with the mid-term census in 1986.

The main purpose of TTS was to collect the household trip information to support the travel demand forecasting activities in the GTA. The household trip information for a 24-hour and 3-hour morning peak period from a representative sample of the households in the GTA was collected. The morning peak period was chosen to minimize the number of non-work trips. The collected data is managed and distributed by the data management group at the University of Toronto. A

relational database EMPRESS/32 version 2.4 is used to store, manipulate and retrieve the database.

a. The 1986 Transportation Tomorrow Survey

The 1986 Transportation Tomorrow Survey was conducted in the Fall [between September 16 and December 13] of 1986. The area covered by this telephone survey as mentioned earlier consists of Metropolitan Toronto, and the Regional Municipalities of Durham, York, Peel, Halton and Hamilton-Wentworth.

An advance letter was mailed to the households to introduce the survey before the actual survey was conducted. The interviewers recorded the data on coding forms and entered the data into computerized forms. Geographic Referencing System or geocoding was used to locate the coordination of the household and the location of all trip origins and destinations. This was the first time geocoding was used. The information collected include the demographic information [household and person levels] and travel information.

The purpose of this survey was to develop a reliable data base of household travel patterns for a typical weekday during this period, collecting the demographic and travel information. This survey was conducted in accordance with the standard procedure prescribed by TATPDCSC. A random sample of 4.1% of dwelling

units in the GTA was selected based on Bell Canada residential billing files. Over 61,000 households were interviewed and completed, based on the 1986 Census count of about 1,470,000 households in the GTA.. Households with unlisted phone numbers were not included in this study. This sample agrees with the minimum sample recommended by the U.S. Bureau of Public Roads Guidelines. To represent the total population in the Greater Toronto Area, each sample household record was given an expansion factor. The factors were defined as the ratio of the number of surveyed household and units in an "aggregation district". The trip information for all persons 6 years of age or older over a 24-hour period was collected and finally trip was defined as one-way movement between two locations for a single purpose.

b. The 1996 Transportation Tomorrow survey

This telephone interview survey is the largest travel survey ever performed in the study area. This survey was conducted in the Fall (Sept.-Dec.), 1996. A total of 115,000 households, that is, 5% of the total households were interviewed successfully. The survey area consists of the GTA surveyed in 1986 plus some additional areas.

The extended areas beyond the boundaries of the 1986 GTA survey include Niagara Region, Waterloo Region, City of Guelph, Wellington County, Town of Orangeville, Simcoe County, City of Barrie, Victoria County, City of Peterborough, and Peterborough County.

Representatives of the households were asked to report on trips made by all members of the household during the day before the interview was conducted. Travel information, demographic information, personal characteristics and modes of travel information collected in this data was the same as that of 1986 TTS survey plus some others such as questions about the work location, employment category, work at home for employed persons, possession of a transit pass, location of place of school and parking availability etc. The expansion procedure was similar to that of 1986 survey.

3.2 Descriptive Analysis of Urban Structure in 1964, 1986 and 1996

The GTA is located on the northwestern side of Lake Ontario, and is defined as the municipality of Metropolitan Toronto and the five regional municipalities of Durham, York, Peel, Halton, and Hamilton-Wentworth (see Figure 3.1). Its land area is about 8173 km² and has an approximate radius of 64 km.

It is important to have some knowledge of the characteristics of the study area, its population, and the differences among various parts of the region which are essential for travel demand. One of the first requirements is to compile an inventory of characteristics such as population, households, labor force, and car ownership. This information can be used graphically to explain the nature of the study area. Analysis can then be carried out in order to determine the relationship

between each variable and the need for travel. Consequently, one can forecast the future needs of the transportation system.

In the following five sections a preliminary examination of the three travel survey databases of 1964, 1986, and 1996 will be discussed in detail. Where possible comparisons of the results of these three surveys and their travel characteristics within the six regions of the Greater Toronto Area will be made.

3.2.1 Demographic and Socio-economic Changes

Travel patterns are considered to change with the socio-economic conditions, land use, transportation facilities, and travel behavior of the individual. The Greater Toronto Area has experienced considerable changes from 1964 to 1986 and from 1986 to 1996. These changes, which have greatly affected the observed travel patterns (i.e., land use, population size, and distribution, employment and labour force distribution, migration patterns within GTA), have had also an impact on trip distribution and trip length. Relative demographic factors that changed over the period of the last 32 years and are considered in this paper are described in the following sections. Each of these factors will be discussed in turn.

a. Population

In 1964 the population of GTA was 2,657,403 and grew to 4,062,943 in 1986 according to 1964 (MTARTS) and 1986 (TTS) surveys. By 1996 the Transportation Tomorrow Survey estimated the population to be 4,927,000 people. It is apparent from Table 3.1 that, since the travel survey of 1964 which was carried out as part of Metropolitan.

Table 3.2a: Distribution of Population of 1964, 1986 and 1996 by Regional Municipality

Regional Municipality	1964 Population Distribution by Region	Percent Distribution	1986 Population Distribution by Region	Percent Distribution	Percent Growth	1996 Population Distribution by Region	Percent Distribution	Percent Growth
Metro	1802670	65.8	2135016	52.6	18.4	2305558	46.8	8.0
Durham	168683	6.2	317886	7.8	88.5	450354	9.1	41.7
York	179708	6.6	344985	8.5	91.9	567689	11.5	64.5
Peel	170906	6.2	577032	14.2	237.6	812512	16.5	40.8
Halton	148488	5.4	264629	6.5	78.2	328264	6.7	24.0
Hamilton	268886	9.8	423398	10.4	57.5	461990	9.4	9.1
GTA	2739341		4062946			4926368		13.1

Source: 1964 MTARTS; and 1986 and 1996 TTS

Toronto and Regions Transport Study, the total population of the study area has grown 48% over the 22 year period (i.e. 1964 to 1986) and 85% from 1964 to 1996.

Over the last 32 years most of the population growth occurred predominantly in the regions outside of Metropolitan Toronto and Hamilton-Wentworth. Of a particular note are the regions of York and Peel, with Peel region having the highest average annual population and highest percentage growth of 238% between 1964 and 1996. However, in 1996 York had the highest percentage growth of 64.5%. As may be seen from Table 3.2a, Metropolitan Toronto still has the major concentration of population.

Durham, York, Peel, and Halton have shown an increased population growth over the past 32 years because of the decentralization. Metro Toronto and Hamilton-Wentworth, however, have had a net outflow of migrants since 1971.

The share of Metro of GTA total population has declined from 66 percent in 1964 to 53 percent in 1986 and 47 percent in 1996. The regional municipality of Peel obtained the largest increase in population - from 6 percent in 1964 to 14 percent in 1986 as shown in Table 3.2a. Thus, the growth has taken place predominantly in the suburbs of the GTA. Over the five years between 1981 and 1986 the average annual population growth rates had been 6.8 and 3.8 percent, respectively. At the time the smallest annual growth of 0.5 and 0.6 percent

occurred in the high density urban regions of Metro Toronto and Hamilton-Wentworth.

b. Household Size

While the population of the GTA increased by 48 percent from 1964 to 1986 and 21 percent between 1986 to 1996 there has been substantial growth in the number of households in GTA, from 712809 in 1964 to 1466074 in 1986, a 106% increase. This represents an average annual growth rate of 4.8 percent. The households in this region continued to increase to 1,805,000 in 1996; a 153% increase from 1964 to 1996. The average annual growth rate in households during the past 32 year period was 4.78%, which is marginally smaller than that of 1986. As Table 3.2b indicates, the average household size [persons/household], one of the major factors affect travel demand and urban density, decreased in 1986 and 1996 compared to 1964. This reduction is the result of the increased proportions in 1 and 2 person households from 26.8 to 50.1 percent between 1964 and 1986, where the proportion of more than 3 persons per household drastically decreased [see Table 3.3]. As one can observe from Figure 3.1 there is a downward trend in household size from 1961 to 1996. The household size in 1996, for example, was the smallest of these three surveyed years at 2.7, while 1986 became the second smallest and 1964 recorded the largest household size of 2.8 and 3.8 respectively [see Table 3.2b].

In 1961 household size ranged between 3.7 and 3.8 persons per household in the regions, while by 1996 this dropped to between 2.5 to 3.2 persons per household. The average household size in all those years was lower in the more urbanized areas of Metro and Hamilton. However, the regional municipalities of Durham, York and Peel kept their high average household size in these years as a result of fewer children per family, the rising divorce rate, improved standard of living, and the departure of young adults at an earlier age from the family home.

**Table 3.2b: Household size (persons per household) by Regional Municipality,
1961-1996**

Municipality	1961	1971	1976	1981	1986	1996
Metropolitan Toronto	3.7	3.2	2.9	2.8	2.6	2.5
Durham	3.7	3.5	3.3	3.1	3.0	2.9
York	3.7	3.6	3.4	3.3	3.3	3.2
Peel	3.8	3.7	3.4	3.3	3.1	3.0
Halton	3.8	3.6	3.3	3.1	3.0	2.8
Hamilton-Wentworth	3.7	3.3	3.0	2.7	2.7	2.6
GTA	3.8	3.4	3.1	2.9	2.8	2.7

Source: 1. Statistics Canada, Census of Canada, 1961-1986

2. 1996 Transportation Tomorrow Survey.

Figure 3.2: Household size for the GTA between 1961 to 1996

Source: Statistics Canada Census (1961-1991) and 1996 TTS

Table 3.3: Distribution of Population by Household Size

Household Size	1964 household Size Distribution		1986 household Size Distribution		Percent Growth 1964 -1986
	Number of Households	Percent Distribution	Number of Households	Percent Distribution	
1	39567	5.6	282736	19.3	614.6
2	151033	21.2	452050	30.8	199.3
3	131649	18.5	275320	18.8	109.1
4	164730	23.1	292411	19.9	77.5
5	112390	15.8	119369	8.1	6.2
6	61034	8.6	31143	2.1	-48.9
7	26676	3.7	9225	0.6	-65.4
8	13947	1.9	2641	0.2	-81.1
9	5765	0.8	691	0.1	-88.0
10 +	6013	0.8	486	0.0	-91.9

Source: 1964 MTARTS and 1986 TTS Databases

3.2.2 Socioeconomic Developments

Urban transportation demand also changed as a result of the socio-economics changes encountered in the study area. The composition of households also changed between 1964 and 1996. Increase in labour force participation has taken place in all the regional municipalities in the GTA for both males and females over the last 32 years.

Some social and economic factors and their pattern of development contributed to the growth of urban transport demand. The continuously flourishing economy in the province created new jobs, many of them being filled by women who have entered in the labour force during the past 25 years in an increasing number. The percentage of employed females which was about 19% in 1964, rose to 49% in 1986. The comparative number for male workers was 57% and 62% in the respective years. The female participation rate in the GTA was slightly more than half of that of the males, that is, the female rate was 45% in 1971 while the male participation rate posted to 78%. By 1986 the rate of the females rose to about 77% of the males and 83% in 1991, despite the fact that the male rate also marginally increased between 1971 and 1996. In the Toronto CMA, the 1971-1991 Census Canada reported that labour force participation rate increased from 45% in 1971 to 79% in 1991 [see Figure 3.3]. The other factor that contributed to

the increase in trip-making is the increase of the proportion of people of working age [16 to 65]. The

proportion of this age for females rose 13.5% between 1986 and 1991, and the male proportion increased to 14.1% in 1991.

Figure 3.3: Labour Force Participation Rate by Gender For Toronto CMA (1971 - 1991)

This dramatic change in the female participation rate accounts for the increase in female participation rate and also for the increase in the proportion of trip-making within the GTA and also for the increase in the proportion of households with more than one person working from 28.4% in 1964 to 49.2% in 1986 [see Table 3.4]. From Table 3.4 one can see the effect of the female participation in the labor force. The evidence is this, in 1964 for example, one

worker households constituted 59.1 percent where the households with 2 members working constituted 22 percent. However, in 1986

the one worker households recorded only 34 percent while the two workers households increased to 37 percent. In general, from 1964 to 1986 the households with two members working showed the greatest growth in the number of households. The percentage of the households with more than one member working increased from 28.4% in 1964 to 49.2% in 1986. Looking at the percentage growth of households with exactly two members working, it grew 252% in the respective years. The households with 3 or more members working have the largest growth and recorded 265% in 1986. The other factor that was parallel to the growth of female participation in labour force

Table 3.4: Distribution of Households by Number of Workers

Number of Workers	1964 Household Distribution		1986 Household Distribution		Percentage Growth
	Number of Households	Percentage Distribution	Number of Households	Percentage Distribution	
0	88626	12.4	250297	17.1	182.4
1	421164	59.1	494436	33.7	17.4
2	153451	21.5	540369	36.9	252.1
3+	49566	6.9	180971	12.3	265.1

Source: 1964 MTARTS and 1986 TTS

and also contributed to the higher travel demand was the driver's license ownership, in which the number of female workers surpassed the male counterparts by recording 44% in 1964 and 78% in 1986 see Table 3.5:

Table 3.5 Percentage of households segmented by gender with driver's license

Percentage of female - Male Workers With Driver's licence	Female	Male
1964	44%	89%
1986	78%	94%

Source: 1964 MTARTS and 1986 TTS

Automobile availability also increased between 1964 and 1996, which caused the high growth of trip-making in the GTA. Between the period of 1980 and 1996 the average annual growth in automobile ownership in the study area increased 8.7%. The average number of automobiles in the household in the GTA increased from 1.0 in 1964 to 1.4 in 1986. The growth in vehicle ownership in the GTA increased from 0.7 million cars in 1964 to 2.1 million cars in 1986, a 190% increase. See Table [3.6]. This Table presents the 1986 vehicle ownership by regional municipality. Despite the relatively small geographical area of Metropolitan Toronto, the general information contained in Table 3.6 clearly demonstrates that Metro has the largest percentage of vehicles and more than half the population, households, and licenced drivers. Nevertheless, the vast majority of the growth in both population and households is now taking place in the regions surrounding

Metro Toronto. Metro Toronto has the lowest level of car ownership [1.22 cars per household and 0.47 cars per person], reflecting its dependence on the TTC in 1986. The highest level of car ownership was in York region which led in vehicles per household [1.89] and in Halton, which led in vehicles per person [0.6]. Please see Table 3.6.

Table 3.6: 1986 Database Car Ownership

Municipality	Number of Vehicles	Total Household	Vehicles per Household	Total Population	Vehicles per Person
Metro Toronto	984840	820700	1.22	2134700	0.46
Durham	179140	106000	1.69	317900	0.56
York	199280	106000	1.89	345000	0.57
Peel	323164	186800	1.73	577000	0.56
Halton	158752	90200	1.76	264600	0.6
Hamilton	212568	156300	1.36	423400	0.52
GTA	2081720	146300	1.42	4062900	0.51

Source: 1986 Census Canada

The distribution of households by vehicle ownership is presented in Table 3.7. This distribution indicates that between 1964 and 1986 there was a decline in the percentage of the households that did not own a vehicle from 20 percent to about 15 percent. This increase of vehicles in these households will eventually increase the number trips made in 1986. Secondly, from Table 3.7 the households with 2 vehicles in 1964 increased from about 16 percent to 33 percent, a 330 percentage growth in 1986. The percentage of households with 3 or more vehicles increased

from 2.1 to 9.8., a percentage growth of 858.7 between these years. Generally, household with multiple vehicles has a tendency to make more trips than those households with 1 or no-vehicles at all. This increase in the auto availability to members of the households is another factor contributing to the high growth of trip-making within the study area.

Table 3.7: Distribution of Households by Number of Vehicles Owned

	1964 Household Distribution		1986 Household Distribution		
Number of Vehicles	Number of Households	Percentage Distribution	Number of Households	Percentage Distribution	Percentage Growth
0	143324	20.1	217461	14.8	51.7
1	443163	62.2	625369	42.7	41.1
2	111332	15.6	479537	32.7	330.7
3+	14988	2.1	143704	9.8	858.7

Source: 1964 MTARTS and 1986 TTS

The 1986 survey indicated that a substantial increase in trip making has occurred. It measured 8.7 million trips a day generated by residents in the GTA in 1986. This compares with approximately 3.7 million trips measured in 1964 by the MTARTS survey.

The amount of travel in the Greater Toronto Area is increasing at a time when the rate of population growth is declining. For example, since the survey in 1964, the

total population of the area has grown 48 % from 1964 to 1986, and over 80 % from 1964 to 1996 or [21 % from 1986-1996].

Table 3.8 presents the average number of morning peak-period trips made per person, or per household. From this table one can see that all the regional municipalities except Hamilton Went-Worth experienced an increase in the average number of morning peak-period trips made per person or per household. Hamilton had the greatest average number of trips for both types of trips considered in 1964, followed by Metro Toronto and Halton. However, in 1986 York and Peel Regional Municipalities had the largest average number of trips by person or by household.

Table 3.8: Trip Rates by Regional Municipality-Morning Peak Period Trips

Region	1964 Trip Rate		1986 Trip Rate		Difference	
	Household	Person	Household	Person	Household	Person
Metro Toronto	0.96	0.25	1.28	0.49	0.32	0.24
Durham	0.60	0.15	1.39	0.47	0.79	0.32
York	0.57	0.14	1.63	0.50	1.06	0.36
Peel	0.78	0.19	1.60	0.52	0.82	0.33
Halton	0.94	0.24	1.41	0.48	0.47	0.24
Hamilton	1.10	0.29	1.10	0.41	0.00	0.12

Source: 1964 MTARTS and 1986 TTS

Interestingly, the average number of trips by household in Hamilton Went-Worth remained unchanged between 1964 and 1986, while its trip rate by person changed

from 0.29 to 0.41. This might be due to the decline in household size in the region between 1964 and 1986.

3.2.3 Analysis of Trip Start Times For Total and Home-based Work Trips in 1964,1986 and 1996 Databases

A trip may be defined as a single journey between a place of origin and a place of destination. Urban trips are normally classified according to two groups of attributes: Home-based trips and non-home-based trips.

The main purpose of studying the temporal distribution of trip making in urban areas is to understand the major characteristics of peaking of trip-making at certain times of the day. This peaking is an increase of volume in a certain segment of the transportation system and consequently will have a major effect on the transportation system. This peak volume is very important because it can be used to deduce the demand patterns of urban travel and the frequency of travel on an average weekday varies with the trip time.

Figure 3.4 shows total trip by trip start time for 1964, 1986, and 1996 data bases, while Figure 3.5 shows the same data expressed as frequency distribution. The distributions illustrated in Figure 3.4 have the characteristics of those observed in many urban areas. In Home-Based Work (HBW) Trips, the pattern of trips is determined by the locations of the residences of workers and the locations of their

work places. The peaking of HBW travel between residential areas and places of employment occurred between 6:00-9:00 a.m. in 1964, and between 7:00-9:00 a.m. in 1986 and 1996 in the morning, and between 4:00 - 6:00 p.m. in the evening. In that case the distribution of the total daily trips will produce a double-peaking system. Secondly the places of employment are concentrated in a few locations such as the central business area district (CBD), thus causing congestion in the major arterials. There is a relatively small peak from 12:00 p.m. to 2:00 p.m. which indicates that people tend to conduct these trips during lunch hours. The information presented in Figure 3.4 identifies the influence of work trips and their importance on the transport study.

Figure 3.4 : Comparison of Trip Start Times for Total Person Distribution in GTA by All Modes for 1964, 1986, and 1996

Source: 1964 MTARTS, and 1986 and 1996 TTS Databases

Figure 3.5: Comparison of Trip Start Time Percent Distribution for Total Trips by All Modes for 1964, 1986, and 1996.

Source: 1964 MTARTS, and 1986 and 1996 TTS Databases

Figure 3.6: Comparison of Trip Start Time Distribution for Home-based Work Trips by All Modes in the GTA for 1964, 1986, and 1996

Source: 1964 MTARTS, and 1986 and 1996 TTS Database

Figure 3.7: Comparison of Trip start Time percent Distribution for Home-Based Work Trips by All modes for 1964, 1986, and 1996

Source: 1964 MTARTS, and 1986 and 1996 TTS Databases

The impacts of the increased trips on worker's departure time for all trip purposes and for home-based work trips are explored below. It is standard practice to use the work trips for planning purposes since it generates the highest traffic volume during the peak periods and it provides the maximum number of trips that can be made during the day. From the peak trips one can determine the capacity of road network and the transit system.

The trip start time distribution for HBW trips is displayed in Figures 3.6 and 3.7. Figure 3.6 shows that in 1964, 1986, and 1996 data, there was an increase in volume in every hour of the day. All the trips considered in all those years

consistently indicated that the strongest peaking is exhibited for morning work trips. Home-based work trips exhibited the typical peaking characteristics.

In general, the peak travel periods for the auto mode are somewhat more spread than the corresponding transit peak periods, probably reflecting flexibility in trip timing provided by the auto mode. Similarly, school trips are generally more peaked (especially during the morning) than work travel [Susan Hanson, 1986]. Since most of the workers start their work early in the morning between (7:00-9:00 a.m.) and return home after an 8:00 hour shift (that is between 4:00-6:00 p.m.) the distribution of the total daily trips will produce a double-peaking system. Secondly, the places of employment are concentrated in a few locations such as the central business district (CBD), thus causing congestion in the major arterials.

Examining the start times of total trips and home-to-work trips in Figures 3.4 and 3.6 respectively, one can infer that the non-work and non-school trips such as home-based shopping trips etc., do not exhibit the two-peak distribution characteristics of the work and school trips [Miller et al, 1990]. Rather, in the case of auto trips, the frequency of trips tends to grow as the day progresses, with the peak occurring in the early evening from 8:00 to 11:00 p.m.. Transit trip-making is even more evenly distributed over the day, with the peak probably actually occurring somewhere around 1:00 PM.

A comparison of the data of 1964 , 1986 and 1996 indicates the start time distribution has changed very little over the past 32 years. Figure 3.4 indicates that the travel volumes increased in 1986 and 1996 periods. Some variation in the trip start time distributions exist among the three travel surveys investigated. The figure shows that in both 1986 and 1996 the peak period shift to the left (started earlier) in both morning and evening peak hours, that indicates that a large number of workers started their work trip earlier because of the increase in volume.

As can be seen from the Figure 3.5 the travel volumes increased for almost every hour of the day. This increase of total travel volume during peak time over this 32 year period also spread in time causing longer peak duration in both mornings and evenings. However, when examining the percentage distribution of trip start time, it is observed that there was still some similarity between 1964, 1986, and 1996 trip start times. The trip start time for Home-based work trips indicated similar characteristics to those of all trips/all modes described above.

3.2.4 The Importance Of Planning District 1 to the GTA

The central business area of most of the urban areas contains the majority of the central retail and commercial employment. It has a high concentration of trips originating or ending in this area. Planning District 1 is the central core of Metro

Toronto. It represents a quarter of the area within the City of Toronto. It covers an area, excluding the harbour, of 20 km² within the 8700 km² GTA (see Figure). It is the business centre of the study area. It is where most of the local and Provincial offices are available. Unlike many other North American cities, Planning District 1 has grown tremendously, even though the suburbs centres also grew quickly.

In 1964 MTARTS suggested that about 37% of the regions employment was in the Toronto Core Area [Province of Ontario, 1966]. Even though the percentage of the GTA home-to-work trips that ended in PD1 are decreasing because of the growth of employment in the suburbs, still the absolute level of employment in this area continues to rise.

Although each Regional Municipality has its commercial, industrial and its separate central core area, each of the four regions surrounding Metro Toronto has made a high proportion of trips to Planning District 1 according to the 1964, the 1986 and the 1996 data bases. In 1964, 1986, and 1996, 30%, 21%, and 20% of the morning peak period trips destined into PD1 respectively. In that case, the central area remains a fundamental concern for transportation planners.

3.2.5 Changes in Transportation Infrastructure

The growth in population, employment, and automobiles in the GTA required the construction of a system of highways in order to maintain satisfactory level of services. Important changes have taken place in the transportation infrastructure in the last 40 years including the construction of subways, rapid transit, expressways, commuter rail facilities and improvements in accessibility which caused increased and longer trips in the GTA.

Both the highway and transit networks changed between 1964 and 1986. Six major highways helped the accessibility of the high travel demand in the study area. This high demand was mainly caused by the growth of auto usage between those years. Highway 401, which was build in the early 1950s and is the main artery of the all highway systems in Metro, provides an east-west access across the north part of the city from Halton to Durham. The QEW and the Gardiner Expressway provide similar access along the Lake shore. Highway 400 and Highway 427 provide north-south access across the western part of the district and Highway 404 and Don Valley Parkway provide again north-south access across the eastern part of Metro. Allen road ends at Eglinton Avenue in the northwest and connects to the 401.

Most of the highways and expressways in GTA were built before 1964. However, the subway system was considerably extended during the 1964-1986 period. In addition the Government of Ontario [GO] Transit service was inaugurated in 1967. Go Transit services the long distance commuters between the regions. Its network and stations are shown in Figure 3.9.

Table 3.9: GO Transit System:

	1986	1996
Passenger carried	26,917	33,529
Route kilometers (km) (Rail & Bus)	1444	1621
Kilometers Operated (x 1000 Km)	15,991	16,442

Source: 1986, 1996 Go Transit Annual Reports.

GO Transit provides commuter bus and train service in the GTA and surrounding areas. It is a crown agency created to carry commuters whose travel requires them to travel through more than one regional municipality. Bus and rail passengers increase at a steady rate: the combined system [bus & rail] carried nearly 26 million people during the fiscal year of 1986-1987, a million trips more than the previous fiscal year. Since the GO Transit was not in operation in 1964, we tabulated the annual ridership, kilometers operated in 1986 and 1996 in Table 3.9. The largest transit system that operates is the Toronto Transit System (TTC) which has a high frequency operation only in Metro. There are 15 other local

transit properties that operate the other regional municipalities outside Metro, but they are smaller in frequency, ridership and network system than the TTC. Figure 3.10 depicts the TTC subway and RT System.

Figure 3.9: The Six Rail Corridors Network and Stations that GO Transit Operate in the GTA.

Source: 1986, 1996 Go Transit Annual Reports.

Figure 3.10: TTC Subway and RT System

Source: TTC Annual Report 1995

Table 3.10: Changes in 1964 - 1996 GTA Transportation System

Table 3.10: Changes in 1964 - 1996 GTA Transportation System

	Year		
	1964	1986	1996
Passengers Trips (million)	275.3	441.0	372.4
Automobiles Registered in GTA (million)	0.541876	1.996	?
Total Vehicles Owned In GTA (million)	0.622005	2.315	?
Transit Total Vehicle Fleet	1802	2720	2757
Transit Total Kilometers Operated (million)	88.3	189.12	?
Transit Route Kilometers	953	1,345	?
Municipality of Toronto's Streets & Roads (Kilometers)	?	4,960	?
Metro Expressway (Kilometers)	86	126	?

Source: 1. Statistics, Road Motor Vehicle Registration, 1980 and 1986.
 2. The Municipality of Metropolitan Toronto, 1988 and 1986.
 3. TTC Annual Report, 1964, 1986 and 1996.
 ? The data was not available

particularly the increase in female participation in the work force, increase in automobile ownership which doubled, for example, between 1964 and 1986, reflect more trips and different trip patterns. Interestingly, the ridership decreased between 1986 to 1996 from 441 million to 372.4. This might be due to the migration of many people to the suburbans and changed their mode of travel from transit to auto. With all those changes in the study area, the question is "are the cross-sectional models of trip distribution " capable of yielding credible travel forecasts in 1986 and 1996?

CHAPTER 4

ANALYSIS METHODS

4.1 Model Calibration Techniques

Several doubly-constrained entropy models will be discussed in this study to investigate their transferability. Each model is specified by writing a set of constraints and then solving for the maximum entropy solution. Entropy maximization is used in this study because of its logical mechanism for generating and testing alternative impedance function specifications. It also reduces the totally unrestricted, ad hoc approach inherent in the gravity model. Lastly, its simplicity is beneficial for the use of powerful, efficient computerized calibration techniques. All of the models considered in this thesis, except the *proportional flow model* which we will use as a base model, possess a set of parameters which represent the sensitivity of trip-making behaviour towards certain system variables [a good example being travel time]. For the purpose of the model derivation, these parameters are assumed to be known, but in reality the models must be calibrated in order to determine or estimate the true behaviour of the model [i.e. the values of the parameters]. These parameter estimates

are obtained through certain procedure searches for that set of parameter values that “best and closely fit” the observed data.

There are several different search procedures that can be used for estimating these parameter values. In all cases, according to Miller and Cubukgil [1981], the calibration process is regarded as consisting of the following components:

1. a definition of what is meant by the term “best fit to observed, historical data”;
2. a procedure or algorithm for finding the “best fit” parameters, given the chosen definition of “best fit”;
3. development of one or more measures for evaluating the results of the calibration; that is, measures which determine how good the model’s fit is in both absolute terms and relative to the fit of alternative models.

For entropy-based models, the maximum likelihood estimation [MLE] parameter estimates are simply found by solving for the set of parameter values which cause the model constants to hold [see Miller and Cubukgil, 1981]. The Newton-Raphson root-finding procedure is used for this task. This procedure has been implemented with a FORTRAN 77 program [calib.f] developed by Professor Eric Miller at the University of Toronto. This program also computes parameter asymptotic t-statistics, as well as variances. The goodness-of-fit measures, such as mean absolute error and log likelihood at convergence, are also used to evaluate how well these models fit to the data.[See Appendix C] for equations for all goodness-of-fit measures.

4.2 Outline of the Model Systems under Investigation

A full description of the structure and the component of models used in this study is presented below. The following subsections outline the models tested in this thesis and the constraints to which they were subjected. The input parameter files and their output results are summarized in Appendix D.

There are some major hypotheses postulated in this study. *The first hypothesis is that, all else being equal, as the travel time from the place of residence to place of work becomes longer, the number of trips is reduced.* Thus, intuitively, one expects that the parameters of all the level of service variables, particularly travel time, will have a negative sign. A higher value of the travel time parameter would indicate that the trip numbers are very sensitive to the travel time or cost, and in that case there will be few long trips. However, a low value of travel time parameter would indicate an insensitivity to the travel cost that leads to a relatively large number of long trips.

The second hypothesis is that disaggregation of the work-trip flows by space or socioeconomic segments is important towards achieving a better understanding of work trip-making behaviour. This will also result in better model performance and enhance model transferability. In order to obtain a model that closely fits the observed data and application contexts, separate models were calibrated and compared with the parent model (the *auto travel time model*).

Three alternative model specifications, based on different assumptions and having different data requirements, were developed in this study. The first is the *proportional flow model* [PFM] which is established as a base for the three models developed. In this model, the future number of trip ends estimated for each origin zone is shared over possible destination zones in proportion to the number of trip ends estimated for each destination zone. Only the future numbers of zonal trip ends are required and no use is made of the base-year trip distribution matrix.

The underlying assumption of the *proportional flow model* is that the size of the flow between the zones depends on the level of activity within those zones. The model does not in any shape or form take account of travel impedance [i.e. the difficulty of traveling] between zones of traffic movement. Because this model is totally insensitive to transport network effects and inappropriate for normal planning applications it was just developed for comparison reasons only. Two categories of mathematical models are used in trip distribution step. *Growth-factor models* (such as Fratar) start with a stable matrix of zonal interchanges for a base year and factor it up to match forecast totals. *Synthetic models* (such as Gravity and Proportional Flow Models) do not require an existing table. Gravity models estimate the interchanges on the basis of travel time and cost. While proportional flow models disregard travel time and cost as an explanatory variables and distribute the trips into the matrix proportionately.

Auto Travel Time Model Estimated for 1964 and Transferred to 1996 [Strategic Transfer]

Figure 4.1: The Specification of 1964 Model and its Transferability into the 1996 Application Context.

The second model is the simple *auto travel time model*. This model uses auto travel time to describe the effect of spatial separation on trip interchange between zones in the study area. Most transportation planning studies report that travel time is one of the most significant variables that affect travel behaviour, particularly work trips [O'Flaherty, 1997].

The third model considered for this study has a "generalized cost" as an impedance function and this is based on the work trip distribution model presented in Miller [1992]. The zone-to-zone [Auto and Transit] travel times [min.] were calculated from EMME/2. The Rail travel time used is the best [minimum] average total weighted travel time [min] across the three rail modes [see model 3 below], and it is estimated from EMME/2. The Walk distance is the centroid-to-centroid straight line distance [km] if this distance is less than or equal to 5.0 km; otherwise the walk mode is not considered available for this O-D pair and this variable is set equal to zero.

The third model is constructed to see the effects of the different impedance functions used in this study, that is, the "auto travel time" and the "generalized cost". The results of this model using the generalized cost impedance function was compared to Model 2, using 1986 data. The third model was neither segmented nor transferred to any other application context.

It became apparent [as can be seen in chapter 5] that this model may not adequately capture or reflect all the travel behaviour of the workers, and travel time per se is not

necessarily representative of the transport service to all parts of a zone [Miller et al, 1990]. One good choice could be to develop a separate model for each trip category in order to clearly visualize the peculiarity of each sub-model.

The procedure used to develop these market segments consists of incrementally segmenting the main morning peak period work trips according to the considered variable and assessing the marginal impact of these variables on model performance [primarily in terms of its goodness-of-fit, as well as the signs and the magnitudes of the estimated parameters]. The objective of this segmentation is to explore the extent to which changes in the specification of the model influenced the explanatory variable and forecasting capabilities of the trip distribution models.

Some factors that were identified in the descriptive analysis to be probable contributors to the differences in trip making were gender, possession of driver's license and occupation [see Figure 4.1]. Some trips concentrated in the central business district [CBD] were investigated separately, due to its extensive transit system and other unique features. While there are distinct spatial patterns in the residential distribution of the central area [PD1] workers, we segmented the second model [simple auto travel time model] into work trips which ended in downtown Toronto [PD1] and those that ended in the remaining part of the GTA denoted here as "Other". This was done to investigate whether each of these regional segmented areas has the potential of being represented by a unique parameter and unique model.

as "Other". This was done to investigate whether each of these regional segmented areas has the potential of being represented by a unique parameter and unique model. This attempt at segmentation did not, however, appreciably improve our understanding of travel behaviour [see Chapter 5].

In 1964 males were typically the family "bread winners" and one might assume had the priority in using the car to travel to work. In addition, many women did not have driver's licenses. By 1986 and 1996, a large number of females had joined the labor force and possessed driver's licenses and possible competition for car was evident. Therefore gender, holding driving licenses and car ownership were expected to play a significant role in explaining changes in the observed spatial interactions among the 1964, 1986 and 1996 databases. Also the female participation rate in labor force increased significantly between 1964 and 1986 but increased marginally between 1986 and 1996. In 1964, the work trips by females recorded in the morning peak period were 26 % of the total trips made in the morning period. This figure increased to about 44 % in 1986 and 46 % in 1996. In order to understand the effect of the new female labor force on the work trips, we segmented the data by male and female workers.

Another segmentation made on the basis of the above observation was that the workers with driver's licenses and at least one car available in their household were modelled separately from those without driver's license.

A third segmentation was made to investigate the wider effect of the socio-economics of workers on their trip-making. The 1964 and 1996 employed labor force was segmented into four distinct occupation groups. White collar workers [Professionals/Management/Technical] formed Group 1; General Office and Clerical workers formed Group 2; Sales and Service workers formed Group 3; and lastly, blue-collar workers [Manufacturing/Construction/Trades] made Group 4 [see Figure 4.1.

All these specifications and segmentations were aimed to update the auto travel time model by making adjustments based on locally 1964 and 1986 available information before transferring to the application contexts. Each given population or segment defined above in terms of space or socio-economic characteristics was modeled separately and tested for the ability to transfer to the application context. Each particular model provides a description of how each group behaves in space. The only difference between the model segments made and transferred from the 1964 context and those from the 1986 context is that of occupation category, for which the information was not available in the 1986 database. The models tested will be discussed in the following three sections.

4.2.1 Model 1: The Proportional Flow model

The *Proportional Flow Model* is a method to distribute trips in a study area. In this model, the trip production of each zone is distributed to the all possible destinations in proportion of the trip attraction of that zones.

The underlying assumption of the model is that the number of trips between the zones depends on the level of activity within those zones. The model does not in any shape or form take account of impedance [i.e. the difficult of traveling] between pairs of zones. This model was developed to estimate traffic zone-to-zone travel volumes during the morning peak period. The simulated traffic volumes were then compared to the observed [survey] trips on a planning district basis. This model is not computed to provide a realistic representation of work trip flows within the GTA. Rather it was developed simply to see the extent to which the “row and column constraints” [i.e. the assumed known zonal trip origins and destinations) determine the model results. The results and the performance of this model were further compared with those of the auto travel time model.

Distribution of home-to-work trips is a long term process, because it can only be changed if the worker changes either his/her residential location or employment location. Otherwise the worker will probably select the same routes to work and the same destination [the work place].

The trip matrix below has a dimension of $m \times n$, where m is the number of origin locations and n the number of destination locations [Adib Kanifani, 1978].

Let O_i be the total number of trips emanating from zone i [trip production of i] , therefore, the row sum of $\{t_i\}$ equals the following:

$$O_i = \sum_j t_{ij} \quad [4.1]$$

Similarly, let the column sum D_j be the total number of trips destined to j [trip attraction];

$$D_j = \sum_i t_{ij} \quad [4.2]$$

Table 4.1 Trip Matrix Table

	1	2	3	4	5	6	O_i
1	t_{11}	t_{12}	t_{13}				O_1
2	t_{21}						O_2
3	t_{31}						O_3
4	t_{41}						
5	t_{51}						
6	t_{61}						
D_j	D_1	D_2					$T = \sum D_j = \sum O_i$

Let T be the grand total of the Matrix (i.e. the total number of trips in the system);

$$T = \sum_i \sum_j t_{ij} = \sum_j O_i = \sum_i D_j \quad [4.3]$$

From the above equation, D_j is the total trips demanded in the trip table [or a destination Matrix]. This destination matrix can be conveniently normalized by dividing all cells by T so that each cell contains t_{pij} , the proportion of the total system demand that is between a particular origin-destination pair i and j .

$$t_{pij} = \frac{t_{ij}}{T} \quad [4.4]$$

Table 4.2 Proportion of Trips in Trip Matrix

	1	2	3	4	5	6	U_i
1	t_{p11}	t_{p12}	t_{p13}	-	-	-	u_1
2	t_{p21}	t_{p22}	-	-	-	-	u_2
3	t_{p31}	t_{p32}	-	-	-	-	u_3
4	t_{p41}	t_{p42}	-	-	-	-	
5	t_{p51}	-	-	-	-	-	
6	t_{p61}	-		-	-	-	
V_j	v_1	v_2	v_3	-	-	-	$=1$

$$U_i = \frac{A_i}{T} = \sum_j t_{pij} \quad [4.5]$$

where A_i is the average total trip production of i

$$V_j = \frac{B_j}{T} = \sum_i t_{pij} \quad [4.6]$$

where B_j is the average total trip destination of j

$$= \sum_i U_i = \sum_j V_j = \sum_{ij} t_{pij} \quad [4.7]$$

4.2.2 Model 2: Auto Travel Time Model

All the models investigated will exclude all intra-zonal trips and K-factors models. The second model explored uses the *auto travel times* as the system variable on all traffic zones. The constraint that is applicable to this model in addition to the constraints [2.6] and [2.7] of section [2.1] is:

$$\frac{\sum_i \sum_j T_{ij} a_{ij}}{T} = \bar{a} \quad [4.8]$$

where

a_{ij} = auto travel time from zone i to zone j [i.e., assumed auto travel time]

T = total observed trips for the system

$\sum_{i=1}^n \sum_{j=1}^n T_{ij}$ = Summation of T_{ij} on i and j.

$\bar{a}$ = observed average auto travel time

The above mentioned constraints are then used to maximize the entropy function,

$$S = - \sum_i \sum_j T_{ij} (\log T_{ij}) \quad [4.9]$$

This equation measures the uncertainty in the system, where the parameter β is dependent upon the observed average auto travel time for the system. Maximizing [4.8] subject to [2.6], [2.7] and yields the following model.

$$T_{ij} = A_i B_j O_i D_j \exp(-\beta a_{ij}) \quad [4.10]$$

where

$$A_i = \left[\sum_j B_j D_j \exp(-\beta a_{ij}) \right]^{-1} \quad [4.11]$$

$$B_j = \left[\sum_i A_i O_i \exp(-\beta a_{ij}) \right]^{-1} \quad [4.12]$$

4.2.3 Model 3: "Generalized Cost" 1986 Model

The main difference between the auto travel time model and this model is the functional form of the impedance function, f_{ij} , used. In this model, a quite complex impedance function is used. One of the main reason, that motivated using this complex impedance function is the multimodal nature of the GTA transportation system. The road system, the vast transit system, and the GO-Rail system all influence the distribution of the origin and destination linkages.

In that case it became apparent that it was necessary to consider all the friction factors that generate the disutility of making a trip. As mentioned earlier in section 3.1, seven modes were considered available in 1986 and 1996 databases. In order to include modal sensitivity within the trip distribution model, some of the modes were combined. The following four modes were explicitly used in this model:

1. Auto-allway [as either a driver or a passenger];
2. Transit-allway (excluding trips using GO-Rail, as well as trips in which an auto is used to access the subway system);
3. GO-Rail trips plus subway with auto access, and
4. Walk-allway.

The following are the rules establishing model feasibility, and the definition of their level of service variables:

1. **Auto:** zone-to-zone auto travel time [minimum] as estimated from EMME/2 [user-equilibrium road network assignment]. The auto drive mode is considered feasible if the worker has a valid driver's licence and the household has at least one car available.
2. **Transit:** zone-to-zone total weighted transit travel time [minimum]; that is in-vehicle travel time plus two times the out-of-vehicle travel time; as estimated from EMME/2 [aggregate transit network assignment]. This travel time is equal to zero if the transit mode is not available for a given O-D pair zone.
3. **Rail:** the travel time is equal to the best [minimum] across the three rail modes of subway with auto access, GO-Rail with transit access, estimated from EMME/2. This rail transit time is equal to zero if no rail was available in that O-D pair.
4. **Walk:** the walk distance measured by the Euclidean [straight-line distance, km] centroid-to-centroid distance represents the impedance variable for the use of this mode.

The above variables and their pertinent constraints were used in this model, in addition to the balancing factors discussed in model 2.

Constraints for this system are as follows:

$$\sum_j \sum_m T_{ijm} = O_i \quad \text{for all } i \quad [4.13]$$

$$\sum_i \sum_m T_{ijm} = D_j \quad \text{for all } j \quad [4.14]$$

$$\frac{\sum_i \sum_j T_{ijm} X_{ijm}}{\sum_i \sum_j T_{ijm}} = X_m \quad [4.15]$$

where

T_{ijm} = trips from i to j by mode m;

X_{ijm} = level of service variable for mode m, evaluated for O-D pair ij;

X_m = average observed value of level of service variable X for mode "m" variable.

The physical interpretation of constraint [4.15] is that the average value of X for mode m predicted by the model for the base year must equal the observed average.

Another constraint was also imposed in this model to make sure that the model correctly predicts the fraction of observed trips using each mode in the system, that is:

$$\frac{\sum_i \sum_j T_{ijm}}{\sum_i \sum_j T_{ij}} = S_m \quad m=1, \dots, (M-1) \quad [4.16]$$

where:

S_m = observed factor of trips mode by mode m;

M = total number of modes;

Only [M-1] such constraints are imposed since in logit models only [M-1] model constants are statistically identifiable.

When constraints [4.15] and [4.16] are imposed on the model in addition to constraints [4.14] and [4.15] and solved for the entropy maximizing set of T_{ij} 's one obtains:

$$T_{ijm} = A_i B_j O_i D_j \exp(\alpha_m + \beta_m X_{ijm}) \quad [4.17.1]$$

$$T_{ij} = A_i B_j O_i D_j \left\{ \sum_m \exp(\alpha_m + \beta_m X_{ijm}) \right\} \quad [4.17.2]$$

$$A_i = 1 / \left[\sum_j B_j D_j \left\{ \sum_m \exp(\alpha_m + \beta_m X_{ijm}) \right\} \right] \quad [4.17.3]$$

$$B_j = 1 / \left[\sum_i A_i O_i \left\{ \sum_m \exp(\alpha_m + \beta_m X_{ijm}) \right\} \right] \quad [4.17.4]$$

Where:

α_m = modal constant for mode m; = 0 for mode M

β_m = level of service parameter for mode m

and the summations in equations [4.17] are over the set of modes which are feasible for O-D pair ij [i.e. $X_{ijm} \neq 0$].

From the above argument regarding the prediction of the observed number of trips by each mode in equations [4.17.1] and [4.17.2], writing mode based constraints such as [4.15] results in an implied mode split model of the form:

$$P_{ijm} = \frac{T_{ijm}}{T_{ij}} = \frac{\exp(\alpha_m + \beta_m X_{ijm})}{\sum_m \exp(\alpha_m + \beta_m X_{ijm})} \quad [4.18]$$

where

P_{ijm} = Proportion of trips from origin i to destination j by mode m

The alternative specific constants, α_m captures any unobserved characteristics not included in the model. They will ensure that these implied probabilities will replicate the aggregate modal shares in the base data.

The other important point that should be noted in designing the trip distribution model impedance function is that the GTA includes two major and one medium-sized work trip commuting systems which have overlapping commuter sheds: the Toronto CMA, the Hamilton-CMA and the Oshawa Census Agglomeration [CA]. Because the Toronto CMA commuter shed overlaps with Oshawa's to the east and Hamilton's to the west, all three must be considered in a Toronto-centered model.

It is not easy to construct an impedance function for this model considering all the above factors. From earlier experimentation [Miller, 1997], it has been found that a single impedance function for each mode included in the model could be used across the GTA, as long as different model constants were used for different major commuter market segments.

The function finally used for this model takes the form

$$f_i = \sum_{m \in M} \exp(V_{im}) \quad [4.19]$$

$$V_{ijm} = \left(\sum_k \alpha_{km} \delta_{ijk} \right) + \beta_k X_{ijm} \quad [4.20]$$

Where:

M_{ij} = set of modes available for work trips from zone i to zone j;

V_{ijm} = “disutility” or “generalized cost” of the m^{th} mode available for O-D pair i,j;

α_{km} = constant terms for mode m for spatial market segment k;

δ_{ijk} = 1 if O-D pair ij is constrained with spatial market segment k; = 0 otherwise.

Four “spatial market segments” are defined in this model:

1. O-D pairs for which the destination zone lies within Planning District 1;
2. O-D pairs for which the destination zone lies within the Regional Municipalities of York, Peel and Halton [excluding Burlington];
3. O-D pairs for which the destination zone lies within the Oshawa Census Agglomeration; and
4. O-D pairs for which the destination zone lies within the Hamilton Census Metropolitan Area.

4.2.4 Travel Time Variable

Auto travel times are obtained from the observed 1986 peak hour auto O-D values. These peak hour O-D values are obtained by taking the product of the peak period auto travel value matrix for all trip purposes and a factor O-D which represents the

percentage of all peak period auto travel occurring in the peak hour. The access time for auto drive was considered negligible. The parking location and walk time to final destination were not available from 1986 survey. The same procedures were used for the calculation of auto travel times in 1964 and 1996, with factors of 0.46 and 0.42 respectively.

Since 1964 data were not geocoded, disaggregate transit assignment was not possible and to maintain the compatibility of all the three matrices, the aggregate assignment procedure was used for all the three data sets of 1964, 1986, and 1996.

Travel demand analysts generally accept the principle that no single model specification is adequate for all travel purposes. The common practice of estimating sub-models to identify which variables have significant effect on travel behavior as an attempt to reduce the problems of specification by a process of market segmentation. In general, the failure of a model to transfer may imply that the model specification used was not adequate to explain the differences in the observed behavior [Ben-Akiva, 1979].

When the model does not appear transferable, the options here are either to improve the model's specification or to segment the travel modes to reduce or eliminate the source of mis-specification. We tested whether these different market segments exhibit significantly different tastes in weighing the importance of the factors hypothesized in page 71 to influence the spatial interaction of the workers.

4.3 Goodness of Fit (Model Performance Measures)

In order to investigate the relative measure of improvement in fit, model number 2 *auto travel time model* [ATTM] was called here the parent model for the sake of our discussion. This parent model was segmented into two or more market segments. The log-likelihood values of these temporary descendent samples were summed up and compared with the log-likelihood of the parent models. This is to examine if these segmented models can improve the goodness-of-fit.

The second model evaluation measure used in this study is the residuals. The difference between the actual and estimated trips in each planning district toward Planning District 1 was noted. The tabulation of these trip prediction errors from each Regional Municipality toward Planning District 1 were aggregated and tabulated at the end of Chapter 5.

The third goodness of fit measure used is the trip length frequency distribution. The observed and estimated traffic volumes in each five minute time intervals defined is compared. The results of the model examined are presented in Chapter 5. Lastly, the statistical significance of the estimated parameters was assessed [t-statistic]. The asymptotic t-statistic is used to test whether any particular parameter β differs from zero or other known value at a chosen level of significance (in our case 5 percent). The critical values for the t-statistic are percentiles of a standardized asymptotic

normally distribution ($N \rightarrow \infty$), which for one tail tests at the frequently used significance level of 0.05 is 1.64. In mode choice the lowest value of rho-square acceptable is 0.2, however, in our case, because of the higher degree of freedom, the values of rho-square are very low (see chapter 5).

4.4 Transferability Tests

4.4.1 Measures of Transferability

There are various transferability measures and tests in application to the model transferability. The following were selected for this work-trip distribution model for evaluating the model transferability.

- 1) Tests of disaggregate prediction
- 2) Tests of aggregate zonal level-prediction.

1. Disaggregate Measures of Transferability

Disaggregate predictive measures are based on the ability of a transferred model to describe in individual observed trips in the application context. These predictive accuracy measures of the transferred models can be formulated as absolute test or relative tests. The absolute measures indicate the ability of transferred model to replicate the observed behaviour. These measures are based on log-likelihood measure.

On the other hand the relative measure indicates the accuracy of the transferred model in comparison with the similarly specified model calibrated in the application context. One relative measure used in this analysis is the goodness-of-fit [rho-square] which describes the degree to which the transferred log likelihood exceeds that of the reference model relative to the log likelihood obtained with [rho-square] a perfect model in the estimation context. The transfer rho-square measure is defined by:

$$\begin{aligned} \text{rho-square} &= \frac{LL_j(\theta) - LL_j(\theta_{pr})}{LL^*j - LL_j(\theta_{pr})} \\ &= 1 - \frac{LL_j(\theta)}{LL_{prj}} \end{aligned} \quad [4.21]$$

where

$LL_j(\theta)$ = log likelihood value of transferred model

LL^*j = log likelihood value of a perfectly specified model in the application context.

LL_{prj} is the log-likelihood value of proportional flow model.

In relative measure, the transfer index [TI] describes the degree to which the transferred model performs in disaggregate fashion relative to the similarly specified locally estimated models in the application context. This is done by examining the degree to which the log likelihood of the transferred model exceeds some reference model [in our case proportional flow model] relative to the estimated model in the application context.

$$TI = \frac{LL_j(\theta) - LL_j(\theta_{pr})}{LL_j(\theta) - LL_j(\theta_{pr})} \quad [4.22]$$

This transfer index has a maximum value of 1 which states that the transfer model is as accurate as a locally estimated model. The negative value of transfer index indicates that the transfer model is worse than the reference model [Koppelman and Wilmot].

2. Aggregate Measures of Transferability

In the planning process, decisions are usually made based on aggregate rather than disaggregate travel. Therefore it is necessary to consider transferability of models in terms of their aggregate accuracy and performance. Absolute and relative error measures in aggregate predictions were also used to evaluate the aggregate transferability test.

The aggregate transfer measures test the ability of a transferred model to reproduce the estimated number of trips produced and attracted in a pair of zones in an aggregate fashion is calculated. We choose the root-mean-square error [RMSE] and the mean absolute percentage error [MAPE] as an absolute measure to test the aggregated prediction accuracy of the transferred model.

First, let us formulate the relative error measure in prediction of trips from i to j .

$$REM_i = \frac{\hat{T}_i - T_i}{\hat{T}_i} \quad [4.23]$$

Where

REM_i = relative error measure in prediction of trips from zone i to zone j.

$\hat{T}_i$ = number of trips predicted to commute from zone i to zone j;

T_i = number of trips observed to commute from zone i to zone j.

In order to evaluate the aggregate predictive accuracy of a model we summarize this measure over a pair of zones i to zone j by means of weighted root-mean-square error [RMSE] measure:

$$RMSE = \left[\frac{\sum_i \hat{T}_i}{\sum_i \hat{T}_i} REM^2_i \right]^{1/2} \quad [4.24]$$

where

$\hat{T}_i$ is the number of predicted trips from i to j.

These error measures here are combined over zones to obtain the aggregate error measure.

The other measure used in an absolute term is the mean absolute percentage error [MAPE]

$$MAPE = \sum_i \left| \frac{\hat{T}_i - T_i}{\hat{T}_i} \right| \quad [4.25]$$

$\hat{T}_i$ and T_i are already defined above.

The use of the mean absolute error criterion means that the cells with higher volumes will have a greater effect than the others in the calibration.

The most important criterion for choosing a forecast method is its accuracy or how closely the forecast predicts the actual events. The objective is to find a forecast such that the future forecasts error $[\hat{T}_i - T_i]$ is as small as possible. The smaller value of each of these two measure forecast accuracy indicates how closely the forecast predicts the observed or actual trip matrices. These statistics indicate the overall predictive power of the model.

The relative transfer measures of aggregate prediction accuracy chosen for our analysis are Relative Aggregate Transfer Error [RATE] and Relative Weighted Mean Absolute Error [RWMAPE]. The RATE is the ratio of the transfer RMSE measure to the local [estimated] RMSE measure. RWMAPE is the ratio of the transferred weighted mean absolute percentage error to that of the local mean absolute percentage error.

$$RATE = \frac{RMSE_j(\theta_i)}{RMSE_j(\theta_j)} \quad [4.26]$$

where

$RMSE_j(\theta)$ is the root-mean-square error of transfer model

$RMSE_j(\theta)$ is the root-mean-square error of local [estimated] model

$$RWMAPE = \frac{RWMAPE_j(\theta_i)}{RWMAPE_j(\theta_j)} \quad [4.27]$$

where

$RWMAPE_j(\theta_i)$ is the relative weighted mean absolute percentage error of the transferred model.

$RWMAPE_j(\theta_j)$ is the relative weighted mean absolute percentage error of the local [estimated] model.

CHAPTER 5

MODEL ESTIMATION AND TRANSFERABILITY RESULTS

5.1 Estimation Results for Different Model Specifications

Parameters were estimated for each of the models derived in chapter 4 using 1964, 1986 and 1996 data, using the Newton Raphson MLE procedure. The estimation results of all the models are presented in Table 5.1 through Table 5.6. A visual inspection of Tables 5.2, 5.4 and 5.6 reveals that all of the 1964, 1986, and 1996 model parameters have the expected negative signs.

One way of measuring how well a particular model fits the existing data on cell to cell basis is to examine the log likelihood value of that model.

5.1.1 ATTM Results and its Comparison with PFM.

The log likelihood values for all the 1964, 1986, and 1996 models are presented in Tables 5.1, 5.3, and 5.5. As may be seen from these Tables the simple auto travel time model has a log likelihood of $[-5.2178 \times 10^7]$ using 1964 data set. Comparison of the log likelihood values of the first two rows in each of these three Tables "*Proportional Flow Model*" [PFM] and simple "*Auto Travel Time Model*" [ATTM] indicate that the ATTM improved the model fit by 7.1% in 1964, 6.25 % in 1986, and 5.7% in 1996. This indicates that travel time is one of the factors that explains the observed patterns of work trips in all the three data sets.

In addition, the first Proportional flow model developed, in which the entire GTA is defined as a single spatial segment, poorly predicted all the zonal trips, particularly, Planning District 1 as both as a destination and origin zone. In addition to the PD1, planning district 46 [Hamilton-Wentworth] was poorly predicted as an origin zone rather than as a destination zone [see Figure 5.1 and Appendix A and B]. The reasons for the poor prediction could be due to the inability of the proportional flow model to adequately account for major changes in land use or the interzonal accessibility.

Figure 5.1 1964 Observed and Predicted Trip Length Frequency Distribution for All Workers Considering the Study area as a One Destination Segment Using PFM.

Table 5.1 1964 Estimation in Log Likelihood values

Model Type	Log Likelihood $\times 10^7$	Improvement in fit (%)
P.F.M.	-0.56142	7.1
A.T.T.M.	-0.52178	
A.T.T.M. Segmented By Space <ul style="list-style-type: none"> • PD1 • Other 	-0.52100	0.15
A.T.T.M. Segmented by Gender <ul style="list-style-type: none"> • Male • Female • Σ 	-0.38647 -0.13014 -0.51661	0.99
A.T.T.M. Segmented by Workers <ul style="list-style-type: none"> • With Driver's Licence • Without Driver's Licence • Σ 	-0.41244 -0.10370 -0.51614	1.1
A.T.T.M. by Occupation Category <ul style="list-style-type: none"> • OCC 1 • OCC 2 • OCC 3 • OCC4 • Σ 	-0.18513 -0.10757 -0.05776 -0.15585 -0.50631	2.96

Table 5.2 1964 Model Parameter Estimates

1964 Model Types	T.T. Parameter	T-Statistic	OATT (min)	Total Trips
A.T.T.M. (All Workers)	-0.12927	-1.38	17.46	490742
A.T.T.M. Segmented by Space:				
• PD1	-0.09717	-2.0	28.38	149668
• Other	-0.14324	-1.17	12.67	341070
A.T.T.M. by Gender				
• Male	-0.12142	-1.33	17.93	360371
• Female	-0.15877	-1.55	13.01	130288
A.T.T.M. by Workers				
• With Driver's Licence	-0.12605	-1.39	17.91	385790
• Without Driver's Licence	-0.14263	-1.30	15.80	104952
A.T.T.M. by Occupation Category				
• OCC1	-0.12272	-1.39	19.20	175820
• OCC2	-0.14427	-1.50	18.20	108872
• OCC3	-0.12930	-1.38	17.50	56562
• OCC4	-0.13298	-1.27	14.90	149488

Table 5.3 1986 Estimation in Log Likelihood values

Model Type	Log Likelihood $\times 10^8$	Improvement in fit (%)
P.F.M.	-0.16090	
A.T.T.M.	-0.15084	6.25
A.T.T.M. Segmented By Space		
• PD1	-0.15069	0.1
• Other		
A.T.T.M. Segmented by Gender		
• Male	-0.08712	0.4
• Female	-0.06313	
• Σ	-0.15025	
A.T.T.M. Segmented by Workers		
• With Driver's Licence	-0.12916	0.6
• Without Driver's Licence	-0.02072	
• Σ	-0.14988	

Table 5.4 1986 Model Parameter Estimates

1986 Model Types	T.T. Parameter	T-Statistic	OATT (min)	Total Trips
A.T.T.M. (All Workers)	-0.10199	-1.18	18.38	1245377
A.T.T.M. Segmented by Space:				
• PD1	-0.08376	-2.05	25.58	263838
• Other	-0.10976	-1.18	16.40	981560
A.T.T.M. by Gender				
• Male	-0.09010	-1.10	19.36	710478
• Female	-0.12617	-1.33	16.45	533412
A.T.T.M. by Workers				
• With Driver's Licence	-0.98494	-1.19	19.10	1061453
• Without Driver's Licence	-0.12598	-1.09	13.96	183923

Table 5.5 1996 Estimation in Log Likelihood values

Model Type	Log Likelihood $\times 10^7$	Improvement in Fit (%)
P.F.M.	1.6702	5.7
A.T.T.M.	1.5748	
A.T.T.M. Segmented By Space		0.133
• PD1	-1.5727	
• Other		
A.T.T.M. Segmented by Gender		0.44
• Male	-0.8642	
• Female	-0.7037	
• Σ	-1.5679	
A.T.T.M. Segmented by Workers		0.43
• With Driver's Licence	-1.4451	
• Without Driver's Licence	-0.1229	
• Σ	-1.5670	
A.T.T.M. by Occupation Category		1.33
• OCC1	-0.71585	
• OCC2	-0.24099	
• OCC3	-0.25467	
• OCC4	-0.34239	
• Σ	-1.5539	

Table 5.6 1996 Model Parameter Estimates

1996 Model Types	T.T. Parameter	T-Statistic	OATT (min)	Total Trips
A.T.T.M. (All Workers)	-0.07876	-1.26	23.51	1283535
A.T.T.M. Segmented by Space:				
• PD1	-0.06374	-2.29	36.07	251206
• Other	-0.08670	-1.24	20.50	1032347
A.T.T.M. by Gender				
• Male	-0.07105	-1.17	24.75	700904
• Female	-0.09538	-1.45	22.02	582607
A.T.T.M. by Workers				
• With Driver's Licence	-0.07886	-1.28	24.02	1176145
• Without Driver's Licence	-0.09890	-1.18	17.98	107365
A.T.T.M. by Occupation Category				
• OCC1	-0.07385	-1.25	25.34	585099
• OCC2	-0.09019	-1.44	23.77	204450
• OCC3	-0.08603	-1.30	21.83	209974
• OCC4	-0.08389	-1.19	20.80	283989

Comparing the auto travel time parameters of the three difference data set models, it is clear from Figure 5.2 that 1964 had the highest travel time parameter. This indicates that the 1964 workers were very sensitive to cost and hence that there were few long trips commuting in the morning peak period. The 1996 model showed the lowest travel time parameter which indicates an insensitivity to cost and hence a relatively large number of long trips was observed.

**Figure 5.2: Auto Travel Time Parameter Values for 1964, 1986, and 1996
Worker Commuting in Morning Peak Period in the Study Area**

Looking at the observed average travel times of the three difference data set models in Figure 5.3, the 1964 model showed the smallest observed average travel times [17.46 minutes] closely followed by 1986 [18.38 minutes], but 1996 recorded the highest observed average travel time [23.51 minutes]. This is not surprising since in 1964 most residential and employment areas were closer to each other particularly in the central part of the city than in 1986 and 1996 and the average travel time tends to increase as the size of the city increases.

Figure 5.3: Observed Average Travel Times for 1964 [MTARTS], and 1986 and 1996 [TTS] Databases for All Workers in the Study Area.

The average trip lengths observed in all the 1964, 1986 and 1996 models are also presented in Table 5.2, 5.4, and 5.6 Column 4. These results indicate that there was a general increase in the average trip length over that time period for all trip categories. The increase was fairly similar in all categories. The increase for all workers taken together was larger than the increase found in other market segmented models. The 1964 gender-based model showed the highest difference of average trip length as compared with the other segmented models. However, in 1986 and 1996, the driver's licence/household car-ownership based model showed to have the highest percentage difference in observed average trip length. That indicates that car-owning workers had higher observed average trip lengths than non-car-owning workers. Gender also became a very important explanatory variable in 1986 and 1996 compared with the 1964 context.

The travel time frequency distributions of the ATTMs for all the work trip models based on the three different data sets are presented in Figures 5.4, 5.5, and 5.6. These figures indicate that generally workers assign a greater disutility to the longer trips. Also note that the trip distribution models for all years consistently underestimate short trips [$0 \leq 10$ minutes] and overestimates longer trips see figures below. One can see the difference between Figure 5.1 and 5.4, when the auto travel time is used in the model as an explanatory variable, the prediction error is getting smaller compared with that of PFM. In that case, the reason that these models are overpredicting longer trips and underpredicting short trips is the missing of the true explanatory variables. The travel time frequency distributions of all the segmented models are shown in Appendix A. Figures 5.7 through 5.9 report these data as cumulative frequency distribution [CDFs].

Figure 5.4: 1964 Observed and Predicted Trip Length Frequency Distribution for all Workers Considering the Study Area as a One Destination Segment Using ATTM.

Figure 5.5: 1986 Observed and Predicted Trip Length Frequency Distribution for All Workers Considering the Study area as a One Destination Segment

Figure 5.6: 1996 Observed and Predicted Trip Length Frequency Distribution for All Workers Considering the Study area as a One Destination Segment

Figure 5.7: 1964 Cumulative Trip Frequency Distribution for All Workers. **Figure 5.8 : 1986 Cumulative Trip Frequency Distribution for All Workers.**

Figure 5.9: 1996 Cumulative Trip Frequency Distribution for All Workers.

However, from the above results, auto travel time alone does not adequately explain all the travel behaviour difference in the model. To further understand the true behaviour of travel patterns in the GTA, it became necessary to investigate the effects of the spatial and socio-economic variables.

5.1.2 Spatial Segmentation Results “Planning District 1 vs. Other”

The procedure used to specify the *auto travel time model* by space, gender, driver’s licence/household auto-ownership, and occupation category includes dividing the data into socioeconomically or geographically homogeneous sub-groups. Then the marginal impacts

of each of these market segments on the model performance is evaluated and compared with that of the simple *auto travel time model* above.

Model calibration results and their implications are discussed as each segmentation of the data is constructed and modelled. Based on the results presented in Table 5.2, 5.4, and 5.6, Columns 6, 31% of total morning peak period work trips in 1964, 21% in 1986 and 20% in 1996 ended their trips in planning district 1. Column 2 of these Tables lists the estimated travel time parameter values for all the models estimated using the three different data sets. When these models were spatially segmented, the PD1 models give a parameter values of -0.09717, -0.08376, and -0.06374 for 1964, 1986, and 1996 respectively. All these PD1 parameters have a larger t-statistics and became highly significant with a 95% level of confidence. These parameters all have the expected signs. The remaining 69% of trips that destined to the other planning districts in the GTA have also the expected signs but their t-statistics are little bit low and are significant at a 90% level of confidence.

The parameter values of PD1 workers are much lower than those of the non-PD1 workers. This implies that the PD1 travelers are willing to travel to work longer distances compared with other travelers in the remaining parts of the GTA. This closely agrees with the observed average travel times of the same models. That is, the travelers to the PD1 have an observed average travel time of 28.38 minutes in 1964 that is two times higher the time traveled by non-PD1 travelers. This figure continued to increase to 36 minutes in 1996 [see Figure 5.10].

Figure 5.10: Observed Average Travel Time for PD1 Workers

Figure 5.10 presents the observed average travel times for PD1 workers for 1964, 1986 and 1996. 1996 workers encountered the longest observed average travel time while 1986 surprisingly showed the shortest average travel time.

Examining the distributions of trip lengths in [min] for 1964, 1986, and 1996 morning peak period home-to-work trips destined to PD1 [see Figures 5.11a] and the equivalent distributions for non-PD1 trips [Figure 5.11b], a distinct difference can be observed. The PD1 trips are not declining constantly with travel time from PD1 as opposed to the non-PD1 trip distribution which describes a fairly smooth aggregate relationship with travel time. The PD1 trips tend to peak for example at 35 [min] in 1964 from work place and then decline [see Figure 5.11a]. More than 75% of the PD1 workers commute less than 40 [min] from downtown [Figure 5.12]. On the other hand the non-PD1 workers tend to peak 10 [min] from work. More than 95% of these workers commute less than 30 [min] in 1964 [see Figure 5.13].

Figure 5.11(a): Trip Lengths Distribution for Workers, destined to PD1 in 1964 **Figure 5.11(b): Trip Length Distribution for Workers Destined to non-PD1 in 1964**

Figure 5.12: Cumulative Observed and Predicted Trips for PD1 workers in 1964

Figure 5.13: Cumulative Observed and Predicted Trips for non-PD1 workers in 1964

Figure 5.14: 1986 Trip Frequency Distribution for Workers that Destined in PD1. **Figure 5.15: 1986 Trip Frequency Distribution for Workers that Destined in non-PD1.** **Figure 5.16: 1996 trip frequency distribution for workers that destined in PD1**

Figure 5.16: 1996 Trip Frequency Distribution for Workers that Destined in PD1 **Figure 5.17: 1996 Trip Frequency Distribution for Workers that Destined in non-PD1**

In terms of the model performances, spatial segmentation did not appreciably improve the model fit in any of the three auto travel time models developed for 1964, 1986 and 1996 data bases, that is, 0.15%, 0.10%, and 0.133% respectively [see last Column of Table 5.1, 5.3 and 5.5].

5.1.3 Socio-Economic Segmentation Results

Segmentation by gender and driver's licence/household auto-ownership marginally improved the model's goodness-of-fit. However, the occupation categories designed for the 1964 and 1996 data bases performed best on the observed databases. Segmentation by gender results in a significant differences in model parameters. Male workers have a lower parameter magnitudes than their female counterparts. Similarly workers who possess driver's licence and also one or more cars available in their households exhibited lower travel time parameters than those without driver's licence. This implies that female workers and workers without driver's licence have a lower tendency to travel for long distances than their counterpart workers. This idea is supported by the observed average travel time [see Tables 5.2, 5.4, and 5.6].

In order to test how well these models fit to the existing data we summed the two log likelihood values from the segmented models of each data set [i.e. male and female or households with driver's licence and those without, etc.] summed them up and compared with the log likelihood of simpler specified auto travel time model. These show marginal improvements in fit that is, about 0.99%, 0.4% and 0.44% for gender segmented models

in these respective years. Similarly the improvement in fit encountered by driver's licence/household auto-ownership model was 1.1%, 0.6% and 0.43% respectively [see the last Column of Tables 5.1, 5.3, and 5.5]. See also the travel time frequency distributions for these models in Appendix B.

Again, comparing the log likelihood values of ATTM with market segmented models yields that the Occupationally disaggregated model gives the best goodness-of-fit to the estimated 1964 and 1996 data with log likelihood values of -0.50631×10^7 and -1.5539×10^7 respectively.

This shows that the improvement in fit increased about 3% and 1.33 % in the respective data sets when compared with the simpler specified ATTMs in 1964 and 1996 respectively. This is followed by a driver's licence/household car-owning based model with log likelihood value of -0.51614×10^7 in 1964, -0.14988×10^8 in 1986, and -1.567×10^7 in 1996. The worst fit was performed by the non-segmented simpler specified ATTM with smallest log-likelihood values in the three data sets.

Examining Figures 5.18 and 5.19, the model parameters of four occupation groups in the 1964 and 1996 data sets indicate that there is roughly an order of magnitude between these sets of coefficients. This order of parameters states that the higher the rank of the occupation category the lower the parameter value. The only exception to this pattern is the occupation category 2 which recorded the highest travel time parameter value (i.e.,

sensitive to travel long distances). The reason might be that in this occupation category about 78% of these trip makers in 1996 and 64% in 1964 were female who put more weight on their travel distance than male workers. The above observation was also supported by the average travel times of the four occupation groups in 1964 and 1996 which again indicated that the observed average travel time of the different occupation groups are relatively increasing according to their level of status/income, which agrees with our prior expectations.

In both 1964 and 1996 models, Consistently, in all the models considered, it was found that white collar occupation group 1 had a higher average travel time than the other three remaining occupation groups. Due to their high income, this occupation group have one or more cars available in their household as opposed to a blue collar worker for example - Occupation group 4). Occupation group 4 showed the lowest average travel time, and the remaining two "intermediate" Occupation groups, group 2 and group 3, lie in between these two extremes [see Tables 5.7a and 5.7b]. These observations tend to fall into an expected order since occupation 3 and occupation 4 are blue collar workers and most are suburban oriented. This results are also generally consistent with the findings of Miller and Cubukgil [1981] in their analysis of 1971 Census Place of Residence - Place of Work linkages in the GTA.

When these occupation categories were further segmented by gender, it was interesting to note that a trend that is intuitively more closer to our a priori expectation became

apparent; that is a consistently increasing weight is placed on travel time as one progresses from higher to lower status/income occupations. [see Figures 5.20 and 5.21 for travel time parameters for both male and female workers by occupation and Tables 5.7a and 5.7b Column 5 for observed average travel times for 1964 and 1996 data sets].

Figure 5.18: 1964 Model Travel Time Parameters Segmented by Occupation Category

From Figures 5.19 and 5.20 one may observe that male workers are dominating the parameter value of where and how far to travel. Within each model males have consistently smaller travel time parameter than female. Interestingly, the blue-collar males in Occupation 4 have even smaller parameter value than the white-collar females in Occupation 1 [Professionals and Managers].

Figure 5.19: 1996 Model Travel Time Parameters Segmented by Occupation Category

Figure 5.20: 1996- Female Model Travel Time Parameters Segmented by Occupation Category

Figure 5.21: 1996 Male Travel Time Parameters by Occupation Category

Table 5.7a: 1964 TTS occupation data segmented by gender

Model	Occ G*	TTP	T-Stat	OATT	PATT	LL x10 ⁷	RMSE	MAPE	Total Trips	Improvement in fit (%)
Occ1	1	-0.12272	-1.39	19.19	19.25	-0.18513	9.32	1.57	175820.3	
Male	1	-0.11915	-1.39	19.96	20.02	-0.15681	10.00	1.59	149898.1	0.62
Female	1	-0.15050	-1.37	14.75	14.81	-0.02496	10.58	1.79	25922.2	
Occ2	2	-0.14427	-1.50	18.19	18.27	-0.10757	25.13	1.51	108871.7	
Male	2	-0.12294	-1.34	18.66	18.74	-0.03918	09.65	1.76	39755.8	1.03
Female	2	-0.16223	-1.64	17.93	18.00	-0.06592	79.55	1.53	69115.9	
Occ3	3	-0.12930	-1.38	17.50	17.57	-0.05776	10.81	1.74	56561.5	
Male	3	-0.12128	-1.33	18.47	18.53	-0.04290	10.48	1.80	42132.3	1.30
Female	3	-0.16278	-1.50	14.66	14.72	-0.01271	11.74	1.73	14429.2	
Occ4	4	-0.13298	-1.27	14.87	14.90	-0.15585	07.45	1.59	149488.3	
Male	4	-0.12699	-1.23	15.16	15.23	-0.13492	07.88	1.65	128584.9	1.02
Female	4	-0.17737	-1.47	13.12	13.16	-0.01835	08.83	1.64	20903.4	

Occ G - Occupation Group
 TTP - Travel Time Parameter
 OATT - Observed Average Travel Time
 PATT - Predicted Average Travel Time
 LL - Log Likelihood at Final Beta
 RMSE - Root Mean Square Error
 MAPE - Mean Absolute Percentage Error
 * Occupation groups are:

1- Professionals/Managements/technicals; 2- General office/Clerical;
 3- Sales and Service; and 4- manufacturing/construction/trades.

Table 5.7b: 1996 TTS occupation data segmented by gender

Model	Occ G*	TTP	T-Stat	OATT	PATT	LL x10 ⁷	RMSE	MAPE	Total Trips	Improvement in fit (%)
Parent	1	-0.07385	-1.25	25.35	25.24	-0.71585	7.665	1.553	585099.2	
Male	1	-0.06526	-1.15	27.03	26.94	-0.39930	8.820	1.680	326828.9	0.62
Female	1	-0.08507	-1.37	23.24	23.34	-0.31210	10.586	1.709	258257.2	
Parent	2	-0.09019	-1.44	23.77	23.89	-0.24099	31.830	1.690	204450.2	
Male	2	-0.07225	-1.19	25.39	25.32	-0.05269	22.042	1.890	45455.1	1.03
Female	2	-0.98159	-1.55	23.31	23.40	-0.18582	11.625	1.720	158995.0	
Parent	3	-0.08603	-1.30	21.84	21.89	-0.25467	16.470	1.775	209973.6	
Male	3	-0.07500	-1.19	23.83	23.74	-0.12920	16.170	1.867	106779.8	1.30
Female	3	-0.10292	-1.45	19.78	19.84	-0.12218	22.180	1.837	1031936	
Parent	4	-0.08389	-1.19	20.78	20.82	-0.34230	13.470	1.682	283959.2	
Male	4	-0.07770	-1.15	21.75	21.85	-0.26833	12.870	1.736	221827.5	1.02
Female	4	-0.11098	-1.34	17.35	17.38	-0.70571	25.970	1.834	62161.5	

Occ G - Occupation Group
TTP - Travel Time Parameter
OATT - Observed Average Travel Time
PATT - Predicted Average Travel Time
LL - Log Likelihood at Final Beta
RMSE - Root Mean Square Error
MAPE - Mean Absolute Percentage Error

* Occupation groups are:

- 1- Professionals/Managements/technicals; 2- General office/Clerical;
3- Sales and Service; and 4- manufacturing/construction/trades.

The estimation results from the graphs and tables discussed so far confirmed the first major hypothesis of this study "i.e., disaggregation of work trip flows by space and by socioeconomics is an important factors in achieving a better understanding of work trip-making behavior".

These variables tend to fall into an expected order and all the calibration results of all the models investigated revealed that consistent differences exist among the segmented groups in terms of their sensitivity to work travel time. In occupation variable, for example, these differences are consistent in agreement with a priori expectations based upon the level of income, social status, and other behavioral attributes of each occupation group. The percentage of improvements in fit are pretty modest for all the models, not all the coefficients are probably statistically different.

5.1.4 Results of Model 3 "Generalized Cost Model"

It should be noted that the 1986 data did not have occupation categories and could not be segmented in this fashion. Transferability tests could not be considered in this model due to time and data limitations. Instead the results of 1986 ATTM model were specifically compared with 1986 "generalized cost" impedance function model results as discussed below.

Goodness-of-fit log-likelihood statistics were issued in this model to check the extent to which the predicted work trip matrix fit the observed trip matrix on cell to cell basis. The log-likelihood value of the 1986 "Generalized cost" model $[-0.15062 \times 10^8]$ was compared with that of ATTM

$[-0.15084 \times 10^8]$. It was found that "Generalized cost" model has marginally a better fit on the observed data set than the *auto travel time model*. See also Table 5.8.

Table 5.8 Home-to-work trip distribution model parameter estimation:

<u>Variable</u>	<u>Parameter Estimates</u>
Auto time	-0.09993
Auto constant	
Toronto CMA	-0.99458
RM3-5 ¹	0.3564
Oshawa CA	0.4625
Hamilton CMA	0.6901
Transit time ²	-0.05843
Rail time	-0.07549
Rail constant ³	-0.0204
Walk distance	-0.01709
Walk constant ³	0.3843.

Notes:

1. Regional Municipalities of York, Peel and Halton [excluding Burlington].
2. Transit is the "base" mode, for which $\alpha_m = 0$ by definition.
3. A single constant is applied to all GTA O-D pairs for which the mode is available.

The overall conclusion that can be made in terms of the model prediction power is that the entropy model overpredicted most of the short trips and underpredicted the long trips. From the prediction error [see Tables in appendix A], we can see that the trips from Metro to PD1 are usually overpredicted, where the long trips were underpredicted.

Like ATTM's prediction errors of this model were computed on Planning District O-D basis. Travel time frequency distribution of this model and its prediction error can also be found in Appendix A.

5.2 Transferability Results

Tables 5.9 , 5.10 and 5.11 present the results of the transferability model tests of the 1964 model transferred to 1986, 1964 to 1996, and 1986 to 1996 respectively. The last test is a relatively short-term model transferability test [10 year period], while the first two tests are long-term transferability tests [22 and 32 year period respectively]. In each transfer test we will discuss the disaggregate and aggregate measures of transferability separately beginning with the 1964 model transferred in 1986 and 1996 respectively followed by 1986 model transferred into the 1996 context.

5.2.1 Disaggregate Measures of Transferability of 1964 Model Transferred to 1986 Context

The prediction results of the 1964 model transferred in the 1986 application context are reported in Table 5.9. It can be seen that in absolute terms the auto travel time model segmented by a driver's licence/household auto ownership gives the best disaggregate predictions on the observed 1986 context with a log-likelihood value of -1.50178×10^8 . The second best transferred model was the auto travel time model segmented by gender with a log-likelihood value of -0.15067×10^7 .

The other interesting observation to be made from Table 5.2 is that this gender model yielded the second best fit in the 1986 estimation context, with log-likelihood of 0.15025×10^8 . The model that resulted in the worst disaggregate prediction on the observed 1986 data is the unsegmented simpler specified ATTM, followed by the spatially segmented model which also performed quite poorly. Their log-likelihood values are -1.5748×10^8 and 1.5727×10^8 respectively. This result particularly agrees with our expectations since the simpler specification model does not usually improve the goodness-of-fit of the model and not capable to transfer

Table 5.9: Results from Transferring 1964 Model to 1986 Application Context

1964 Model Types	Absolute and Relative Disaggregate Transfer Measures			Aggregate Transfer Measures				
	$LL_{64}(\theta_{64}) \times 10^7$	TI	$Rho_{64}(\theta_{64})$	$Rho_{64}(\theta_{64})$	$RMSE_{64}(\theta_{64})$	$MAPE_{64}(\theta_{64})$	RATE	RMAPE
A.T.T.M. (All Workers)*	-1.51168	0.97	0.063	0.061	21.90	1.31	2.80	0.971
A.T.T.M. Segmented by Space:								
• PD1 & Other	-1.51124	0.96	0.063	0.061	40.45	1.78	4.11	0.970
A.T.T.M. by Gender								
• Male	-0.8742	0.94	0.055	0.052	24.72	1.51	2.91	0.973
• Female	-0.6325	0.98	0.076	0.074	95.93	1.46	4.77	0.975
• Σ	-1.5067	0.96	0.064	0.064	55.20	1.49	3.70	0.97
A.T.T.M. by Workers								
• With Driver's Licence	-1.29450	0.97	0.062	0.059	21.80	1.37	2.68	0.972
• Without Driver's Licence	-0.20728	0.99	0.064	0.063	11.67	1.63	1.29	0.991
• Σ	-1.50178	0.97	0.062	0.060	20.30	1.41	2.47	0.97

* ATT M - Auto Travel Time Model

Table 5.10: Results from Transferring 1964 Model to 1996 Application Context

1964 Model Types	Absolute and Relative Disaggregate Transfer Measures				Aggregate Transfer Measures			
	$LL_{\theta_4}(\theta_4) \times 10^3$	TI	$Rho_{\theta_4}(\theta_4)$	$Rho_{\theta_4}(\theta_4)$	$RMSE_{\theta_4}(\theta_4)$	$MAPE_{\theta_4}(\theta_4)$	RATE	RMAPE
A.T.T.M. (All Workers)	-0.15907	0.83	0.057	0.048	87.31	1.29	11.91	0.956
A.T.T.M. Segmented by Space:								
• PD1 & Other	-0.15894	0.84	0.058	0.048	167.59	1.28	17.83	0.957
A.T.T.M. by Gender								
• Male	-0.08752	0.76	0.050	0.038	113.10	1.47	12.01	0.959
• Female	-0.07122	0.84	0.068	0.057	264.31	1.45	23.96	0.959
• Σ	-0.15874	0.80	0.058	0.047	181.73	1.46	17.43	0.96
A.T.T.M. by Workers								
• With Driver's Licence	-0.14591	0.84	0.057	0.048	75.66	1.32	9.83	0.959
• Without Driver's Licence	-0.01236	0.91	0.060	0.054	184.36	1.73	7.56	0.959
• Σ	-0.15827	0.85	0.057	0.049	84.75	1.35	9.63	0.96
A.T.T.M. by Occupation Category:								
• OCC1	-0.07236	0.80	0.051	0.041	67.73	1.50	8.80	0.968
• OCC2	-0.02430	0.87	0.062	0.054	657.2	1.65	20.65	0.976
• OCC3	-0.02565	0.89	0.064	0.057	109.26	1.73	6.63	0.976
• OCC4	-0.03459	0.85	0.066	0.056	148.57	1.62	11.03	0.962
• Σ	-0.15690	0.84	0.058	0.049	186.30	1.59	10.83	0.97

1986 Model Types	Absolute and Relative Disaggregate Transfer Measures				Aggregate Transfer Measures			
	LL ₉₅ (θ_{64}) $\times 10^7$	TI	Rho(θ_{95})	Rho ₉₅ (θ_{95})	RMSE ₉₅ (θ_{95})	MAPE ₉₅ (θ_{95})	RATE	RMAPE
A.T.T.M. (All Workers)	-1.5784	0.96	0.057	0.055	19.79	1.31	2.70	0.97
A.T.T.M. Segmented by Space:								
• PD1 & Other	-1.5764	0.96	0.058	0.056	27.50	1.30	2.93	1.30
A.T.T.M. by Gender								
• Male	-0.8662	0.96	0.050	0.048	20.89	1.50	2.22	0.98
• Female	-0.7069	0.96	0.068	0.065	43.71	1.48	3.96	0.98
• $\Sigma =$	-1.5731	0.96	0.058	0.055	31.25	1.49	1.30	0.98
A.T.T.M. by Workers								
• With Driver's Licence	-1.4481	0.97	0.057	0.055	17.43	1.34	2.27	0.98
• Without Driver's Licence	-0.1232	0.97	0.060	0.058	79.41	1.74	3.26	0.98
• $\Sigma =$	-1.5713	0.97	0.055	0.056	22.61	1.74	2.35	0.98

The relative transfer index [TI] developed by equation 4.22 for the entire 1986 data set models are reported in Table 5.9, 5.10, and 5.11, column 3.

The unweighted transfer index [TI] varies from 0.94 to 0.99. The driver's licence/household auto-ownership segmented models consistently performed better than all other models with the highest weighted transfer index value of 0.97, followed by the gender segmented model. This indicates that the 1964 model segment defined by driver's licence/ household auto ownership provides the highest amount of information obtained from the locally estimated 1986 context, followed by the gender-based model. The spatially segmented model performed worst with a weighted transfer index value equal to 0.96.

Table 5.9, column 5 contains the results of the last absolute measure of disaggregate transferability used in our analysis, that is, the transferred rho-square values. It describes the degree to which the log likelihood of each transferred model exceeds that of the base [proportional flow model] relative to the degree to improvement in log likelihood with a perfect [predicts all zonal trips correct] local model. Its upper bound is the local rho-square measure [see column 4 of Tables 5.9, 5.10, and 5.11]

The 1964 driver's licence/household auto-owner-trip segmented model is the principal or is the most important of all the transferred models from the travel behaviour perspective in terms of the predictive power in the 1986 data set.

5.1.2 Aggregate Measures of Transferability

Examination of the aggregate prediction capability of transferred models, in contrast to the disaggregate prediction measure: in absolute terms, the aggregate measure of transferability in Table 5.9 columns 6 and 7 showed that a less well specified not segmented auto travel time model gives the best aggregate predictions. Table 5.9 columns 6 and 7, set out the absolute measures of Root-mean-square and Mean Absolute Percentage Errors respectively. Column 8 and 9 set out the relative measure of aggregate prediction accuracy. Column 8 presents the Relative Aggregate Transfer Error [RATE], and column 9 presents unweighted Relative Mean Absolute Percentage Error [WRMAPE].

Both RATE and WRMAPE indicate that the model segmented by driver's licence yields the best aggregate prediction of spatial interaction of workers. The RATE values vary between 2.47 for the driver's licence segmented model to 4.11 for the spatially segmented model.

Again the second best model that gives the lowest error is the simpler specified ATTM followed by the gender-based model. This is contrary to the results calculated when the weighted disaggregate transfer measure was used as a criterion for transferability. This may be due to the aggregation errors encountered in the zonal trips.

When mean absolute percentage error is used as measure of transferability, the relative transfer measures [see Tables 5.9, 5.10, and 5.11 Column 9] indicate that all the transferred models have relatively smaller error compared with that found by the use of local models.

In general, the segmentation of the 1964 model by gender and driver's licence undoubtedly improved the model's fit to the observed base case and also the transferability.

5.2.3 Results of 1964 Model Transferred to 1996 Context

The results from Table 5.10 columns 2 and 3 show that the occupation category model has the best fit on the 1996 observed data set with log-likelihood value of -0.1569×10^8 . This is followed by the market segment defined by driver's licence with log-likelihood value of -0.15827×10^8 , followed by the gender segmented model log-likelihood value equal to -0.15874×10^8 .

Again occupation has the maximum value of transfer index [TI] of -0.84, followed by driver's licence. The auto travel time model without any segmentation recorded the third best model transferred into the 1996 application context. These results show that socio-economic market segmentation, particularly occupation category and driver's licence/household car-ownership improved the transferability test. The occupationally

segmented model appeared to be the single most important model in improving the specification and the transferability tests for the whole of the models transferred to 1986 and 1996 as an application contexts.

The aggregate transfer measure yields mixed results which are worse than 1964-1986 and 1986-1996 model transferability tests performed. and this might be attributed to the aggregate bias and the longer time frame [32-year period].

In general the different market segments exhibit significantly different tests in weighting the importance of the factors that influence the spatial interaction of workers. The above segmentation of the data into some homogeneous sub-models eliminated marginally the source of mis-specification.

It is apparent from Table 5.10 that the better specified [segmented] models showed higher transfer index [TI] than the model without segmentation.

5.2.4 Results of 1986 Model Transferred to 1996 Context

In Table 5.11 column 2 sets out the log-likelihood values of the 1986 transferred models which have the same specifications as those discussed above between 1964 models transferred to 1996 and 1986. Column 2, column 3, and 5 report the disaggregate transfer measures. From the results in column 2, it is observed that the log-likelihood values increases as the data is segmented by the socio-economic variables of gender and driver's

licence. Consistent with the 1964 transferred model discussed earlier, the 1986 auto travel time model segmented by driver's licence/Household auto-ownership appeared to yield the best disaggregate predictive accuracy with log-likelihood of 1.5713×10^7 , followed by the gender-based model with log-likelihood of 1.5731×10^8 . When checked against the 1996 estimated models, gender performed the best goodness-of-fit, while driver's licence model which performed the best disaggregate prediction on the observed 1996 recorded the second best in the estimated model.

The unweighted range of Transfer Index [TI] is between 0.96 to 0.97. All the other models except the driver's licence/household auto ownership-based model have the Transfer Index value of 0.96, while the driver's licence model recorded the value of 0.97.

The RATE values varies from 2.35 for the driver's licence model to 3.01 for the gender segmented model. Gender segmented model is one of the well specified models which estimated well on the 1996 context. It is also observed that the market segmentation adjusts a model's coefficients and improved the performance of the transferability.

The immediately evident result after observing Table 5.11, column 2 and 3 indicates that gender and driver's licence segmentation models improved the transferability.

Again the aggregate measures present mixed results, but with smaller prediction errors than from the longer-term tests.

Examining the impact of the market segmentation leads to the conclusion that the disaggregate predictions measures indicate that the models are transferable, but the aggregate measures yield mixed results due to the aggregation bias inherited from the models.

In general driver's licence/household auto-ownership model proved to be the single most important model in improving the transferability tests for the models transferred from the 1986 to 1996 contexts.

The overall conclusion would be that there is very small improvement from the worst to best models, that is, the log likelihood -1.51168 versus -1.50178 [i.e. 0.65% improvement]. Generally the gravity model seems to be good, at least relative to the performance of "base year" model.

CHAPTER 6

SUMMARY, CONCLUSIONS AND FURTHER RESEARCH

6.1 Summary of Results

Trip distribution is the second step of the analytical tools that form the core of the strategic urban transportation modelling process. A key assumption in the use of trip distribution models is that the model parameter estimates obtained by using cross-sectional data will remain stable over time.

The intent of this thesis is to investigate the validity of these assumptions and the temporal transferability of trip distribution models. An empirical study of both strategic and tactical forecasts was carried out using the travel behaviour databases collected in the 1964 Metropolitan Toronto And Regional Transportation Survey [MTARTS], the 1986 Transportation Tomorrow Survey [TTS] and the 1996 TTS, all for the Greater Toronto Area [GTA].

The models estimated from the 1964 morning peak-period data were used to predict the 1986 and 1996 trip interchanges respectively [long-range transferability tests]. That is,

forecasts over 22 and 32 year time periods were undertaken. The models estimated from the 1986 data were also used to predict the 1996 travel behavior [a relatively short-range, 10-year transferability test,].

Different model specifications, ranging from proportional flow model [PFM] to a market segmentation of auto travel time model [ATTM], were tested to investigate the relationship between the model specification and the transfer effectiveness.

This study's results can be divided into three main groups:

1. Description analysis of GTA trends;
2. Model estimation and specification test results;
3. Transferability test results.

Each of these are discussed in the following subsections.

6.1.1 GTA Trends

The GTA study area changed substantially in terms of population, labour force participation, transportation system and employment between 1964 and 1996. There were important increases in socio-economic differentiation of the resident labour force and jobs associated with this growth which have implications for work trip commuting patterns. Parts to note include the following.

1. Passenger vehicle registrations between 1964 and 1986, increased 268 percent, while total vehicle registrations similarly increased 272 percent.
2. The relative the GTA population increased 80 percent between 1964 and 1996, while number of households increased by 153 percent over this same time of period, resulting in a overall decrease in average household size of 30 percent. These increases in population and auto ownership and decreases in household size coincided with the migration of many to the suburban areas of the GTA, outside Metropolitan Toronto, where the housing is much cheaper and more abundant than in downtown Toronto.
3. An other major important factor that contributed to travel change is the change in socioeconomic status of females in the society. Their work force participation increased by 36 percent between 1971 and 1986, and about 5 percent between 1986 and 1991. In 1991 more females possessed a driver's licence than in 1964 and in 1986.
4. The blue-collar labour force resident in many of the inner residences of Metropolitan Toronto, decreased due to departure of the manufacturing industries from the high density central area [CBD] to the lower cost, less populated suburban areas.

Technology also changed in that heavy machines and the increase in the form of automated machines forced these industries to be relocated in outer districts of Metro and neighboring regions. On the other hand the central business area became

service, commercial and information oriented area. These differential changes in job, labour force technology, structural changes in employment sources affected the employment distribution and travel patterns. The magnitude and directions of these changes also dramatically changed the trip length and the observed travel patterns of the study area.

5. The region changed from monocentric in 1964 to polycentric in 1986 and 1996.

6.1.2 Model Estimation and Specification Test Results.

1. Calibration results [see prediction errors in appendix B] indicated that the doubly constrained entropy models of 1964, 1986 and 1996 journey-to-work data bases underestimated longer commuting trips and overestimated the short trips.
2. Spatial segmentation provided little improvement in model performance
3. Socioeconomic segmentation yielded small but important improvement in model explanatory power. Occupational segmentation [alone and/or combined with gender segmentation] resulted in the most significant improvement in model fit.
4. Occupation status influences were varied and rank-ordered. The lower the status of the occupation category of a trip maker the higher the travel time sensitivity. The lower status, blue-collar occupation category 4 is ,for example, more travel time sensitive than all other occupation categories.
5. Use of a "generalized cost" impedance function which is sensitive to level-of-service

variables for the full range of travel modes available to give O-D pairs results in improved model explanatory power relative to use of a simple auto travel time impedance function.

6.1.3 Transferability Test Results

In terms of model transferability, alternative model specifications were tested to investigate which model specification gives the best predictive performance in the application contexts. Both aggregate and disaggregate transfer measures were used to evaluate the transferability effectiveness. In disaggregate terms, tests examined the ability of the estimated model to reproduce household level trips in the application context. On the other hand, in aggregate prediction performance, the ability of the estimation model to replicate the observed aggregate trips interchanges in the application context was tested. The findings of this investigation point out possible interrelations between the socioeconomic characteristics of the trip maker and the transportation service attributes serving the area. Specific results include the following:

1. In general, the final results indicate that both long term and short term transferred models provide very useful information about the application contexts, namely 1986 and 1996 contexts.
2. The disaggregate transfer measure results indicate that the 1986 morning peak-period trips yielded very good prediction of the 1996 application context. The prediction

was good from the standpoint of transferred log-likelihoods and the relative transfer index [TI] values.

3. In general, the different model specifications tested give relatively different transferability results. Theoretically, these transfer models did not exhibit direct transferability. However, the relatively short-range [10 year period] from 1986 to 1996 transferability test showed better results compared with the long-range forecasts in the disaggregate level.
4. The transfer index [TI] values of the 1964 transfer models relative to the local 1986 models range from 0.96 to 0.97. The TI values of the same 1964 models relative to the local 1996 models range also from 0.80 to 0.85, where the TI values of 1986 transfer models relative to the local 1996 models range from 0.96 to 0.97. Thus, all the transferred models are able to provide at least 80 percent of the log likelihood values of maximum likelihoods locally estimated in the application contexts.
5. In short-range forecasting [1986 to 1996] and long-range test [1964 to 1986], the unweighted relative transferability index [TI] indicates that the driver's licence/household auto-ownership based model consistently performed better than all other models [excluding the occupationally segmented model], followed by the gender segmented model. This suggests that of all the models transferred from the 1986 context, the driver's licence/household auto-ownership based model provides the largest amount of information component that can be obtained from 1996 application context. The simpler specified ATTM performed worst. However, when the occupationally segmented model was available for

- use [1964-1996], it had the highest [TI] value, followed by the driver's licence/household auto-ownership model.
6. In all cases, the fitting model in the estimation context proved to also be the best forecasting model, when judged by the TI measure.
 7. From the results of the relative aggregate transfer measure, the driver's licence/household auto-ownership model again gives the smallest relative aggregate transfer error [RATE] in short-range transferability test with the value of 2.35, followed by the simpler specified ATTM with the value of 2.70. The spatial segmented model performed worst.
 8. In the long-term transferability test [1964 to 1996] the results indicated that the occupation category model performed the best fit in estimation context in 1964 and also give the best model transfer performance in 1996. The absolute transfer measures of transferability suggest that the model with the best fit in the estimation context, namely the occupation and driver's licence category model, gave the best transfer predictive performance in 1996 and 1986 respectively.
 9. The relative transferability test indicate the transferred models from 1964 to 1996 do provide very useful information in the application context. In these aggregate predictive terms, all the models transferred from 1964 to 1986 were able to provide at least 97 percent of the log-likelihood value provided by a locally estimated 1986 model. On the other hand, the 1964 model transferred into 1996 context was able to provide at least 77 percent of the log-likelihood estimated locally in 1996.
 10. The RATE values range from 9.62 of driver's licence/household ownership

segmented model to 17.83 of the spatially segmented model of the models transferred from 1964 to 1996. This relative aggregate measure gives the largest aggregate errors. The reason might be the longer time difference [1964 to 1996] in which most of the travel related factors have changed in the GTA. This result should not be viewed as negative, but rather credited to the superior performances of the 1986 and 1996 models in prediction on the 1986 and 1996 data sets respectively. In general, the RMSE error measure had larger magnitude the better the model specification.

11. Finally, from these outputs it can be inferred that in both aggregate and disaggregate transfer measures, the shorter the transfer time, the better the transfer performance. The relatively short-term transferability [1986 to 1996] showed the best results whereas the long-term transferability [1964 to 1996] measured the worst predictive performance.
12. It is interesting to see that in all the three data sets, the models that have high goodness-of-fit are the well-specified models and thus are those with more effective in transfer. In particular occupationally segmented model is consistently the most-effectively transferred.

6.2 CONCLUSIONS AND DIRECTION FOR FUTURE RESEARCH

In disaggregate terms, consistent with the findings of the other researchers, model transferability is found to improve with improved model specification, that is the model

with best the estimation context specification gave the best transfer predictive performance. However, aggregate measures showed more mixed results, possibly due to the effects of the aggregation process.

The findings of this investigation point out possible interrelations between the socioeconomic characteristics of the trip maker and the transportation service attributes serving the area.

Detailed data concerning household travel and travel characteristics were obtained for the analysis of this study, but the richness of this detail is washed out by the aggregation.

The degree to which these units [households] are not homogeneous results in a loss of disaggregate detail. If sufficient detail is lost, then a full explanation of the results is not obtained.

Further research is necessary using a data base collected for the purpose of examining some of the hypotheses expressed here in which both the range and the extent of socioeconomic and transportation variables can be enlarged.

The number of cars owned by a particular household has implications for the households' frequency of trip-making. When this variable is aggregated by areal unit and a degree of heterogeneity exists, the relationship between travel and vehicle ownership at the household level is clouded.

Contextual characteristics that affect the transfer effectiveness in both estimation and application contexts should further researched and possible sources of transfer bias should be investigated.

Additional research should be undertaken to explore the relationship between transfer effectiveness and model specification.

Useful extents to the current analysis include:

1. K - Factors should be added to the model to investigate their effects on model fit and the stability over time.
2. Time series data (panel data) may give very good information on how model parameters may change over times.
3. Further investigation of the effects of socioeconomic variables may improve the model performance.
4. The performance of the "generalized cost" model with respect to temporal transferability should be investigated.

REFERENCES

- Anas A. [1981] The Estimation of Multinomial and Testing of a Simulation Model of Residential Location, *Environment and Planning* 7A, PP. 899-920.
- Anas A. [1983]. Discrete Choice Theory, Information Theory and the Multinomial Logit and gravity Models. *Transportation Research*. 17B, PP. 13-23.
- Atherton, T.J. and Ben-Akiva, M., [1976]. Transferability and Updating of Disaggregate Travel Demand Models. *Transportation Research Record* 610, TRB, National Research Council, Washington, D.C. PP. 12-18.
- Badoe, Daniel A. and Miller, J. Eric, [1995]. Analysis of the Temporal Transferability of Disaggregate Work Trip Model Choice Models. *Transportation Research Record* No. 1493. PP. 1-11.
- Badoe, Daniel, A. [1994]. An investigation into the Long Range Transferability of Work-Trip Discrete Mode Choice Models. *Ph.D. Dissertation*
- Batty, M. [1976]. Urban Modelling: Algorithms, Calibrations, Predictions. *Cambridge University Press, Cambridge, U.K.*
- Batty, M. and Mackie. [1972]. The Calibration of Gravity, Entropy, and Related Models of Spatial Interaction. *Environment and Planning* 4, PP. 405-233
- Ben-Akiva, [1981]. Issues in Transferring and Updating Travel-Behaviour Models. *In New Horizons in Travel-Behavior Research [Stopher, P.R. et al., eds.]*, Lexington Books, Lexington, Massachusetts
- Ben-Akiva, M. and Lerman, S.R. [1985]. Discrete Choice Analysis, Theory and Application to Travel Demand. MIT Press, Cambridge, Massachusetts.
- Black, J.A. and Salter, R.J. [1975]. A Statistical Evaluation of the Accuracy of a Family of Gravity Models. *Proc. Institution of Civil Engineers* 59, PP. 1-20.
- Dehghani, Y., Kousheshi, B., Sievert, R. and McKearney, T. [1981]. Transferability and An Analysis of Prediction Errors in Mode Choice Models for Work Trips, *Presented at the 60th Annual Meeting of the Transportation Research Board, Washington, D.C.*
- Duffus, L.N., Alfa, A.S., and Soliman, A.H., [1987]. The Reliability of Using the Gravity Model for Model Forecasting Trip Distribution. *Transportation Vol. 14*, No. 3, PP. 175-192

- Edens, H.J., [1970]. Analysis of a Modified Gravity Model. *Transportation Research*, vol. 4, P.51-62.
- Emmerson, P. [1983]. Accuracy of Trip Forecast Based on Simple Distribution Models. Transport and Road Research Laboratory, *TRRL Laboratory Report* 1044, Crowthorne, Berkshire.
- Erlander S. [1981]. Entropy in Linear Programming. *Mathematical Program.* 21, 137-151.
- Evans A.W. [1970] Some Properties of Trip Distribution Methods. *Transportation Research*, 4, PP. 19-36.
- Evans A.W. [1971]. The Calibration of Trip Distribution Models with Exponential or Similar Cost Functions. *Transportation Research* 5, PP. 15-38
- Evans, S.P., and Kirby, H.R. [1974]. A Three-Dimensional Furness Procedure for Calibrating Gravity Models. *Transportation Research*, 8, PP. 105-122.
- Galbraith, R.A. and Hensher, D.A. [1980]. Intra-Metropolitan Transferability of Mode Choice Models. *Journal of Transport Economics and Policy*, 16. P. 7-29.
- Gun, H.F., Ben-Akiva, M.E. and Bradley, M.A. [1985]. Tests of the Scaling Approach to Transferring Disaggregate Travel Models, *Transportation Research Record* 1037. PP. 21-30.
- Hensher, D.A. and Johnson, L.W. [1977]. A Two-Period Analysis of Commuter Mode Choice: The Prediction Capability of Individual Choice Models, *The Logistics and Transportation Review Volume 13*, No. 4, PP. 361-375
- Hutchinson, B.G. [1984]. The journey to Work to the Greater Toronto Area 1971-1982. *The Transportation Group, Department of Civil Engineering, University of Waterloo*, Waterloo Ontario.
- Hutchinson, B.G., and Smith, D.P. [1979]. Empirical Studies of the Journey to Work in Urban Canada. *Canadian Journal of Civil Engineering*, 6, PP. 308-318.
- Hutchinson, B.G., and Smith, D.P. [1981]. Goodness of Fit Statistics for Trip Distribution Models. University of Waterloo, December.
- Jaynes, E. T., [1957]. Information Theory and Statistical Mechanics I., *Physical Review*, vol. 106, PP. 171-190.

- Kanifani, Adib., [1983]. *Transportation Demand Analysis*. University of California, Berkeley.
- Kirby, H.R. [1974]. Theoretical Requirements for Calibrating Gravity Models. *Transportation Research* 8, PP. 97-104.
- Kitamura, R. [1990]. Panel Analysis in Transportation Planning: An Overview, *Transportation Research* Volume 24 A, No. 6, PP. 401-415.
- Koppelman, S.F. [1976]. Guidelines for Aggregate Travel Prediction Using Disaggregate Choice Models. *Transportation Research Record* 610, TRB. PP. 19-
- Koppelman, S.F. [1976]. Methodology for Analyzing Errors in Prediction with Disaggregate Choice Models. *Transportation Research Record* 592, TRB. PP. 17- 23.
- Koppelman, S.F. and Wilmot, C.G. [1982]. Transferability Analysis of Disaggregate Choice Models. *Transportation Research Record* 895, TRB, National Research Council, Washington, D.C. PP. 18-26.
- Louviere, J.J. [1981]. Some Comments on Premature Expectations Regarding Spatial, Temporal, and Cultural Transferability of Travel Choice Models. *New Horizons in Travel Behaviour Research*, Lexington, MA.
- Manheim, M.L., [1979]. *Fundamentals of Transportation Systems Analysis*, Vol. 1, Basic Concepts, MIT Cambridge, Massachusetts.
- Manski, C.F., and McFadden, D., [1981]. *Structural Analysis of Discrete Data with Econometric Applications*, MIT Press, Cambridge, Massachusetts.
- McCarthy, P.S., [1982], Further Evidence on the Temporal Stability of Disaggregate Travel Demand Models, *Transportation Research*, Volume 16B, No. 4, PP. 263-278.
- McCoomb, L.A. [1986]. Analysis of the Transferability of Disaggregate Demand Models Among Ten Canadian Cities. *Transportation Development Centre, Transport Canada*, Montreal, Quebec, Canada.
- McFadden, D., [1973]. Conditional Logit Analysis of Qualitative Choice Behavior. *In Frontiers in Econometrics*. Academic Press, New York.
- Meyer, M.D. and Miller, E.J., [1984]. *Urban Transportation Planning: A Decision-Oriented Approach*, New York: McGraw-Hill.
- Miller, E.J. and Cubukgil, A. [1981]. Land use Trends and Transportation Demand forecasting: *An Empirical and Theoretical Investigation*, research report no. 75, Toronto: University of Toronto-York University Joint program in Transportation

- Miller, E.J. et al. [1984]. An Occupationally Disaggregate Analysis of Census Data for the Greater Toronto Area, 1971-1981, Final Report to the Demand Forecasting Office, Policy Planning Branch, Ontario Ministry of Transportation and Communications, Toronto: University of Toronto-York University Joint Program in Transportation. July.
- Miller, E.J., Steuart, G.N., Jea, D., Hong, J., [1990]. Understanding Urban Travel Growth in the Greater Toronto Area: Vol. II
- Miller, E.J. and Cheah, L.S. [1992]. Development of an Operational Peak-Period Mode Split Model for Metropolitan Toronto: Volume I: Final Report, Report to the Transportation Division, Metropolitan Toronto Planning Department, Toronto: University of Toronto Joint Program in Transportation, August.
- Miller, J.E. [1997]. The Greater Toronto Area Travel Demand Modelling System, Version 1.0, Vol. II.
- O'Flaherty, C.A. [1997]. Transport Planning and Traffic Engineering.
- Ortuzar, J. de D., and Willumsen, L.G., [1990]. Modelling Transport, New York; John Wiley and Sons.
- Parody, T.E. [1977]. An Analysis of Disaggregate Mode Choice Models in Prediction. *Transportation Research Record* 637, TRB, National Research Council, Washington, D.C. PP. 51-57.
- Pluym, W. K. [1979]. Transportation Planning. *New Developments in Modeling Travel*. PP. 320-343
- Road and Transportation Association of Canada., [1974]. *Technical Publication No. 4*
- Shannon, D. A., and Weaver, W., [1963]. The Mathematical Theory of Communications, University of Illinois Press.
- Sidkar, P.K., and Hutchinson, B.G. [1981]. Empirical Studies of Work Trip Distribution Models. *Transportation Research*, 15A, PP. 233-243.
- Silman, L.A. [1982]. The Time Stability of a Disaggregate Modal-Split Model. *Environment and Planning A*, Vol. 13.
- Smith, D.P., and Hutchinson, B.G. [1981]. Goodness of fit Statistics for Trip Distribution Models. *Transportation Research*., 15A, PP. 295-303.
- Snickars, Folke and Weibull, Jorgen W., [1977]. A Minimum Information Principle,

Theory and Practice. Regional Science and Urban Economics No. 7 PP. 137-168

Stopher, P.R. and Meyburg, A.H., [1975]. Urban Transportation Modelling and Planning.

Stopher, P.R. and Wilmot, C.G. [1979]. Work Trip Mode Choice Models: A south African Datum Point. *Transportation Engineering Journal of the American Society of Civil Engineering*.

Talvitie, A. and D. Kirshner. [1978]. Specification, Transferability and the Effect of Data Outliers in Modelling the Choice of Mode in Urban Travel. *Transportation*, Vol. 7.

Theil, H., [1967]. Economics and Information Theory, Amsterdam, North Holland, Publ. Co.

Train, K. E., [1979]. A Comparison of the Predictive Ability of Mode Choice Models with Various Levels of Complexity. *Transportation Research Volume* 13A, No. 1, PP. 11-16

Volet, P. and Hutchinson, B.G., [1986]. Explanatory and Forecasting Capabilities of Trip Distribution Models. *Canadian Journal for Civil Engineering*, Vol. 13, No. 6, PP. 666-673

Watson, P.L. and Westin, R.B. [1975]. Transferability of Disaggregate Mode Choice Models. *Regional science and Urban Economics*, vol. 5. PP. 227-249.

Webber, m. J., [1977]. Pedagogy Again: What is Entropy?, *Association of American Geography Annal*, Page 254-266.

Williams, H.C.W.L., [1977]. On the Formation of Travel Demand Models and Economic Measures of User Benefit. *Environment and Planning A*, Vol. 9, PP. 285-344.

Wilson, A.G. [1967]. A Statistical Theory of Spatial Distribution Models. *Transportation Research*, Vol. 1, PP. 253-269.

Wilson, S.R. [1976]. Statistical Notes on the Evaluation of Calibrated Models. *Transportation Research* Vol. X, PP. 343-345.

Wilson A.G. [1969]. The Use of Entropy Maximizing Models in the Theory of Trip Distribution, Mode Split and Route. *Journal of Transport Economics and Policy*, 3. PP. 108-126.

Wilson A.G. [1970]. Entropy in Urban and Regional Modeling (Pion, London)

Appendix A

Trip Length Frequency Distributions

Figure A1: 1964 Observed and Predicted Trips for Male Workers

Figure A2: 1964 Observed and Predicted Trips for Female Workers

Figure A3: 1964 Observed and Predicted Trips for Workers with Driver's Licence

Figure A4: 1964 Observed and Predicted Trips for Workers without Driver's Licence

Figure A5: 1964 Observed and Predicted Travel Time Frequency Distribution for Occupation Category 1

Figure A6: 1964 Observed and Predicted Travel Time Frequency Distribution for Workers in Occupation Category 2

Figure A7: 1964 Observed and Predicted Travel Time Frequency Distribution for Workers in Occupation Category 3

Figure A8: 1964 Observed and Predicted Travel Time Frequency Distribution for Workers in Occupation Category 4

Figure A9: 1986 Trip Frequency Distribution for Male Workers.

Figure A10: 1986 Trip Frequency Distribution for Male Workers.

Figure A11: 1986 Observed and Predicted Trip Frequency Distribution for Workers with Driver's Licence.

Figure A12: 1986 Observed and Predicted Trip Frequency Distribution for Workers without Driver's Licence

Figure A13: 1996 Observed and Predicted Trip Frequency Distribution for Male Workers

Figure A14: 1996 Observed and Predicted Trip Frequency Distribution for Female Workers

Figure A15: 1996 Observed and Predicted Trip Frequency Distribution for Licenced Workers

Figure A16: 1996 Observed and Predicted Trip Frequency Distribution for non-licenced Workers

Figure A17: 1996 Observed and Predicted Trip Frequency Distribution for Occupation Category 1 Workers

Figure A18: 1996 Observed and Predicted Trip Frequency Distribution for Occupation Category 2 Workers

Figure A19: 1996 Observed and Predicted Trip Frequency Distribution for Occupation Category 3 Workers

Figure A 20: 1996 Observed and Predicted Trip Frequency Distribution for Occupation Category 4 Workers

Appendix B

Prediction Errors

1964 Models Summary Prediction Errors

A.T.T.M.		Spatially Segmented Model	
From Municipality To PD1		From Municipality To PD1	
Metro	-1739	Metro	-6748
Durham	-45	Durham	289
York	-631	York	310
Peel	-1035	Peel	1726
Halton	290	Halton	703
Hamilton	117	Hamilton	633
Grand Total		Grand Total	-3087

Gender Segmented Model [Male]		Segmented Mode [Female]	
From Municipality	To PD1	From Municipality	To PD1
Metro	-1196	Metro	-609
Durham	-55	Durham	01
York	-543	York	0
Peel	642	Peel	265
Halton	213	Halton	0
Hamilton	88	Hamilton	0
Grand Total		Grand Total	2

Driver's Licence Model
From Municipality To PD1

Metro	-1286
Durham	-82
York	-577
Peel	820
Halton	204
Hamilton	50
Grand Total	-871

No Driver's Licence Model
From Municipality To PD1

Metro	-222
Durham	188
York	0
Peel	0
Halton	0
Hamilton	0
Grand Total	-410

Segmented by Occupation (Occ 1)
From Municipality To PD1

Metro	-1741
Durham	0
York	-227
Peel	0
Halton	0
Hamilton	0
Grand Total	-1968

Occ 2 Model
From Municipality To PD1

Metro	-944
Durham	0
York	-173
Peel	270
Halton	0
Hamilton	0
Grand Total	-847

Occ 3 Model**From Municipality To PD1**

Metro	13
Durham	0
York	0
Peel	-139
Halton	0
Hamilton	0
Grand Total	-126

Occ 4 Model**From Municipality To PD1**

Metro	-95
Durham	0
York	0
Peel	123
Halton	0
Hamilton	0
Grand Total	2

1986 Models Summary Prediction Errors**A.T.T.M.****From Municipality To PD1**

Metro	15097
Durham	-4056
York	-3123
Peel	-3675
Halton	-4023
Hamilton	391
Grand Total	-171

Spatially Segmented Model**From Municipality To PD1**

Metro	-1973
Durham	-494
York	-805
Peel	780
Halton	-781
Hamilton	726
Grand Total	-973

Gender Segmented Model (Male)		Segmented Model (Female)	
From Municipality To PD1		From Municipality To PD1	
Metro	8945	Metro	7437
Durham	-2074	Durham	-1747
York	-1941	York	-1762
Peel	-2081	Peel	-1977
Halton	-2891	Halton	-1448
Hamilton	-107	Hamilton	-270
Grand Total	-149	Grand Total	233

Driver's Licence Model		No Driver's Licence Model	
From Municipality To PD1		From Municipality To PD1	
Metro	13443	Metro	15097
Durham	-3796	Durham	-4056
York	-2603	York	-3123
Peel	-3243	Peel	-3675
Halton	-3819	Halton	4923
Hamilton	-175	Hamilton	-391
Grand Total	-193	Grand Total	-171

1986 "Generalized Cost" Model

From Municipality	To PD1
Metro	7563
Durham	-1204
York	-2710
Peel	-457
Halton	-2888
Hamilton	-3060
Grand Total	-636

1996 Models Summary Prediction Errors**A.T.T.M.****From Municipality To PD1**

Metro	24383
Durham	-7469
York	-8301
Peel	-2252
Halton	-6173
Hamilton	-394
Grand Total	-206

Spatially Segmented Model**From Municipality To PD1**

Metro	832
Durham	-650
York	-4133
Peel	2996
Halton	-2042
Hamilton	1253
Grand Total	-1744

**Gender Segmented Model (Male)
(Female)**

From Municipality To PD1

Metro	12492
Durham	-3119
York	-4259
Peel	-970
Halton	-4186
Hamilton	-1253
Grand Total	-1744

Segmented Model

From Municipality To PD1

Metro	13428
Durham	-4431
York	-4505
Peel	-1888
Halton	-2262
Hamilton	-340
Grand Total	2

Driver's Licence Model

From Municipality To PD1

Metro	23303
Durham	-7011
York	-7737
Peel	-2156
Halton	-5976
Hamilton	-356
Grand Total	-67

No Driver's Licence Model

From Municipality To PD1

Metro	1547
Durham	0
York	-216
Peel	-256
Halton	0
Hamilton	0
Grand Total	-1075

Segmented by Occupation (Occ 1)**From Municipality To PD1**

Metro	12176
Durham	-3698
York	-4114
Peel	-1392
Halton	-3421
Hamilton	-115
Grand Total	-564

Occ 2 Model**From Municipality To PD1**

Metro	4225
Durham	-1620
York	-1524
Peel	-261
Halton	-740
Hamilton	-167
Grand Total	-87

Occ 3 Model**From Municipality To PD1**

Metro	3010
Durham	-627
York	-782
Peel	-488
Halton	-828
Hamilton	-0
Grand Total	-289

Occ 4 Model**From Municipality To PD1**

Metro	3032
Durham	-729
York	-813
Peel	-413
Halton	-3930
Hamilton	0
Grand Total	614

APPENDIX C

Goodness-of-Fit Equations

Coefficient of Determination (R^2):

$$R^2 = 1 - \frac{\sum_i \sum_j (T_{ij} - T_{ij}^*)^2}{\sum_i \sum_j (T_{ij} - \bar{T})^2}$$

Chi-Square (χ^2):

$$\chi^2 = \sum_i \sum_j \frac{(T_{ij} - T_{ij}^*)^2}{T_{ij}^*}$$

Log Likelihood Ratio (DL):

$$DL = \sum_i \sum_j T_{ij} \left[\log\left(\frac{T_{ij}}{T}\right) - \log\left(\frac{T_{ij}^*}{T}\right) \right] \quad T_{ij}, T_{ij}^* \neq 0$$

Standard Deviation (σ):

$$\sigma = \left[\frac{\sum_i \sum_j (T_{ij} - T_{ij}^*)^2}{(n^2 - 1)} \right]^{1/2}$$

Phi (Φ):

$$\frac{\Phi}{T} = \frac{1}{T} \sum_i \sum_j T_{ij} \left| \ln \frac{T_{ij}^*}{T_{ij}} \right| \quad T_{ij}, T_{ij}^* \neq 0$$

Normalized Phi:

$$\text{NORMALIZED } \Phi = \frac{\Phi}{T}$$

Mean Error (MERR):

$$\text{MERR} = \frac{\sum_i \sum_j (T_{ij} - T_{ij}^*)}{n^2}$$

Mean Absolute Percentage Error (MABSERR):

$$\text{MABSERR} = \frac{\sum_i |T_{ij} - T_{ij}^*|}{\sum_i n}$$

where

T_{ij} = Observed flow between i and j

T_{ij}^* = Predicted flow between i and j

$\bar{T}$ = Mean trip interchange value

n = Number of zones

APPENDIX D

Input Parameter Files

1404 7 46	NZONE,NMODE,NPD
1986 TTS Work Trips, GTA Model: 05-	25b from GTAMODEL is the base.
40	NITER
0.005	SMIN
0.005	XLIMIT
0	TPRINT
0	INTRA
0	LOGSUM
0	BEST
0 0	PD1,PD2
0	NPDS
0	NPD1
0 0 0.0	SHORT,SHVAR,VSHMAX
3 3	NUV,NUVPD
5 1	UVAR,VBASE
6 3	
7 4	
8 17 39	
9 18 24	
10 40 46	
1	MDEF
250. 10. 0	MINOBS,MINERR,PRTOPS
zn1770-46.dat	
0	NKFAC
0	ASPLIT
2	NMCOMP
3 3 4 5	NYCP,YCP: All rail modes
2 1 7	Auto drive + auto passenger
1 0 0 0	NCLASS,NCLINT,INOUT,CONOUT
0 0 0.0	KSPLIT,VSPLIT,VSMAX
10 1 46 1 46	
a-time	-0.1024 1 1 9
ac-base	-0.7527 1 5 9
ac-RM3-5	0.3162 1 8 9
ac-OCMA	1.0304 1 9 9
ac-HCMA	0.9687 1 10 9
t-time	-0.0761 1 2 2
rb-time	-0.0830 1 3 8
rb-const	-1.3781 1 6 8
w-dist	-1.7172 1 4 6
w-const	0.6147 1 7 6
2 2	NVAR1,NVAR2
/milstuff/elmi/atime86.abdi	
/milstuff/elmi/trttot86.311	
/milstuff/elmi/rbest	
1 (2i5,f10.5)	
wdist.dat	
1 (2i5,f6.3)	
/milstuff/elmi/worktrips.dat	

IMAGE EVALUATION TEST TARGET (QA-3)

APPLIED IMAGE, Inc.
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989

© 1993, Applied Image, Inc.. All Rights Reserved