

MODAL SPLIT ANALYSIS  
FOR THE  
JOURNEY TO WORK

by

W.N. Aplin, B.E. (Civil) (Melb).

-----

A thesis submitted in partial fulfilment of the requirements for the Degree of Master of Science in the Faculty of Engineering, University of Cape Town.

The copyright of this thesis is held by the University of Cape Town.  
Reproduction of the whole or any part may be made for study purposes only, and not for publication.

Department of Civil Engineering,  
UNIVERSITY OF CAPE TOWN

May. 1974.

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

DECLARATION OF CANDIDATE

I, Neil Aplin, hereby declare that this thesis is my own work and that it has not been submitted for a degree at another University.

  
.....

W.N. APLIN

May, 1974

# TABLE OF CONTENTS

## VOLUME I

	Page
Synopsis	i
Acknowledgements	iii
Statement of the Problem	v
Justification for Research	vi
 <u>CHAPTER I: The Role of Modal Split in the Transportation Planning Process.</u>	
1.1 The Transportation Planning Process	1
1.1a Problem Definition	2
1.1b Planning in the South African Context	9
1.1c Goals and objectives	11
1.1d The Survey Phase	12
1.1e Analysis and Model Building	12
1.1f Forecasting, Evaluation and Iteration Phases	14
1.2 The Role of Modal Split in Transportation Planning	14
References: Chapter One	R.1.1
 <u>CHAPTER 2: Modal Split Analysis - A review of Literature.</u>	
2.1 The "Theory" of Modal Split	17
2.2 Demand for Travel	19
2.3 A Formal Definition of Modal Choice	20
2.2.1 Utility	20
2.3.2 Price	20
2.3.3 Income	20
2.4 Modal Choice and Indifference Theory	21
2.4.1 The value of Time	24
2.5 Beimborn's Impedance Minimization Model	26
2.6 De Donnea's Micro Economic Models of Mode Choice	29
2.7 Other Economic Based Models	32
2.7.1 Wilson's Entropy Maximizing Models	32

	Page
2.7.2 Kain's Econometric Models	33
2.8 Behavioural or Psychological Models	35
2.8.1 Perceived Costs and Times	35
2.8.2 Ewing's 'Psychophysical' Model	37
2.8.3 Demetsky and Hoel's User Perception Model	39
2.8.4 Golob's Attribute Measures	40
2.9 Empirical Modal Split Models	42
2.9.1 Wilson's Linear Regression Models for Modal Choice	42
2.10 A Critique of Modal Split Models	43
2.10.1 The Economic Based Models	44
2.10.2 Behavioural Models	46
2.10.3 Empirical Models	47
2.10.4 Summary	48
References: Chapter 2	R.2.1

### CHAPTER 3: Surveys - Designs and Methods

3.1 Reasons for the Surveys	49
3.2 Survey Objectives	50
3.3 Socio Economic and Personal Trip Information	51
3.4 Design of the Questionnaire Form	53
3.4.1 Criteria for Questionnaire Design	54
3.4.2 Formulation of the Questions	59
3.4.3 Provision for Encoding Data	61
3.5 The Questionnaire Survey Method	63
3.5.1 The Pre-Test Survey	63
3.5.2 The Pre-Test Survey Response	65
3.6 Selection of Sample	66
3.7 Final Survey Method	66
3.8 Questionnaire Response	69
3.9 Coding the Data	69
3.10 Additional Surveys	72
3.10.1 Car Time Survey	73
3.10.2 Parking Survey	74
3.10.3 Train Times Survey	75
3.10.4 Bus Times Survey	75
3.10.5 L.O.S. Index for Trains	76

3.10.6	L.O.S. Index for Buses	79
3.11	Summary	79
	References: Chapter Three	R.3.1

CHAPTER 4: Analysis of Data: Factors Affecting  
Modal Choice

4.1	Data Handling Methods	82
4.2	Modal Split Results	82
4.3	Inconvenience Factors in Public Transport	87
4.3.1	Waiting Times	91
4.3.2	Transfers	91
4.3.3	Seat Availability	92
4.3.4	Walking Distance to or from stop of station	92
4.3.5	Public Transport arrives too early/late for work	93
4.3.6	Lack of Shelter	94
4.4	Reasons for not Using Car for the Work Journey	94
4.4.1	Roads are too crowded	98
4.4.2	The Car Trip is too Expensive	98
4.4.3	Unsuitable Parking Arrangements at Workplace	99
4.4.4	The Car is used by Another Member of the Family	99
4.4.5	Travel by Car Pool	100
4.5	Reasons for Choosing to Use Car for the Journey to Work	101
4.5.1	Car Trip is Quicker	101
4.5.2	The Employee Uses his Car During Working Hours	105
4.5.3	The Car Trip is Cheaper	106
4.5.4	Drive Children to School	106
4.5.5	Other Reasons Relating to the Inconvenience of Public Transport	107
4.6	Male-Female Stratification in Modal Split	109
4.7	Modal Split as a Function of Distance	110
4.8	Economic Index	115
4.9	Travel Costs as a Function of Distance	117

	Page
4.9.1 Cars	117
4.9.2 Bus Costs	121
4.9.3 Train Costs	122
4.10 Sampling Bias	122
4.11 Influencing Factors in Modal Choice Decisions	127
References: Chapter 4	R.4.1

#### CHAPTER 5: Model Development

5.1 Model Development	129
5.2 Design of Model	130
5.2.1 Linear Least Squares Regression	130
5.2.2 Assumptions in Linear Regression Techniques	131
5.2.3 Design Objectives for the Model	134
5.3 The BMD02R Programme	134
5.4 Model Specifications	137
5.5 Zonal Regression Analysis	137
5.6 Zonal Regression Equations	139
5.7 Regression Equations at the Personal Level	141
5.8 Interpretation of the Equations	144
5.9 Use of the Models	149
References: Chapter 5	R.5.1

#### CHAPTER 6: Discussion and Conclusions

6.1 The Implications of the Model Equations	150
6.2 Conclusions	156
References: Chapter 6	R.6.1

### VOLUME II

#### Appendices

A.1.1 A Definition of a Free Market and Perfect Competition	A.1.1.1
A.2.1 Utility and Disutility	A.2.1.1

A.2.2	Budget Constraints, Utility Maximization and Substitution	A.2.2.1
A.3.1	Questionnaire Survey Form	A.3.1.1
A.3.2	The Firms' Circulars to the Staff	A.3.2.1
A.3.3	Map of Survey Zones	A.3.3.1
A.3.4	Questionnaire Coding Forms	A.3.4.1
A.3.5	Computer Programme for O.D. Survey	A.3.5.1
A.3.6	Plots of Frequency of Car Times from O.D. Study	A.3.6.1
A.3.7	Location of O.D. Survey Stations	A.3.7.1
A.3.8	Car Isochrone Map	A.3.8.1
A.3.9	Walking Time Contours for CBD	A.3.9.1
A.3.10	Train Isochrone Map	A.3.10.1
A.3.11	Line Haul Bus Service: Isochrone Map	A.3.11.1
A.3.12	Formulae for Level of Service Indices	A.3.12.1
A.4.1	Modal Split Analysis at the Workplace	A.4.1.1
A.4.2	Preference Figures at the Workplace	A.4.2.1
A.4.3	Car Pool and Parking Figures at the Workplace	A.4.3.1
A.4.4	Listing of Computer Programme MD2	A.4.4.1
A.4.5	Listing of Computer Programme RTCCBD	A.4.5.1
A.4.6	Measurement of Perceived Times	A.4.6.1
A.4.7	Graphs of Relationships between Independent and Dependent Variables	A.4.7.1
A.4.8	Listing of Datafile RDATA	A.4.8.1
A.5.1	Technique of Least Squares	A.5.1.1
A.5.2	Input Requirements for BMD02R Programme	A.5.2.1
A.5.3	Statistical Tests of Significance of Regression Equations	A.5.3.1
A.5.4	Listing of BMD02R Programme Output of Simple Correlation Matrices	A.5.4.1
A.5.5	Significance Test of Multiple Linear Regression Equations and BMD02R Output	A.5.5.1
A.6.1	Elasticity of Demand	A.6.1.1

## SYNOPSIS

The choice of travel mode for the journey to work is an important aspect in the planning of adequate transportation systems in urban areas. This choice process is complex and consequently a generalized theoretical basis for modal choice is difficult to construct. Nevertheless, empirical modelling of modal choice behaviour has enabled transportation planners to predict future travel demands for different modes. Furthermore, such models have led to a body of knowledge which has allowed researchers to explore the modal choice decision on a more theoretical basis.

This study involves the analysis of the modal choice process for White commuters in Cape Town. The investigation of the role of modal split in transportation planning is provided to illustrate the relevance of this sub process in the overall transportation planning process. An investigation of some of the theoretical and applied literature in this field indicates, that to obtain suitable and simple planning tools for modal split analysis, an empirical approach to modelling is probably the best alternative. The theoretical approaches are still in an embryonic stage and require more research before they offer a practical solution to modal split modelling.

The data collection technique used in this study involved the distribution of a questionnaire survey to a sample of employees at their workplace. The technique provided an excellent response rate and can be performed with a minimum of resources. Other detailed travel time studies are described and once again an appreciable amount of data was able to be collected with a minimum of funds and manpower. The surveys described in Chapter 3 are the first of their type performed, on such a scale, in Cape Town.

The data collected was analysed to determine the influential factors

in the modal choice process. Some of these factors were found to be significantly different from those obtained in studies performed overseas. Throughout this work emphasis is placed on findings from this study which appear to be particular to South African conditions.

A multiple linear regression model is derived which simulates the modal split between car users and train users in Cape Town's southern suburbs. Particular attention is given to the analysis of the regression model as there is evidence of misinterpretation of these models in other studies. The implications of the model for planning are discussed. It is demonstrated that such models can be used to determine the price elasticity of demand for public transport. Also an estimate of the money value of time can be obtained from the models. To the author's knowledge this study represents the first approach to modelling modal choice behaviour in Cape Town.

ACKNOWLEDGEMENTS

I wish to express my gratitude to the following:

Professor W.H. King, Department of Civil Engineering, University of Cape Town, under whose supervision this research was conducted.

Mr. B. Floor, Transport Research Centre, University of Stellenbosch, whose assistance, both financial and technical, helped to ensure the success of the questionnaire survey.

Mr. Eric Pas, Department of Civil Engineering, University of Cape Town, for his helpful comments and advice throughout this work.

The managements of all the firms who, so obligingly, assisted with the questionnaire survey.

Messrs. R. Del Mistro, P. Tomlinson and K. Sturgeon of the Cape Town City Engineer's Department for their cheerful co-operation with several aspects of this research.

Mr. Terry Griffin and Mr. Colin Goldberg for their advice on several computing problems.

The staff of the Department of Mathematical Statistics, University of Cape Town, for their advice on statistical problems.

Mrs. Ensor and Mrs. Swan of the Computer Centre, University of Cape Town, for their assistance with the punching of data cards.

The Country Roads Board, Victoria, Australia for leave of absence to undertake this study.

The Rotary Foundation, for the scholarship awarded to the author, enabling him to complete this work.

Miss Pat Nicholson, for her patient drafting of the drawings for this manuscript.

Miss Lesley Jennings, for her conscientious typing of this thesis.

Mr. Charles Basson, for the printing of the manuscript.

The Argus for their kind permission to reproduce the photograph shown in Chapter 3.

Finally, I would especially like to convey my thanks to my wife, Norma, whose patient understanding and help during this work have been much appreciated.

STATEMENT OF THE PROBLEM

The purpose of this thesis is to investigate the causal factors influencing the commuter's choice of travel mode for the journey to work. This is known as modal split and constitutes one stage in the transportation planning process.

A preliminary definition of modal split can be stated as follows:-

Modal split is the result of the commuter's choice of mode of travel for a journey, given that the decision to make that journey has already been made.

The problem involves the collection of data concerning commuters' travel habits, transport system characteristics and land use parameters. This data provides a basis for analysis and for the development of a mathematical model to explain the modal split process. The purpose of this thesis is to develop such a model for the journey to work in Cape Town.

JUSTIFICATION FOR RESEARCH

The prediction of future choice of mode of travel is probably the most crucial yet least understood of the phases in transportation analysis. To predict future modal split, it is essential to understand the present choice behaviour of commuters. Furthermore, accurate modal split information is one of the basic requirements for rational transport investment decisions and an essential tool in the formulation of a balanced urban transportation proposal.

Research in this field can also be justified for a particular region on the basis of the rather suspect practice of the local adoption of modal split models which have been developed for other regions. There is evidence that such a practice ignores basic differences in urban parameters between regions which can grossly distort the explanatory capabilities of models developed for other urban areas.

As there appears to have been no detailed study of modal split for Cape Town, the particular object of this study is to investigate the modal choice process for White commuters in Cape Town.

## CHAPTER ONE

### THE ROLE OF MODAL SPLIT IN THE TRANSPORTATION PLANNING PROCESS

In this chapter several basic concepts are presented which provide a necessary perspective for the remainder of the thesis. Before exploring modal split in detail it is essential to recognise its role in the overall transportation planning process.

A description of this planning process is provided with particular reference to the problem recognition stage in general and several South African problems in particular.

A brief discussion of the methodology of transport planning is outlined and finally the role of modal split in the planning process is discussed.

#### 1.1 THE TRANSPORTATION PLANNING PROCESS

Transportation has had a major influence in the structuring of cities and regions. Rees<sup>1</sup> notes that in early Roman towns streets were designed to accommodate a rider carrying his lance crossways. Similarly modern urban freeways are designed to cater for expected vehicular flows and movements.

However, with an increasing sophistication in transport technology and a corresponding increase in traffic volumes, the science of transportation planning has undergone significant transformations in the last 15 years.

The rapid evolution of the motor car and the failure of earlier traffic planning techniques to predict future travel

demands has prompted the evolution of more rigorous planning methods. The conventional planning process consists of six basic stages.

Assuming the decision to plan has been made, these phases are:

1. Problem Definition
2. Formulation of Goals and Objectives
3. Survey - Data Acquisition
4. Analysis and Model Building
5. Forecast
6. Evaluation.

Within each of these stages various sub-activities occur and these are shown schematically in figure 1.1.

The decision to plan necessitates a recognition of the need to plan which in turn connotes the identification of the problems.

Fuller<sup>2</sup> notes the importance of this problem definition phase.

"A problem adequately stated is a problem fundamentally ripe and potential of solution."

#### 1.1a PROBLEM DEFINITION

The existence of a transportation problem in most large cities is apparent to most people, but the ability to recognize and define the causes of such problems is not so evident.

There are several well documented trends which have led to the transportation problems experienced today. Firstly the rapid growth in urbanization has concentrated an increasing proportion of the world's population in urban areas.

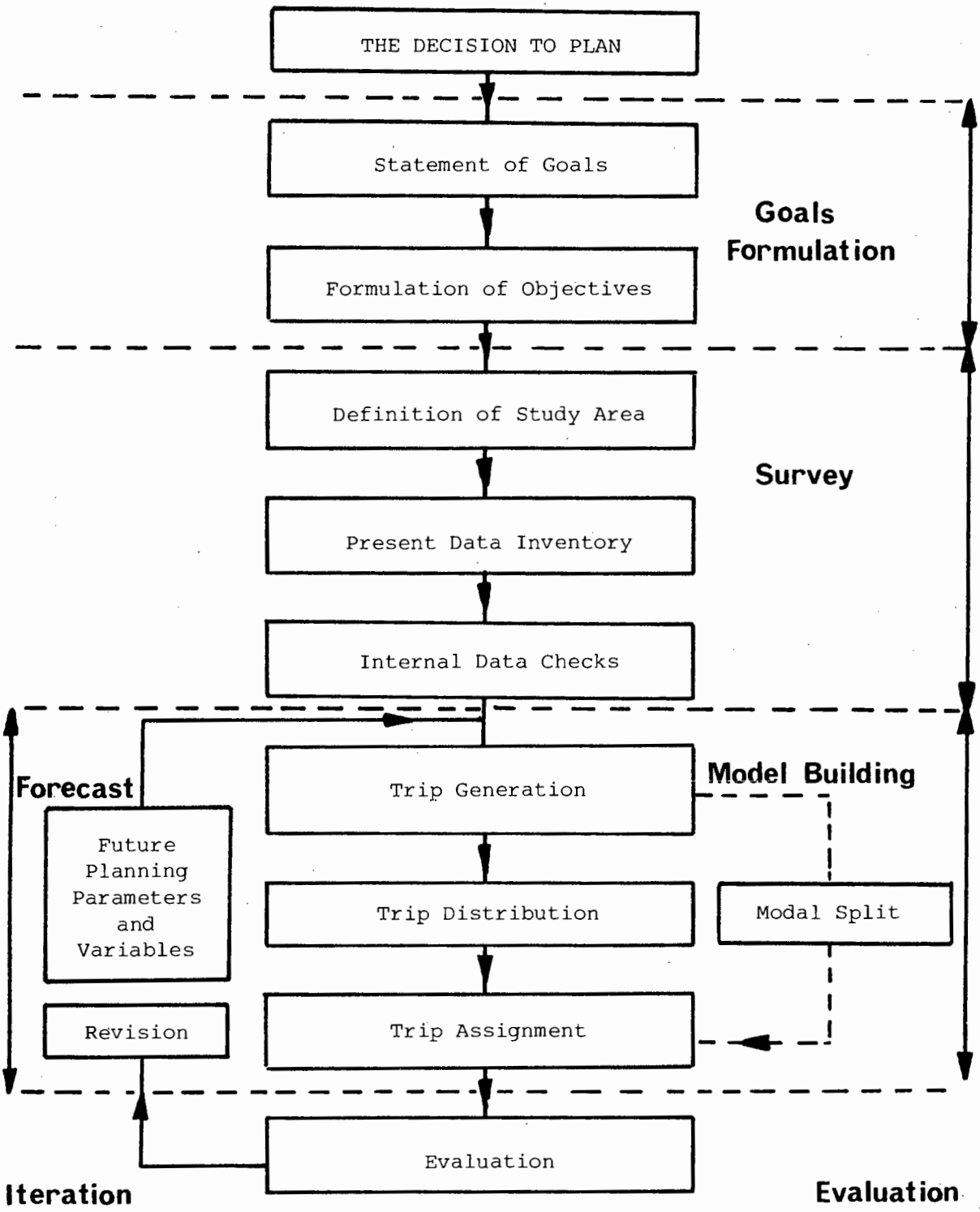


Fig 1-1

The Transportation Planning Process

Davis<sup>3</sup> notes that in Britain in 1801, only 10% of the population lived in cities containing 10000 people or more in size, whilst in 1951, 69% of Britain's population were situated in cities of these sizes.

The situation in South Africa is similar, census data<sup>4</sup> indicate that in 1950, 74,8% of the white and coloured population were in urban areas, whereas in 1970, 82,4% were located in urban areas.

Clark<sup>5</sup> terms such locational choices between cities and regions as macro-locational choices and the choices of location within a city as micro-locational choices.

Macro-locational theory serves to explain the phenomenon of increasing population concentration which is described above as urbanization. Similarly, micro-locational theory describes what is commonly referred to as urban sprawl.

Clark<sup>6</sup> postulates a model for micro-location within cities of the form

$$Y = Ae^{-bx} \quad (1.1)$$

where Y is the gross residential population (persons/hectare)

x is the distance from the city centre

A is the hypothetical city centre density where x = 0

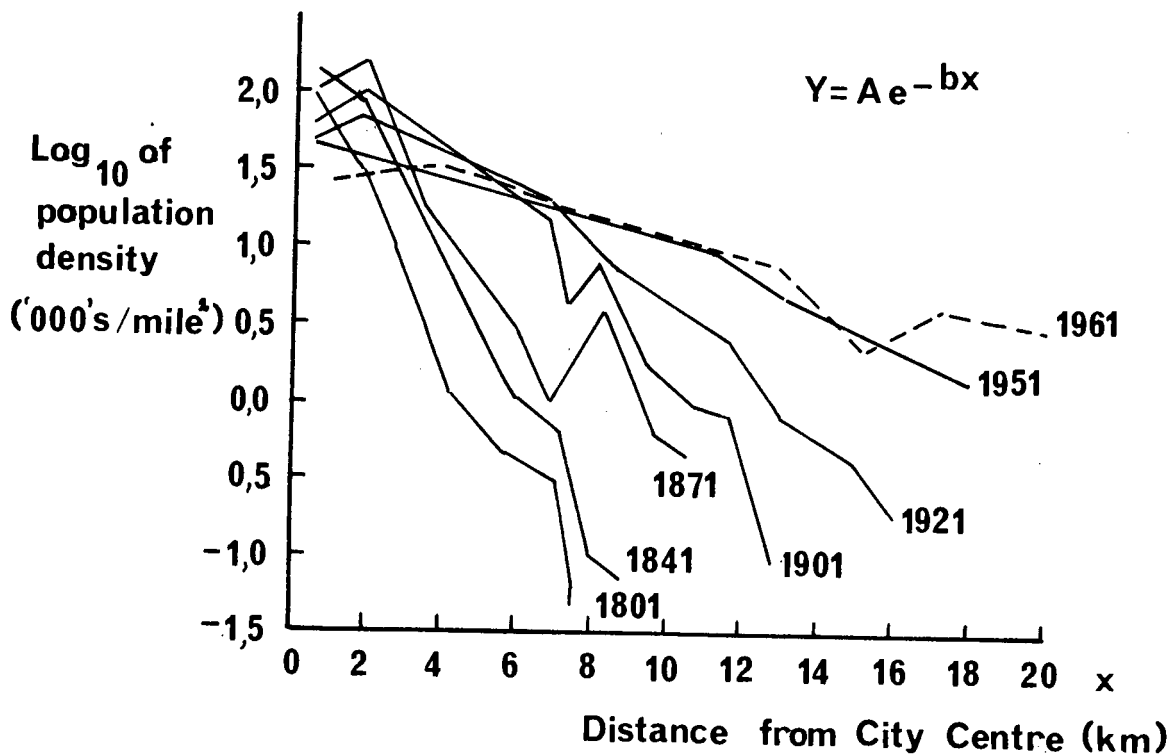
b is a coefficient which is an inverse measure of population dispersal.

Figure 1.2<sup>7</sup> shows a plot of population density versus distance from the centre of London and illustrates the sprawl of London's population from 1801 to 1961.

## Distance vs Population Density

### London 1801 - 1961

**Fig 1.2**



**Values of A and b**

Year	A	b
1801	269	1.26
1841	279	0.94
1871	224	0.61
1901	170	0.37
1921	115	0.27
1951	62	0.20
1961	57	0.17

**Figure 1-2 The Relationship between Residential Density and Distance from Centre of London 1801 - 1961**

The corresponding  $b$  values\* from this graph for London in 1801 are  $b = 1,26$ , and for 1961,  $b = 0,17$ . Thus at the micro-locational level one finds the tendency for decreasing concentration.

The implications of this dispersal for the transport planner are substantial. One consequence is an increase in the average length of the work trip in cities. Paterson<sup>8</sup> cites the case of the increase in the average length of the work trip to Melbourne's central business district (CBD). Between 1961 and 1966 the average distance travelled increased 7% from 9,9 km to 10,6 km. Similar figures provided by the U.S. Bureau of Public Roads<sup>9</sup> show an increase in the median length of the work trip of between 3,6% and 18,7% between 1941 and 1951 in American cities.

This increase in the trip length necessitates the provision of increased transport facilities to cater for these trends in urban dispersal.

A further contributory factor to the transportation problem is the rapid growth in car ownership. The figures in Table 1.1<sup>10</sup> indicate the increase in car ownership in the Cape Province of South Africa.

In the eight year period 1962-70 the number of vehicles compounded at 10% per annum for the white community and at 16% per annum for the coloured community.

These growth rates are significant when compared with the anticipated growth rate of vehicles in Melbourne (a city

---

\* Clark has "fitted" straight lines to the curves shown in figure 1.2. He has converted these slopes from units of  $\log_{10}$  (density)/mile to units of  $\log_e$  (density)/mile.

This model is used in this chapter merely as a graphical illustration of urban dispersal.

TABLE 1.1Vehicle ownership in the Cape Province

Date	Location	Licenced Motor Vehicles			
		White	Coloured	Bantu	Total
30/6/62	Total for Cape	206724	10729	4678	242687
	Vehicles per 1000 persons	164	7	unknown	
30/6/70	Total for Cape	379928	39113	11527	457818
	Vehicles per 1000 persons	345	23	8	

of similar size to Cape Town) of 5,5% per annum for the period 1964-1985<sup>11</sup>.\*

From these three concepts of urbanization, urban dispersal and rapid growth in vehicle ownership stem further problems.

Congestion of urban arterial roads in peak periods is a common phenomenon and is a result of many factors. Some of these are:

- (i) The vehicle growth rate
- (ii) The major radial flow of traffic in C.B.D. orientated cities
- (iii) The acute peaking of work and school trips
- (iv) The low person per vehicle ratio
- (v) The decline in public transport patronage.

---

\* The City Engineer's Department of the City of Cape Town adopted Melbourne's growth figures as an indication of the planning requirements for Cape Town. In 1957 the two cities had similar vehicle populations, but in 1970 Cape Town had 249000 vehicles and Melbourne had 110000<sup>12</sup>. This illustrates the problems resulting from the injudicious adoption of overseas findings for local planning.

This decline in the use of public transport (p.t.) is related to urban dispersal. Conventional urban mass transit forms operate most efficiently in high density residential and business areas. Obviously urban dispersal has a detrimental influence on the efficiency of such fixed transit systems.

The decline in the use of public transport is well documented throughout the most Western cities. In the U.S. there has been an absolute decline in p.t. patronage in recent years, as evidenced in Table 1.2<sup>13</sup>.

TABLE 1.2

Total transit passengers in the United States by mode at 10-yearly intervals 1935-1963  
(Millions)

Year	Railway*	Trolley Coach	Motor Bus	Grand Total
1935	9512	96	2618	12226
1945	12124	1244	9886	23254
1955	3077	1202	7250	11529
1963	2313	657	6425	9395

Whereas in South Africa Floor<sup>15</sup> provides figures which indicate that there has been an absolute increase in the numbers of suburban rail passengers since 1963/64, but the number of journeys per white person has declined steadily in recent years.

These figures are shown in Table 1.3.

This apparent reversal of overseas trends in South Africa can be explained by the large proportion of lower income

TABLE 1.3

Suburban passenger train journeys in South Africa  
1962/63 to 1966/67

Year	Total Journeys (Millions)	1st Class Journeys* (Millions)	Total Relative† Patronage Journeys/Person	1st Class Relative† Patronage Journeys/Person
1962/63	317,8	101,3	42,6	37,6
1963/64	360,4	107,7	47,0	39,0
1964/65	400,2	113,1	50,9	40,0
1966/67	423,4	116,7	52,4	40,3
1967/68	434,4	113,8	52,3	38,4

\* 1st Class travel is reserved for White commuters.

† To obtain these ratios, the author used the 1960 and 1970 census figures and interpolated urban population figures for the years 1962 to 1967. The calculated growth rates were 2,7% per annum and 2,4% per annum for total urban and white urban populations respectively.

non-whites who are not able to afford private vehicles. Such a person is termed a transit captive. These figures are a further illustration of the inadvisability of adopting overseas findings without due regard to the South African situation.

Having investigated some of the problems in urban transportation generally, it is opportune to examine further unique problems in the South African context.

#### 1.1b PLANNING IN THE SOUTH AFRICAN CONTEXT

In his monograph, Floor<sup>15</sup> refers to the problems of adopting solutions to South Africa's problems from overseas studies.

"Indeed in several instances solutions (from overseas)

have been regarded as so obvious ..... that the preliminary investigations to establish and define the problem have received scant attention."

There are several problems peculiar to South Africa which should be appreciated before applying methodologies from abroad.

Firstly locational trends in urban areas are different from those found in the United States or Britain.

In traditional locational theory the basic assumption is that the land market exists under conditions of perfect competition in a free market.\*

In practice such a theory fails to some extent or another. Controls like zonings, policy decisions on land development and large industrial and shopping firms' investment decisions distort the above assumptions. However, with certain adjustments the original assumptions serve as a basis for locational theory. But in South Africa the situation is further affected by other factors so that the basic assumptions are invalid. Rees<sup>16</sup> writes:

"Social mores compound the problem [of a free land market] even further and where these are transmitted into formal laws, as in South Africa, in the form of group area legislation the 'free' land market is distorted out of all recognition."

For example, Clark's equation (1.1) cited earlier, would, if applied in South Africa, give completely erroneous values, because it has been developed for cities where the free marketing of land exists.

---

\* For a description of the terms free market and perfect competition, the reader is referred to Appendix 1.1.

One other factor which is particular to South Africa, perhaps in degree only, is the percentage of captive riders in the cities. The study of latent demand, that is the demand for transportation by captives, is a field which has received scant attention in most overseas studies. Hoel et al.<sup>17</sup> comment that:

"....., little, if anything, has been done to estimate transit system benefits for the disadvantaged members of the community."

It is a field of study which deserves close attention in South Africa, but to the author's knowledge, no such study has been undertaken in the Republic.

Having noted several of the basic problems in urban transportation, and the inadvisability of adopting foreign solutions to local problems, the next phase in the planning process is that of the goals formulation.

#### 1.1c GOALS AND OBJECTIVES IN PLANNING

The formulation of goals and objectives is most important, but is often ignored by many writers<sup>18</sup>. The failure to state adequately one's goals often has serious repercussions. The cancellation of proposed freeway projects in Boston, Toronto, London, Adelaide and Edmonton<sup>19</sup>, to name a few, is a result of social pressures. Society values certain environmental qualities, and safeguards for these should be incorporated in the original goals formulation.

Furthermore, this goals formulation needs to be consistent within the context of active urban planning. Otherwise serious contradictions could occur in final plans and the implementation of these plans.

McLoughlin<sup>20</sup> differentiates between goals (or ideals) and the identification of objectives (or detailed design statements). He notes that goal statements are, by their nature, "somewhat vague and general", whereas objectives are "more detailed" and provide "operational measures" of the rate of progress towards the goals.

The third step in the planning process is that of survey or data collection.

#### 1.1d THE SURVEY PHASE

The purpose of the survey phase is to provide the essential data for the modelling phase. Initially a survey area needs to be defined, and information concerning existing land use, traffic patterns, transport facilities, commuter characteristics and other planning parameters needs to be collected for analysis.

A more detailed discussion of survey methods is provided in Chapter 3.

#### 1.1e ANALYSIS AND MODEL BUILDING

The purpose of model building is to determine a mathematical form which will simulate observed patterns and (within the bounds of certain assumptions) provide a tool for predictive purposes.

The lack of success of the modelling phase in many instances has been noted by a number of authors<sup>22, 23</sup>. This is partly due to the complexity of the problem to be modelled and is discussed further in Chapter 2.

There are four sub-models in the modelling stage. They are:

1. Trip Generation.
2. Trip Distribution.
3. Modal Split.
4. Trip Assignment.

A detailed description of each model is not warranted in this work, but a brief outline of each is provided in order to explain the role of modal split in this modelling phase.

Trip generation is measured by the number of trips in a unit of time generated by (attracted to or produced by) a unit of area such as a zone, or a household, or a unit of activity such as a shopping centre or factory. The method usually adopted to construct such a model is to isolate variables likely to influence trip generation and relate them to the measured trips generated per household (say) by means of linear regression techniques.

The next stage can either be that of modal split or trip distribution. (The position of modal split in this modelling phase is considered in the next section.) Once the number of trips produced by or attracted to a zone has been determined from the trip generation model, it is essential to know the origins and destinations of these trips, i.e. the trip ends.

Synthetic models\* such as the gravity<sup>24</sup> model and the intervening opportunities model<sup>25</sup> are used to simulate the distribution of the trips.

Modal split then apportions the trip-makers to the available modes and thus at this stage one has determined the

---

\* The term synthetic model indicates that the measured trip ends, i.e. origin and destinations (O.D.) are not used to establish the model (as is the case with regression models). However, this O.D. data is used to calibrate the model's coefficients<sup>24</sup>.

number of trips to be catered for, their origins and destinations, and their mode usage. All that remains is to assign each of the trips to a particular route on the existing transportation network.

This final stage is known as trip assignment and generally assumes route choice to be determined on a minimum route time or cost basis.

#### 1.1f FORECASTING, EVALUATION, AND ITERATION PHASES

These models should be so developed that, by forecasting future values for their constituent variables, future travel demand can be predicted. An obvious assumption therefore is that the model's logical basis is non-transient and that the constituent variables are able to be predicted.

The next phase is to evaluate the network model in terms of the original goals.

Furthermore, in this evaluation phase these original variable forecasts must be re-examined in the light of the proposed future plans. For example, a residential growth rate may be markedly different from that originally envisaged if the proposed future network has altered the original bases of the forecasting assumptions.

Consequently it is essential to re-evaluate each proposal in the context of the original forecasting assumptions. This re-evaluation process is known as the iteration phase, and it is a phase which is often ignored in transport studies.

#### 1.2 THE ROLE OF MODAL SPLIT IN TRANSPORTATION PLANNING

Modal split is not a defineable single phase, but is a product of many of the traveller's decision processes.

It is partly for this reason that modal split can be considered either before or after the trip distribution stage.

To consider modal split after trip generation involves an analysis of the variables peculiar to the origin (e.g. journey purpose, mode availability, income). Whilst after trip distribution the choice can be analysed in terms of O.D. characteristics (e.g. relative times and costs of the trip, availability of parking and levels of service).

It is assumed in both cases that the decisions to make a trip and the trip purpose have already been made and therefore two decision processes can be described as follows:

- (i) at the trip generation stage the traveller makes the decision as to the mode he will use to make the trip. He then chooses his destination to suit that mode and to satisfy the purpose of his trip, or
- (ii) after trip distribution, the traveller has decided on his trip destination and then makes his choice of mode to suit that destination.

Hutchinson<sup>26</sup> refers to the former process as captive modal split and the second process as choice modal split. It is apparent that a captive traveller will first decide on a mode which is available to him, and then decide on a destination which is best served by that available mode, whereas a person with a choice of modes is more likely to decide where he wants to go and then choose a mode which best serves that destination.

The method adopted in this thesis is to investigate the modal split process independently of the planning process, in an attempt to examine all the causal factors at both the trip generation and distribution stages.

The method assumes that all the commuters have chosen their destinations. An analysis is then made to define the choice of mode process for these commuters.

REFERENCES : CHAPTER ONE

1. Rees, B.J., "Group Areas Legislation - Its Impact on Urban Structure and Movement Patterns with Special Reference to Cape Town".  
Proc. 5th Quinquennial Convention S.A.I.C.E.,  
Transportation and the Environment, August 1973,  
Paper 3.2; p. 3.2.02.
2. Fuller, B., "Comprehensive Design Strategy", World  
Design Science Decade, 1965-1975, Phase II (1967),  
Document 5, p. 39.
3. Davis, K., "The Urbanization of the Human Population",  
A Scientific American Book, 1966.
4. Rees, B.J., Op Cit, p. 3.2.07. (Source: Department of  
Statistics - Census Data.)
5. Clarke, C., "Economics of Urban Areas", Proc. Tewksbury  
Symposium, University of Melbourne, July 1970,  
Chapter 1, Paper 1, p. 1-37.
6. Ibid, p. 1-4.
7. Ibid, p. 1-5.
8. Paterson, J., "The Dynamics of Urban Change", Proc.  
Tewksbury Symposium, University of Melbourne, July  
1970, Chapter 1, Paper 2, p. 1-37.
9. Lapin, H.S., "Structuring the Journey to Work",  
Pennsylvania U.P., 1964, p. 40.
10. Rees, B.J., Op Cit, p. 3.2.09.
11. Guerin, N.S., "Proposed Road Network", Proc. Symposium  
Balanced Urban Transportation in Melbourne,  
I.E. Aust., Victoria Division H. + T.E. Branch,  
April 1970, p. 1.
12. City Engineer's Department - City of Cape Town,  
Metropolitan Area - Population and Vehicle Data;  
Drawing Ref. No. TPR 4435, September 1971  
(unpublished).
13. Meyer, J.R., Kain, J.F. and Wohl, M., "The Urban  
Transportation Problem", John Wiley, New York, 1966,  
p. 57.
14. Floor, B.C., "The Urban Transportation Problem in South  
Africa - an Introductory Monograph", Transport  
Research Centre, University of Stellenbosch, 1968,  
p. 35. (Source: S.A.R. & H., Annual Reports  
1962/63 - 1966/67.)
15. Floor, B.C., Ibid, p. 7.

16. Rees, B.J., Op Cit, pp. 3.2.05-06.
17. Hoel, L.A., Perle, E.D., Kansky, K.J., Kuehn, A.A., Roszner, E.S. and Nesbitt, H.P., "Latent Demand for Urban Transportation", Transportation Research Institute, Carnegie-Mellon University, Pittsburgh, 1969.
18. Lane, R., Powell, T.J., Prestwood-Smith, P., "Analytical Transport Planning", Duckworth, London, 1971, p. 18.
19. Hutchinson, B.G., "A Critical Assessment of the Urban Transportation Process", Proc. 5th Quinquennial Convention S.A.I.C.E., Transportation and the Environment, August 1973, p. 5.7.03.
20. McLoughlin, J.B., "Urban and Regional Planning - A Systems Approach", Faber and Faber, London, 1970, p. 106.
22. Hutchinson, B.G., Op Cit, pp. 5.7.03/04.
23. Brown, R.J. and Burton, R.W., "Some Conceptual Issues of the Transport Problem", Proc. 5th Quinquennial Convention S.A.I.C.E., Transportation and the Environment, August 1973, p. 6.1.10.
24. Clark, N.F., "Techniques and Procedures for Transport Planning", Traffic Engineering Practice, 2nd edition, Department of Civil Engineering, University of Melbourne, 1969, Chapter 7, p. 7.12.
25. Creighton, R.L., "Urban Transportation Planning", University of Illinois Press, Chicago, 1970, pp. 56-58.
26. Hutchinson, B.G., Op Cit, p. 5.7.03.
27. Lipsey, R.G. and Steiner, P.O., "Economics", 3rd edition, Harper and Row, New York, 1972, p. 54.

## CHAPTER TWO

### MODAL SPLIT ANALYSIS - A REVIEW OF THE LITERATURE

This chapter contains an account of past research into modal split analysis. The approach is one of examining several mode choice models for their methodologies and theoretical bases.

The models are discussed under three headings:

1. Economic Models,
2. Behavioural Models, and
3. Empirical Models.

A critique of the different methods is given at the end of the chapter.

#### 2.1 THE "THEORY" OF MODAL SPLIT

To apply the word theory to modal split research is pretentious. For, as Wheeler<sup>1</sup> notes,

"..... generalizations from empirical data on the journey to work have not led to a body of principles from which theory might be generated."

Nevertheless, there has been a growing volume of research aimed at reconciling modal choice in terms of well-established economic theory of consumer choice.

The trend is a promising one but the nature of commuters' choice of travel mode involves some unique problems which still require further study.

Furthermore, modal choice is not simply an economic choice, but involves attitudes and perceptions as further

subjective variables in the decision process. The problem is complex and consequently early studies ignored the question of searching for common patterns in the decision process and relied rather on empirical descriptions of the process.

There is the further question of whether the model should be aggregated (at, say, a zonal level) or left discrete (at the person level). Past practice has been to aggregate at a zonal level, but contemporary trends, especially in the behavioural models, have been towards little or no aggregation. This question is further discussed in Chapter 5 in relation to the model developed in this thesis.

One common feature of all the models is the recognition and isolation of causal factors. Several writers have constructed exhaustive lists of possible influential factors. Wilson<sup>2</sup> has listed 72 possible factors and Golob et al.<sup>3</sup> no less than 91. Despite such formidable numbers, these variables can generally be grouped under four headings. These are:

- (i) Trip Characteristics,
- (ii) Traveller or Commuter Characteristics,
- (iii) Transport System Characteristics, and
- (iv) Origin and Destination Zonal Characteristics.

The problem of the inclusion of variables will be discussed more fully when considering regression models. For the present it is sufficient to appreciate that the inclusion of particular variables is a result of the modeller's hypotheses and does not follow some inveterate rule.

The best means of understanding modal choice "theory" is

to investigate some of the models developed by other writers and the basic assumptions contained in these models.

## 2.2 DEMAND FOR TRAVEL

The economist has been using models to explain consumer behaviour for many years, so to apply these to modal split behaviour is a natural extension of these studies. However there is one basic difference between the demand for travel and the demand for more explicit goods and services (e.g. food, cars, entertainment). Whereas the consumer receives satisfaction from the consumption of a chocolate, he receives no satisfaction from travelling per se. A commuter makes a trip which contributes to his satisfaction at the destination but the trip itself provides him with no satisfaction.\* The economist defines this demand for travel as a derived demand.

The term demand (for transportation) itself is a source of debate amongst economists. Truett and Balek<sup>4</sup> consider that planning models lack the necessary basis of "a suitable definition of demand for transportation - a definition that reflects - the causative factors which create the demand". For the purposes of this thesis the author has adopted the following definition proposed by Beesley and Roth<sup>5</sup>:

"By 'demand' is meant the amount which, on the average, a given quantity of vehicles (or persons) will be prepared to pay rather than forego the opportunity of making the trip."

---

\* An exception might be a leisurely trip made to view the scenery; but in the context of the work journey, this assumption holds.

## 2.3 A FORMAL DEFINITION OF MODAL CHOICE

As a preliminary to a formal definition of modal choice it is necessary to define certain terms used in economics.

### 2.3.1 UTILITY

The concept of utility is an important one and the reader is referred to Appendix 2.1 for a detailed derivation of utility. For the purposes of this thesis the utility of an object will be considered as "the maximum sacrifice which each consumer (traveller) would be willing to make in order to acquire the object"<sup>6</sup>.

Utility is relative to the consumer and its units of measurement are arbitrary. The economist has opted for the term 'utile' as a measure of utility. Disutility is a negative measure of marginal utility (see Appendix 2.1).

### 2.3.2 PRICE

Throughout this thesis the word price is used to denote "a vector quantity incorporating a number of components which contribute positively to the traveller's total disutility (money outlays, travel times, inconveniences, and so on)"<sup>7</sup>. Consequently, where the terms "influencing factors/variables" or the like have been used earlier, these can now be regarded as components of the generalized price vector.

### 2.3.3 INCOME

Furthermore, the characteristics of the traveller can be considered as components of the "income" vector - a vectorial measure of the ability to purchase.

Kemp<sup>7</sup> now defines modal choice in these generalized terms as:

"the demand for one mode  $j$  (measured by, say, trips per capita in unit time), as determined by the price of that mode, the relative prices of competitive modes, and the "income" or "total expenditure" of the average consumer under consideration."

This formal definition is a necessary and sufficient statement for the purposes of this thesis.

#### 2.4 MODAL CHOICE AND INDIFFERENCE THEORY

Consumer choice theory considers the economic man as a "rational consumer" who maximizes his utility with respect to each consumption activity. He does this within his budget constraints (see Appendix 2.2 for a description of budget constraints).

Furthermore, for a two component income vector there will be a convex curve (with respect to the origin) representing combinations of the two components which have equal utilities to the consumer (Appendix 2.2).

Such a curve is known as an indifference curve and a series of such curves constitute an indifference map. Several writers<sup>8,9,10</sup> have discussed modal split in relation to indifference theory.

An indifference map for a traveller is constructed in figure 2.1a with income vector components of salary and leisure time for the axes.

The traveller attempts to maximize these income components from the effects of the travel price components. If the origin of the indifference curves of figure 2.1a is

transformed to the point O' and the axes now made to represent components of the price vector, the resultant curves (fig. 2.1b) are convex, representing loci of equal disutility. Such curves are referred to as impedance curves<sup>8</sup>.

Several things should be noted about these curves. Firstly in the indifference map (fig. 2.1a) a person would rather be above any particular curve, that is the further a curve is away from the origin, the higher the utility it represents. As can be seen from Appendix 2.2, his position is determined by his budget constraint function and in fact his utility is maximized, where his budget constraint line is tangential to an indifference curve.

This provides one with a more rigorous definition of the term captive traveller, mentioned earlier. If each mode of travel were represented by a point in this impedance space (fig. 2.1b), it is apparent that certain budget constraints would preclude the use of some of these modes. For example, if a traveller had a budget constraint represented by line A-B in figure 2.1a, he would be unable to use the car as a mode of transport. He would therefore be considered a transit captive.

In the impedance situation, the person will choose the mode with the least impedance - that is, he would rather choose a mode on a curve closer to the origin.

The same modes are marked on figure 2.1b, and it can be seen that BUS<sub>2</sub> offers the greatest impedance, and car the least. But there are other components to the price vector and these could be included by considering additional dimensions. If, for example, a third component of the price vector of, say, comfort were to be plotted, a series of

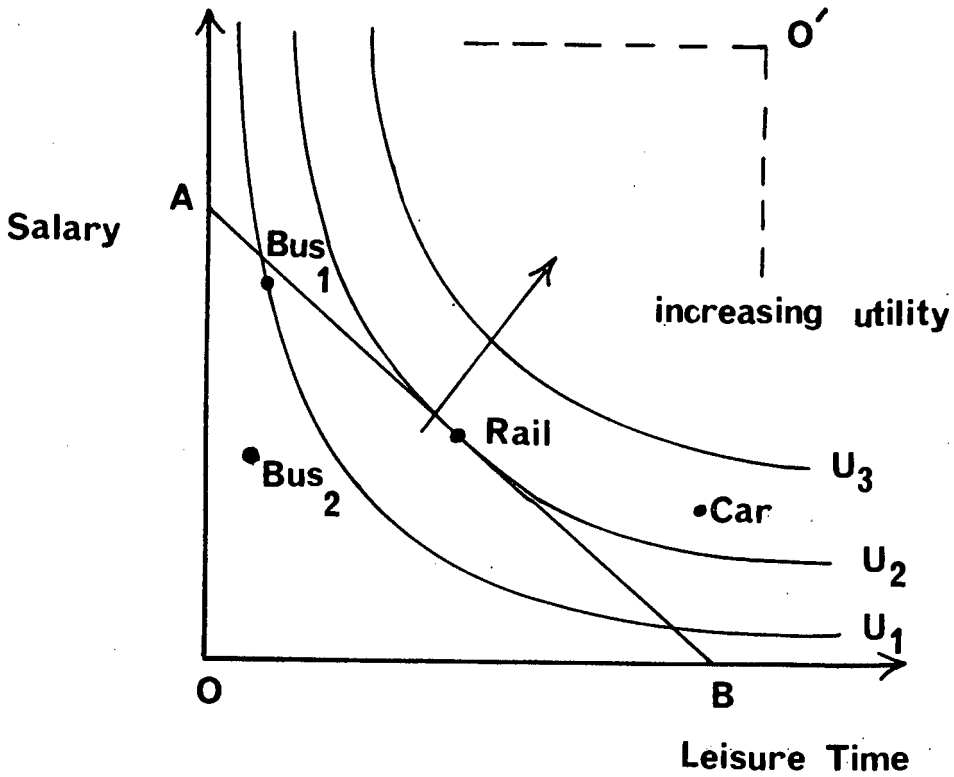


Fig 2-1a

An Indifference Map  
Salary vs Leisure Time

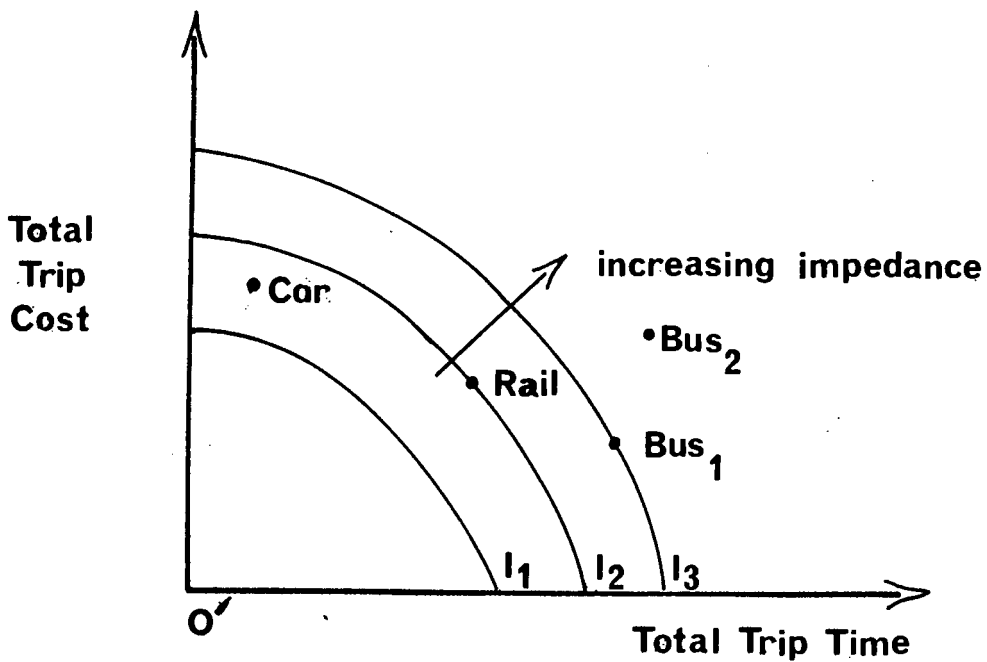


Fig 2-1b

A Family of Impedance Curves

surfaces would result.

#### 2.4.1 THE VALUE OF TIME

A feature of the indifference curves is their convex shape. As is shown in Appendix 2.2, the slope of these curves is a measure of the marginal rate of substitution (MRS) of salary for leisure time (for the case considered here).

This raises a further concept - that of the money value of time. It is a notion which is either explicitly or implicitly inferred in all the studies into modal choice.

Intuitively one recognizes that a business man regards "time as money" and that he would place some value on being able to travel on a faster (but perhaps more costly) mode.

But the housewife, on a shopping trip, would not ascribe the same value to time saved as the businessman. Thus we find that time value in monetary terms is income and trip purpose dependent.

Thus the MRS of salary and leisure time will differ for different commuters. Consequently the shapes of the curves will differ also. So for a common trip purpose one could draw hypothetical impedance curves for various income groupings as shown in figures 2.2a-c. In other words there is an interaction between a component of the income vector and the shape of the impedance curve in terms of price vector components.

From figures 2.2a-c it can be seen that the value a person places on his time savings will affect his choice of mode. The mode impedances remain the same but the curves alter the ranking of their selection.

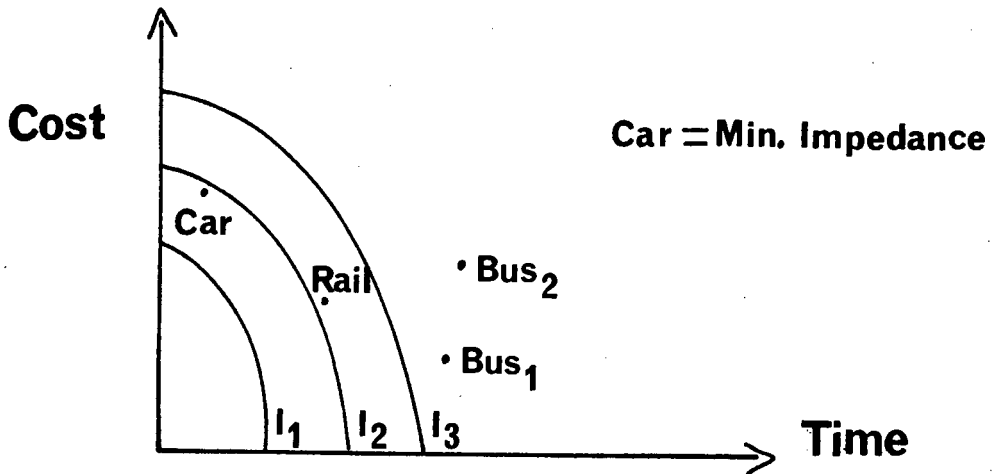


Fig.2-2a Impedance Curves High Value of Time

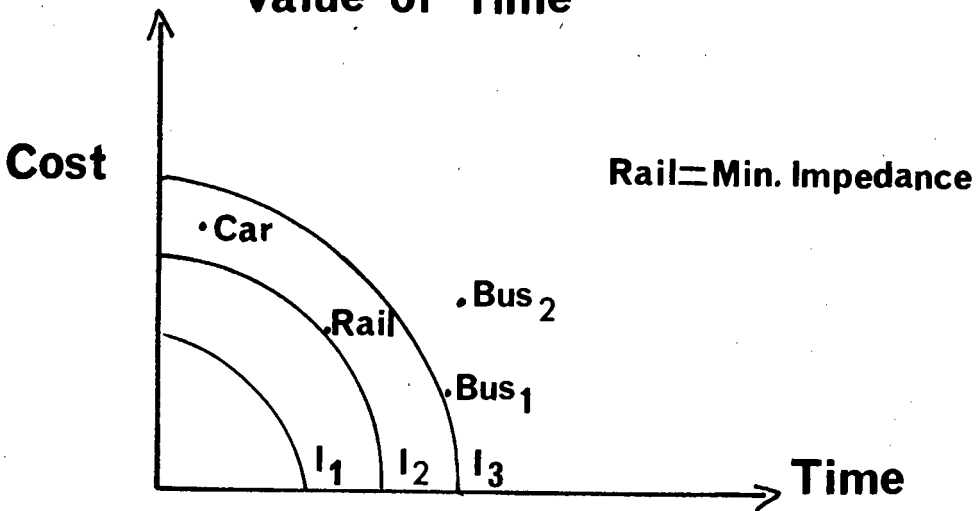


Fig.2-2b Impedance Curves Medium Value of Time

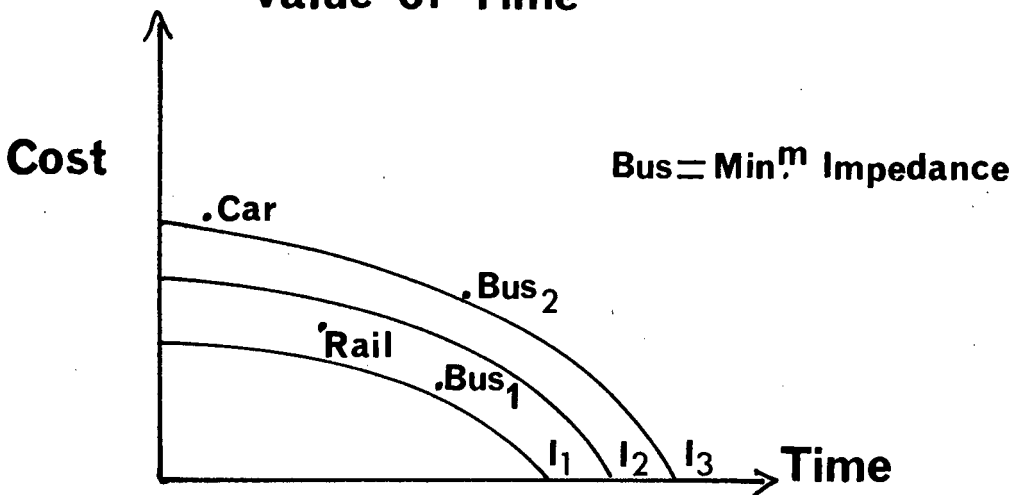


Fig. 2-2c Impedance Curves Low Value of Time

From figure 2.2a a person who values time highly would choose car as a first preference, whereas a person with a low value of time (figure 2.2c) would choose bus as his first preference.

## 2.5 BEIMBORN'S<sup>8</sup> IMPEDANCE MINIMIZATION MODEL

Beimborn, working on the theory developed in the previous sections, has formulated a model on the assumption that "a traveller selects the travel alternative that minimizes his concept of impedance".

Initially he develops an impedance function of the form:

$$I_{ijkm} = \left[ \text{Time}_{ijm}^\alpha + \left( \frac{\text{cost}_{ijm}}{V_k} \right)^\alpha + \text{other price vector components} \right]^{\frac{1}{\alpha}} \quad (2.1)$$

where

$I_{ijkm}$  = impedance of choice m between points i and j for traveller k,

$\text{Time}_{ijm}$  = total trip time between points i and j via choice m,

$\text{Cost}_{ijm}$  = total trip cost between points i and j for choice m,

$V_k$  = value of time for traveller k, in Rand's/hour,

and  $\alpha$  = a coefficient where  $\alpha \geq 1$  and has the same value everywhere in the equation.

Other components of the price vector could be included in equation (2.1), each divided by a  $V_k$  term representing the relative value of that component to traveller k.

In terms of the generalized price vector the impedance function could be written as:

$$I_{ijkm} = \left( \sum_{\ell=1}^n \left( \frac{P_{\ell m}}{V_{\ell k}} \right)^\alpha \right)^{\frac{1}{2}} \quad (2.2)$$

where

$P_{\ell m}$  = the  $\ell^{\text{th}}$  component of the  $n$  dimensional price vector  $P$  for mode  $m$ ,

$V_{\ell k}$  = the value of component  $\ell$  for traveller  $k$ .

Equation (2.1) now represents a family of impedance curves\* which pass through two points as shown in figure 2.3. The value of  $\alpha$  is necessarily greater than or equal to unity in order to obey the law of diminishing MRS (see Appendix 2.2).

Beimborn postulates that the impedance function per se is not a good predictor of modal choice. People's perceptions of modal price vectors may result in a mode being chosen which does not have the minimum impedance function. Obviously if one could determine what these perceived price vectors were for each traveller, the impedance function would probably suffice as a modal choice indicator.

However, to measure people's perceptions is difficult, and in an attempt to avoid this, Beimborn develops a choice function to explain the probability that mode A is preferred to mode B (Pr [A over B]).

Beimborn's choice function is:

$$\text{Pr [A over B]} = \frac{1}{2} \left[ 1 - e^{-\left( \frac{K_{AB}}{\text{SLOPE}} \right)^2} \right] + 0,5 \quad (2.3)$$

where

$$K_{AB} = I_{ijKA} - I_{ijKB}$$

---

\* Beimborn<sup>8</sup> refers to these curves as a family of indifference curves but in fact the curves represent impedance.

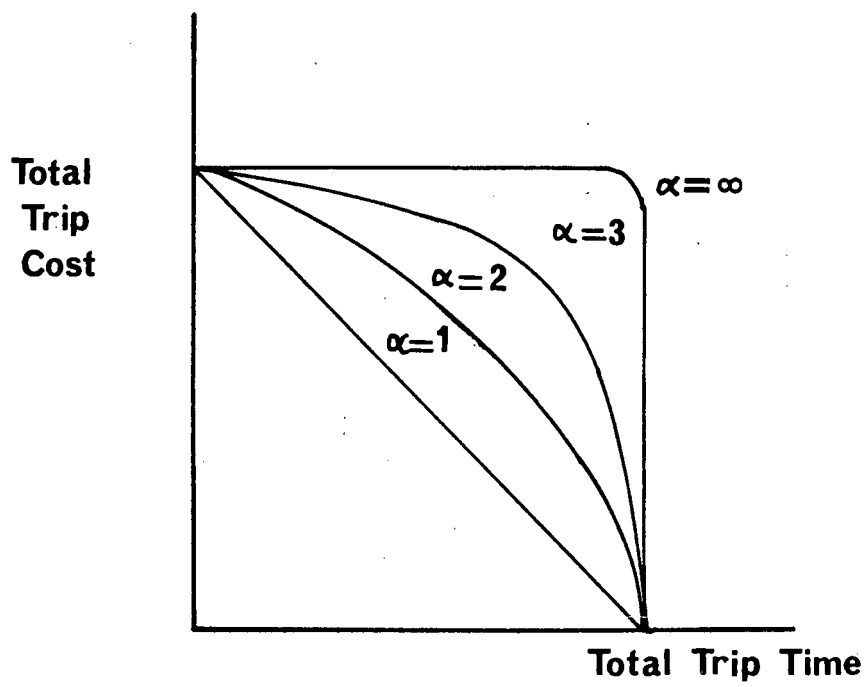


Fig 2-3 Plot of Eqn 2-1 for Various  $\alpha$

and SLOPE is a curve parameter.

SLOPE is a measure of the gap between the difference between the real impedances of A and B and the difference in the perceived impedances of A and B. If the actual and perceived differences in impedance are the same, then SLOPE = 0 and Pr [A over B] = 1.

If the problem is extended to m modes, the m simultaneous equations can be solved for the respective probabilities as demonstrated in equation (2.4):

$$\begin{aligned}
 \text{Pr [A over B]} &= \frac{\text{Pr[A]}}{\text{Pr[A]} + \text{Pr[B]}} \\
 \text{Pr [A over C]} &= \frac{\text{Pr[A]}}{\text{Pr[A]} + \text{Pr[C]}} \\
 &\vdots \\
 &\vdots \\
 &\vdots \\
 \text{Pr [A over M]} &= \frac{\text{Pr[A]}}{\text{Pr[A]} + \text{Pr[M]}} \\
 \\
 \sum_{M=A}^M \text{Pr[M]} &= 1.0
 \end{aligned}
 \tag{2.4}$$

Each probability represents the proportional mode useage. There are significant deficiencies in the model and these will be discussed in the final section of this chapter.

## 2.6 DE DONNEA'S<sup>12</sup> MICRO ECONOMIC MODELS OF MODE CHOICE

De Donnea has developed three models of modal choice behaviour. They involve advanced economic theory and are included in this work as an indication of one approach to economic modelling. The models are not discussed in detail but the assumptions and conclusions are worth considering.

The writer considers the traveller as maximising his utility under two constraints 'a budget constraint and a time constraint'. Furthermore, the consumer is considered as a 'production unit' of consumption activities which require two inputs (goods and services) and (time).

$$A_i = f[X_{ki}, t_i] \quad (2.5)$$

where

$A_i$  = level of the  $i^{\text{th}}$  consumption activity (in our case travel),

$X_{ki}$  = quantity of good or service  $k$  ( $k = 1, n$ ) used in the production of activity  $i$  (e.g. time, money, comfort)

and  $t_i$  = time used in the production of activity  $i$ .

A utility function is then formulated which provides a ranking for subjective preferences of the consumer.

The utility function can be written as:

$$U = U[A_i, L(t_i)] \quad (2.6)$$

where

$L(t_i)$  = the (dis)satisfaction which proceeds from the circumstances under which time is spent to produce a given activity  $i$ .

$A_i$  is from equation (2.5).

The consumer maximises this utility function subject to two constraints - income and time. The income constraint is:

$$Y = \sum_{k=1}^n \sum_{i=1}^m P_k X_{ki} \quad (2.7)$$

where

Y = the income of the consumer per period,  
 $P_k$  = the market price of the goods or service k  
 and  $X_{ki}$  = the quantity of goods or service k in the  
 production of i.

The time constraint is:

$$T_c = \sum_{i=1}^m t_i \quad (2.8)$$

where

$T_c$  = total consumption time per period,  
 $t_i$  = time consumed for the production of consumption  
 activity i.

The utility function (2.6) will be maximized subject to  
 the constraints (2.7) and (2.8) when the following is  
 maximized:

$$V = U\left(A_i, L(t_i)\right) + \lambda \left[ Y - \sum_{k=1}^n \sum_{i=1}^m P_k X_{ki} \right] + \mu \left[ T_c - \sum_{i=1}^m t_i \right] \quad (2.9)$$

where  $\lambda$  and  $\mu$  are Lagrangian multipliers representing the  
 'marginal utility of income' and the 'marginal utility of  
 time' respectively.\*

Taking partial derivatives of (2.9) with respect to  $X_{ki}$   
 and  $t_i$  yields the necessary conditions for maximization, i.e.:

$$\frac{\partial V}{\partial X_{ki}} = \frac{\partial U}{\partial A_i} \cdot \frac{\partial A_i}{\partial X_{ki}} - \lambda p_k = 0 \quad (2.10)$$

---

\* Marginal utility is discussed in Appendix 2.1, but the  
 proof of the above statement is not presented. It can be  
 found in Intriligator Ref. No. 11.

and

$$\frac{\partial V}{\partial t_i} = \frac{\partial U}{\partial A_i} \cdot \frac{\partial A_i}{\partial t_i} + \frac{\partial U}{\partial L} \cdot \frac{\partial L}{\partial t_i} - \mu = 0 \quad (2.11)$$

From equation (2.10) an interesting result is found, namely that

$$\frac{\partial U}{\partial A_i} \cdot \frac{\partial A_i}{\partial X_{ki}} = \frac{\partial U}{\partial A_j} \cdot \frac{\partial A_j}{\partial X_{kj}} = \lambda P_k \quad (2.12)$$

In other words, the marginal utility of a given good must be the same irrespective of the activity. Thus the marginal utility of comfort is the same for a bus as it is for a train or a car. This result is implicitly assumed in other studies<sup>16</sup>.

Another conclusion which De Donnea reaches is of interest but does not warrant detailed derivation. He<sup>12</sup> concludes

"that there does not exist a fixed theoretical relationship between the marginal value of time and the wage rate".

This conclusion has important significance in the arguments about the money value of time used by Beimborn earlier and other writers.

The micro-economic models developed by De Donnea and others provide one with a more logical approach to modelling, but the mathematics is complicated. This aspect will be further discussed in section 2.10.2 of this chapter.

## 2.7 OTHER ECONOMIC BASED MODELS

### 2.7.1 WILSON'S ENTROPY MAXIMISING MODELS<sup>13</sup>

This method is included here for its relevance to the

question "What is the purpose of models?", which is discussed later in section 2.10.2. Wilson adopts a gravity function to estimate the trips between two nodes subject to two trip end constraints and an overall budget constraint. It uses variational calculus techniques to maximize the objective function. This maximization is "a most probable distribution of trips between zones for w states" and is a distribution used in statistical mechanics for molecular behaviour in gases where w is defined in statistical mechanics as entropy. Thus the term entropy maximization is adopted.

Wilson then derives one generalized model for the three sub-models of planning, i.e. distribution, modal split and assignment. He also takes into account perceptions in costs.

Although complicated, the model is a significant step towards eliminating the rather arbitrary segregation of distribution, modal split and assignment into separate decision phases.

#### 2.7.2 KAIN'S<sup>14</sup> ECONOMETRIC MODEL

Kain has developed a nine equation econometric model to explain various decisions concerning modal choice. He considers these choices as stemming from decisions relating to consumption of residential land and to automobile ownership.

It differs from the previous models insofar as it is an empirical model using data from the Detroit Area Traffic Study conducted in 1953.

The two equations which are of direct relevance to this thesis are of identical structure and can be written as:

$$M_j^P \text{ or } M_j^a = \gamma + \gamma_1 A_j + \gamma_2 R_j^S + \gamma_3 L_j + \gamma_4 Y_j + \gamma_5 S_j + \gamma_6 N_j \quad (2.13)$$

where

$M_j^{P*}$  = the percentage of white workers employed in zone  $j$  who use public transit for the work trip,

$M_j^{a*}$  = the percentage of white workers employed in zone  $j$  who travel to work in their cars,

$A_j^*$  = mean car ownership of white workers in zone  $j$ ,

$R_j^{S*}$  = the percentage of white workers employed in zone  $j$  residing in single family units,

$L_j$  = the level of transit service at the  $j^{\text{th}}$  workplace. Coach miles of transit service in each workplace zone divided by the number of acres in workplace zone  $j$ ,

$Y_j$  = the mean family income of white workers employed in zone  $j$ ,

$S_j$  = the percentage white workers employed in zone  $j$  who are male,

$N_j$  = the percentage of  $j$ 's workers who are single wage earners,

and  $\gamma_i$  =  $i^{\text{th}}$  partial regression coefficient.

(N.B.: Obviously the  $\gamma_i$  values will differ for  $M_j^P$  and  $M_j^a$ .)

The variables indicated with an asterisk are endogenous variables and the remainder are exogenous. The endogenous variables are predicted by earlier equations in terms of the exogenous variables. Kain has built up the model on the assumption that automobile ownership decisions only follow

after previous decisions about the level of residential space consumption, by the consumer's income and by his preference for automobile ownership.

Kain's model is an attempt to explain modal choice in terms of the relationship between urban transportation and land-use parameters and in 1961 was a bold step in transport modelling.

The four types of economic model described here are a selection from the various approaches to modal choice modelling. At this stage their importance is in their interpretation of the consumer and his decision making process. Summarising, the models examined were an impedance minimization model (Beimborn), a utility maximization model with equality constraints (De Donnea), an entropy type model with equality constraints of trip ends and expenditure (Wilson) and an empirical regression equation model linking residential space and commuting decisions.

## 2.8 BEHAVIOURAL OR PSYCHOLOGICAL MODELS

A person's decision behaviour is a result of his perceptions and attitudes to various alternatives. Behavioural models attempt to explain the modal choice decision in terms of these perceptions and attitudes.

Three such models will be briefly discussed here, but firstly an investigation of perception of times and costs needs to be made.

### 2.8.1 PERCEIVED COSTS AND TIMES

An examination of times and costs as components of the generalized price vector is all very well, but it is actually

the consumer's perception of these components which is relevant.

Recently several writers have examined this aspect of perception. Hotchkiss and Hensher<sup>15</sup> use the concept of 'perception space' which inherently involves the notion of an 'attitude'.

This theory proposes that one does not perceive duration independently of what endures. Time perception becomes a reality when associated with disutility. The car driver who passes bus queues at every stop has a perception of waiting time far in excess of the actual waiting time incurred. Thus the motorist overestimates alternative mode times. This is substantiated in the author's survey of commuters in Cape Town (see Chapter 4).

Hensher<sup>16</sup> uses the term perceived time units (PTU's) as distinguished from clock time units (CTU's). Similarly, he provides figures demonstrating that perceived car running costs by car drivers are usually less than the actual car running costs supplied by the manufacturer. But the public transport user is far more likely to know his fare. The question which concerns modal choice behaviour is the perceived difference in costs and times between modes. In other words, how well does the car user's estimate or perception of alternative mode times agree with the actual times? (The reader may notice that this question was the reason Beimborn felt compelled to formulate his choice function equation (2.3).)

Steele<sup>17</sup> and Sterne<sup>18</sup>, in an analysis of the author's data for Cape Town, found that both car users and train users underestimated their door to door travel times.

Three behavioural models which attempt to account for travellers' perceptions will now be investigated.

### 2.8.2 EWING'S<sup>19</sup> 'PSYCHOPHYSICAL' MODEL

Ewing, in an attempt to cater for perceptions in modal choice behaviour, uses Steven's<sup>32</sup> power law to relate stimulus and sensation. This law is written as

$$\Psi = c\phi^b \quad (2.14)$$

$\Psi$  = magnitude of the sensation,

$\phi$  = magnitude of the stimulus,

$c$  = an empirical constant,

$b$  = fundamental empirical constant.

The variation in  $b$  seems to reflect genuine perceptual differences, whereas the variation in  $c$  is purely a result of the units of measurement employed.

One further concept, that of adaptation level, is postulated. Adaptation is related to past histories of stimulation. A threshold level of stimulus  $\phi$  is defined as  $\phi_0$ , and equation (2.14) becomes:

$$\Psi = c(\phi - \phi_0)^b \quad (2.15)$$

The use of the power law can then be used to interpret the effect of an increase in a particular travel attribute on the usage of that mode. This Ewing terms "psychological impact", and compares it to perceived disutility; it is more analogous to the economist's concept of elasticity of

demand.\*

He then defines a tolerance ( $Y_{tk}$ ) limit for the disutility of travel, i.e. the point at which travel becomes unacceptable. This he averages for the community for each particular component of the price vector (e.g. a mean waiting time which the "average" commuter would find unacceptable).

Ewing then postulates a perceived disutility function  $Z_{nj}$  for mode  $j$  of:

$$Z_{nj} = \sum_k (Y_{jk}/Y_{tk})^{b_k} \quad (2.16)$$

where

$Y_{jk}$  = the value of component  $k$  of the price vector for mode  $k$

and  $b_k$  = the power law constant for the "average traveller" for attribute  $k$ .

Thus perceived disutilities for each mode can be determined and mode usage assigned on the basis of minimizing the perceived disutility function. This disutility function has been aggregated for a zone, and Ewing has adopted values for  $b_k$  and  $Y_{tk}$  for analysis of a Dial-A-Bus service in Columbia, Maryland.

Ewing notes that equation (2.16) negates having to distinguish between waiting time, travel time and congestion times as having different disutilities, one need only adjust the tolerance limit for travel time.

---

\* Price elasticity in economics is, broadly defined, the change in the demand or supply of a good with a unit change in the market price of that good.

See Chapter 6 for a more detailed coverage of elasticity.

### 2.8.3 DEMETSKY AND HOEL'S<sup>20</sup> USER PERCEPTION MODEL

Demetsky and Hoel postulate a binary choice model in terms of a logistic function as follows:

$$P(T_k) = \frac{e^{I_k}}{1 + e^{I_k}} \quad (2.17)$$

where

$P(T_k)$  represents the probability of a trip  $T$  occurring on mode  $k$ .

$$I_k = X'b^k = b_0^k + b_1^k x_1 + \dots + b_m^k x_m$$

where

$b_m^k$  =  $m^{\text{th}}$  partial regression coefficient,

$x_m$  = a subset of  $X'$ . (The explanatory variables.)

The writers then note that in determining the explanatory variable of time "we do not know how tripmakers perceive the relative times of competing modes and make value judgements". To cater for this deficiency, they adopt a function which they call the modal preference rating (MPR). The MPR is "an index which specifies the tripmaker's perception of the service rendered by any particular mode on a zero to one continuum".

To obtain these MPR values, a psychological scaling method\* is proposed, whereby a measure of attitudes can be expressed in terms of the MPR function. However, the writers use information obtained from an attitudinal survey conducted by NCHRP<sup>21</sup> in 1968, concerning modal attributes for use in their formulation of MPR values for various modes.

---

\* See the next section for a description of this technique.

These MPR values are further subdivided into component values for various modal attributes. Table 2.1 gives the writers' derived component values of MPR for mass-transit work trips.

TABLE 2.1<sup>21</sup>

Subdivision of MPR for Mass-Transit  
Work Trips into Component Values

Component	Value
Time	1,32
Cost	0,60
Convenience	0,60
Comfort	0,93
Others	0,95
Total	4,40

These values can now be used as weighting factors in the model. In other words, a preference rating can be directly included into the constituents of the generalized price vector.

2.8.4 GOLOB'S<sup>22</sup> MODAL ATTRIBUTE MEASURES

Considering modal choice to be a function of the perceived disutility of each of the modal attributes, Golob derives a criterion for binary choice. For mode 1 to be chosen instead of mode 2, then:

$$\sum_{i=1}^N I_i (A_{i,k}^1 - A_{ik}^2) < 0 \quad (2.18)$$

where

$A_{ik}^j$  = measure of disutility for attribute  $i$ , as perceived by user  $k$ , for mode  $j$ ,

$I_i$  = importance users place on attribute  $i$ .

The problem then becomes to assign values to  $A_{i,k}$  and  $I_i$ . To measure the disutility values, which Golob assumes as a measure of dissatisfaction, a survey was designed to measure the differences between attribute satisfactions for different modes. The technique used was "a semantic differential scaling technique with explicit anchor phrases". This means that an attribute is ranked (scaling) from, say, poor to excellent (anchor phrases) in seven steps (see figure 2.4).

#### SEMANTIC SCALING FOR ATTRIBUTE SATISFACTION

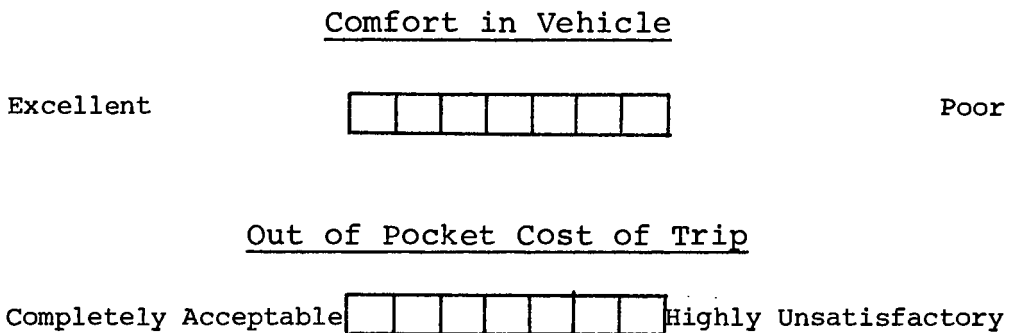


figure 2.4<sup>23</sup>

Thus for two modes, a value of  $A_{ik}^1 - A_{ik}^2$  can be derived and used in the utility function.

From these three behavioural models one can see the approaches to the decision process as a perception problem. However, it should be noted that the underlying assumptions in these behavioural models are still based on the economic concepts of utility maximization.

## 2.9 EMPIRICAL MODAL SPLIT MODELS

Several researchers have taken a more pragmatic approach to modelling modal split. Instead of postulating a hypothesis for the decision process they propose certain influencing factors and relate these factors to observed modal split figures empirically.

The usual method of establishing the relationship is by the technique of multiple least squares regression. The influencing factors are termed the independent variables and are assumed to be linearly related to the dependent variable, in most cases the percentage use of the mode. This technique is used in this thesis and is discussed in detail in later chapters.

### 2.9.1 WILSON'S<sup>24</sup> LINEAR REGRESSION MODELS FOR MODAL CHOICE

Wilson postulated a series of probable factors influencing modal choice and then collected information concerning many of these factors. The surveys involved are not detailed here but are discussed in Chapter 3.

Wilson then performed a preliminary analysis of this survey data to determine which variables were the more important and thus which variables should enter the regression equation.

Several models for Coventry based on this approach were developed; one, for two industrial CBD firms, has the form:

$$\begin{aligned}
 Y = & -0,0163 + 1,477 \log x_1 - 3,347 \log x_2 - 0,7641 x_3 \\
 & -0,1813 x_4 - 0,0135 x_5 - 0,0152 x_6
 \end{aligned}
 \tag{2.19}$$

where

$Y$  = percentage of all trips by public transport,

$x_1$  =  $\frac{\text{time of PT trip}}{\text{time of car trip}}$

$x_2$  =  $\frac{\text{cost of trip for PT}}{\text{cost of trip for car}}$

$x_3$  = car ownership level in zone

$x_4$  = family size index for zone

$x_5$  = economic class index for zone

and  $x_6$  = trip length

This type of model is relatively simple to use and is readily understood by practising planners. But it does lack a rational basis of approach which the models cited earlier exhibit.

Having examined the three types of models it is appropriate to investigate their relative merits and deficiencies.

## 2.10 A CRITIQUE OF MODAL SPLIT MODELS

Before presenting any critical arguments concerning models, one should ask "What is required of the model?". Firstly, the model needs to be able to simulate present travel behaviour and also should be a tool for forecasting travel behaviour.

Immediately problems arise, there are inherent assumptions in most models concerning the static nature of demand, i.e. the explanatory variables will have the same causal effects in the future. Another point is that many of the causal factors included may not be relevant to future conditions. An example of such a variable is the level of service variable Kain includes in his model. To consider

such a level of service index in the context of demand-response systems\* is not viable. Thus the assumption of constant transport technology is not always justified. Several researchers<sup>25,26</sup> have proposed the notion of abstract modes to overcome this problem. This method involves simply noting the respective attributes of a mode and then describing them by these attributes rather than by their actual names. But the method is merely a device and does not really cater for new technology.

The other criticisms of the models are less general. The range of modelling methodologies presented in this chapter afford one the opportunity of examining many of these criticisms.

#### 2.10.1 THE ECONOMIC BASED MODELS

Beimborn's model is suspect in several respects. Firstly the assumption that impedance is independent of the type of trip and the origin and destination of the trip has been shown by Wilson<sup>27</sup> to be erroneous for Coventry. Secondly the coefficient  $\alpha$  is difficult to determine (if not impossible for a three dimensional case) and the term SLOPE is presumably "estimated" as the writer provides no explanation of how it can be determined. This seems to defeat the purpose of the choice function (equation (2.3)), which is to assign travellers to a mode by accounting for their perceptions of costs and times. But SLOPE is a measure of "the sensitivity of a traveller's perception in

---

\* Demand response systems are flexible route mini-vehicles available on demand (usually by telephoning a central computer operated system). They collect and deliver travellers at their desired destination.

impedance" and yet is still only estimated!

De Donnea's micro-economic models contain three assumptions that man behaves "rationally" (in the economic sense) in order to minimize his disutility, that disutilities for various activities are additive and that the disutility associated with an attribute is a linear function of the attribute's magnitude.

It is this last assumption which is of some doubt. Researchers recognize that waiting time has a higher disutility than travelling time. But does the first minute of waiting time have the same disutility as the thirtieth minute of waiting time? Here Ewing's threshold limit may provide a solution to this problem. In other words disutility is proportional to the magnitude of the activity until the threshold level is reached, after which the activity becomes unacceptable.

Wilson's entropy-gravity models have been described by Beckmann and Golob<sup>28</sup> as "metaphysical methods". Their main criticisms are that the causal relationship between consumer choice decisions and entropy maximization modelling are obscure and that the derivation is obtained from a "noncritical adaptation from Kinetic Gas Theory".

Another critical feature of this method is that the mathematics involved seems unnecessarily complex for the marginal returns when compared to less complicated models. It also has the disadvantage along with De Donnea's model of being difficult for practising planners to understand and therefore implement. The notion of "the simpler the better" is quite important in this regard.

Kain's econometric model is subject to criticism on two

main issues. Firstly his data bases for calculation of the regression coefficients are suspect. He developed this model in 1961 and used OD data from the 1953 Detroit Area Traffic Study, income from 1949 U.S. census figures and level of transit service data from a survey conducted in 1958. Kain offers no explanation as to how he reconciles these disparate data bases.

Another critical feature, his hypothesis that economic rent is proportional to the distance from major work places; this is not well founded. Locational behaviour is very dependent upon the city concerned, and to use this somewhat tenuous hypothesis as the basis of his nine models is doubtful. Even assuming the hypothesis is justifiable, he uses a proxy variable for location rent  $P_j$  (which equals 11,5 minus the zone  $j$ 's distance to the CBD) which is very difficult to estimate. Finally his regression equations have very small values of  $R^2$  (0,26; 0,44) and standard errors of the estimates of the income variable's coefficient (equations (2.1a), (2.1b), (2.1c))<sup>29</sup> more than three times the value of the coefficients themselves.

#### 2.10.2 BEHAVIOURAL MODELS

The almost general recognition that perception is an important factor in behavioural choice emphasizes the need for such behavioural models, but the measurement of perception and attitudes is prone to subjectiveness.

Ewing's adoption of Steven's psychophysical law, in order to relate the effects of certain stimuli to their sensory effects, raises considerable problems of measurement.

Although accepting that the law is in fact valid, the

author is doubtful about the measurement of sensation. Ewing<sup>31</sup> quotes Stevens' 'b' value for duration of time. On examining Stevens'<sup>32</sup> method of measuring this value one notes that 16 subjects were asked to estimate various durations of time from 0 to 4 seconds. Although the author is not conversant with such studies in psychophysics, the above justification for extrapolations to waiting times in the order of minutes for different people seems somewhat of large step.

Furthermore, the concept of money as a stimulus involves the notion of secondary sensations. In other words, money can only produce sensations insofar as it provides opportunities for other (secondary) sensations. The author could not determine how Stevens measured the b value for money.

The papers of Golob, Demetsky et al. again contain measurement difficulties. Semantic scaling techniques produce problems of interpretation. There is evidence that the statistics so derived are non-parametric (i.e. not belonging to a definite distribution). Also there is conjecture about the validity of scaling techniques; in other words, can or (more importantly) does a respondent give realistic subjective responses on a seven point scale and what justification is there for assuming the scale is linear?

### 2.10.3 EMPIRICAL MODELS

Empirical models lack theoretical justification. They rely on mathematical tools, of regression particularly, to determine causative relationships. Many researchers have come under criticism for the indiscriminate use of regression

techniques. Millar<sup>33</sup> perhaps summarizes the situation in discussing a paper presented by Bell involving linear regression modelling techniques. Millar noted:

"I have no objection to multiple regression analyses provided that people fit sensible models.

Unfortunately ..... too often the only justification (for linear regression models) is that there happens to be a convenient computer programme ..... for fitting linear regressions."

More detailed discussion of Wilson's model will be given in Chapter 5.

#### 2.10.4 SUMMARY

In this chapter three categories of models (Economic, Behavioural and Empirical) have been examined, which provide a spectrum of approaches and methodologies. The author is of the opinion that the theoretical approach to modelling is still in the formative stages. Probably the "best" of such models are the utility maximization type as consumer choice behaviour modelling has undergone rigorous testing in the field of economics.

Eventually a merging of utility theory and psychophysics may be the answer. The attempts of Golob and Demetsky and Hoel are steps in this direction.

Meanwhile one relies on the empiricist to furnish answers to present problems and it is conceivable that a better understanding of modal choice behaviour will stem from such studies.

REFERENCES : CHAPTER TWO

1. Wheeler, J.O., "Research on the Journey to Work", Council of Planning Librarians, Exchange Bibliography No. 65, January 1969, p. 2.
2. Wilson, F.R., "Journey to Work - Modal Split", Maclaren and Sons, London, 1967, pp. 19-22.
3. Golob, T.F., Canty, E.T., Gustafson, R.L. and Vitt, J.E., "An Analysis of Consumer Preferences for a Public Transportation System", Transportation Research, Vol. 6, 1972, pp. 85-87.
4. Truett, J.B. and Balek, A.J., "On the Need for a Definition of Demand for Transportation", HSGT, Vol. 6.
5. Beesley, M.E. and Roth, G.J., "Restraint of Traffic in Congested Areas", The Town Planning Review, Vol. 33, 1962-63, p. 187.
6. Dupuit, J., "On the Measurement of the Utility of Public Works", Annales des Ponts et Chaussées, 1884, 2nd series, Vol. 8. Reprinted in English in "Transport", Penguin Books, 1968, Ed. D. Munby, p. 28.
7. Kemp, M.A., "Some Evidence of Transit Demand Elasticities", Transportation, Vol. 2, 1973, p. 26.
8. Beimborn, E.A., "Impedance Minimization Travel Choice Process", Transport Engineering Journal, Proc. A.S.C.E., May 1972, pp. 177-190.
9. McGillivray, R.G., "Mode Split and the Value of Travel Time", Transportation Research, Vol. 6, December 1972, pp. 304-318.
10. Rassam, P.R. and Ellis, R.H., "Analytical Estimation of Transit Access Impedances", Transportation Engineering Journal, Proc. A.S.C.E., May 1972, pp. 191-
11. Intriligator, M.D., "Mathematical Optimization and Economic Theory", Prentice Hall, N.J., 1971, pp. 28-39.
12. De Donnea, F.X., "Consumer Behaviour, Transport Mode Choice and Value of Time, Some Micro Economic Models", Regional and Urban Economics Vol. 1, No. 4, 1972, pp. 355-382.
13. Wilson, A.G., "The Use of Entropy Maximising Models in the Theory of Trip Distribution, Mode Split and Route Split", Journal of Transport Economics and Policy, 3, 1, 1969, pp. 108-126.

14. Kain, J.F., "A Contribution to the Urban Transportation Debate an Econometric Model of Urban Residential and Travel Behaviour". The Review of Economics and Statistics, London, 1964, pp. 55-64.
15. Hotchkiss, W.E. and Hensher, D.A., "The Value of Time Spent in Travelling : An Australian Case Study", unpublished.
16. Hensher, D.A. quoted by Bowen, W., "Modal Split", B.Comm.(Hons.) Thesis, University of Melbourne, 1971, unpublished, p. 5.
17. Steele, R.H., "Isochrone Map and Level of Service Index", Final Year Thesis, Department of Civil Engineering, University of Cape Town, 1973, unpublished, p. 72.
18. Sterne, K.B., "Isochronal Map and Level of Service Indices", Final Year Thesis, Department of Civil Engineering, University of Cape Town, 1973, unpublished, p. 45.
19. Ewing, R.H., "Psychological Theory Applied to Mode Choice Prediction", Transportation, Vol. 2, No. 4, December 1973, pp. 391-409.
20. Demetsky, M.J. and Hoel, L.A., "Modal Demand : A User Perception Model", Transportation Research, Vol. 6, No. 4, December 1972, pp. 293-316.
21. Ibid, p. 301.
22. Golob, T.F., "The Survey of User Choice of Alternate Transportation Modes"
23. Ibid, Appendix A, p. 180.
24. Wilson, F.R., Op Cit.
25. Beimborn, E.A., Op Cit, p. 178.
26. McGillivray, R.G., Op Cit, p. 309.
27. Wilson, F.R., Op Cit, p. 163.
28. Beckmann, M.J. and Golob, T.F., "A Critique of Entropy and Gravity in Travel Forecasting", Traffic Flow and Transportation, ed. G.F. Newell, Elsevier, New York, 1972, p. 109.
29. Kain, J.F., Op Cit, p. 60.
30. Heggie, I.G., "Transport Engineering Economics", McGraw-Hill, London, 1972, p. 10.
31. Ewing, Op Cit, p. 404.

32. Stevens, S.S., "On Psychophysical Law", Psychological Review, No. 64, 1958, p. 155.
33. Millar, A.J. discussing a paper by Bell, G., "The Distributions of Work Trips in Urban Areas : Some Considerations" - Paper No. 605, Proc. 5th Conference A.R.R.B. Canberra, 1970, p. 150.

## CHAPTER THREE

### SURVEYS : DESIGNS AND METHODS

This chapter deals with the design and methodologies of surveys, conducted by the author, concerning travel behaviour for white commuters in Cape Town.

The main study was a questionnaire survey to obtain information concerning the travellers' socio-economic characteristics and the details of their trips to work. The design and execution of this survey is discussed. Further surveys on trip characteristics such as travel times and parking times are also discussed.

Special attention is given to problems encountered in these methods and possible solutions are suggested to avoid such problems in the future.

#### 3.1 REASONS FOR THE SURVEYS

The author examined several methods of modelling modal choice behaviour. Some of these methods have been discussed in Chapter 2.

After ascertaining that no attempt had been made to investigate this type of travel behaviour in Cape Town and that suitable data did not exist for developing such a model, a decision was made to formulate an empirical model of modal split for Cape Town.

An empirical model was chosen for several reasons. Such a model

(i) is relatively easy to formulate,

- (ii) is easily understood by practising engineers,
- (iii) provides an information base for further study,
- (iv) is within the resources of a single researcher, and
- (v) only requires data which is relatively easy to obtain.

Information concerning the likely causal factors in mode choice were required from a cross section of commuters. Detailed train and bus loadings, times, capacities, frequencies, and car times were also required. This data provided the essential information from which the models could be formulated.

### 3.2 SURVEY OBJECTIVES

The following initial decisions were made regarding the objectives of this survey.

Firstly the survey was limited to white commuters. This was done for several reasons. As a foreigner to South Africa, the author was unfamiliar with the non-white communities, and discussions with transport planners indicated that special survey techniques are required to ensure satisfactory response rates in such surveys.\* Another reason for excluding non-white commuters was that the problem was more one of latent demand for travel than a modal choice decision. It should be acknowledged that such a decision to exclude a large sector of the community detracts from the viability of the model as a planning tool. As mentioned in Chapter 1, non-white commuting is a field of research which requires urgent study in this country.

---

\* A personal discussion with Mr. R. Jackson, Chief Engineer (Road Traffic and Planning), City Engineer's Department, Johannesburg, confirmed that special problems were encountered by interviewers in non-white areas.

The second decision made was to consider modal choice in the context of the work trip only. The transportation systems in Cape Town are generally only used to full capacity in the peak periods, i.e. for the journey to and from work. Consequently design capacities of networks are usually related to peak flows and therefore initial studies should concentrate on such journeys.

The study area was initially broadly defined as the greater Cape Town area. But subsequent events required that the study area for the model development be the southern suburbs only. The reasons for this drastic reduction in the survey area will be explained in later sections of this chapter.

With these decisions made, the author conducted a survey of existing models and techniques. The methodology of Wilson<sup>1</sup> greatly influenced this work and considerable reference is made to his thesis.

As mentioned earlier, all likely causative factors need to be investigated for possible inclusion in the model. Of Wilson's 72 factors, many could be excluded on the basis of the modelling experiences of earlier researchers. Information concerning the remaining variables had then to be obtained. The following sections describe the methods by which this was accomplished.

### 3.3 SOCIO-ECONOMIC AND PERSONAL TRIP INFORMATION

It was apparent that, in order to obtain socio-economic information and data on individuals' tripmaking habits, some method of direct measurement was required. In other words the commuters would need to be asked questions.

There are two basic approaches to such a problem. One can either conduct an interview study or a questionnaire study.

Both methods have their relative merits and deficiencies. One vital requirement of any such study is that it results in a good response rate. Sellitz et al.<sup>2</sup> comment on the two types of survey. The interview survey is where a set of questions are asked of a person and the replies noted by an interviewer. The interview technique has the advantages of flexibility and that answers can be "followed up" by the interviewer. The response rate is usually high with such a survey. The interview has a further advantage of being able to ascertain something about the non-respondents, e.g. do they all live or work in a particular area? The disadvantages of such interview methods are:

1. the interviewer can affect the responses,
2. the technique involves a good deal of time and/or manpower and consequently is expensive,
3. the conditions under which answers are given are not standard,
4. the anonymity of the respondent is not preserved.

The questionnaire techniques involve issuing (by mailing, distribution by hand or other methods) a set form of questions. The questionnaire method has the following advantages. It is:

1. less expensive in time and labour than the interview survey,
2. simpler to administer,
3. relatively easy to survey a wide area,
4. more standardized in its replies,
5. able to preserve anonymity, and
6. provides the respondent with time for thoughtful replies.

However, it has the disadvantages of being:

1. liable to a low response rate,
2. unable to classify the non-respondents,
3. ambiguous to or misunderstood by some people,
4. inflexible with the consequence that "follow up" questions are not possible, and
5. highly structured which may preclude important conditional replies (e.g. "I used a bus to come to work because my car broke down and the train was delayed" - would not be a response likely from a questionnaire).

In this present instance an interview survey was ruled out on the basis of available resources, so a questionnaire survey was adopted for this study.

Wilson<sup>1</sup> conducted a questionnaire survey in Coventry at the place of work, and obtained reasonable response rates. It was decided therefore to adopt a similar approach for this survey.

#### 3.4 DESIGN OF THE QUESTIONNAIRE FORM

The design of the questionnaire form was done with two main objectives in mind. Firstly to obtain the required data, and secondly to achieve a "satisfactory" response rate. These two aims are often counter-productive.

Goode and Hatt<sup>3</sup> present a detailed discussion on questionnaire design, and the author has used their suggestions in the formulation of the questionnaire.

### 3.4.1 CRITERIA FOR QUESTIONNAIRE DESIGN

#### (a) AUSPICES OF THE QUESTIONNAIRE

In previous studies performed overseas, it has been recognized that the sponsorship of a survey has had an important role to play in securing a good response. It was felt that, for this survey, joint sponsorship by both the University of Stellenbosch and the University of Cape Town would induce wider support than either university acting separately.

The Transport Research Centre of the University of Stellenbosch agreed to co-sponsor this survey with the Department of Civil Engineering at the University of Cape Town, and generously offered to finance the printing of the questionnaires. Further sponsorship was considered, but ruled out as having doubtful advantages.

#### (b) LENGTH OF THE QUESTIONNAIRE

The length of a questionnaire (both real and apparent) affects the response rate, and this proved to be one of the more difficult aspects with which to deal.

To obtain the necessary data and to cater for the two official languages (English and Afrikaans), it was necessary to produce a thirteen page form printed on both sides. A copy of the final form is presented in Appendix 3.1. The effect of receiving such a weighty form would have deterred many people from completing the form. In an attempt to avoid this, the questionnaire was divided into four sections, a general section to be completed by everyone, and three other sections only one of which needed to be completed by each

informant. Also the covering introduction to the form stressed that filling out of the form would take only 5 to 10 minutes. These times were checked beforehand using a sample of people with different backgrounds.

(c) THE ATTRACTIVENESS OF THE FORMAT

The attainment of this criterion was largely limited by the finance available for printing. Even with this constraint the author, on reflection, feels that the layout could have been improved. For instance, Wilson<sup>2</sup> comments on the use of bold lines to separate questions:

"very distinct heavy lines create the effect of short sections rather than a continuous long list of questions."

The basic format was:

1. a covering note on the front cover,
2. instructions for completing the form on page two,
3. a general section - pages 2 to 4,
4. Section A, for non-car owners - pages 5 to 7,
5. Section B, for car owners not using their own vehicles - pages 8 to 10, and
6. Section C, for car users - pages 11 to 13.

The two languages were arranged so that neither language group appeared to have any preference.

(d) THE COVERING NOTE

The covering note was constructed to be brief, simple, and to elicit support. The sponsors of the survey were noted and an altruistic appeal, that such a survey was of importance to the respondents, was made to elicit better response. The front cover presented reasons why the survey was being

conducted and also assured the respondent that the replies were confidential.

(e) QUESTION FORMATS

Goode et al.<sup>4</sup> refer to open ended and structured questions. The majority of questions in this survey were structured, i.e. "tick the appropriate box" questions. Two of the exceptions to this form were the preference questions in Sections A and C.

In these questions space was provided for other preferences not specifically listed. The advantage of highly structured questions is that they provide a standardized reply, but they do not provide for unforeseen alternatives.

If repeating the survey, the author would make some of the questions direct reply rather than "tick the box". A good example of this mistake is the questions concerning "the time taken to come to work".

These questions required marking boxes encompassing ten minute intervals, i.e. 0 to 10 minutes; 11 to 20 minutes, and so forth. By providing this limited choice, one created some uncertainty in the minds of the respondents and also the interval of ten minutes was too large and required interpolation in subsequent analyses.\* A better method would have simply been to ask for the commuter to write his trip time.

---

\* The author "copied" this question from Wilson's survey. This brings two questions to one's mind: Did Wilson recognize the same difficulties in interpolation, and if so, what is the value of publishing a work with an obvious flaw without commenting on such shortcomings?

The questions requiring written answers were: occupation, address, railway station used and the fares of public transport users.

A brief and simple set of instructions were placed at the beginning, and as only about 1,3% of the returned forms were incorrectly answered, these instructions seemed to have been adequate.

All the questions related to the trip to work that day. A basic premise here was that the survey days were "typical weekdays".

(f) THE CHARACTERISTICS OF THE RESPONDENTS

It was recognized that a wide cross-section of employees would constitute the sample population, consequently the questions were phrased in simple terms with a minimum of ambiguity. It was further endeavoured to phrase the questionnaire in the idiomatic characteristics of each language to minimize the possibility of offending the sensibilities of either language group. The Afrikaans translation was made by the staff of the Transport Research Centre at Stellenbosch.

(g) THE SUPPORT OF THE MANAGEMENT

The survey technique is described in detail in section 3.5. Briefly the procedure involved approaching a number of firms and soliciting their cooperation with the survey. This involved an agreement to the distribution of the forms to the firms' employees during working hours.

Full management cooperation is obviously essential with such a survey technique. In the majority of instances a

circular from the management was distributed to all staff members either before the survey day or with the questionnaire form. The circulars indicated the management's support of the survey and requested full cooperation in completing the forms. Samples of such circulars are presented in Appendix 3.2.

(h) INDUCEMENTS TO RESPOND

Surveys<sup>5</sup> conducted in the United States have offered material inducements (e.g. free samples, money) to persuade the informant that he ought to respond. However, the only implied inducements in this survey were that the respondent's answers would be:

- (a) in his own interest,
- (b) in his community's interest, and
- (c) of benefit to transportation planning in his city.

(i) ANONYMITY AND SENSITIVE QUESTIONS

Protection of the respondents' anonymity increases the likelihood of a good response. Also the less sensitive questions asked the better the response rate is likely to be. These two criteria were unavoidably conflicting in this questionnaire, and provided perhaps the weakest part of the questionnaire.

Two questions were considered potentially detrimental to the response rate - the income and the full address questions. These were placed last in the general section on the premise that "it seems to be more difficult for an informant to break off an interview or questionnaire than to refuse to begin at all".<sup>6</sup> The income question was qualified by statements

indicating that the questions were strictly confidential and, further, that "this question is optional". The author is doubtful as to the value of this optional statement. Only 17% of the sample refused to answer the income question and it is likely that if the optional aspect were omitted an even lower refusal rate may have resulted. The missing income figures were inferred from the occupation status (question 7) and in the author's estimation this method was sufficiently accurate considering the wide salary ranges adopted. (A range of approximately R200-00 per month was used.)

One wanted to be able to assign each person to a survey zone (see section 3.9). To do this, question 8 asked for the respondent's full address which could then be used to determine in which zone the person lived. Several informants interpreted this as an infringement of their anonymity and about 10% refused to state their complete address. However, only about 1% of the sample population could not be placed in a zone.

#### 3.4.2 FORMULATION OF THE QUESTIONS

Goode and Hatt<sup>7</sup> write "the inclusion of every item (in a questionnaire) should be defensible on the grounds that the researcher can logically expect the answer to be significant for his central problem". In other words if the respondent is unable to see the relevance of a question to the study he is unlikely to be convinced that he should answer. Furthermore, the collection of irrelevant data is fruitless and indeed wasteful.

For this reason the author investigated several models developed elsewhere and examined the causal variables used

VARIABLE	TIME	COST	COMFORT	TRIP DISTANCE	TRIP PURPOSE	SEX/AGE	CAR AVAILABILITY	INCOME	RESIDENTIAL DENSITY	LOS <sup>†</sup>	OTHER
Researcher											
Adams <sup>9</sup>	a			*			*		*	*	*
Bowen <sup>10</sup>	*	*	*		n/a	*		*		*	*
Kain <sup>11</sup>						Sex	*	*	*	*	*
Lave	*	*	*	a	n/a	*	n/a	a			
Lisco	*	*			n/a	*	n/a	*			
Quarmby	*	*			n/a						*
Stopher	*	*			n/a						
Thomas	*	*	*	*	n/a	Sex	n/a	*			*
Tomazinis <sup>12</sup>					+		*		*	*	*
Warner	*	*		*	+	*	n/a	*			
Wilson <sup>13</sup>	*	*		*	+	Sex	*	*			*

a = Indirectly Used  
\* = Used as a variable  
† = Level of Service  
+ = Used to stratify sample

TABLE 3.1 INDEPENDENT VARIABLES USED IN PREVIOUS STUDIES<sup>8</sup>

and then derived the types of questions which needed to be asked in order to construct these variables, which are listed in Table 3.1.

The reasons for the inclusion of most questions are self evident. However, several questions require some explanation.

Question 4 asked the number of children under 16 years of age. This was an indication of the number of dependants in a family, as sixteen years is the minimum school-leaving age in the Cape.

Questions 10 and 11 in Section A were to analyze how travellers perceived times and costs for alternative modes. These questions were often left unanswered.

Question 4 in Section B was phrased to find out the reasons why people did not use their cars to come to work.

In Section C, questions 7, 8 and 9 were aimed at determining the car users' perceptions of alternate travel times and transfers.

The questionnaire was designed to find the basic information as shown in Table 3.2.

#### 3.4.3 PROVISION FOR ENCODING DATA

It was proposed that the data obtained from the questionnaire survey would be analyzed by the Univac 1100 computer at the University of Cape Town. Consequently, provision had to be made to allow easy transfer of data from the forms onto coding sheets and then onto punched cards. Originally it was envisaged punching the data directly onto cards from the forms. However, after consultation with staff at the computer centre, it was realized that this would be far

TABLE 3.2BASIC INFORMATION OBTAINED FROM QUESTIONNAIRE SURVEY

<u>INFORMATION</u>	<u>QUESTION No.</u>
<u>General Section:</u>	
Sex; Household Status	1
Car Ownership	2
Drivers' Licences/Household	3
Children under 16 (Dependants)	4
Mode of Travel (on Survey Day)	5 & 6
Occupation	7
Address	8
Income	9
<u>Sections A, B &amp; C:</u>	
Usual Mode of Travel	A1; B3; C3
Estimated Total Travel Time	A2; B56; C6
Transfers	A3,5; B5a; C9
Fares	A4,66; B6,7
Train Station (Embarkation)	A6a
Inconvenience Factors (P.T.)	A9; C10
Inconvenience Factors (Car Travel)	B4
Alternative Mode Times	A10a; B5f; C7b
Shelter Provision (P.T.)	B5e
Parking Provision	B1; C1
Parking Cost	C2
Effect of Weather on Travel Mode	B2; C5

too time consuming for the punch operators. Consequently a form was designed such that each "tick the box" question was given a number as indicated by the digit under the marked box. This number was transferred to the corresponding box in the right hand column (see Appendix 3.1). The questions requiring direct responses required separate coding methods and these are discussed later in this chapter.

All the data from the right hand column was then transferred to coding sheets for punching.

### 3.5 THE QUESTIONNAIRE SURVEY METHOD

The proposed survey technique involved distributing the questionnaire forms to employees at their places of work on preselected survey days. The selection of these work places is discussed in section 3.6.

The firms were approached and asked if they would cooperate in such a survey. In every instance the firm approached agreed to cooperate. There were two methods of distribution proposed. They were, either to have members of the survey team hand out and collect the forms, or to have the firms distribute and collect the forms themselves.

It was further intended to distribute and collect the forms on one day in order to ensure that the answers all concerned travel on the same day.

Before any such survey is conducted, a pre-test of the questionnaire should be performed.

#### 3.5.1 THE PRE-TEST SURVEY

A pre-test survey is used to check the design of the questionnaire and to test the survey method. This survey

should be conducted with all the rigour of the final test. For this purpose a selection of fifty employees of Norwich Union Life Assurance Company offices in Cape Town were sampled. The management sent a circular to all the staff members as discussed in section 3.4.1. No indication was given to the employees that this was a test run. The forms were distributed and collected by the author on one day. The day was Thursday, 19th July 1973.

The pre-test method provided the author with some valuable information for the final survey. The following points were noted:

1. The management's circular used the statement "... any information divulged will not be disclosed to the management". The Administration Superintendent<sup>14</sup> had received some feedback from the staff who were concerned about the term "information divulged". This term apparently had sinister connotations for some people and the author attempted to "screen" the firm's pre-survey memoranda in the final test.
2. The author attempted to find a brief and satisfactory 'spiel' to be used while distributing the forms. A method which seemed adequate and was used by the survey team in the full survey included the following statements:
  - (i) "Good morning";
  - (ii) "Could you please give this form your attention, and any assistance you could give would be appreciated";
  - (iii) "The questionnaire is anonymous";
  - (iv) "It should only take about 5 - 10 minutes to

complete";

(v) "I shall return this afternoon to collect the forms";

(vi) "Thank you".

There was no need to explain who one was or what one was doing as the staff had already been informed of the survey and generally were expecting the forms.

3. One should avoid tea-breaks for distribution and collection as it requires re-visits.
4. It was possible to note the numbers and sexes of those who refused to accept a questionnaire. This approach was adopted during the final survey.
5. No attempts at persuasion were made with those who were reluctant to accept forms.
6. It was decided to select bilingual members of the final survey team where possible.

The distribution and collection seemed to run smoothly and the author concluded that about 150 forms could be adequately distributed by one person in a day.

### 3.5.2 THE PRE-TEST SURVEY RESPONSE

There were 51 people sampled and 2 blank forms were returned, there was one refusal to accept a form, and two unusable returns. This gave an overall response rate of 94% and a usable response rate of 90%.

Such response rates were considered as a vindication of the questionnaire. There were several alterations needed, and these were incorporated into the final draft which was then ready for printing.

### 3.6 SELECTION OF SAMPLE

The author could not find detailed employee figures for sample selection purposes. Consequently an attempt was made to choose a variety of firms covering various occupations and locations (CBD and non-CBD). It was proposed to detect major sampling biases by comparing the zonal populations, from figures which the City Engineer's Department in Cape Town had, with those obtained from the survey. The firms selected and their locations and type of industry are noted in Table 3.3.

The locations of the CBD firms are shown in figure 3.1.

### 3.7 FINAL SURVEY METHOD

The management of all the firms agreed to cooperate in principle with the survey. However, some firms wished to distribute and collect the forms themselves whilst others insisted that their employees be given a full day to complete the forms, i.e. the forms should be collected the following day. Table 3.4 lists the various distribution methods for each firm.

The questionnaire forms were delivered to the firms prior to the survey day and circulars (where these were issued) were usually issued to the staff the day before the survey.\* The survey was conducted over several days mainly during the week commencing Monday, 12th August 1973.

A total of 2140 questionnaires were distributed.

---

\* Several firms attached the circular to the questionnaire forms.

TABLE 3.3EMPLOYMENT CENTRES SELECTED FOR QUESTIONNAIRE SURVEY

<u>Central Business District Firms</u>		<u>Sample Size</u>
1.	Caltex Oil (S.A.) H.O.                      Accounts Department	91
2.	City Engineer's Department, Cape Town City Council.                      All Departments	470
3.	C.L.F. Borckenhagen & Louw.                      Quantity Surveyors	20
4.	Davidson & Ewing (Pty) Ltd.                      Medical Aid Society	30
5.	ESCOM.                      Various Departments	140
6.	Freight Services (Pty) Ltd., Foreshore.                      Shipping Agents	74
7.	Liebenberg & Stander.                      Consultant Engineers	63
8.	Mobil Oil H.O.                      Administration	200
9.	Nico Malan Theatre.                      Theatre Personnel	57
10.	Ninham Shand & Partners.                      Consulting Engineers	149
11.	Norwich Union (General).                      General Insurance	20
12.	Norwich Union (Life).                      Life Assurance	51
13.	Safmarine Ltd.                      Shipping Agents	26
	TOTAL:	1391
<u>Non Central Business District Firms</u>		<u>Sample Size</u>
1.	Cape Town City Tramways (Woodstock)                      Bus Company	126
2.	Freight Services (Observatory)                      Workshops	6
3.	Reckitt & Colman (Pty) Ltd. (Pinelands)                      Administration and Marketing	92
4.	Reids (Observatory)                      Car Sales	17
5.	Reids (Salt River)                      Workshops	16
6.	Mobil Oil (Rondebosch)                      Marketing Distribution	47
7.	S.A. Mutual Insurance (Pinelands)                      Various Departments	200
8.	Southern Life Association (Newlands)                      Various Departments	250
	TOTAL:	754
	GRAND TOTAL:	2145



Figure 3.1

LOCATION OF CENTRAL BUSINESS DISTRICT FIRMS

Legend

A	Sanlam Building	Borckenhagen & Low Caltex (S.A.) (Pty) Ltd. Liebenberg & Stander Ninham Shand & Partners
B	Foreshore Building	Davison and Ewing
C	Nico Malan Theatre	
D	ESCOM	
E	Norwich House	Norwich Life Assurance Co. Norwich General Insurance Co.
F		Freight Services (Pty) Ltd.
G	BP Centre	Safmarine
H	Mobil House	Mobil (S.A.) (Pty) Ltd.
J	City Park	Cape Town City Engineer's Department
K	Cape Town Railway Station	
L	Bus Terminus, Darling Street	

The large areas devoted to parking in the CBD are evident from the above figure.

FIRM	No. FORMS DISTRIBUTED	No. USABLE RETURNS	% USABLE RESPONSE	NON-USABLE RETURNS	DISTRIBUTED AND COLLECTED ON ONE DAY	DISTRIBUTED BY FIRM	DISTRIBUTED BY SURVEY TEAM
Caltex	91	91	100,0	0	No	Yes	
C.C. Engineer's Dept.	470	320	68,0	7	No		Yes
Borckenhagen & Louw	20	19	95,0	0	Yes	Yes	
Davidson & Ewing	30	22	73,5	0	No	Yes	
ESCOM	140	109	78,0	3	No	Yes	
Freight Services	74	64	86,5	1	No	Yes	
Liebenberg & Stander	63	55	87,3	1	No	Yes	
Mobil H.O.	200	125	62,5	1	No	Yes	
Nico Malan Theatre	57	50	87,5	4	No	Yes	
Ninham Shand	149	133	89,1	1	No		Yes
Norwich General	20	13	65,0	0	Yes	Yes	
Norwich Life	51	46	90,0	2	Yes		Yes
Safmarine	26	23	88,8	1	No	Yes	

**TABLE 3.4 SURVEY RESPONSE FIGURES FOR 13 CBD FIRMS**

### 3.8 QUESTIONNAIRE RESPONSE

The overall response rate was 81,5% and the usable response rate was 80,4%. Perhaps a more important measure of the response rate is the median response rate. From the figures shown in Tables 3.4 and 3.5 this was about 87,5%.

These overall results, when compared with Wilson's results, are quite satisfactory. Wilson<sup>15</sup> obtained an overall usable response rate of 59,7%. He decided that if the response rate was less than 80% he would conduct a non-response survey to detect if his original sample was biased. Briefly, this method involved distributing a second questionnaire asking the non-respondents to complete questions concerning their reasons for non-response, their sex, car ownership and mode of travel. He then performed statistical checks to see if his original sample was biased in terms of the respondents' sex, car ownership status, modes of travel and the respondents' marital status.\*

With the response rates obtained in this survey the author considered such a non-response survey was unnecessary.

The figures tend to show that there is a decreasing response rate with increasing sample size. This is probably due to many factors including the methods the firms used for distribution.

### 3.9 CODING THE DATA

The coding of the data turned out to be a major undertaking, taking over 200 man-hours to compile. The basic

---

\* A weakness in Wilson's non-response survey was that he was unable to test if there was bias amongst income groups, as this was the question most people gave as their reason for not responding.

FIRM	No. FORMS DISTRIBUTED	No. USABLE RETURNS	% USABLE RESPONSE	NON-USABLE RETURNS	DISTRIBUTED AND COLLECTED ON ONE DAY	DISTRIBUTED BY FIRM	DISTRIBUTED BY SURVEY TEAM
C.T.C. Tramways	126	99	78,5	0	Yes		Yes
Ft. Services (Obs)	6	6	100,0	0	No	Yes	
Reckitt & Colman	92	92	100,0	0	Yes	Yes	
Reids (Obs)	17	17	100,0	0	No	Yes	
Reids (Salt River)	16	12	75,0	0	No	Yes	
Mobil (Rondebosch)	47	46	97,8	1	No	Yes	
S.A. Mutual	200	178	89,0	4	No	Yes	
Southern Life	250	202	81,0	3	No	Yes	

TABLE 3.5 SURVEY RESPONSE FIGURES FOR 8 NON-CBD FIRMS

encoding process is described earlier. However, there are several points which need to be clarified at this juncture.

Firstly, each address had to be allocated to a zone as the analyses which follow are performed at a zonal level. The Cape Town Municipality was divided into the same zones used by the Advance Planning Section of the City Engineer's Department<sup>16</sup>. These zones were chosen for several reasons. Firstly, the City Engineer's Department were collecting data for these zones and it seemed a wasteful exercise to collect data on a different zonal basis. Secondly, the zones were selected as homogeneous (with respect to socio-economic and housing characteristics) "environmental cells" and homogeneity within zones is a basic premise in future analyses. And finally there were population figures available for each of the zones.

Similarly, the zones in the Constantia area were chosen to correspond with the Provincial Administration's<sup>17</sup> planning cells for this area. Other zones outside these areas were chosen by the author. The zonal divisions are shown on the map in Appendix 3.3.

Each zone was given a number and these numbers are marked on the map in Appendix 3.3. Three columns were provided for the zone number for coding purposes.

Occupation was coded on the stratified system used by the U.S. Bureau of the Census<sup>23</sup>. This code involves two card columns; one as an occupation code 0 to 9 and the other as an industry code 0 to 9 with two additional alphanumeric classifications if required.

Question 6A in Part A asked for the name of the station where the train travellers commenced their journey. For

this purpose each station was given a code number and two columns were allowed for this purpose.

A sample of the coding sheets for each of the three categories is given in Appendix 3.4. One error which involved a major change during the analysis stage was in coding the workplace. The author only provided one column for the workplace code (column 1). As there were 21 workplaces this posed some problems. The second column was either a '1', '2' or a '3' indicating that either Part A, B or C had been answered. On inspection, it was recognized that this information could be determined by examining questions 2 and 5 in the general section. Consequently the data was edited, after storing the information on computer discs, so that columns 1 and 2 represented the workplace codes.

The analysis of the questionnaire survey data is discussed in the next chapter.

### 3.10 ADDITIONAL SURVEYS

In addition to the information obtained from the questionnaire survey, there were further surveys to determine:

1. Car Times from zones to workplaces;
2. Train and Bus times;
3. Level of Service Indices for public transport;
4. Walking and waiting times, and
5. Parking facilities available in the CBD.

The author was fortunate in having the above surveys performed by six final year students.\* This work was

---

\* These students were Messrs. Beyers, Heegar, Rowan, Stanger, Steele and Sterne.

performed under the author's direct supervision and the overall supervision of the final year thesis project was undertaken by Professor King of the Civil Engineering Department. These projects were devised so that the data obtained could be used for this thesis. As each of the surveys themselves constituted a final year thesis, a brief description only can be given of the survey techniques.

### 3.10.1 CAR TIME SURVEY

Rowan and Steele<sup>18</sup> produced an isochrone map for private vehicles travelling to the CBD from the southern suburbs. To determine the car times on the major arterial roads the group of six conducted an origin and destination time study for a period of nine days. This method involved reading the last three digits of the vehicle registration number into a tape recorder and recording the time at two minute intervals. The data was then punched onto cards and verified.

The author wrote a computer programme to analyze this data. The programme can be found in Appendix 3.5. In brief, the programme selected a vehicle number at an origin station and searched the destination arrays for the same number. When this pair were matched their corresponding times were subtracted and thus the time for the vehicle to move from the origin to the destination was thus determined. However, there were a large number of spurious matchings due to a lack of total definition of the number plate. These false matchings could be reduced by excluding "impossible" trip times, but there still remained a large number of matchings which were duplicated. This was due to major gaps in the origin cordons. Neither time nor manpower allowed

the team to capture all the possible influent vehicles. To overcome this problem, frequency charts of the vehicle times were drawn and the modal value proved to be a good indication of the trip time. This can be explained by the fact that the roads at which the origin cordons were placed carried high numbers of vehicles and the probability of the modal time value occurring at a time above or below the expected trip time was quite small. Appendix 3.6 gives some examples of these frequency charts. Appendix 3.7 lists the origin and destination sites for this survey.

A floating vehicle study was conducted to determine vehicle times from different zones to the origin stations used in the OD study. The method involved driving a car with an observer who measured the times and distances. The driver attempted to maintain an average lane speed, that is, he passes as many vehicles as pass him. These floating vehicle times were also used to check the times obtained from the origin and destination time study. In Appendix 3.10 the isochrone map prepared by Rowan and Steele is given. The broad isochrones resulted from large standard deviations in the measured times. The time for a vehicle to travel from a zone to the CBD in the peak period (7.30 a.m. - 8.30 a.m.) was very much dependent upon when, during this period, the trip commenced. Thus the isochrones produced represent a spread of travel times. The times include the time to unpark the car at the home.

### 3.10.2 PARKING SURVEY

Rowan and Steele conducted a parking survey in Cape Town's CBD and produced the isochrone map for walking times shown in

Appendix 3.9.\* These walking time figures are used in subsequent analyses as part of the door to door travel time.

### 3.10.3 TRAIN TIMES SURVEY

Stanger and Sterne<sup>19</sup> investigated travel times on the southern suburbs railway line (Simonstown - Cape Town). They determined walking times by timing a random selection of people walking to and from a station. They also measured the waiting times of peak hour commuters at railway stations in the morning peak. The trains were found to run to schedule on this line. With this information they were able to produce the isochrone map for trains shown in Appendix 3.10. Again, due to insufficient time, only the southern suburbs line could be surveyed. Thus the isochrones around the Southfield railway line are not necessarily accurate. The waiting time used in the isochrone map preparation was taken as one half the average train headway and this agreed closely with the observed waiting times. The average walking speed was measured as 4,8 km/hr. This value was thought to have an upward bias due to the measurement techniques which ignored delays due to traffic signals and street crossings. The writers therefore adopted a walking speed of 4 km/hr.

### 3.10.4 BUS TIMES SURVEY

Beyers and Heeger<sup>20</sup> produced an isochrone map (see Appendix 3.11) for line-haul services† from the southern

\* The contours are rectangular in shape on this map. The writers propose that this method accounts for the rectangular layout of the walking routes.

† A Line-Haul Service is one which conveys passengers to a central area along a major route. The collector service, on the other hand, is a service which collects passengers on minor routes and delivers them to a line haul terminal either a rail station or a bus stop.

suburbs to Cape Town's CBD. They also produced an isochrone map for collector bus services in the southern suburbs for CBD commuters. They also produced a third map which showed isochrones for composite travel, i.e. for people using a bus and then transferring to a train.

Beyers and Heeger also adopted a 4 km/hr walking speed and they used an average time of 3 minutes. To measure the bus times they rode each bus at least once in the peak period.

### 3.10.5 L.O.S. INDEX FOR TRAINS

Stanger and Sterne also produced a level of service index for the train service to the southern suburbs. Level of service has been used in previous surveys and is defined by Wohl and Martin<sup>21</sup> as "the total difficulty (a traveller experiences) of making the trip at that particular time".

Various formulae for such indices have been developed and some of these are mentioned in Appendix 3.12. The formula developed by Stanger and Sterne in conjunction with the author is

$$I_j = \sum_{n=i-1}^{i+1} I_{jn} = \sum_{n=i-1}^{i+1} \frac{S_n}{\sqrt{D_{jn}}} \quad (3.1)$$

where

$I_j$  = LOS index (trains) for zone j,

$I_{jn}$  = LOS for zone j due to station n's effect alone,

$S_n$  = Train Service at station n

$\frac{(\text{No. stopping trains/peak period}) \times (\text{No. seats available/p.p.})}{1000}$

$D_{jn}$  = Distance from the population centroid to station n,

$i$  = the station closest to zone  $j$ 's centroid.

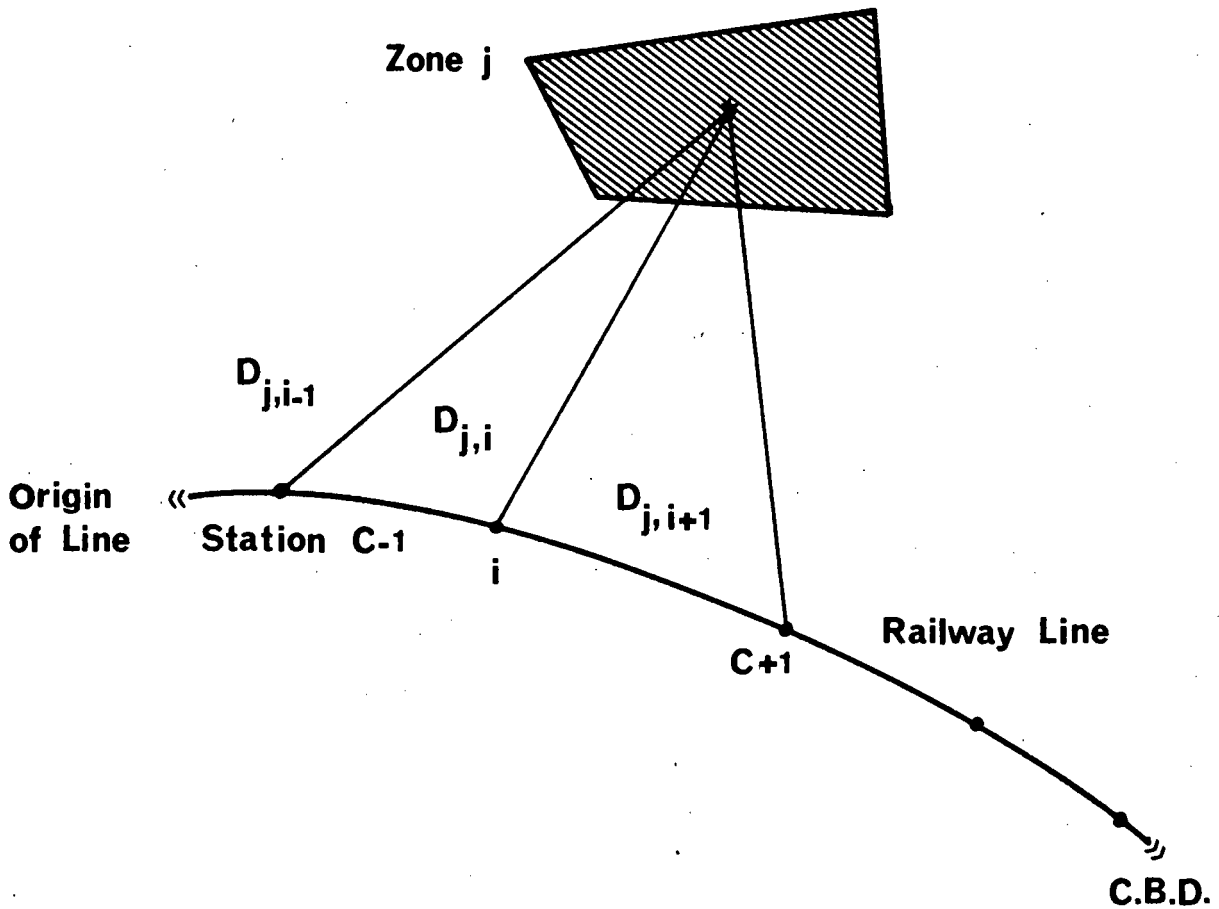
The population centroid of a zone is defined as a point where all the residents of that zone could be considered to be residing. This point can be established by cutting out a piece of cardboard in the shape of a zone, and by weighting high residential areas with putty one can find the centroid by hanging the cardboard in a vertical plane. This is done by supporting the piece at different points with a pin. After each operation a vertical line is drawn through the pin. The intersection of these lines represents the population centroid.

In the author's estimation this index is superior to those listed in Appendix 3.12. The number of available seats, the distance to the station and the number of stopping trains are variables which have a more realistic effect on service than say the number of coach miles in 24 hours or the number of seats per bus. If there are no stations in a zone or the train does not stop in peak periods in that zone, or there are no seats available in a bus, the earlier formulae yield zero LOS indices.

The concept of using the three nearest stations was developed to cater for some stations being by-passed by express trains and also where there may be some ambiguity as to which station serves a particular zone. Figure 3.2 gives a schematic explanation of the above formula.

To measure the seat availability the writers manned each station from St. James to Salt River and noted the numbers embarking and disembarking for each train. On the assumption that each weekday produced the same ridership patterns, a loading diagram and therefore a seat availability

## Schematic Development of Equation L.O.S. Index for Trains



- \* Zone Centroid
- Railway Station

**Fig 3.2** Level of Service for Trains and the Spatial Distribution of Zones

schedule could be constructed.

Table 3.6 shows the LOS index for the various zones in the southern suburbs for the Simonstown - Cape Town train line.

### 3.10.6 L.O.S. INDEX FOR BUSES

Heeger and Beyers collected data for a LOS index for buses. The index they chose was not suitable for the purposes of this thesis. However, from the available data the author was able to construct a LOS index for the southern suburbs. This is discussed further in Chapter 4.

### 3.11 SUMMARY

The surveys described in this chapter represent the most comprehensive study of this type that has been conducted in Cape Town.

The studies have several weaknesses especially in the sampling methods. However, data collection on this scale is usually carried out by a survey team under the auspices of some metropolitan planning group. The data collection and coding phases described in this chapter represent over fifty man-weeks of work and the materials costs of this study were less than R400-00. These figures illustrate that data collection does not necessarily involve huge outlays of money.

However, there are some large data losses in this project. Limitations on available manpower and funds meant that time studies and LOS studies could not be performed for the entire questionnaire study area. To complete travel time studies for this whole area would have required a large survey team and considerable resources. These resources

TABLE 3.6

L.O.S. INDICES FOR SOUTHERN SUBURBS  
SIMONSTOWN TO CAPE TOWN LINE

Zone No.	L.O.S. Index	Zone No.	L.O.S. Index
17	50	59	91
18	60	60	79
19	51	61	55
20	42	62	29
21	74	63	36
22	53	64	53
23	54	65	71
24	55	66	76
25	65	67	51
26	45	68	66
27	55	69	57
28	90	70	47
29	48	71	31
32	32	72	39
33	47	73	51
34	80	74	50
35	60	75	59
36	55	76	48
37	39	77	58
40	45	78	40
41	30	79	40
42	44	80	34
43	60	81	40
44	69	82	22
45	61	83	93
46	44	84	90
49	30	85	44
50	39	86	35
51	71	87	35
52	51	88	44

were not available to the author, and consequently the analysis for the model development is restricted to the southern suburbs.

It should also be noted that this data was collected prior to speed restrictions resulting from the "oil crisis". The figures for times and levels of service refer to the situation existing prior to these restrictions. Figures supplied by Mr. B. Floor<sup>22</sup>, Head of the Transport Research Centre of the University of Stellenbosch, indicate that there has been an increase in suburban train useage of about 25% since speed restrictions were introduced. Consequently LOS and time figures will have altered since the restrictions and any further studies conducted by the author in these areas could not be applied to the questionnaire data. Thus the models developed in subsequent chapters refer to the southern suburbs only.

REFERENCES : CHAPTER THREE

1. Wilson, F.R., "Journey to Work - Modal Split", Maclaren and Sons, London, 1967.
2. Sellitz, R., Jahooda, M., Deutsch, M. and Cook, S.W., "Research Methods in Social Relations", Henry Holt Rev. Edn., Vol. 1.
3. Goode, W.J. and Hatt, P.K., "Methods in Social Research", McGraw-Hill, Kogakusha, Tokyo, 1952, pp. 132ff.
4. Ibid, p. 134.
5. Ibid, p. 177.
6. Ibid, p. 137.
7. Ibid, p. 135.
8. Michaels, R.M. and Watson, P.L., "Social and Psychological Factors in Urban Transport Mode Choice", Research Report, The Transportation Centre, Northwestern University, 1971, p. 5.
9. Wilson, F.R., Op Cit, pp. 11-12.
10. Bowen, W., "Modal Split", B.Comm.(Hons.) Thesis, University of Melbourne, 1971, unpublished, pp. 2-6.
11. Kain, J.F., "A Contribution to the Urban Transportation Debate : An Econometric Model of Urban Residential and Travel Behaviour", The Review of Economics and Statistics, London, 1964, pp. 56-57.
12. Tomazinis, A.R., "Modal Split Model in the Penn-Jersey Transportation Study Area", Highway Research Board Record No. 165, 1967, p. 59.
13. Wilson, F.R., Op Cit, p. 187.
14. Author's discussion with Mr. R. Sharpe, Administration Superintendent, Norwich Union Life Assurance Co., 19/7/1973.
15. Wilson, F.R., Op Cit, p. 65.
16. Information supplied by Messrs. K. Sturgeon and P. Tomlinson of the Advance Planning Section, City Engineer's Department, Municipality of Cape Town.
17. Information supplied by Mr. B. Glasson, Department of Urban and Regional Planning, University of Cape Town, formerly of Cape Provincial Administration.
18. Rowan, D. and Steele, R.H., "Isochrone Map and Service Index", Two Final Year Theses, Department of Civil

Engineering, University of Cape Town, 1973,  
unpublished.

19. Stanger, N. and Sterne, K.B., "Isochrone Map and Service Index", Two Final Year Theses, Department of Civil Engineering, University of Cape Town, unpublished.
20. Beyers, O. and Heeger, E.A.J., "Isochrone Map and Service Index", Two Final Year Theses, Department of Civil Engineering, University of Cape Town, unpublished.
21. Wohl, M. and Martin, B.V., "Traffic Systems Analysis", McGraw-Hill, New York, 1967, p. 117.
22. Floor, B.C., Head, Transport Research Centre, University of Stellenbosch, discussion with the author on 21/2/1974.
23. Public Administration Service, "Conducting a Home Interview Origin-Destination Survey", Procedure Manual prepared by U.S. Bureau of Public Roads, U.S. Department of Commerce, Chicago, 1954.

## CHAPTER FOUR

### ANALYSIS OF DATA : FACTORS AFFECTING MODAL CHOICE

The data obtained from the surveys described in Chapter 3 is analysed with a view to determining which factors influence the choice of mode of travel for the journey to work. The factors examined include sex, distance from home to workplace, LOS, costs of travel and location of the workplace. An analysis of the preference questions is also made. Finally, an investigation of the possible biases in the sampling technique and other limitations concerning the data collection is provided.

#### 4.1 DATA HANDLING METHODS

The data from the questionnaire survey was initially stored in the computer with each workplace's data contained in a separate element of a data file. In this way the data could be analyzed at the workplace level or aggregated depending on the requirements of the particular analysis. Programmes were then written to establish basic tables of data. Some of the results of these preliminary analyses are now discussed.

#### 4.2 MODAL SPLIT RESULTS

The analysis of the sample population's modal choices was first performed. Figures 4.1 and 4.2 show the percentage use of each mode for CBD and non-CBD firms respectively. A breakdown of these figures for each workplace is given in

Appendix 4.1. If two or more modes were used for the journey to work the major mode only was coded, e.g. a linked trip, where a commuter used a bus to get to a railway station to catch a train would be coded as a train trip.

For the non-CBD firms 55,0% of the workers use cars to come to work, as compared with 42,7% of the CBD employees. This is probably a result of three factors:

1. public transport facilities do not serve non-CBD firms as well as they serve CBD firms.
2. roads are less crowded in decentralized areas.
3. parking is more easily obtainable at non-CBD workplaces.

It can also be observed that trains, a fixed route system, do not appear to provide as good a service to non-CBD workplaces as they do to the CBD, whereas buses which have a flexible routing system, attract proportionally more non-CBD patrons than CBD patrons.\*

Car passengers constitute the third highest mode of travel yet vehicle occupancy is still less than 1,3 persons per vehicle. These figures were obtained by adding the number of car drivers, car passengers and other passengers, where these were mentioned in the questionnaire (e.g. car pools, taking children to school, etc.) and dividing by the number of car drivers. The figures for the CBD and non-CBD are given on figures 4.1 and 4.2. The overall (global) breakdown of modal split is given in figure 4.3 and shows a similar pattern to that described above.

---

\* The non-CBD figures include Cape Town City Tramways whose employees have free bus travel passes. However, if this firm is excluded from the analysis the above observation is still true. See Appendix 4.1 for detailed figures.

## Journey to Work - Modal Split

### C.B.D. Figures

Modal Split Analysis for C.B.D. Population for the Work Journey

Vehicle Occupancy = 1,27 Persons / Vehicle

Sample Size 1072  
13 C.B.D. Firms

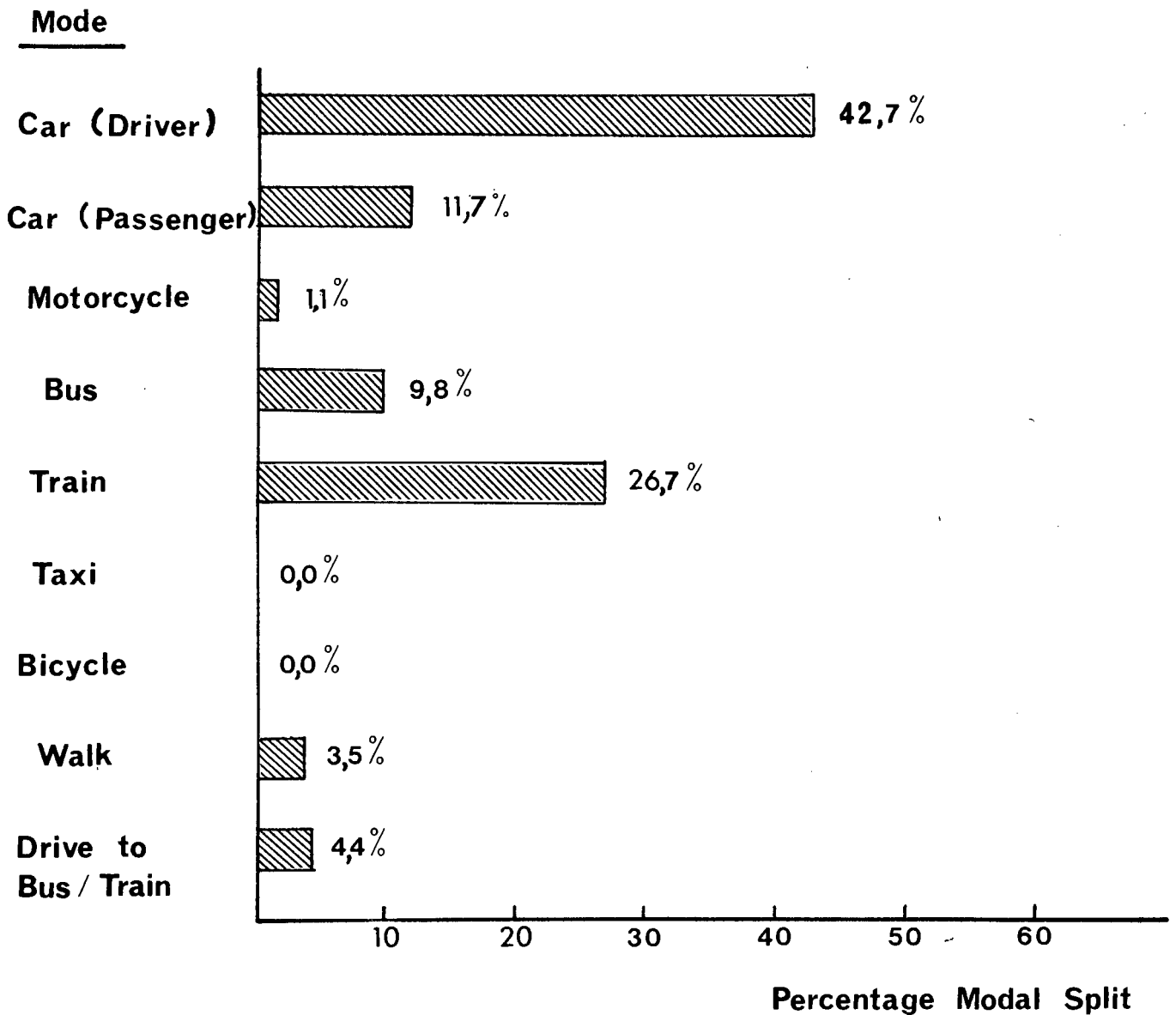


Fig 4.1

Modal Split Analysis for the Work Journey  
in Cape Town - C.B.D. Figures

## Journey to Work - Modal Split

### Non C.B.D. Figures

Modal Split Analysis for Non C.B.D. Population  
for the Work Journey in Cape Town

Vehicle Occupancy = 1,25 Persons/Vehicle

Sample Size 649  
8 Non C.B.D. Firms

#### Mode

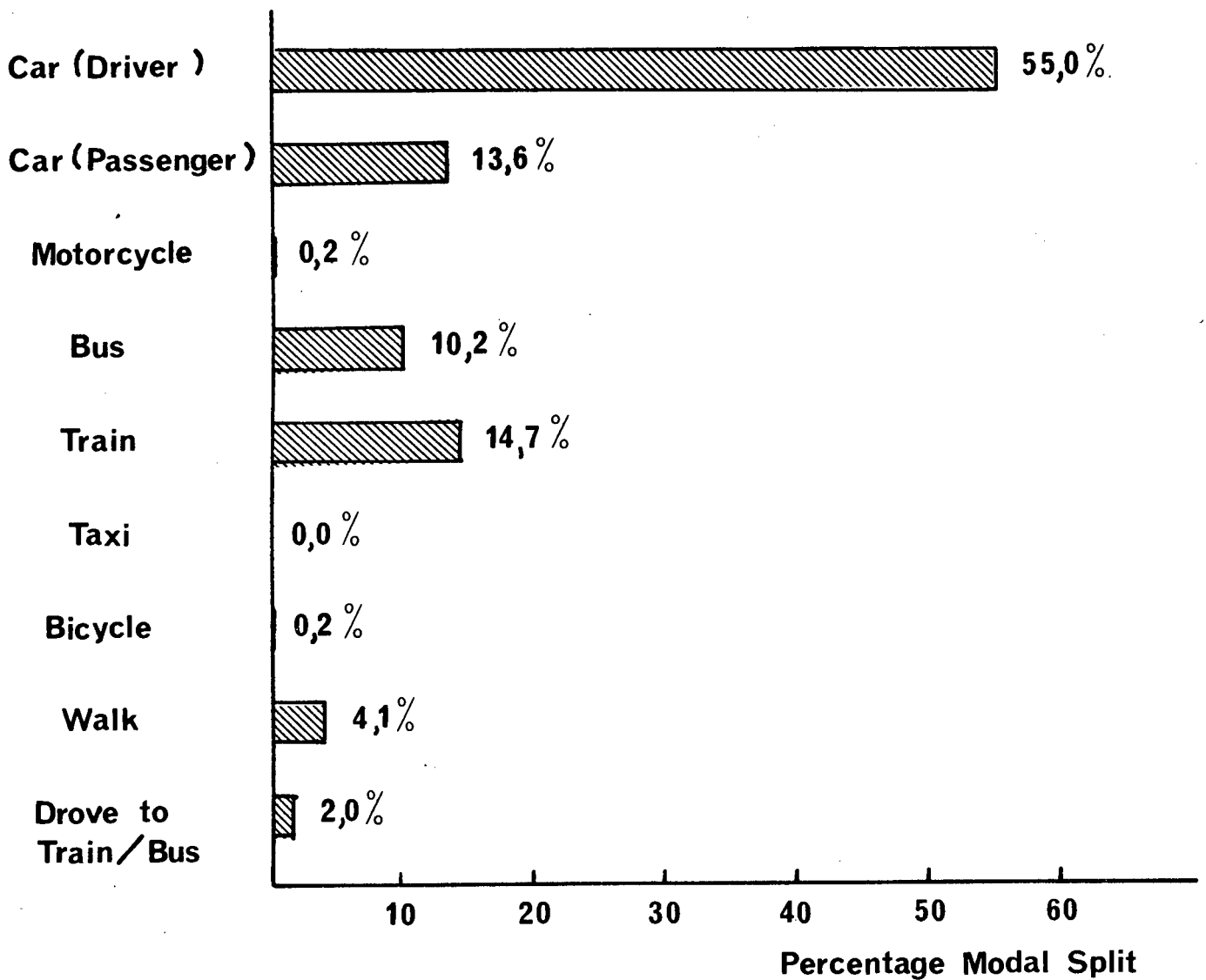


Fig 4.2 Modal Split Analysis for the Work Journey  
in Cape Town - Non C.B.D. Figures

## Journey to Work - Modal Split

### Global Figures

Modal Split Analysis for Sample Population for the Work Journey

Vehicle Occupancy = 1,26 Persons / Vehicle

Sample Size 1721  
8 Non C.B.D. Firms  
13 C.B.D. Firms

### Mode

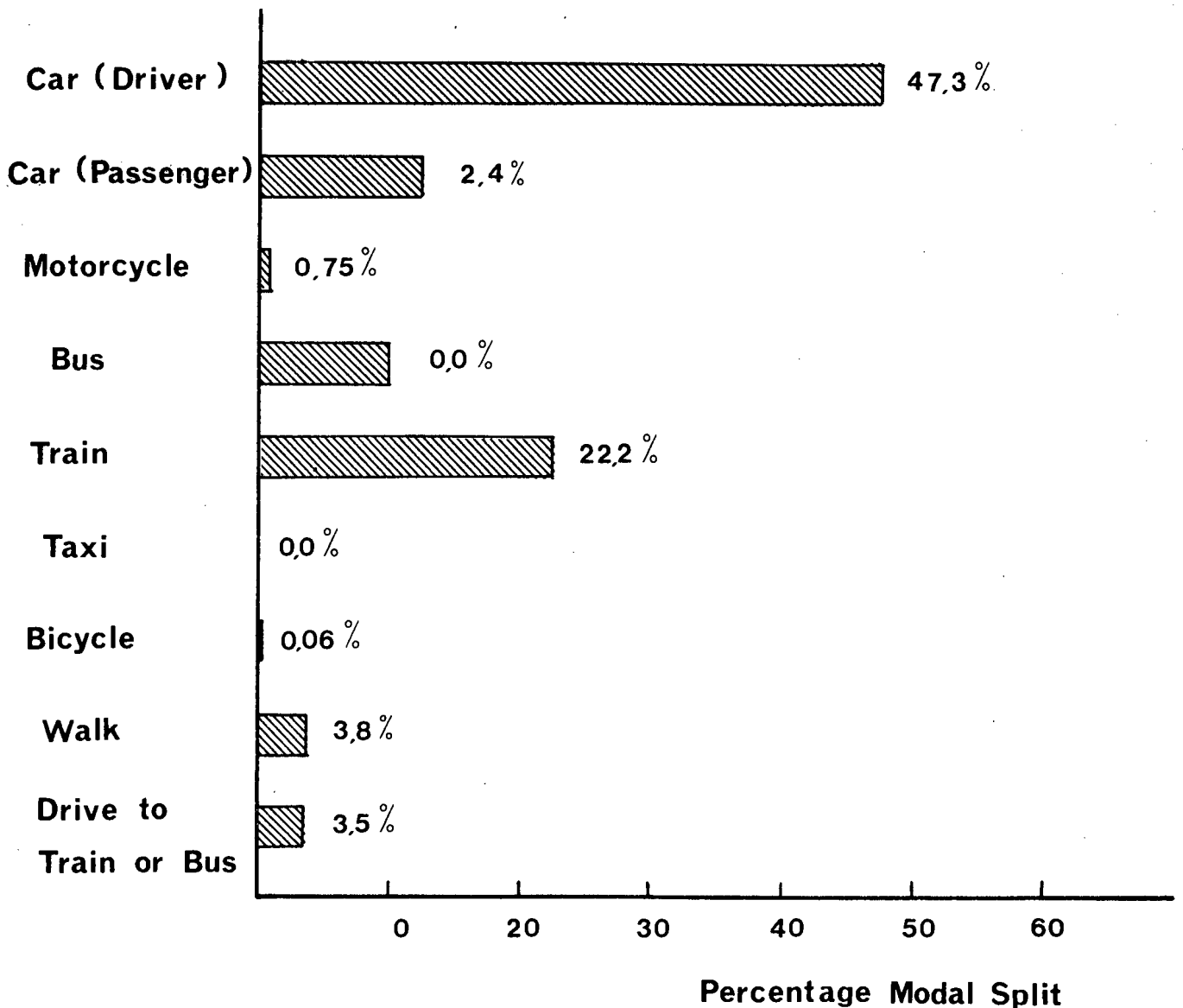


Fig 4.3

#### 4.3 INCONVENIENCE FACTORS IN PUBLIC TRANSPORT

In Part A of the questionnaire survey non-car owners were asked to list what they considered to be most inconvenient about using public transport. In the coding process space was provided for three reasons. However, if more than three reasons were given the author "judged" from earlier responses which three should be entered. It was a rare occurrence to find more than three preferences listed.

Again a computer programme was written which added each of the preferences for all the non-car owners at each workplace. These are listed in Appendix 4.2. The measure of "percent primacy reported"<sup>1</sup> was used to accommodate the respondents who gave more than one answer to this question. The percent primacy factor is simply an index for measuring the relative influence of the different factors. The percent primacy factory can be written as

$$PPF_i = \frac{N_i}{M} \times 100 \quad (4.1)$$

where

$PPF_i$  = Percent Primacy Factor for Preference  $i$ ,

$N_i$  = Number of persons stating reason  $i$  as a preference,

$M$  = Total number of preferences stated.

The bar diagrams shown in figures 4.4, 4.5 and 4.6 list the percent primacy ratings for each of the inconvenience factors relating to public transport in Cape Town. Again one notes basic differences between the CBD and non-CBD commuters in figures 4.4 and 4.5.

**Inconvenience Factors Relating to use of Public Transport in Cape Town**

Percent Primacy of Inconvenience Factors as Reported by

Non Car Owners

13 C.B.D. Firms  
Combined.  
Sample Size 287  
No. of Preferences 407

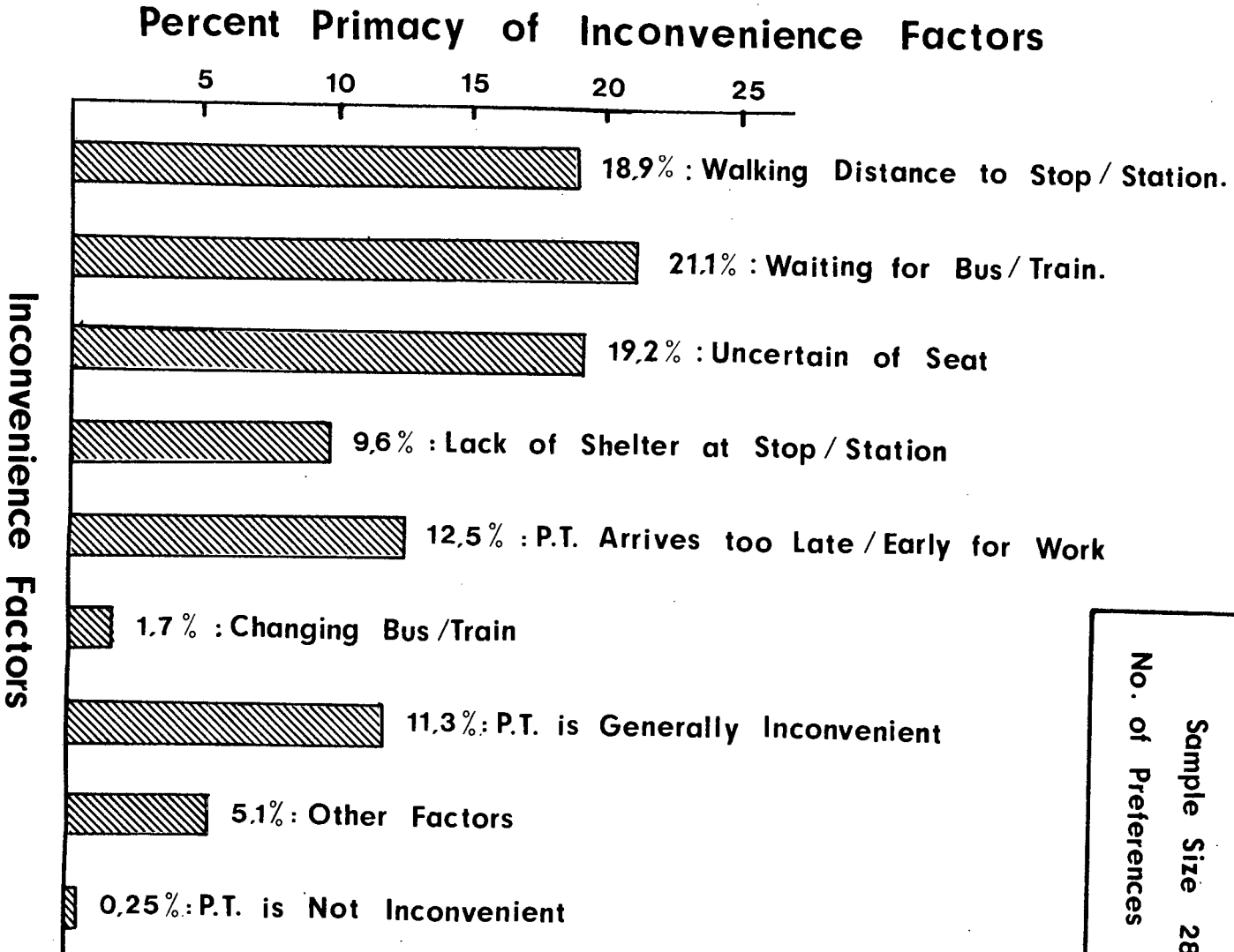


Fig.4.4 Inconvenience Factors for Public Transport Non Car Owners—13 C B D Firms

# Inconvenience Factors Relating to use of Public Transport in Cape Town

Percent Primacy of Inconvenience Factors as Reported by Non Car Owners

6 Non C.B.D. Firms  
Combined:  
Sample Size : 187  
No. of Preferences: 287

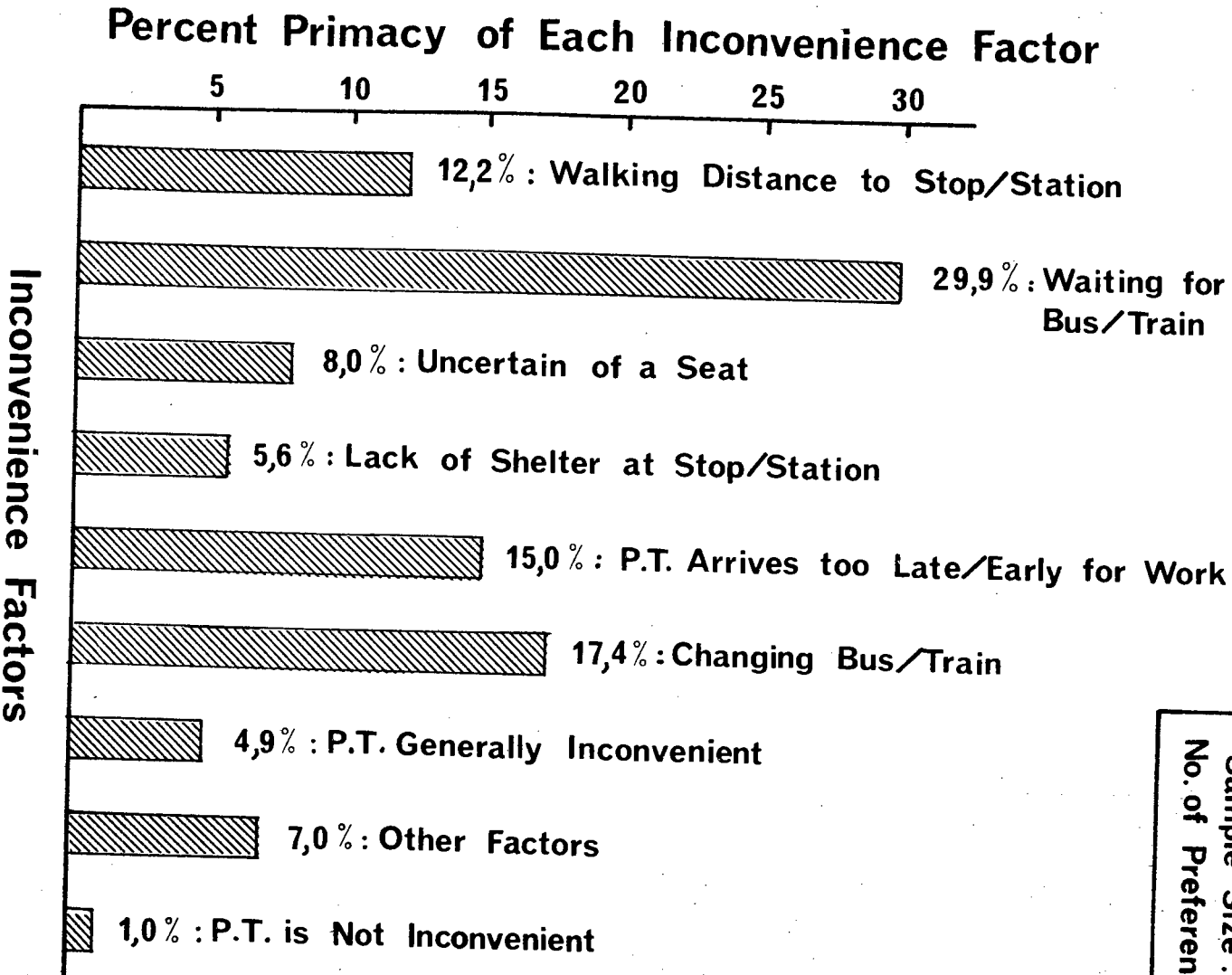


Fig.4.5 Inconvenience Factors for Public Transport Non--Car Owners - 6 Non C.B.D. Firms

## Inconvenience Factors Relating to use of Public Transport in Cape Town

Percent Primacy of Inconvenience Factors  
as Reported by Non Car Owners

Global Total of  
19 Firms :  
Sample Size: 474  
No Preferences: 694

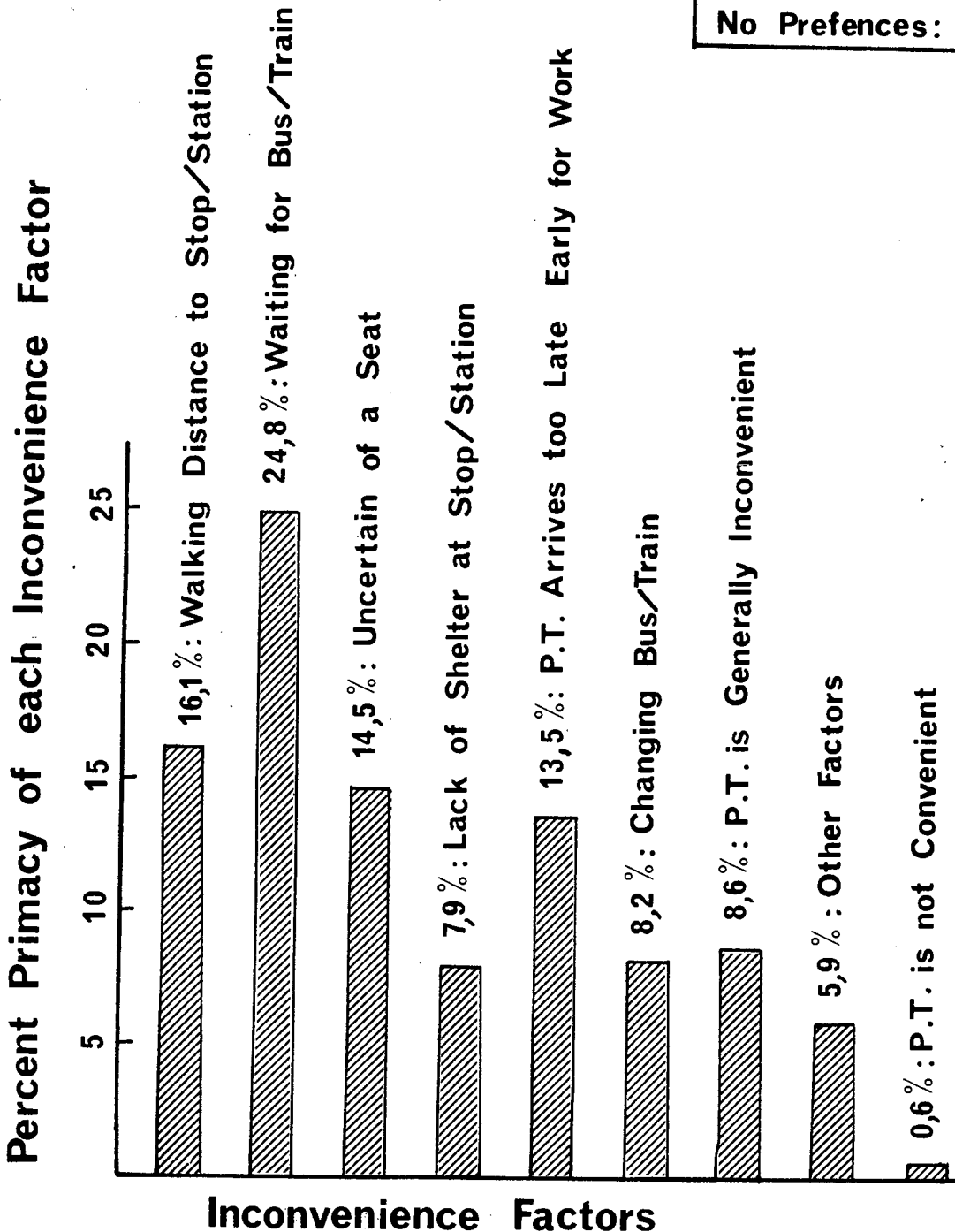


Fig 4.6 Inconvenience Factors for Public Transported Non Car Owners - Global Figures

#### 4.3.1 WAITING TIMES

This factor was rated as the most inconvenient aspect of public transport travel by both groups.

The hypothesis that public transport does not serve decentralized zones as well as the CBD is supported by these figures. Train and bus frequencies are less for decentralized areas hence waiting time is rated more highly by the non-CBD commuters (29,9%) than the CBD commuters (21,1%).

The appeals made for firms to decentralize to avoid transport problems (usually in the CBD) are not always well founded. Public transport is not well suited to serving decentralized workplaces and, from the above figures, decentralization would probably increase the number of car users and adversely affect public transport usage and consequently affect the service provided.

This aversion to waiting also supports the theory that people attach a higher disutility to waiting time than they do to actual travel time. The waiting times measured in the time studies discussed in Chapter 3 indicate that waiting times were not excessive (3 to 6 minutes) and yet they are regarded as the prime inconvenience factor for public transport.

#### 4.3.2 TRANSFERS

The inconvenience of transferring vehicles has a large importance amongst non-CBD travellers (17,4%) but is a negligent factor for CBD commuters (1,7%). This discrepancy is easily appreciated by examining the numbers of transfers reported by each group. For car owners using other modes (Section B of the questionnaire), 3,0% of the CBD commuters said they made transfers whereas 19,8% of the non-CBD firms'

employees made transfers. (The detailed figures are given in Appendix 4.2). Similarly, for car users (Section C) CBD workers estimated an average of 0,26 transfers per trip if they travelled by public transport compared with 1,01 (Appendix 4.2) for non-CBD commuters.

The comments made in section 4.2.1 concerning decentralization are supported by these results concerning the inconvenience of having to make transfers during the journey to work.

#### 4.3.3 SEAT AVAILABILITY

Seat availability is obviously more of a problem for the CBD routes than for the non-CBD routes. It was rated as the second most inconvenient factor (19,2%) by CBD commuters. This is a vindication of the LOS index for trains which was developed in Chapter 3 and which included "the number of seats available" in its formulation (equation (3.1)).

#### 4.3.4 WALKING DISTANCE TO OR FROM STOP OR STATION

This inconvenience factor was rated as the third highest factor, 18,9%, by the CBD commuters but only ranked fourth by the non-CBD commuters (12,2%). This is probably due to:

1. the locations of the CBD survey centres. These were primarily in the foreshore area\* where there are no bus services and the train station is about 5 to 10 minutes' walk. (Furthermore, the walking conditions in this area are not pleasant with little shelter from a strong

---

\* On present trends this is fast developing into Cape Town's central business district. The Financial Mail<sup>2</sup> of September 7th, 1973 commented that "Thibault Square will be the centre of Cape Town within 10 years".

prevailing wind.), and

2. that commuters to the CBD are more inclined to use public transport than non-CBD commuters (figures 4.1 and 4.2) and are therefore likely to be prepared to walk further to a station or a bus stop. Thus walking distance is a relatively greater inconvenience for CBD travellers than for non-CBD travellers.

Walking distance is also related to time and like waiting time has a relatively high perceived disutility when compared with travel time and consequently is rated highly as an inconvenience factor.

#### 4.3.5 PUBLIC TRANSPORT ARRIVES TOO EARLY/LATE FOR WORK

This factor was ranked third (15,0%) and fourth (12,5%) by the non-CBD and CBD groups respectively. These results are difficult to interpret, especially for the CBD centres where public transport operates with headways of less than 10 minutes in the morning peak. Furthermore the trains run to schedule and the buses keep reasonably consistent schedules<sup>3</sup>.

One possible explanation is that an employee is highly unlikely to consistently arrive late for work and thus his only alternative is to arrive early. It would appear that the time spent at work before working hours has a high disutility to the employee. Several firms surveyed used "flexi-time" for their office hours (see Appendix 4.2) and it was felt that this inconvenience factor would rate less highly in these firms than in other firms with fixed working hours. But there appears to be no such trend in the figures shown in Appendix 4.2, thus the above explanation may not be valid.

#### 4.3.6 LACK OF SHELTER

This factor was not particularly significant for either group (9,6% and 5,6%). In Section B of the questionnaire public transport commuters were asked if they had shelter provided at their bus stop or station, and 81,2% indicated that they were provided with shelter. These figures indicate that the provision of shelter is adequate and that it is not regarded as a major inconvenience by commuters.

The other factors were seen to be relatively unimportant. However, one question was given a new interpretation during the coding process. The inconvenience factor, "public transport vehicle is unclean", was not considered a major inconvenience by the respondents. But many replies included comments that "public transport was generally inconvenient" and so this factor was included separately in lieu of the "public transport is unclean" factor. Amongst the "other" comments were "smoking should be prohibited", "too many non-white buses" and "buses are too bumpy".

#### 4.4 REASONS FOR NOT USING CAR FOR THE WORK JOURNEY

In Section B of the questionnaire, people who owned cars and did not use them to come to work were asked their reasons for not using their own cars to make the journey. The charts in figures 4.7, 4.8 and 4.9 show the percent primacy rating for each of these reasons for the CBD, non-CBD and total samples. (See Appendix 4.2 for detailed figures for each workplace.)

These graphs again demonstrate basic differences between the CBD and the non-CBD commuters' preferences.

## Reasons for Not Using Private Vehicle for Journey to Work in Cape Town

Percent Primacy Figures of Reasons as Reported by Car Owners

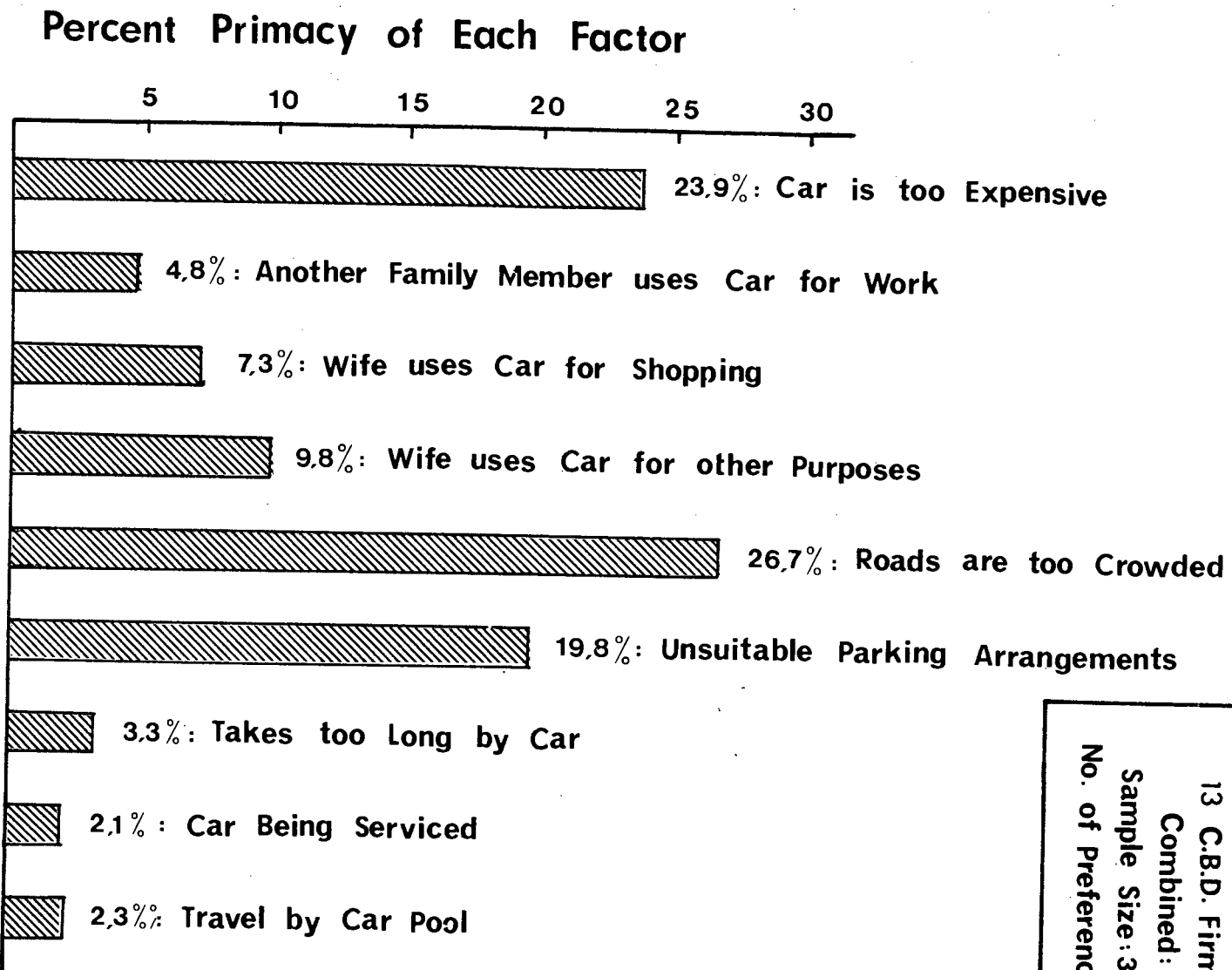


Fig.4.7 Reasons for Not Using Private Vehicle for the Journey to Work in Cape Town — 13 C.B.D. Firms

## Reasons for not Using Private Vehicle for Journey to Work in Cape Town

Percent Primacy Figures of Reasons as Reported by Car Owners

6 Non C.B.D. Firms  
Combined:  
Sample Size: 110  
No. of Preferences: 134

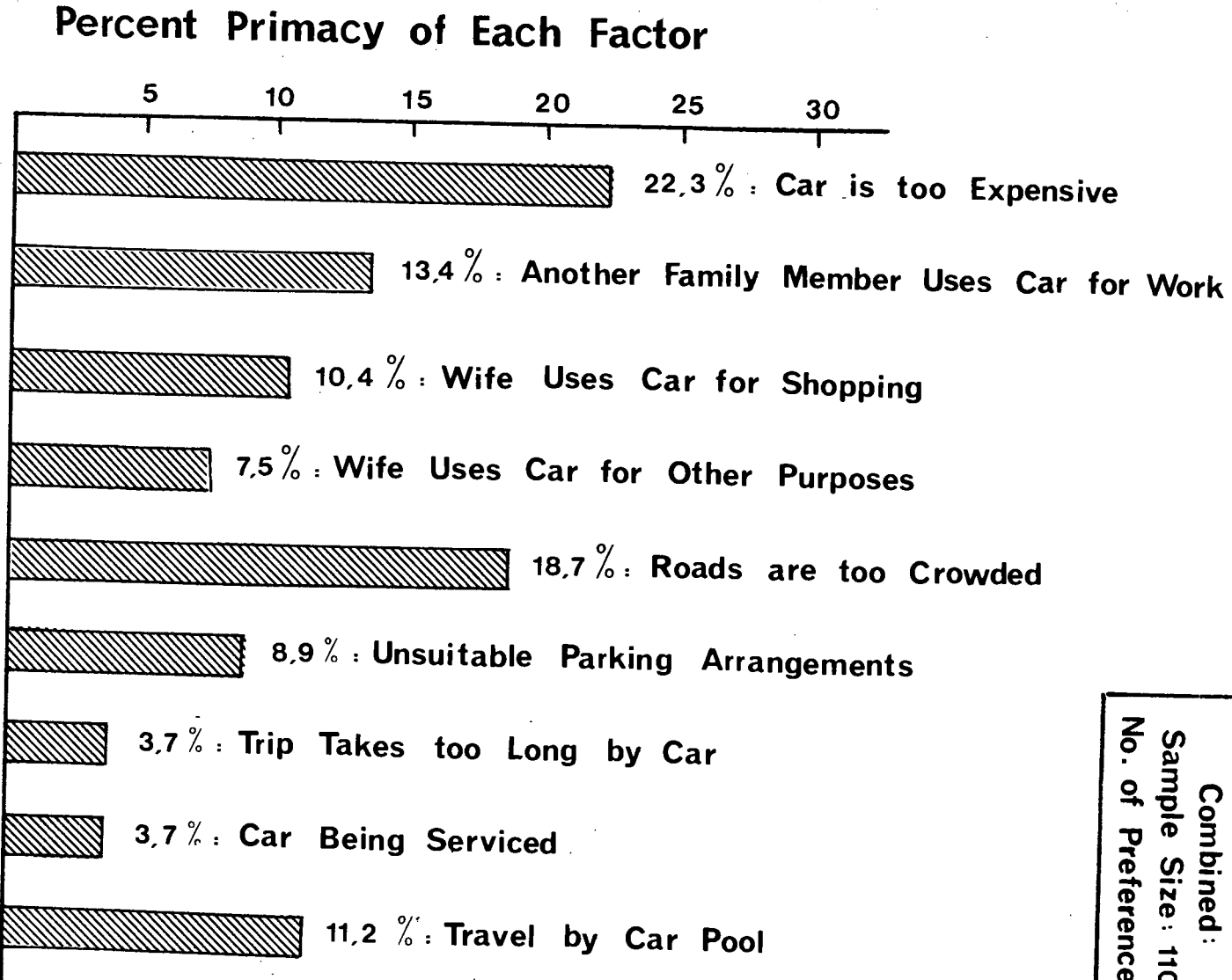
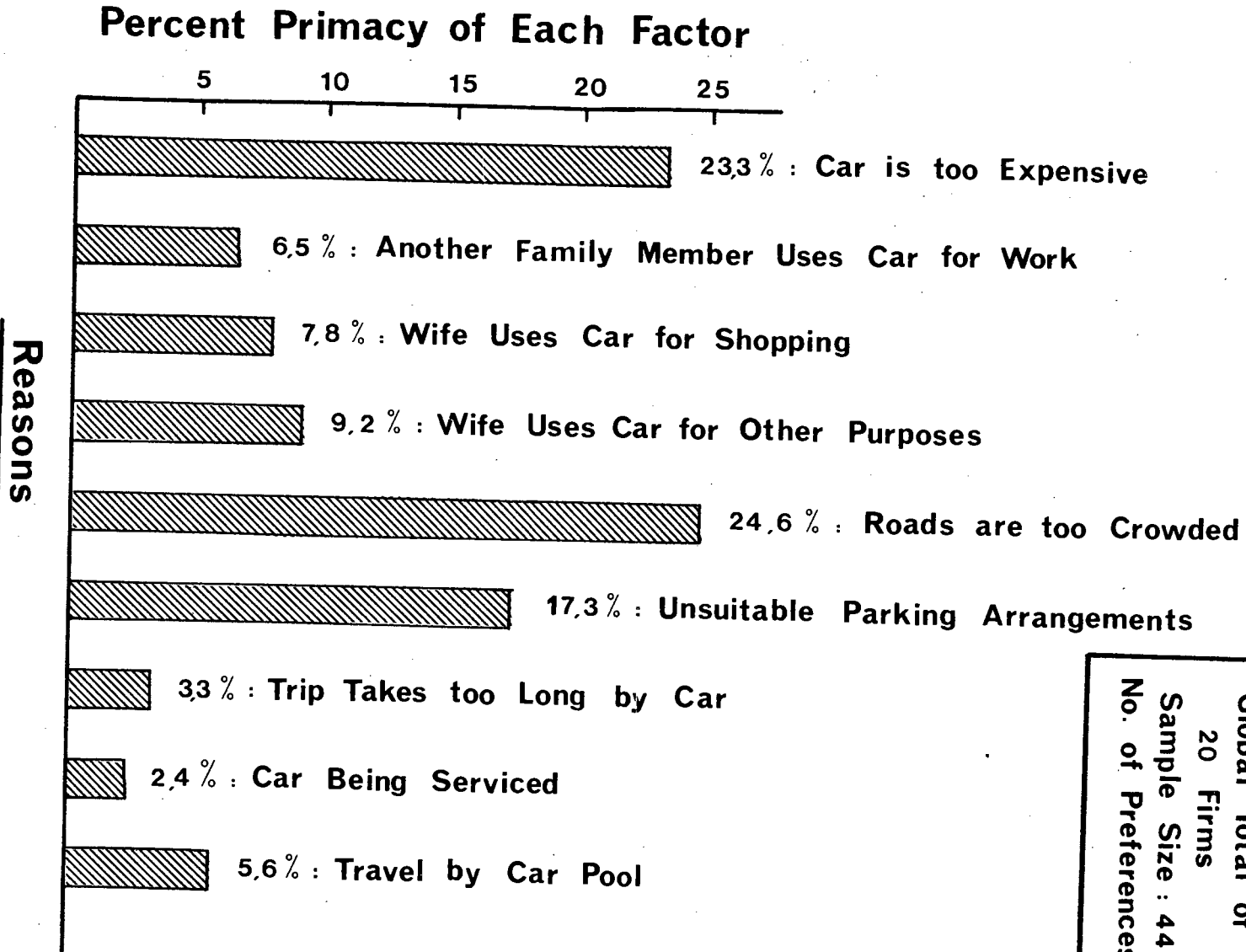


Fig4.8 Reasons for Not Using Private Vehicle for the Journey to Work - 6 Non C.B.D. Firms

## Reasons for Not Using Private Vehicle for Journey to Work in Cape Town

Percent Primacy Figures of Reasons as Reported by Car Owners

Global Total of  
20 Firms  
Sample Size : 446  
No. of Preferences 665



**Fig. 4.9** Reasons for Not Using Private Vehicle for the Journey to Work in Cape Town – Global Firms

#### 4.4.1 ROADS ARE TOO CROWDED

The differences in the car usage between the CBD and the decentralized areas is partially explained by this reason. For the CBD commuters this was the prime reason (26,7%) for leaving one's car at home whereas it was ranked second (18,7%) amongst the non-CBD workers. Both groups perceived a high disutility towards congestion which is, of course, related to time delays.

This factor also gives an indication of the effect of freeway improvement projects. As has been observed in most urban areas road improvements are quickly nullified by increasing traffic. One of the reasons for this is the large number of motorists who are attracted to these new facilities because congestion has been eased (albeit temporarily). This suppressed traffic<sup>4</sup>, as it is termed, is quite significant if the figures resulting from this question are any indication.

#### 4.4.2 THE CAR TRIP IS TOO EXPENSIVE

This is a measure of the cost variable relating to modal choice. Both groups regarded this factor as an important reason not to travel by car (CBD 23,9% and Non-CBD 22,3%). This also tends to contradict Hensher's<sup>5</sup> observations that the motorist drastically underestimates the true costs of driving his car. However, it is still likely that he does so but not to the extent that he considers car travel is cheaper than public transport fares. In fact, only 6,2% of car users regarded car travel as cheaper than public transport (see Section 4.5, figure 4.12).

#### 4.4.3 UNSUITABLE PARKING ARRANGEMENTS AT WORKPLACE

There is a large discrepancy in the percent primacy factors for this reason between the CBD and the non-CBD groups. The percent primacy factor for unsuitable parking facilities in the non-CBD responses was 8,9% compared to 19,8% for the city workers. This again helps explain the differences in car useage shown in section 4.2. The above figures are explained by the figures listed in Appendix 4.3, which show that only 14,0% of the CBD respondents stated that they had free parking at their workplace, whilst 85,5%\* of non-CBD respondents were able to park for no charge.

#### 4.4.4 THE CAR IS USED BY ANOTHER MEMBER OF THE FAMILY

This reason was divided into three sections. These sections were, that the commuter does not use his own car because:

1. another member of the family uses the car to go to work,
2. the wife uses the car for shopping purposes, or
3. the wife uses the car for other purposes.

These reasons really relate to the car availability for car owners. People who own cars may not necessarily have them available for their exclusive use. A decision is made as to who should have the use of the vehicle. In other words, is a commuter's perception of alternative modes of travel favourable enough that he is prepared to forego using his car to drive to work in order that another member of the household may have use of the car? And more importantly if that

---

\* If Cape Town City Tramways' respondents are eliminated from these non-CBD figures the percentage of people who have free parking increases to 98,8%.

household can afford a second car, will the husband continue to use an alternative mode or will he revert to driving his own car now that the household has the second vehicle available? One can not really answer this question from the percent primacy figures shown in the figures 4.7 and 4.8. However, one notices that the non-CBD employees rated these three reasons for not using their cars at 31,3% (total) whilst the CBD commuters responses represented a percent primary figure of 21,9% for these reasons. One may infer, therefore, that with an increase in the car availability for all households, all other factors remaining constant, a proportionally higher number of the non-CBD commuters will drive to work than the CBD commuters. This reinforces the argument which has been developed in this chapter - that a reduction in any of the components of the car user's price vector will result in an increase in car usage which will be proportionally larger for the non-CBD sector\*.

#### 4.4.5 TRAVEL BY CAR POOL

At the time this survey was conducted, car pools were illegal under the provisions of the 1930 Motor Carrier Transportation Act. Consequently it is possible that the response to this question may be biased. However the difference between the CBD and the non-CBD respondents is again evident. For the non-CBD firms 13,6% (PPF 11,2%) of the car owners not using their cars were members of a lift club. But only 3,6% (PPF 2,3%) of those working in the CBD

---

\* Stated more precisely this means that the price elasticity of demand for car use is greater for the non-CBD commuters than for the CBD commuters. The concepts of elasticity are discussed in Chapter 6.

were car pool members. The survey showed that only 4,6% of the car owners were members of car pools\*. (See Appendix 4.3.)

The car pool was therefore ignored as a separate mode in the modal split models. But if future trends indicate an increase in the use of car pools this method of travel should be considered separately. It is obvious that car travel becomes proportionally less expensive with the more people participating in such a scheme. However, car pools do detract from the privacy, flexibility and independence which a car provides for the single driver.

#### 4.5 REASONS FOR CHOOSING TO USE CAR FOR THE WORK JOURNEY

In section C of the questionnaire car users were asked to give their reasons for using their cars to come to work. Bar charts showing the percent primacy figures for each reason are given in figures 4.10, 4.11 and 4.12 for CBD, non-CBD and global totals respectively.

##### 4.5.1 THE CAR TRIP IS QUICKER

The main reason car users gave for driving their cars to work was that it was quicker to travel by car than by using alternative modes. Again time is an important factor in the choice of travel mode. There appears to be little difference between the CBD (28,3%) and the non-CBD (31,1%) primacy figures for this factor. From the isochrone maps shown in

---

\* In March 1974 the S.A. parliament amended<sup>6</sup> the 1930 Motor Carrier Transportation Act to legalize car pools. The Argus<sup>7</sup> newspaper is sponsoring a computerized data bank to organize lift clubs or car pools as an attempt to reduce car usage and thus to save fuel.

## Reasons For Choosing to Use Private Vehicle For The Journey to Work in Cape Town

Percent Primacy Figures for Reasons as Reported by Car Users

13 C.B.D. Firms  
Combined  
Sample Size : 456  
No. of Preferences 812

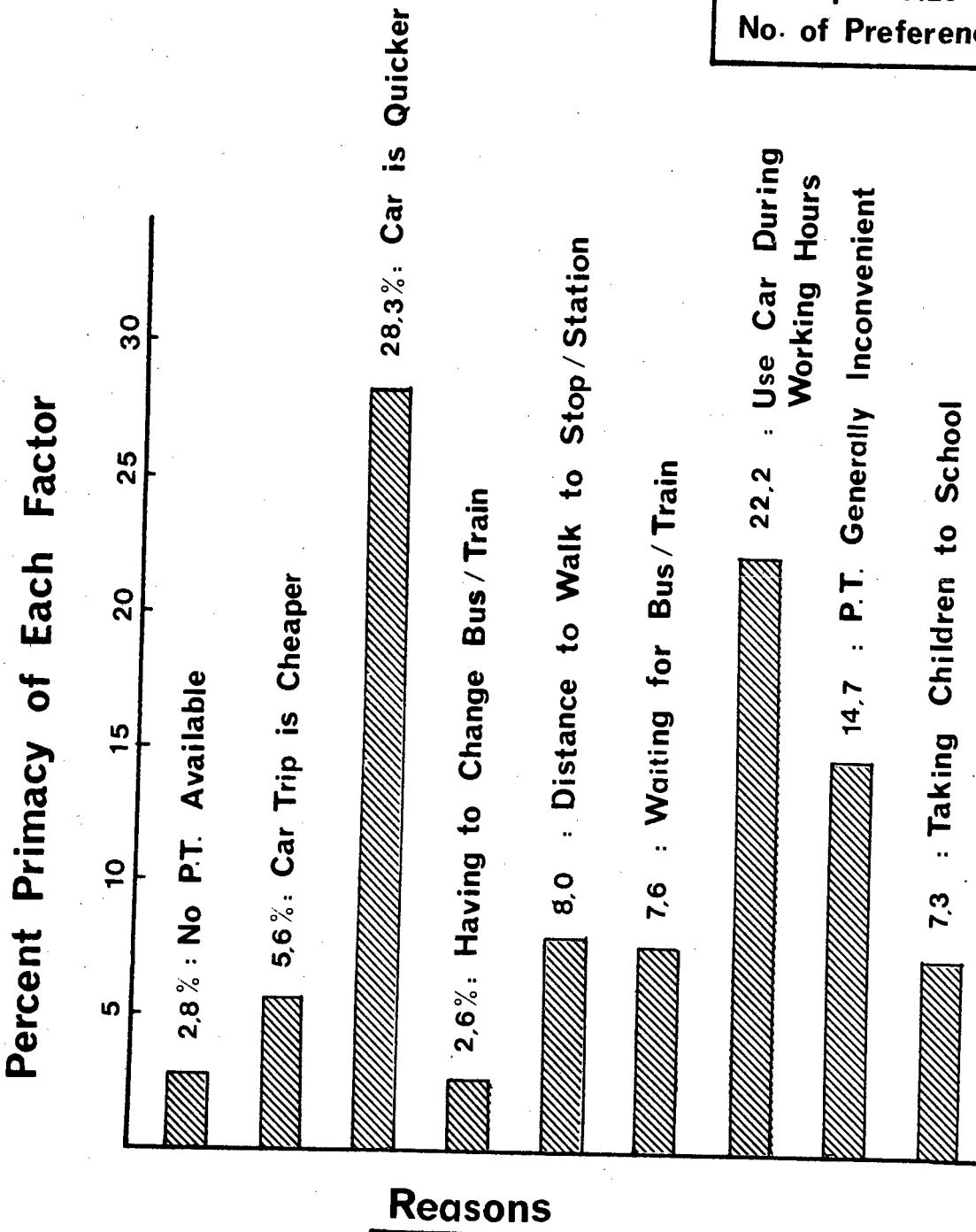


Fig 4.10 Reasons for Using Private Vehicle for the Journey to Work in Cape Town — 13 C.B.D. Firms

## Reasons for Choosing to use Private Vehicle for The Journey to Work in Cape Town

Percent Primacy Figures for Reasons as Reported by Car Users

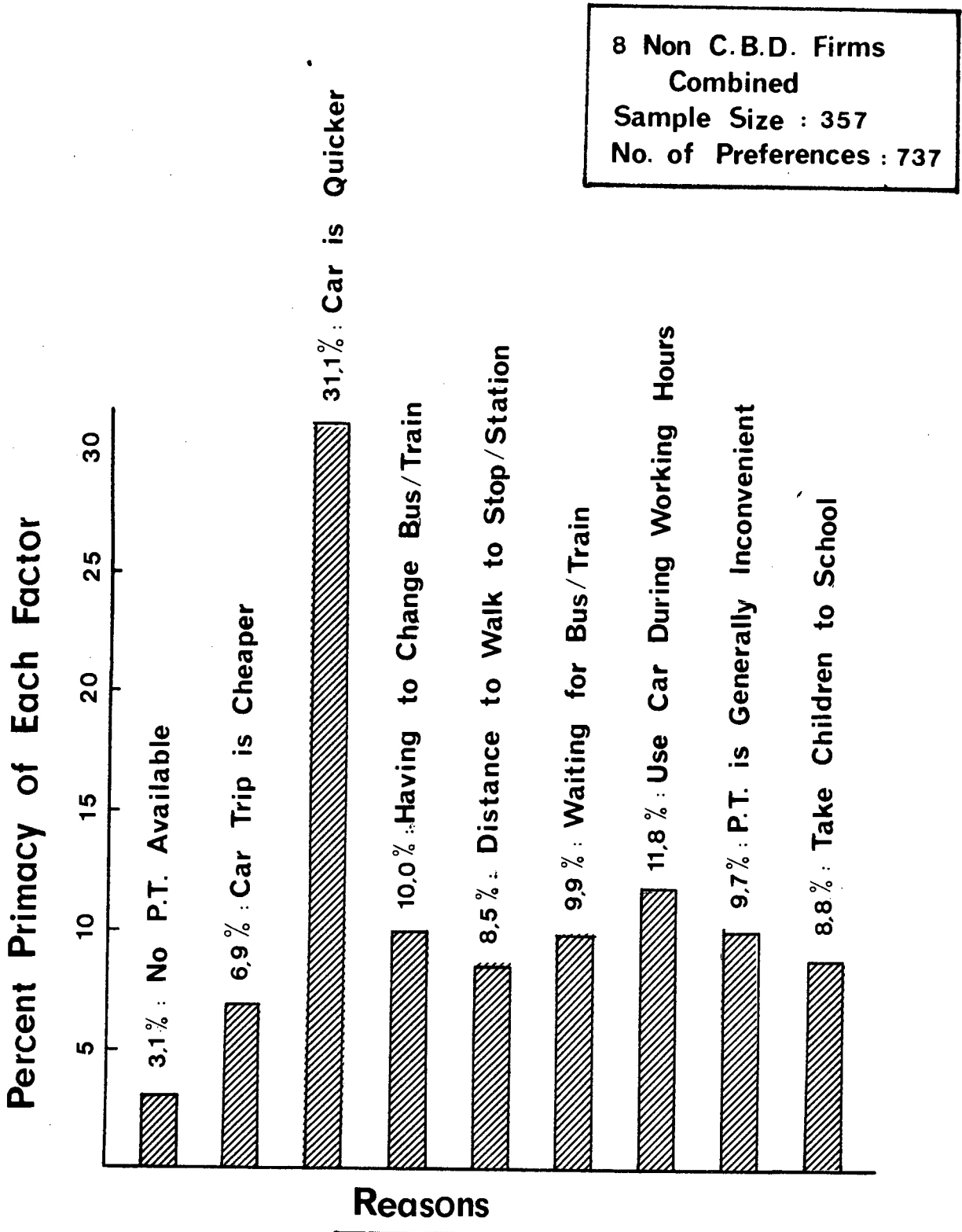


Fig. 4.11 Reasons for Using Private Vehicle for the Journey to Work in Cape Town - 8 Non C.B.D. Firms

**Reasons for Choosing to Use Private Vehicle  
for the Journey to Work in Cape Town**

Percent Primacy Figures for Reasons as Reported by Car Users

Global Figures  
Sample Size 813  
No of Preferences 1549

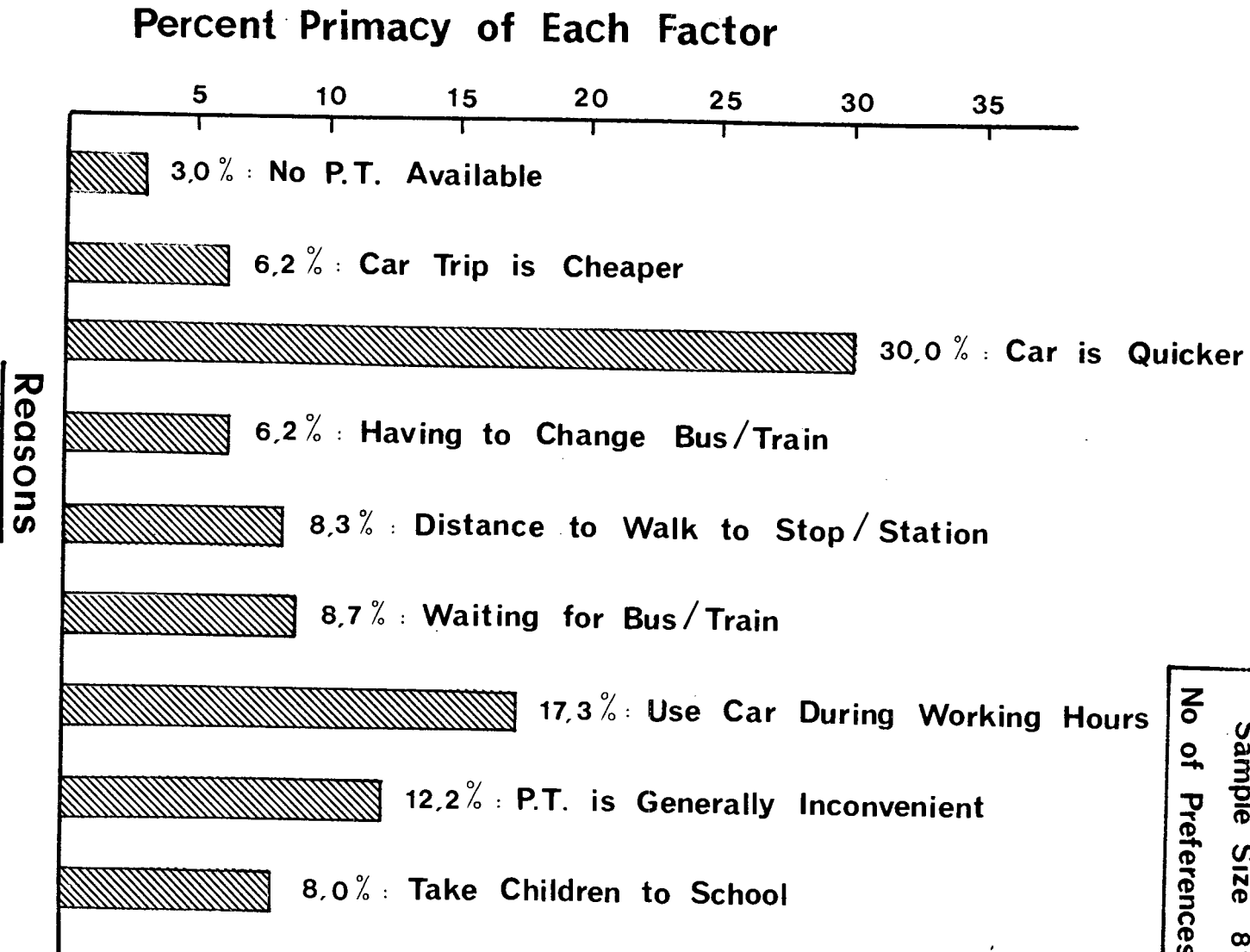


Fig 4.12 Reasons for Using Private Vehicle for the Journey to Work in Cape Town – Global Figures

the appendices to Chapter 3 it can be seen that the car times are lower than either the bus or train times at all points. But, as was noted earlier, the car driver probably over-estimates the times of public transport and this would explain to some extent the predominance of the responses for this reason.

#### 4.5.2 THE EMPLOYEE USES HIS CAR DURING WORKING HOURS

This question was redefined during the coding stage to mean that the car was used during, or immediately after, working hours. Several respondents noted that they needed their car to go to night classes at university or to sports practice immediately after work. Such reasons were coded under 'uses car during working hours'. The large discrepancy between the CBD figures of 22,2% and the non-CBD figure of 11,8% can probably be attributed to the difference in the nature of the firms sampled. In the city area a number of engineering firms (consultants and the Cape Town City Engineer's Department) were sampled. These firms require many of their employees to visit sites in their own vehicles. Whereas the non-CBD workplaces consisted mainly of insurance companies and the like, whose employees are not usually required to leave their offices during working hours. It is interesting to note that Wilson obtained very similar figures of 20,8% (CBD) and 5,3% (non-CBD) for his survey in Coventry. It is probably correct to assume that at least 15 to 20%\* of

---

\* It should be remembered that the figure of 22,2% quoted earlier is a percent primacy rating. In fact this p.p.r. represents 39,9% of the CBD car drivers in the survey. Consequently this figure of 15% to 20% is a conservative estimate.

the CBD employees who use their car to come to work are required to use it during working hours. This is a factor which certainly influences the modal choice decision and it is one which has been ignored in past surveys. This group of commuters are in effect car captives and should be considered separately in modal split studies.

#### 4.5.3 THE CAR TRIP IS CHEAPER

This reason only rated 6,2% as an overall percent primacy rating. As was noted earlier in this chapter, people involved in car pools, who share driving costs, may in fact have a cheaper travel cost than if each used public transport. As the number involved in such schemes and their methods of sharing costs was unknown, drivers in car pools are assumed to incur the same costs as other drivers for the purposes of the model development.

#### 4.5.4 DRIVE CHILDREN TO SCHOOL

The percent primacy rating for this factor was 8,0% overall, which represents 15,2% of all the car users in the survey. This figure, when related to the number of husbands and wives who drive cars to work and who have children less than sixteen years of age, becomes 34,5% of this group\*.

This aspect of dual purpose trips is a field which needs further exploration. There are hidden costs and benefits involved here which to the author's knowledge have not been explored in the context of modal split. The motorist who takes, say, two children to school on his way to work incurs

---

\* 74,0% of the car users were either husbands or wives and 59,2% of these had children less than sixteen years of age.

extra fuel costs which are not accounted for when analyzing his choice decision. Furthermore, a simple difference between public transport costs and car costs is misleading in this case. This motorist would need to pay for his children's bus or train fares to school if he did not use his own car.

The author has no such similar figures for work journeys overseas. However, from observations in Australia\* and Europe there does appear to be a higher proportion of children delivered to (and collected from) schools in Cape Town than in these other countries. These conclusions are purely subjective but nevertheless future surveys conducted in South Africa should consider collecting more detailed information concerning these trips. This is a further demonstration of the need to perform local research which can explore these special situations.

#### 4.5.5 OTHER REASONS RELATING TO THE INCONVENIENCE OF PUBLIC TRANSPORT

The above reasons related to the convenience of using a car. The remaining factors were concerned with the negative aspects of the public transport alternatives. If one compares these figures with those for section B shown in figures 4.7 to 4.9, it can be seen that the rankings are of the inconveniences of waiting, walking and transfer times are comparable. The car users' perceptions of alternative modes were discussed in Chapter 2. The figures in Table 4.1 demonstrate that car users do in fact perceive the

---

\* One likely reason for this observation is that schools in the south eastern states of Australia commence at 9 a.m., thus avoiding the morning work peak. Consequently it is less convenient to deliver children to school on the way to work.

inconvenience of travel by public transport with similar priorities to those shown by commuters who use public transport. But the relative magnitudes of these disutilities are still a subject for conjecture. In Table 4.1 the rankings of certain inconveniences of public transport are listed with the percent primacy figures for the car owner using public transport and the car user.

TABLE 4.1  
PRIORITY RANKINGS OF INCONVENIENCE FACTORS  
FOR PUBLIC TRANSPORT

	CBD FIRMS		Non-CBD FIRMS	
	Car Users	Car Owners Using Alt. Modes	Car Users	Car Owners Using Alt. Modes
INCONVENIENCE FACTORS	PRIORITY RANKINGS WITH P.P.R. IN BRACKETS			
Walking Distance to Stop	2 (8,0)	2 (18,9)	4 (8,5)	3 (12,2)
Transferring Bus or Train	4 (2,6)	4 (1,7)	1 (10,0)	1 (29,9)
Waiting for Bus or Train	3 (7,6)	1 (21,1)	2 (9,9)	2 (17,4)
PT is generally inconvenient	1 (14,7)	3 (11,3)	3 (9,7)	4 (4,9)

The percent primacy figures do not lend themselves to direct comparisons as they relate to different population groups facing different reasons in their respective questions. However it does appear that the "waiting" factor is not regarded by the car user in such a low regard as it is by the non-car user. However, a more direct measurement of perceived times is discussed in Appendix 4.6.

The percent primacy rating for "no public transport is available" was 3,0%. This represents 5,7% of the car users. These figures indicate that Cape Town is relatively well

served with public transport. Wilson's<sup>8</sup> survey in Coventry resulted in percent primacy ratings between 9,0% and 14,9% for this factor relating to bus service in Coventry. (There is no rail service in Coventry.)

#### 4.6 MALE-FEMALE STRATIFICATION IN MODAL SPLIT

Previous studies have noted that the commuters' sex is an important factor in modal split. The sample population was divided into male and female and modal split figures were produced for each group. For purposes of comparison these figures are listed in Table 4.2.

TABLE 4.2

MALE-FEMALE MODAL SPLIT COMPARISON  
GLOBAL SAMPLE POPULATION

MALES : Sample Size 1099

Mode:	Car Driver	Car Passenger	Motor Cycle	Bus	Train	Taxi	Bicycle	Walk	Drive to PT
No.:	644	71	10	71	220	0	1	40	42
%:	58,6	6,5	0,9	6,5	20,0	0,0	0,1	3,6	3,8
FEMALES : Sample Size 622									
No.:	169	142	3	103	162	0	0	24	19
%:	27,2	22,8	0,5	16,6	25,9	0,0	0,0	3,9	3,1

Female commuters make proportionally higher use of public transport (42,5%) than do males (26,5%).

In addition, 58,6% of the male commuters drive cars to work as compared to only 27,2% of the female commuters. These figures are related to the car ownership rates and the age distributions of the male and female commuters. It was

decided to test this variable in the regression equation as it appeared to have a significant effect on the modal choice process.

#### 4.7 MODAL SPLIT AS A FUNCTION OF DISTANCE

Several writers have examined the distance of the work trip as an influencing variable in the modal choice process. Mitchell and Clark<sup>9</sup> formulated a generalized cost function which was a function of the distance of the trip. The author chose zones 1 to 88 for this analysis. A population centroid was established for each zone and rectangular coordinates were given to each centroid. Similarly coordinates were determined for each workplace. A computer programme was written which checked if each respondent lived in zones 1-88, and if so, for each mode, a direct distance was calculated from the origin zone to the workplace centre. In Appendix 4.4 the programme MD2 (Mode Distance 2) is listed with an explanation of its use. The results of this analysis are shown in figures 4.13, 4.14, 4.15 and 4.16\* for CBD and non-CBD centres. The three major modes (car, bus and train) are shown in figure 4.13 (for the CBD) and figure 4.15 (for the non-CBD). The shapes of the curves show an irregular behaviour but the basic tendencies are evident. The bus provides an attractive mode of travel initially but becomes decreasingly popular the greater the trip length becomes. It is interesting to note that the decline with distance of bus patronage is much more rapid for the non-CBD firms. As

---

\* The results in these graphs are corrected for sample bias. An explanation of this procedure is given in section 4.10.

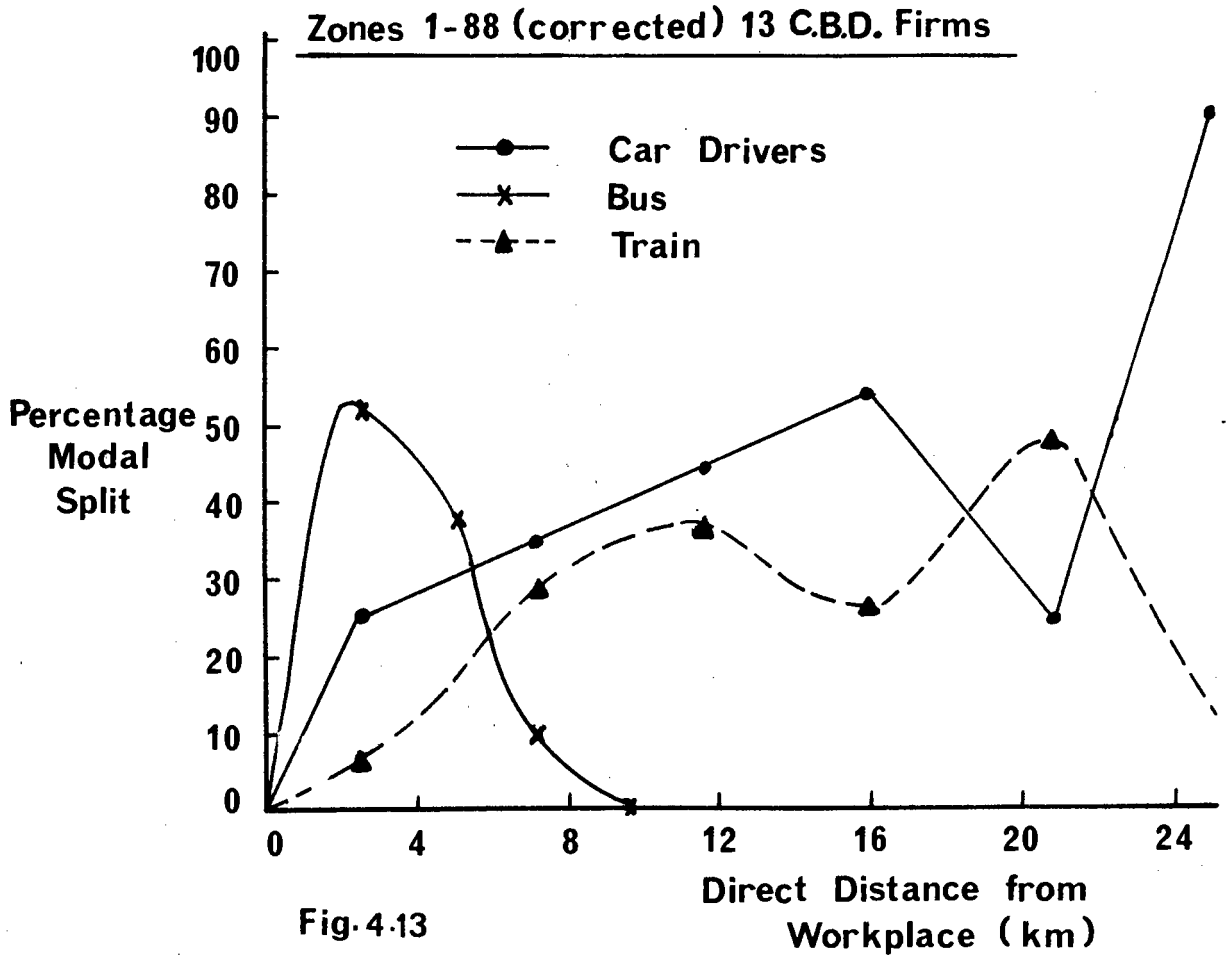


Fig. 4.13

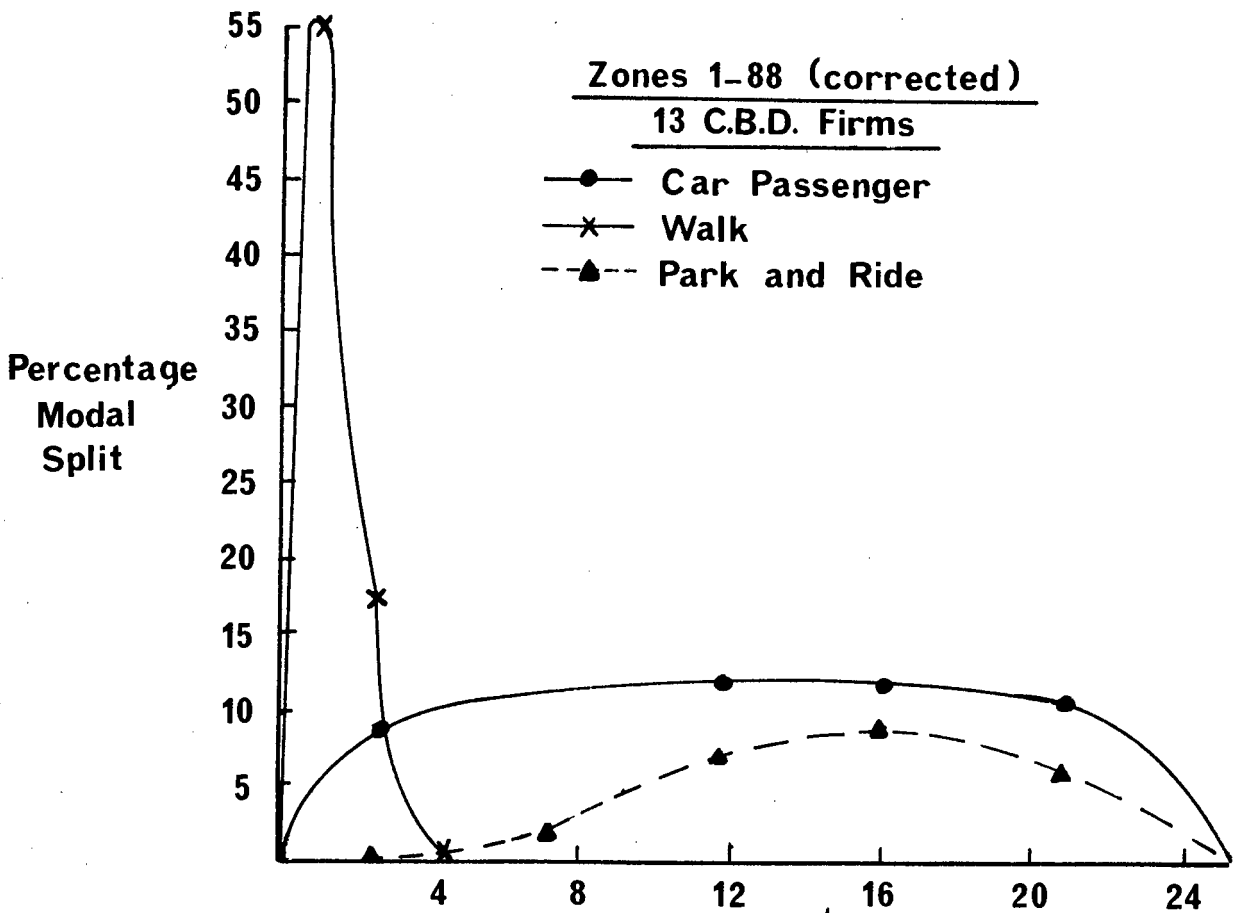


Fig. 4.14 Modal Split vs Distance from Workplace (C.B.D.)

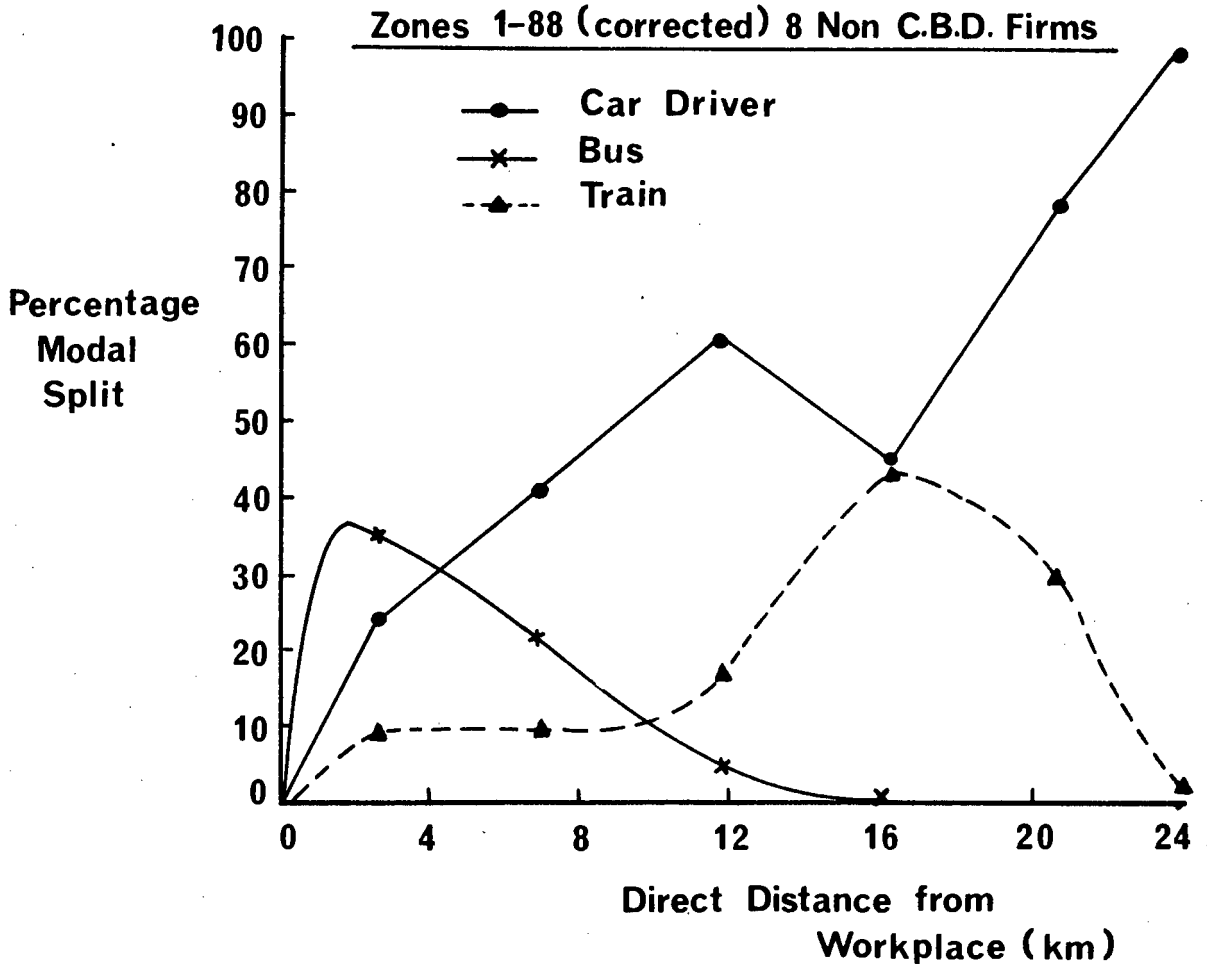


Fig. 4.15

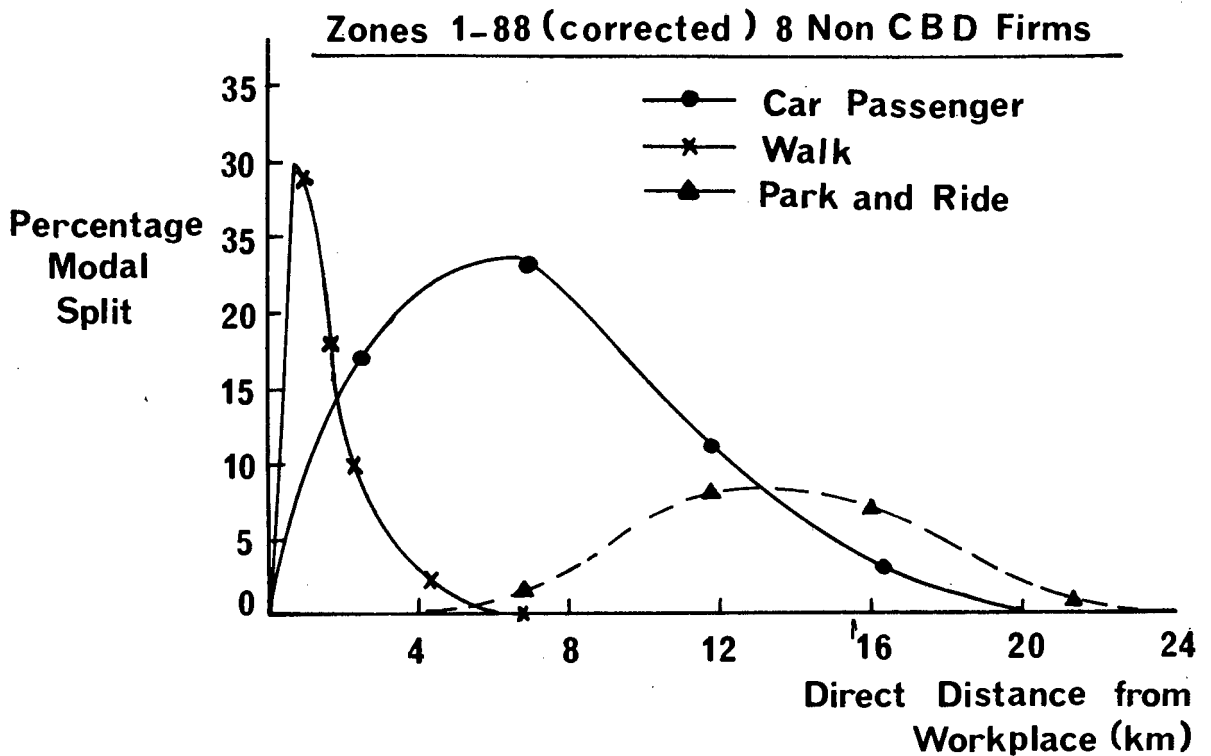


Fig.4.16

Modal Split vs Distance from Workplace  
(Non C.B.D.)

mentioned in Chapter 3, the bus company provides a line haul service along the main arterial routes. But for suburbs away from these routes the buses are used as a collector system. The zones included in this analysis (zones 1-88) are primarily in the Atlantic suburbs (Camps Bay, Sea Point), the central suburbs (Gardens, Vredehoek), and the southern suburbs (Observatory to Muizenberg). In the southern suburbs the line haul service for buses effectively terminates at Wynberg\* (see zonal map in Appendix 3.11). This is a possible explanation of the cut off point for buses at the 9 km point in figure 4.13.

Trains provide a more consistent service over a wider range of distances; however, there are no train routes in the Atlantic suburbs nor in the central suburbs. These suburbs have short direct distances to the CBD and hence the lower percentage train usage for the short distances is displayed in these graphs. The train curve has a minimum point at 16 km which requires some explanation. An "employee centroid" of the CBD firms was found by weighting each firm's location by the number of employees it had in zones 1 to 88 and locating the centroid of those points. An arc representing 16 km was struck on the zonal map. The arc coincided with the zones which were situated in areas away from the railway routes and in relatively high socio-economic areas (e.g. Tokai, Constantia Neck, Zeekoevlei). It would be expected, therefore, that rail usage would diminish in these areas and car useage would increase. The geography of Cape Town is such that Table Mountain is a barrier to urban sprawl away

---

\* There are some services which run from Diep River but these are few and do not really match the services offered by the train.

from the southern railway line and towards the mountain. Hence the zones are very close to the railway line around Mowbray and Newlands but become further away around Constantia. This different accessibility to the railway line is reflected in the dip in the curve at the 16 km point. This geographical dispersion is also illustrated by the sharp rise in train usage at about the 21 km point. This point coincides with the Lakeside-Muizenberg region where the sea and the mountains leave only a small area of land for residential use and the railway line. This means that people living at this distance from the CBD are very close to the railway line. Furthermore, zone 77, near Lakeside, had the second lowest economic index for all the zones 1 to 88. It is highly likely therefore that these two factors (high accessibility to trains and low socio-economic rating) would account for the rise in train travel at this point. In addition, the arterial roads in this area are relatively poor when compared with the suburbs closer to the city which have access to the Van der Stel Freeway. This would help to account for the drop in car drivers at this point also. This point described above is also evident at the 16 km mark in figure 4.15 for non-CBD commuters and in fact they represent the same areas. The distance from zone 77 to the "employee centroid" of the non-CBD firms is very close to 16 km.

The trends for these three modes can be considered:

1. for trains, as a gradual rise to a maximum of about 40% mode usage with a relatively steep decline thereafter;
2. for buses, as a steep rise to around 40% to 50% and then a decline to zero usage. This decline is much steeper for the CBD than for the non-CBD firms;

3. for car drivers, as a steady increase in usage with increasing distance.

In figures 4.14 and 4.16 the three modes of walking, riding as a car passenger, and "park and ride"\* are plotted against distance from the workplace. The patterns for the CBD and the non-CBD situations are similar except for the car passenger mode. For the CBD group this mode has a relatively constant percentage usage (11%). Whereas for the non-CBD firms the percentage of car passengers increases rapidly to about 23% and then gradually declines with increasing distance. The "park and ride" curve indicates that there is a minimum distance from the origin to the workplace before the commuter considers driving his car to a railway station or a bus stop. For authorities planning any large scale parking facilities at railway stations or at bus termini a detailed study should be performed to see if such schemes are justified for suburbs closer to the city.

#### 4.8 ECONOMIC INDEX

A commuter's income plays a large role in his modal choice behaviour. As was seen in Chapter 2, his budget constraint which is highly related to his income determines the modal choices attainable by him. In fact the budget constraint represents his disposable income (i.e. the funds a consumer has to spend on any item after savings and taxes are subtracted). It was considered that income was not a good measure of the budget constraint. Consequently an economic

---

\* "Park and ride" is the American term for driving a car to a station or a bus stop and then commuting by public transport.

index was developed which was considered to be a more realistic measure of the budget constraint. The form of this index is

$$E.I_1 = \frac{I_H}{n + 2} \quad (4.2)$$

$$E.I_2 = I_H$$

where

$E.I_1$  = economic index for a household where the commuter is married,

$E.I_2$  = economic index for a household where the commuter is single,

$I_H$  = income of the household,

and  $N$  = number of children under 16 years of age.

In other words if the commuter is married his income is divided by  $(n + 2)$ , where  $n$  is the number of dependent children (i.e. children under the minimum school leaving age). It was implicitly assumed that (a) the wife was a dependant and the commuter's income was the sole income for any household. The single commuter was considered to have no dependants. For the zonal analysis a mean economic index for the zone was calculated from equation (4.3):

$$EI_j = \frac{1}{M} \sum_{i=1}^M \frac{I_i}{n_i + 2} \quad (4.3)$$

where

$EI_j$  = mean economic index for zone  $j$ ,

$I_i$  = income for household  $i$ ,

$n_i$  = number of children under 16 years in household  $i$ ,

M\* = number of households sampled in each zone.

#### 4.9 TRAVEL COSTS AS A FUNCTION OF DISTANCE

##### 4.9.1 CARS

To determine travel costs for car users the running costs of a medium size car were used. These figures were obtained from the Automobile Association's surveys published in August 1973<sup>10</sup>. These figures are shown in Table 4.3 and are based upon an assumed mileage of 16,000 km per year.

TABLE 4.3

VEHICLE RUNNING COSTS  
(cents per kilometre)

		1972		1973	
TYPE OF VEHICLE	FIXED	DEPRECIATION	RUNNING	TOTAL	TOTAL
Small vehicle	1,947	1,557	2,283	5,787	6,214
Medium vehicle	2,496	2,228	2,643	7,367	7,671
Large vehicle	3,397	3,329	2,892	9,618	10,037

The author adopted the running costs plus the fixed costs for the small car for the purposes of the analysis. This value was 4,4 cents/kilometre. As has been mentioned earlier, the motorist does not really consider the full costs of running his vehicle when choosing to drive to work. The elimination of the depreciation costs is somewhat arbitrary,

---

\* It was assumed that the number of households sampled in a zone equalled the number of respondents in each zone, i.e. no two respondents were from the same household.

but does reflect a more realistic travel cost. The depreciation costs would still be incurred whether or not the car was used for the work trip.

The adoption of the smallest car was also arbitrary, but the costs quoted in Table 4.3 do not consider persons repairing and servicing their own vehicles. Also the Automobile Association's figures include garaging and parking costs which may not be incurred by every motorist. Furthermore, parking costs are considered separately in the analysis presented in Chapter 5. Consequently the figure of 4,4 cents/kilometre is probably a realistic one in the absence of more detailed information.

Thus we assume that car costs are directly proportioned to the distance travelled. The direct distances from each workplace to an origin centroid can be calculated as explained earlier. However, it is the actual route distance which is of importance in this regard. A series of shortest route distances were scaled off a map from several zones to various workplaces in the CBD. These were matched against the direct distances and a simple linear regression analysis was performed on these two variables. The resulting equation for the southern suburb zones was found to be

$$D = 0,096 + 1,48d \quad (4.4)$$

$$(r = 0,989)$$

where

D = the route distance, and

d = the direct distance,

r = coefficient of linear correlation.

This relationship between direct distance and route distance has a high positive correlation as shown in figure 4.17.

### Plot of Direct Distance vs Route Distance

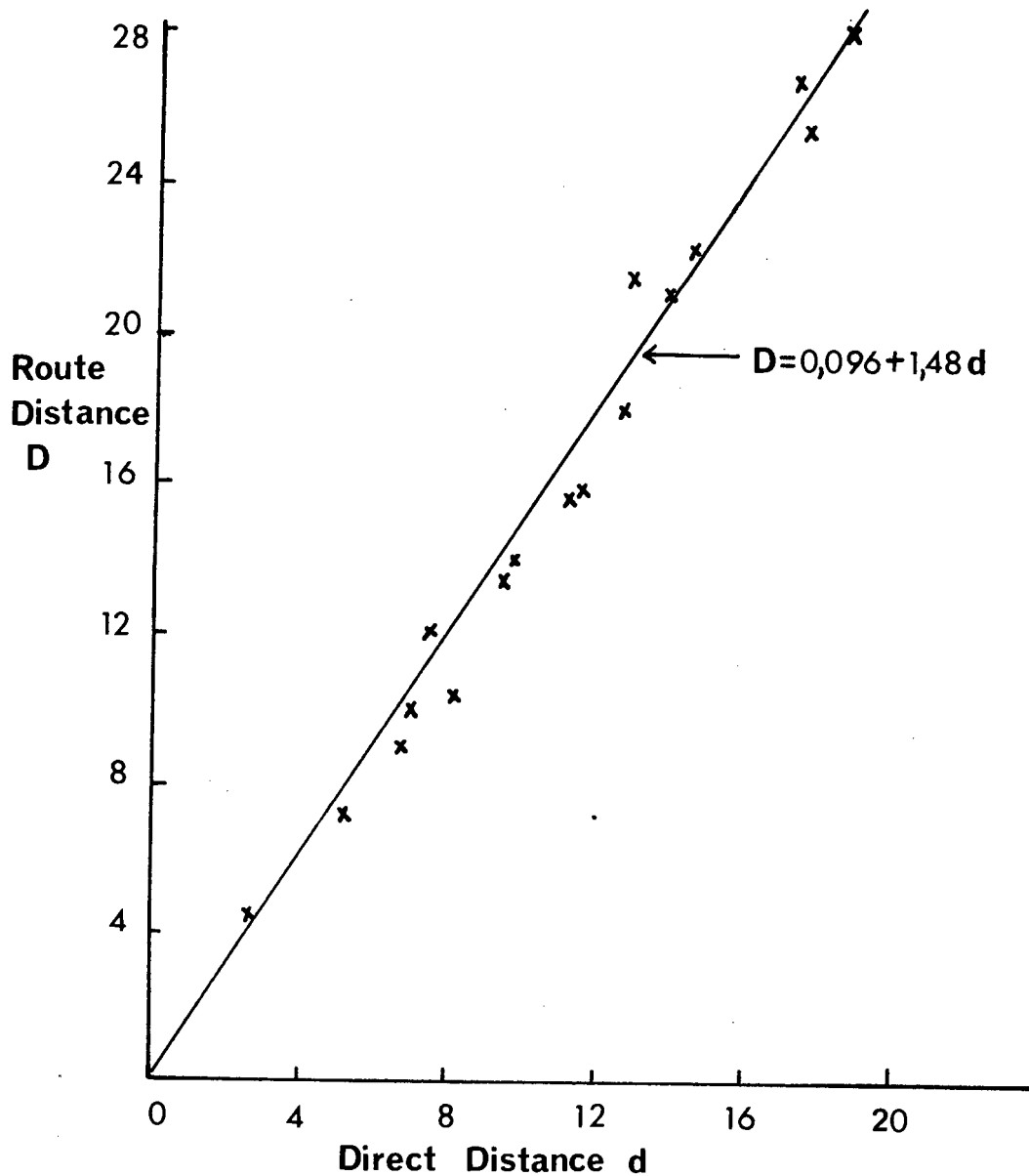


Fig. 4.17

With this equation it was possible to allocate a cost for a journey by car from any zone in the southern suburbs to any workplace in the CBD. A  $147 \times 21$  array was established which contained a list of the direct distances from each origin zone to each workplace. This array was stored in a datafile named ODDM (Original Destination Distance Matrix). A programme was written to establish an array of all the variables to be entered in the regression programme analysis (see Chapter 5). This programme RTCCBD (Regression of Train and Car users for the CBD) is shown in Appendix 4.5. The ODDM array was read in as data and each car user checked for his zonal code number (kH) and his workplace code number (kW). Thus the direct distance for this commuter from his zone to his workplace is given by ODDM (kH, kW). His vehicle running costs are therefore given by

$$C_V = D \times 4,4 = (\text{ODDM}(kH, kW) \times 1,48 + 0,96) \times 4,4 \quad (4.5)$$

where

$C_V$  = vehicle running costs,

$D$  = route distance,

and  $4,4$  = cost to run car in cents/km.

For the purposes of this analysis the car users costs were calculated as:

$$TC_V = C_V + PC/2 \quad (4.6)$$

where

$TC_V$  = total vehicle running costs for one way trip,

$PC$  = daily parking costs.

The daily parking costs are divided by two as the costs for the journey to work and the journey home are assumed to be the same and the parking costs are shared between these two trips.

#### 4.9.2 BUS COSTS

Bus costs were analyzed for the major southern suburbs routes and it was evident that there was not a linear relationship between distance and cost as is demonstrated in figure 4.18<sup>11</sup>.

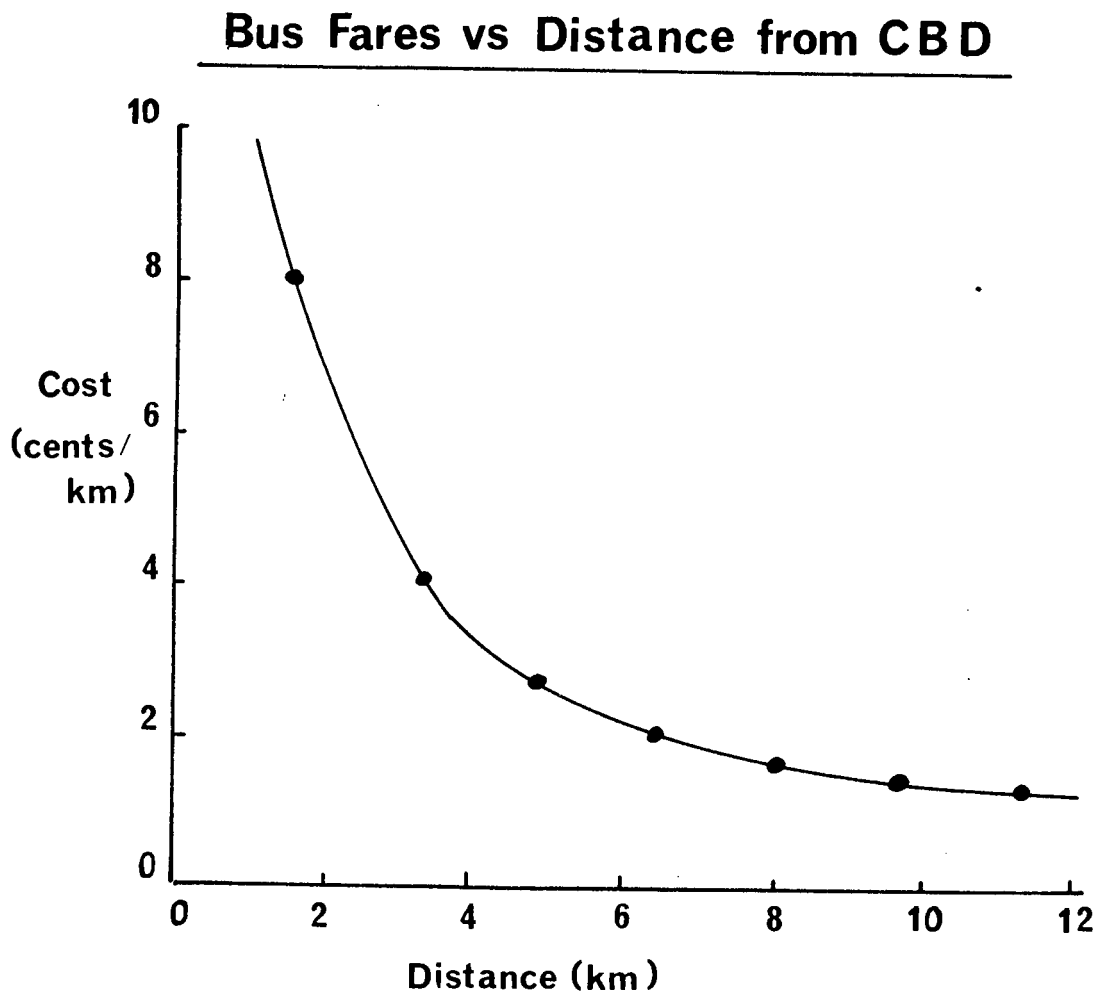


Fig. 4.18

Consequently, an average bus fare was given to each zone. This assumed the commuter used a line haul service.

#### 4.9.3 TRAIN COSTS

A list of train fares was obtained from the railways and it was again evident that a non-linear cost distance relationship existed. The single fare chosen for analysis was one tenth the cost of a weekly ticket\*, (i.e. assuming the commuter only worked a five-day week). Again a standard fare was allocated to each zone. The method by which this was done was similar to the derivation of the LOS index. The fares of the three nearest stations to a particular zone were taken and the mean fare was used for that zone, i.e.

$$F_j = \frac{1}{3} \sum_{i=i-1}^{i+1} f_i \quad (4.7)$$

where

- $F_j$  = the train fare for zone  $j$ ,
- $f_i$  = the train fare from station  $i$ ,
- $i$  = the closest station to zone  $j$ .

#### 4.10 SAMPLING BIAS

As mentioned in Chapter 3, this survey did not sample a random selection of commuters. Consequently there will be some bias in the sample population. In an attempt to detect this bias, the author investigated the zones numbers 1 to 88 and compared the total population of these zones with the sample population selected from the same zones. It was

---

\* The cost of a monthly ticket is exactly four times the cost of a weekly ticket.

found that the sample population in these zones was 0,56% of the total population in zones 1 to 88.

An array ZCOM (Zonal COMparisons) was formulated as shown in the following equation:

$$ZCOM(I) = \frac{0,56 \times P_I}{100 \times S.P_I} \quad (4.8)$$

where

ZCOM(I) = the correction factor for zone I,

$P_I$  = the population in zone I,

$S.P_I$  = the sample population selected from zone I.

Thus by dividing the sample size of zone I by ZCOM(I) one obtains a new sample size in each zone which represents 0,56% of the population in that zone. Hence the corrected sample represents a constant proportion of the population in each zone.

This method has several shortcomings; firstly one is assuming that the ratio of the employees to the total population in every zone is constant and, secondly, that the original sample has a representative cross section of income groupings, socio-economic classes and occupations. These assumptions may be wrong. But unless a good deal more information concerning these factors can be obtained on a zonal basis, sampling methods will have these deficiencies. The census surveys collect much of this data on an enumerator tract basis. However, the author, despite several attempts, was unable to obtain this detailed information from the authorities. In fact, at the time of this survey (1973) much of the aggregated census figures from the 1970 census were still not published.

However, Pas<sup>12</sup> conducted a home questionnaire survey in Cape Town in October 1973 to obtain a data base for a trip generation study. His sampling technique was random insofar as he chose his households randomly from a directory listing all residential addresses. In Table 4.4 some comparative figures have been assembled for Pas' survey and the author's survey. Pas chose a different survey area than the one described in this work. The author has assembled data from his survey on the same basis as Pas' survey for the comparison shown below.

TABLE 4.4

COMPARISON OF SURVEY RESULTS

MODAL SPLIT ANALYSIS

MODE:	CAR	CAR PASSENGER	BUS	TRAIN	MOTOR CYCLE	TAXI
Aplin	46,8%	12,2%	13,3%	26,3%	1,1%	0,0%
Pas	59,4%	16,6%	9,7%	14,2%	0,0%	0,0%

Vehicle Occupancy

Aplin : 1,26

Pas : 1,28

There is a variation between these two surveys; this is probably due to:

1. Pas' survey only included households (i.e. flats and houses), whereas the author's survey included commuters residing in hotels, boarding houses and hostels. It is likely, therefore, that the author's survey included a higher proportion of non-car owners than Pas' sample.

Thus explaining to some extent the higher car usage in his figures, and

2. The author's sample is likely to be biased due to the sampling techniques adopted.

An analysis was made using the ZCOM correction for the modal split in zones 1 to 88. The results are listed in Table 4.5 together with the modal split figures without corrections. The only significant difference in these figures is that the non-corrected figures yield a higher train usage to bus usage ratio than does the corrected version.

TABLE 4.5

COMPARISON OF CORRECTED AND ORIGINAL  
MODAL SPLIT FIGURES - ZONES 1-88  
(All figures expressed as percentages)

MODE	Car	Car Passenger	Motor Cycle	Bus	Train	Taxi	Bicycle	Walk	Park and Ride
Original	45,7	11,6	1,1	11,9	22,2	0,0	0,1	4,5	3,0
Corrected	44,8	11,2	1,1	16,2	19,9	0,0	0,1	4,4	2,5

Vehicle Occupancy

Original : 1,25

Corrected : 1,25

These corrected figures are only "corrected" within the assumptions stated above. One needs to recognize several points with respect to the data and the analyses performed so far. These points are:

1. the survey data is probably biased,
2. with the information available to the author, it is

difficult to detect this bias. Furthermore, to cater for any bias requires detailed zonal information which was unavailable at the time of doing this survey (see page 123), and

3. much of the original sample population cannot be used in the model building phase as there is no supplementary information regarding travel times and LOS indices for suburbs other than the southern suburbs (see Chapter 3).

However, the "corrected" figures shown in Table 4.7 indicate that the qualitative analyses made in this chapter appear to be valid. A decision was made at this stage to construct a model for the population of employees as sampled from zones 17 to 88. It is in these zones that detailed information regarding travel times and LOS indices has been collected. It was further recognized that there were very few (less than 10) commuters who used bus as their major mode from these zones to the CBD\*. It was decided, therefore, to concentrate on the car versus the train modal choice for the journey to the CBD. The non-CBD trips require additional information on times and levels of service which again would require a large survey.

These limitations are, in fact, an indication of the size of the problem. However, the data collected has not been useless. As has been seen in earlier sections of this chapter, valuable information regarding the travel behaviour of the CBD and the non-CBD commuters has been derived. Furthermore, a recognition of some of the travel characteristics which are unique to South Africa has been made (e.g. the large

---

\* Heeger<sup>13</sup> notes that in his surveys of bus routes in the southern suburbs only 560 white commuters used buses from the southern suburbs to the CBD between the hours of 7.15-8.30 a.m.

proportion of the linked trip of "home to school to work" requires special attention in future studies). The male-female modal split figures provide information on mode usage which is of benefit to future studies. Most important, like all first studies, the author has presented the problems which will be of benefit for future studies. The derivation of the model in Chapter 5 affords one the opportunity of examining the causal factors in a more quantitative fashion.

#### 4.11 INFLUENCING FACTORS IN MODAL CHOICE DECISIONS

Some of the factors which have been found to affect modal choice decisions have been ascertained (qualitatively at least) from the preference questions in the questionnaire. These factors have been:

1. Time - waiting, travelling and transfer times,
2. Distance from origin zone to workplace,
3. Costs - fares, vehicle running costs and parking fees,
4. The sex of the commuter, and
5. The location of the workplace - CBD or non-CBD.

Further influential variables which have been postulated are:

- i. The economic status of the household represented by the economic index,
- ii. The LOS offered by the available public transport,
- iii. The family or household size, and
- iv. The percentage of non-car owners in a zone (at the zonal level) or whether or not a commuter owns a car (at the person level).

The computer programme RTCCBD, which is presented in Appendix 4.5 established a data file which contained for zones

17-88 the dependent variable (percentage of commuters travelling by train) and the values of the independent variables for each of the zones. If a zone had no residents\* then it was eliminated from the analysis. The values of the independent variables are shown in Appendix 4.8.

To test the likely effects of these variables, graphs were plotted of the dependent variable against each of these independent variables. These graphs are shown in Appendix 4.7. The trends for the independent variables of car ownership, the economic index, the LOS for trains and the percentage females in the zone are quite definite. The other relationships are not immediately discernible.

It was decided to enter all of these variables into the regression model and to investigate their effects on the equations developed. The development of the multiple linear regression model is discussed in Chapter 5.

---

\* Some of the zones represented non-residential areas, e.g. cemeteries and sports fields.

REFERENCES : CHAPTER FOUR

1. Wilson, F.R., "The Journey to Work Modal Split", Maclaren and Sons, London, 1967, p. 31.
2. Financial Mail Supplement, "Changing Face of Property", 7th September, 1973.
3. Beyers, O., Heeger, E.A.J., Stanger, N., Sterne, K.B., "Isochrone Map and Service Index", Six Final Year Theses, Dept. of Civil Engineering, Univ. of Cape Town, 1873 (Unpublished).
4. Clark, N.F., "Traffic Engineering Practice", Dept. Civil Engineering, Univ. of Melbourne, 1969, p.7.3.
5. Hensher, D.A., quoted by Bowen, W., "Modal Split", B.Com.(Hons) Thesis, Univ. of Melbourne, 1971, (Unpublished) p. 5.
6. Government Gazette, Amendment of Section 1 of the Motor Carrier Transportation Act, 1930, Vol. 105, No. 4214, 13th March, 1974.
7. The Argus, 11th March, 1974.
8. Wilson, F.R., Op Cit, pp. 137-141
9. Mitchell, C.G.B. and Clark, J.M., "Modal Choice in Urban Areas", Transportation Research, Oct. 1973, pp. 71-85.
10. Motoring Argus, "Sharp Rise in Cost of Running - A.A.", results of Automobile Association's survey, Cape Town, 1st August, 1973.
11. Beyers, O., "Isochrone Map and Service Index", Final Year Thesis, Dept. Civil Engineering, U.C.T., 1973 (Unpublished).
12. Sterne, K.B., "Isochrone Map and Service Index", Final Year Thesis, Dept. Civil Engineering, U.C.T., 1973 (Unpublished), pp. 44-51.
13. Pas, E.I., "A Study of Residential Trip Generation for Cape Town", M.Sc. Thesis, Dept. Civil Engineering, U.C.T., 1974 (To be submitted).
14. Heeger, E.A.J., "Isochrone Map and Service Index", Final Year Thesis, Dept. Civil Engineering, U.C.T., 1973 (Unpublished), p. 28.
15. Steele, R.H., "Isochrone Map and Service Index", Final Year Thesis, Dept. Civil Engineering, U.C.T., 1973 (Unpublished), pp. 65-74.

## CHAPTER FIVE

### MODEL DEVELOPMENT

This chapter contains the development of model equations to simulate the modal split between car drivers and train passengers in the southern suburbs. The technique used is multiple least squares regression. The equations are developed at the zonal level and at the personal level. An outline of the statistical procedures for interpretation of the model equations is provided together with an examination of two of the equations developed.

#### 5.1 MODEL DEVELOPMENT

The variables which are likely to influence modal choice between train and car travel have been investigated in Chapter 4. These variables have been listed and stored in the datafile RDAT. As mentioned earlier, the mathematical tool chosen for the modelling process was multiple linear regression. The general reasons for this choice were stated in Chapter 2. However, the analyses performed in Chapter 4 have introduced one further concept. That is, that different levels of modal choice vary, even when related to the same factors, within the greater Cape Town area. For this reason the model developed in this chapter will refer to CBD trips only.

If it is accepted that the model should only apply to a specific region in the urban area, the next consideration concerns what population should be included in the model.

Firstly, the model can be applied only to those persons who have a modal choice. This method entails defining a captive, and although some cases are obvious (e.g. a person without a driver's licence with only a bus service available), others are not easily defined. In Chapter 4 some of these "border-line captives" were noted (e.g. the person who is required to use his car during working hours)\*. The alternative method is to work with the entire sample population. Thus the model would attempt to determine the level of use made of particular modes of travel by the total population.

The method adopted for this work involves considering the entire sample population of workers (in zones 17-88) (Appendix 3.3), and developing a model to evaluate the level of use of the train system in the revised survey area. The destination area (in this case the CBD) is an implicit stratification of land use for the model. The model is developed at two levels; firstly at a zonal level, and secondly at a personal level.

## 5.2 DESIGN OF THE MODEL

### 5.2.1 LINEAR LEAST SQUARES REGRESSION

The basic form of the multiple linear regression equation is shown in equation (2.19). The concept of regression is that some variable, Y (the "response" or independent variable), responds to changes in some other variable/s, X (the "explanatory" or independent variables).

---

\* In fact, even the non-car owner is not necessarily a public transport captive. He may be able to afford a car, but has taken the decision not to own one. Thus car travel is still within his budget constraints.

A common form of regression analysis is the least squares regression, whereby a response surface is fitted to observed data by using the principle of least squares<sup>1</sup>. A description of this technique is given in Appendix 5.1. Linear least squares methods have been widely used in transportation modelling. As a result of this extensive use, a wide range of sophisticated computer programmes have been developed which can provide a series of "best fit" equations, automatically selecting suitable combinations of the independent variables.

A range of statistical programmes developed at the University of California<sup>2</sup> were part of the library of package programmes available at the computer centre at the University of Cape Town. These statistical programmes included a multiple stepwise linear regression programme (BMDO2R - BioMeDical Regression series) which was used for the model development. The use of this package routine requires very little knowledge of the statistical background to regression or to the mathematical techniques used in the programme. However, Douglas and Lewis<sup>3</sup> offer a warning with regard to the interpretation of regression equations.

They write:

"..... but the correct analysis and interpretation of a multiple regression analysis requires considerable statistical expertise".

It is opportune, therefore, to explore some of the assumptions contained in least squares regression methods.

#### 5.2.2 ASSUMPTIONS IN LINEAR REGRESSION TECHNIQUES

The prime assumption is that there exists some linear relationship between the dependent variable Y and the

independent variables ( $X_1, X_2, \dots, X_m$ ). The regression model can be expressed by equation (5.1) below:

$$Y = a_0 + a_1X_1 + \dots + a_mX_m + e \quad (5.1)$$

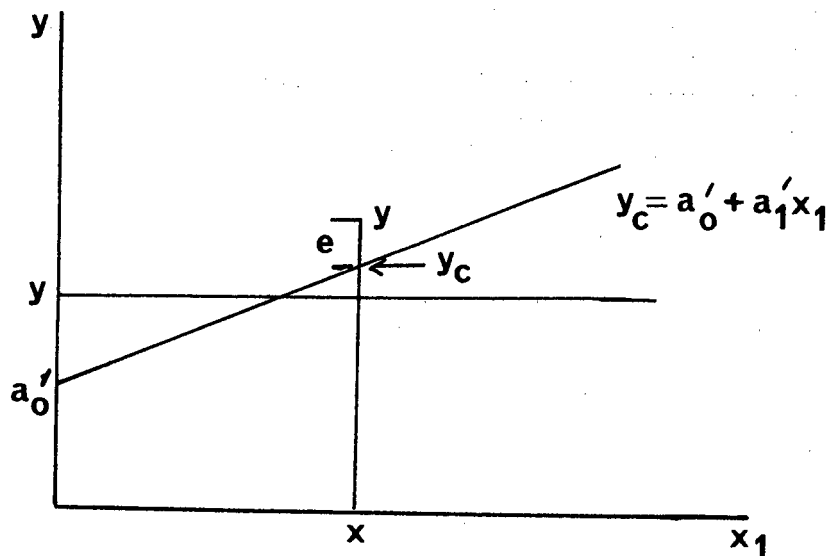
where  $a_0, a_1, \dots, a_m$  are the model parameters, and

$e$  is the stochastic disturbance or error term in the equation,

$Y$  is the dependent variable,

$X_i$  is the  $i^{\text{th}}$  independent or response variable.

The term  $e$  can be best explained by reference to figure 5.1 below, which shows its relevance on a two dimensional chart.



**Fig 5.1      The Error Term e**

This error term arises due to the inability of the model to account for any excluded variables which may influence  $Y$ .

The model parameters are estimated by using the principle of least squares and the estimated line or regression equation is

$$Y_c = a'_0 + a'_1 X + \dots + a'_m X_m \quad (5.2)$$

where  $a'_0, a'_1, \dots, a'_m$  are the least squares estimators of the unknown parameters  $a_0, a_1, a_m$ ,

$a'_0$  is the intercept term, and

$a'_1, a'_2, \dots, a'_m$  are the partial regression coefficients.

A second basic assumption in least squares regression is that the stochastic disturbance terms should have:

1. zero mean and covariance,
2. a constant variance, and
3. a normal distribution.

The BMD02R programme gives a plot of these error terms versus each independent variable so that a check on the above criteria can be made to this plot. Douglas et al.<sup>3</sup>

note that if the variance in the error terms is not constant, there can be serious overstatements of the accuracy of the regression equation.

Another assumption is that the effects of the independent variables are additive and it is assumed that there is no multi-collinearity between the variables. The BMD02R routine produces a matrix of simple correlations between all the variables (see Appendix 5.4). If two variables enter the regression equation which are highly correlated, it is sometimes reflected in a value of a sign of a regression coefficient which is contrary to expectation.

Finally, the least squares technique can take no account

of errors of measurement in the independent variables.

### 5.2.3 DESIGN OBJECTIVES FOR THE MODEL

The model's design objectives have been outlined in Chapter 2, viz. simplicity, the ability to simulate modal split and the ease of measurement of the variables. Furthermore, the variables should be able to be predicted for forecasting purposes and, most importantly, the resulting model should be realistic in terms of intelligent expectations. Least squares regression is solely a mathematical tool to test "goodness of fit" of two or more variables. The model's equation can provide no assessment of actual causal behaviour between two variables. It is for this reason that the influencing variables were discussed in detail in Chapter 4. The relationships discovered in the preliminary analysis provide a check on the appropriateness of the model.

### 5.3 THE BMD02R PROGRAMME

A detailed description of this programme is given in the BMD manual (reference number 2). However, a brief description of how this programme operates is warranted at this stage. This programme is a stepwise multiple linear regression routine. This means that a sequence of multiple linear regression equations are computed in a stepwise fashion by the addition or deletion of an independent variable at each step. The variable entered or deleted is the one which makes the greatest reduction in the error sum of squares. The error sum of squares is the sum of the squares of the stochastic disturbance terms  $e$  in equation (5.1). This term  $e$  is often

called the unexplained error. The residual sum of squares ( $S_E$ ) is another term for the error sum of squares. From figure 5.1, shown earlier, the residual sum of squares is given by:

$$S_E^2 = e^2 = (Y - Y_C)^2 \quad (5.3)$$

where the sum is taken over all the sample observations,

$Y$  = the observed value of the dependent variable, and  
 $Y_C$  = the estimated value of the dependent variable from the regression equation.

Also, from figure 5.1, the explained error term is given by  $(Y_C - \bar{Y})$  and the regression sum of squares ( $S_R$ ) is defined as being the sum of the squares of the explained error terms, i.e.

$$S_R = \Sigma (Y_C - \bar{Y})^2 \quad (5.4)$$

where  $\bar{Y}$  is the mean value of the independent variable.

Finally, the total sum of squares ( $S_{YY}$ ) is defined as:

$$S_{YY} = \Sigma (Y - \bar{Y})^2 = S_E + S_R = \Sigma (Y - Y_C)^2 + \Sigma (Y_C - \bar{Y})^2 \quad (5.5)$$

Thus the variable added at each step must make the largest increase in the regression sum of squares. The added variable also increases the coefficient of multiple determination  $R^2$  which is defined as:

$$R^2_{Y.1,2,3\dots n} = \frac{S_R}{S_{YY}} = \frac{\Sigma (Y_C - \bar{Y})^2}{\Sigma (Y - Y_C)^2} \quad (5.6)$$

and is a measure of the improvement in the error sum of squares

brought about by the regression equation. This improvement is measured from the hyperplane which passes through the points  $(Y_1, X_1, X_2, X_3)$

In addition, variables can be forced into the regression equation and a zero intercept value may also be selected. The input requirements for the programme are listed in Appendix 5.2. As noted in Appendix 5.2, a minimum F value needs to be specified for the deletion or inclusion of each variable. The programme will continue until the F values become too low for any further inclusion of additional variables. The F value chosen for the problems analyzed in this work was 2,5, which represents<sup>4</sup> a 1% level of significance for  $v_1 = 9$  and  $v_2 = 200$ .  $v_1$  is the number of degrees of freedom for the regression sum of squares and is equal to the number of variables which have entered the regression.  $v_2$  is equal to the degrees of freedom for the residual sum of squares and can be expressed as:

$$v_2 = N - v_1 - 1 \quad (5.7)$$

where  $N =$  the number of cases.

This value of F was chosen for both the zonal and personal regressions. At the personal level, this value represents a level of significance of about 5%.

In Appendix 5.2 it has been noted that an F-level is mentioned in the BMD manual as being required for input. However, if this level of significance is used as input, meaningless equations result. It was found that an F-value needed to be used. This causes some problems in specification as the correct F-value can only be chosen (for a specific level

of significance) after the equation has been established and hence  $v_1$  and  $v_2$  are known.

#### 5.4 MODEL SPECIFICATIONS

The interpretation of the model equations required certain specifications as to what could be considered acceptable results. For the zonal analysis the following specifications for the results were established:

1. The coefficient of multiple correlation,  $R$ , should be not less than 0,65, and preferably greater than 0,9.
2. The entire regression equation must be significant at the 1% level of significance (see Appendix 5.3 for a discussion of this test).
3. The partial regression coefficients should be significantly greater than zero at the 1% level of significance (Appendix 5.3).
4. Collinearity between two independent variables should be less than the mean of the simple correlation coefficients of these two independent variables and the dependent variable.

#### 5.5 ZONAL REGRESSION ANALYSIS

The variables discussed in Chapter 4 for use in the modelling stage were supplemented by some additional transgenerated variables. A list of these variables and the transformations is given below in Table 5.1.

The transgenerated variables numbers 12 to 14 were used to test which measure of "time difference" appears to have the greatest influence on modal choice. Wilson<sup>6</sup> adopted the

TABLE 5.1VARIABLES USED IN ZONAL REGRESSION

<u>DEPENDENT VARIABLE</u>	<u>LABELS</u>	<u>No.</u>
Percentage of Zonal Sample* travelling by train	PCTTRN	1
<u>Independent Variables</u>		
Direct distances from home to work	DISTHW	2
Percentage females in zonal sample	PCTFEM	3
Train LOS for zone	TRNLOS	4
Train times (zonal centroid to workplace)	TTIMES	5
Car times (zonal centroid to workplace)	CTIMES	6
Train costs for trip	TCOSTS	7
Car costs for trip	CCOSTS	8
Economic Index (averaged for zone)	EINDEX	9
Family Size (averaged for zone)	FMSIZE	10
Percentage of zonal sample who are non-car owners	PERNCO	11
<u>Independent Transgenerated Variables</u>		
Difference in travel times (5-6)	TIMEDF	12
Ratio of travel times $\left(\frac{5}{6}\right)$	TIMERO	13
Log <sub>10</sub> of the travel time difference (log <sub>10</sub> <sup>12</sup> )	LOGTDF	14
Difference in travel costs (8-7)	COSTDF	15
Ratio of travel costs $\left(\frac{8}{7}\right)$	COSTRO	16
Log <sub>10</sub> of travel cost difference (log <sub>10</sub> <sup>15</sup> )	LOGCDF	17
Square root of the family size ( $\sqrt{10}$ )	ROOTFS	18
New Economic index (9×18)	NEWECI	19
New LOS for trains (4×2)	NEWLOS	20

\* Zonal Sample refers to the population of car users and train users in each zone.

logarithm of the time ratio whereas Demetsky and Hoel<sup>7</sup> conclude that time differences are better suited to modal choice analysis than time ratios. Thus three measures were entered to check which produced a better  $R^2$  statistic. The same reasoning was used to justify the creation of the variables numbered 15 to 17 relating to cost differences. The variables labelled NEWECI and NEWLOS were defined as:

$$\text{NEWECI} = \frac{\text{EINDEX}}{\sqrt{\text{FMSIZE}}} \quad (5.8)$$

$$\text{and NEWLOS} = \text{TRNLOS} \times \text{DISTHW} \quad (5.9)$$

where EINDEX and TRNLOS have been defined in Chapter 4. These were introduced to examine their effects on the  $R^2$  statistic.

The simple correlation matrix for these variables is shown in Appendix 5.4.

## 5.6 ZONAL REGRESSION EQUATIONS

Several different sub-problems were examined and a wide variety of regression equations resulted. A list of several such equations is given below in Table 5.2. The subscript of the independent variables refers to the number of the variable as listed in Table 5.1. The figure in brackets under the partial regression coefficient is the standard error of the regression coefficient. The F Ratio and the standard error of the estimate (SEE) of the regression equation are also listed. These terms are explained in Appendix 5.3.

TABLE 5.2

ZONAL REGRESSION EQUATIONSSignificant Equations

$$1. \text{ PCTTRN} = 63,589 + 1,749x_2 - 17,597x_{10} + 0,678x_{11}$$

$$(0,678) \quad (3,650) \quad (0,204)$$

$$R = 0,7988 \quad \text{S.E.E.} = 15,978 = 35,8\%$$

$$\text{F Ratio} = 15,869 \quad S_{YY} = 19045,8$$

$$2. \text{ PCTTRN} = 21,794 + 55,088x_{17} - 19,391x_{10} - 0,007x_{19}$$

$$(20,080) \quad (4,017) \quad (0,003)$$

$$R = 0,7402 \quad \text{S.E.E.} = 17,859 = 40,0\%$$

$$\text{F Ratio} = 10,905 \quad S_{YY} = 19045,8$$

$$3. \text{ PCTTRN} = -0,996 + 0,375x_3 + 0,618x_4$$

$$(0,155) \quad (0,257)$$

$$R = 0,6516 \quad \text{S.E.E.} = 19,784 = 44,2\%$$

$$\text{F Ratio} = 10,331 \quad S_{YY} = 19045,8$$

Marginal Equations

$$4. \text{ PCCTRN} = 40,592 + 0,371x_{15} - 1,094x_{12}$$

$$(0,157) \quad (0,336)$$

$$R = 0,5535 \quad \text{S.E.E.} = 21,722 = 48,6\%$$

$$\text{F Ratio} = 6,182 \quad S_{YY} = 1904,8$$

$$5. \text{ PCCTRN} = 13,870 - 17,458x_{10} + 0,597x_{11} + 44,409x_{17} - 0,003x_{19}$$

$$(3,774) \quad (0,244) \quad (18,952) \quad (0,003)$$

$$R = 0,7954 \quad \text{S.E.E.} = 16,404 = 36,8\%$$

$$\text{F Ratio} = 11,195 \quad S_{YY} = 1904,8$$

The first three equations in the above table satisfy the specifications set out in section 5.4. The equations listed under the heading 'marginal equations' satisfy only one of these specifications and are listed for discussion purposes. The first equation is probably the "best" equation from a statistical viewpoint. It has the largest R value, the smallest S.E.E., the largest F ratio, and all the partial correlation coefficients are significant at the 1% level (see Appendix 5.5). The listing for this programme is given in Appendix 5.5.

This equation is also simple insofar as there are only three independent variables, all of which are relatively easy to measure. The variables, distance from home to work, average family size in the zonal sample and the percentage of non-car owners in the sample are all origin characteristics. The significance of these equations will be discussed later in this chapter. It is opportune at this stage to examine the regression equations derived at the person level.

#### 5.7 REGRESSION EQUATIONS AT THE PERSON LEVEL

The variables for the regression analysis at the person level were basically the same as those for the zonal analysis (Table 5.2). However, at the person level the data needed to be discrete. Thus a number of the variables became dummy or proxy variables. That is they could only take the value one or zero. Such variables are listed below in Table 5.3.

TABLE 5.3DUMMY VARIABLES USED IN THE PERSON EQUATIONS

Original Variables		
VARIABLE	LABEL	NUMBER
<u>Dependent Variable</u>		
Probability Commuter Uses Train	PCTTRN	1
<u>Independent Variables</u>		
The Commuter is Female	PCTFEM	3
*The Commuter owns a Car	PERNCO	11

\* N.B. This variable now refers to car ownership not non-car ownership as before.

The above three variables all have the value one if the commuter uses the train, is a female, and owns a car. The correlation matrix for the person level regression is given in Appendix 5.4. It should be noted that the dependent variable is now the probability that a commuter is a train traveller. Obviously the observed probabilities are either one or zero.

In Table 5.4 the regression equations at the personal level are listed with the pertinent statistics as before. The R statistic will be naturally lower for the personal level equations than the aggregated zonal equations. This is due to the fact that in the zonal equations only the variation between the zonal observations needed to be accounted for, whereas in this case the variations between single observations needs to be explained.

TABLE 5.4

PERSONAL REGRESSION EQUATIONSSignificant Equations

$$1. \quad \text{PCTTRN} = 0,38576 - 0,45857x_{11} + 0,36267x_{17} - 0,00004x_{19}$$

$$(0,07130) \quad (0,10674) \quad (0,00002)$$

$$R = 0,4057 \quad \text{S.E.E.} = 0,4591 = 88,5\%$$

$$\text{F Ratio} = 24,625 \quad S_{YY} = 94,601$$

$$2. \quad \text{PCTTRN} = 0,23326 + 0,16174x_3 + 0,06536x_{16} - 0,00005x_{19}$$

$$(0,06259) \quad (0,01501) \quad (0,00002)$$

$$R = 0,3288 \quad \text{S.E.E.} = 0,4743 = 91,2\%$$

$$\text{F Ratio} = 19,096 \quad S_{YY} = 94,601$$

$$3.* \quad \text{PCTTRN} = -0,62156 - 0,03670x_2 + 0,00470x_4 - 0,00004x_9$$

$$(0,00867) \quad (0,00186) \quad (0,00002)$$

$$- 0,498525x_{11} + 0,99256x_{17}$$

$$(0,06806) \quad (0,18963)$$

$$R = 0,4458 \quad \text{S.E.E.} = 0,4508 = 86,5\%$$

$$\text{F Ratio} = 18,502 \quad S_{YY} = 94,601$$

$$4. \quad \text{PCTTRN} = 0,79759 - 0,51410x_{11} + 0,00235x_{15}$$

$$(0,06691) \quad (0,00067)$$

$$R = 0,3920 \quad \text{S.E.E.} = 0,4614 = 88,8\%$$

$$\text{F Ratio} = 34,138 \quad S_{YY} = 94,601.$$

$$5. \quad \text{PCTTRN} = 0,30103 - 0,51595x_{11} + 0,37023x_{17}$$

$$(0,06696) \quad (0,10727)$$

$$R = 0,3915 \quad \text{S.E.E.} = 0,4616 = 88,8\%$$

$$\text{F Ratio} = 34,027 \quad S_{YY} = 94,601$$

\*

N.B.

This equation includes two variables which are highly correlated  $x_2$  and  $x_{17}$ , hence the sign for  $x_2$  is opposite to the sign for the simple correlation between  $x_1$  and  $x_2$  (see Appendix 5.4).

The first equation appears to be the "best" equation and only involves three variables; the car ownership variable, the log of the cost difference of travel, and the "new" economic index. In Appendix 5.5 a listing of the steps for this equation is given together with the tests of significance for the various statistics.

### 5.8 INTERPRETATION OF THE EQUATIONS

As mentioned earlier, interpretation is probably the most difficult phase. Taking the first equation in Table 5.2 as an example, the following interpretations can be made. The equation is:

$$\text{PCTTRN} = 63,859 + 1,749x_2 - 17,597x_{10} + 0,678x_{11} \quad (5.10)$$

where

- PCTTRN = the percentage of the sample of car and train users in a zone using train.
- $x_2$  = distance from home to work.
- $x_{10}$  = average family size for the zonal sample, and
- $x_{11}$  = percentage of non-car owners in the zonal sample.

The partial regression coefficients are a measure of the change in percentage train usage for a unit change in the independent variable providing that all the other variables remain constant. In other words, if the average family size for a particular zone remained constant and the percentage of commuters to the CBD who owned cars decreased by 10%, the predicted increase in train usage would be 6,78% for that zone.

A more practical measure of the relative influence of the response variables is the standard partial regression coefficients or  $\beta$  coefficients for short. The  $\beta$  coefficient is defined as

$$\beta_i = a_i \frac{\sigma_i}{\sigma_{y_c}} \quad (5.11)$$

where

- $\beta_i$  = standard partial regression coefficient for the  $i^{\text{th}}$  response variable,
- $a_i$  = the partial regression coefficient for the  $i^{\text{th}}$  response variable,
- $\sigma_i$  = the standard deviation of the  $i^{\text{th}}$  response variable,
- $\sigma_{y_c}$  = the standard deviation of the dependent variable.

The BMD02R output includes a listing of the standard deviations for all the variables, thus the coefficients can be readily calculated. The  $\beta$  coefficients for equation (5.10) are calculated as shown below:

$$\beta_{x_2} = \frac{1,749}{4,706} = 0,372$$

$$\beta_{x_{10}} = \frac{17,597}{0,874} = -20,134$$

$$\beta_{x_{11}} = \frac{0,678}{14,915} = 0,045$$

These  $\beta$  values indicate that for a one standard deviation change in the distance from the zone to the workplace (4,706 km) there is a change in PCTTRN of 0,372 times its standard

deviation ( $0,372 \times 25,196 = 9,372\%$ ).

The other equations in Table 5.2 are of interest insofar as they incorporate different causal variables. The effect of the cost difference is shown in the second equation in this table (variable  $x_{17}$ ). If the cost difference between car and train fares doubles then the percentage train use will increase by  $(55,088 \times 0,3010) = 16,58\%$ , again assuming all other variables are kept constant. The implications of this figure for pricing and subsidy policies are very important and will be examined more fully in Chapter 6.

From the same equation the effect of the average income of the commuters is seen in the coefficient of the variable  $x_{19}$ . If the average family income increases by R1000-00, *ceteris paribus\**, then train usage would decrease by 7%.

Equation number three in Table 5.3 shows the effects of the percentage females in a zone and the LOS index for trains.

Similar analyses can be performed for the other equations. The interpretation of the equations at the person level requires some further explanation.

The use of dummy variables in these equations affects the interpretation of the partial regression coefficients. The dependent variable can be regarded as a continuous variable representing the probability that a particular commuter is a train traveller. However, certain of the independent variables are dummy variables. For instance the first equation in Table 5.4 reads:

---

\* *ceteris paribus* is the economists' term for "all other things remaining constant"<sup>8</sup> - from the Latin "all other things equal".

$$\text{PCTTRN} = 0,38576 - 0,45857x_{11} + 0,36267x_{17} - 0,00004x_{19} \quad (5.11)$$

where

- PCTTRN = the probability that a commuter is a train traveller,
- $x_{11}$  = a dummy variable which equals 1 if the commuter owns a car and 0 if he does not own a car,
- $x_{17}$  =  $\log_{10}$  cost difference between car travel and train travel,
- $x_{19}$  = a measure of the commuter's economic status (see equation (5.8)).

The interpretation of the dummy variable  $x_{11}$  is as follows: if a person does own a car, i.e.  $x_{11} = 1$ , then there is a decreased probability, of 0,4586, that he will use the train as opposed to driving his car, when compared to a person who does not own a car, *ceteris paribus*. Equation number three in Table 5.4 has been included to illustrate an important point in the interpretation of regression equations. This equation has the highest R statistic of all the equations and all the partial correlation coefficients are significant at the 1% level. However, it has a negative coefficient for the variable  $x_2$  (the trip distance) whereas the intelligent expectation for this coefficient is positive. The simple correlation coefficient between  $x_2$  and the dependent variable is positive (see Appendix 5.4). This provides an immediate signal of instability of the equation and the possibility of multicollinearity between two or more of the independent variables.

Indeed, on inspection, it was found that the simple correlation coefficient between  $x_2$  and  $x_4$  was 0,490, whereas the simple correlation coefficients between these variables and the dependent variable were 0,042 and 0,091 respectively. This contradicts the fourth specification in section 5.4.

Wilson<sup>6</sup> in his work on modal split modelling for Coventry made a thorough investigation of his survey data and found<sup>9</sup> that the percentage public transport usage ( $y$ ) was related to the ratio ( $x$ ) of bus times to car times as shown below:

$$y = ax^{-b} \quad (5.12)$$

where  $a$  and  $b$  are empirical constants. However, in his model equations<sup>9</sup> this variable  $x$  (the ratio of bus times to car times) enters with a positive partial correlation coefficient, which is opposite in sign to intelligent expectations. Wilson explains this "by the process of interaction between variables, as the influence rates for different variables change from one area to another". This is a good description of the instability of an equation due to multicollinearity. Wilson's equations have very high  $R$  statistics (0,99) due to the fact that seven variables are incorporated in the equations\*. But due to this high multi-collinearity, the interpretation of his partial regression coefficients in these models is meaningless. As an example of the unrealistic coefficients obtained under unstable conditions, his model<sup>9</sup> for non-CBD industries on

---

\* The  $R$  statistic will normally increase with the addition of further variables.

radial roads in Coventry has a partial regression coefficient of 0,8007 for the variable  $\log \left( \frac{\text{cost of bus travel}}{\text{cost of car travel}} \right)$ . The dependent variable was the percentage of commuters using buses to travel to work.

Now, assuming that the logarithmic function is to the base 'e' (the writer never specifically mentions this detail), then a change in the cost ratio from 1,0 to 2,0, i.e. doubling the bus fares without altering the car costs, would only decrease the bus usage by 0,56%, ceterus paribus. This is highly unrealistic and is a further indication of the need to examine regression equations critically.

#### 5.9 USE OF THE MODELS

The models derived for this study provide a basis for planning policy. There are several limitations, however, which require mentioning.

1. The models have been developed for a specific urban area (basically the southern suburbs).
2. The models only refer to the choice between the car driver and train rider.
3. They have been developed for work trips to the central business district of Cape Town, and
4. The models are based on data collected from a survey which may contain biases.

The model equations also provide a basis for interpreting the probable effects of changes in the response variables. For example, changes in train fares or times or an increase in car ownership will have marked effects on transit usage. The partial regression coefficients in these equations provide a quantitative measure of the expected influences these changes will have on train usage.

REFERENCES : CHAPTER FIVE

1. Yamane, T., "Statistics - An Introductory Analysis", Harper, New York, 2nd Edition, 1969, p. 339ff.
2. Dixon, W.J. (Ed.), "Biomedical Computer Programs", University of California Press, Los Angeles, 1968, 2nd Edition.
3. Douglas, A.A. and Lewis, R.J. "Trip Generation Techniques", Traffic Engineering and Control, November 1970, p. 363.
4. Yamane, T., Op Cit., p. 882.
5. Dixon, W.J., Op Cit., p. 237.
6. Wilson, F.R., "Journey to Work - Modal Split", Maclaren, London, 1967, p. 160.
7. Demetsky, M.J. and Hoel, L.A., "Modal Demand : A User Perception Model", Transportation, December 1972, p. 298.
8. Lipsey, R.G. and Steiner, P.O., "Economics", Harper and Row, New York, 1972, 3rd Edition, p. 48.
9. Wilson, F.R., Op Cit., p. 187.

CHAPTER SIXDISCUSSION AND CONCLUSIONS

Some of the information which can be derived from the modal choice models developed in the previous chapter is briefly discussed. The final section deals with the more important conclusions from the research described in this work.

6.1 THE IMPLICATIONS OF THE MODAL EQUATIONS

The implications of the equations developed in Chapter 5 for transportation planning policy are worth considering at this juncture. A brief examination of some of the more important points will illustrate the worth of modal split modelling in planning, over and above its traditional role as a forecasting tool.

It has been often argued that if public transport were subsidised (e.g. by the government) then fares could be reduced or services improved. The result of this would be a higher usage of public transport. An increase in patronage would then negate the need for the original subsidy and public transport would become a self supporting system.

Meyer et al.<sup>1</sup> comment on the subsidy of rail fares in the U.S.A. by saying

"Despite subsidies of this nature and a relative increase in automobile commuting costs, railroad commuting has decreased some 20% since 1947."

It is in debates such as these where modal split models can offer some insights into the likely effects of changes in

certain causal variables such as fares. Equation number two in Table 5.2 shows that the partial regression coefficient for the variable  $x_{17}$  representing the  $\log_{10}$  (cost car trip - cost train trip) is 55,088. Consider for example a reduction in the single train fare from Rondebosch to Cape Town of 2,0 cents from 10,0 cents to 8,0 cents. The car costs for such a trip, allowing for 20 cents per day parking, would be about 54 cents (one way). Thus the original cost difference is 44 cents and the cost difference after the fare reduction is 46 cents. The expected increase in train usage caused by the fare reduction is then given by:

$$\Delta PCTTRN = (+)55,088 (\log_{10}46 - \log_{10}44) \quad (6.1)$$

where

$\Delta PCTTRN$  is the increase in the percentage train usage.

This results in a 2,45% increase in train ridership, *ceteris paribus*. This is a measure of the price elasticity of demand (see Appendix 6.1). The loss ratio (LR) can be expressed as:

$$LR = \frac{\Delta q}{\Delta p} \cdot \frac{p_1}{q_1} \quad (6.2)$$

where

$\Delta p$  is the change in price of a good,

$p_1$  is the original price,

$\Delta q$  is the change in the demand for that good due to the change in price (*ceteris paribus*),

$q_1$  is the original quantity demanded.

Thus, by assuming the values of  $p$  and  $q$  to be equal to 10 cents and 42,2\* respectively, the value of LR for train usage is

$$LR = \frac{2,45}{2} \times \frac{10}{42,2} = 0,29.$$

This value is comparable with values of between 0,14 and 0,34 which Kemp<sup>2</sup> has derived for surface rail in New York. The concepts of price elasticity of demand are important to public transport planners. One needs to know the benefits likely to accrue from any policy decisions such as the subsidy of rail fares.

The same procedure as above could be adopted for an analysis of the effects of a road tax or congestion toll<sup>3,4,5</sup> on roads used in the peak period. One could estimate from equation two in Table 5.2 the likely reduction in car usage which a certain toll would bring about. If the objective of a road toll is to decrease car usage by a certain amount this can be checked beforehand by use of the model.

A further use of such models is in the prediction of the effects of changes in travel times of public transport. The use of bus priority lanes<sup>6</sup> in urban areas to increase the running speeds of buses is being widely advocated as a means of inducing a wider support for bus travel. By use of an equation for bus and car choice the effect of the decrease in bus times could be predicted. Another useful purpose of these models is in the calculation of a money

---

\* This value is taken by adding the number of car users and train users travelling to the CBD and finding the proportion, of this group, on each of the two modes.

value of time. Equation number four in Table 5.2 reads:

$$\text{PCTTRN} = 13,870 + 0,371x_{15} - 1,094x_{12} \quad (6.3)$$

where

$x_{15}$  = cost difference between train and car, and

$x_{12}$  = the time difference between train and car.

By converting the partial regression coefficients to standard partial regression coefficients the relative effects of time ( $\beta_T$ ) and money ( $\beta_m$ ) can be found. The standard partial regression coefficients are

$$\beta_T = -1,094 \times \frac{12,522}{25,196} = 0,543$$

$$\beta_i = -0,371 \times \frac{26,807}{25,196} = 0,395$$

Thus for a change in  $x_{12}$  of 12,522 mins (i.e. one standard deviation) there is an  $(0,543 \times 25,196)$  change in PCTTRN, i.e. 13,68%. Similarly for a change of 26,807 cents in  $x_{15}$  there is an  $(0,395 \times 25,196)$  change in PCTTRN, i.e. 9.952%. Thus the ratio ( $\delta$ ) of the change in cost to the change in time to bring about an equal change in PCTTRN is:

$$\delta = \frac{26,807 \times 13,68}{12,522 \times 9,952} = 3,0 \text{ cents/min}$$

In other words, for the commuters in the sample of car drivers and train users, a change in one minute of travel time is equivalent to a 3 cents change in travel cost. The imputed money value of time is therefore R1-80 per hour for the work trip. This compares with values obtained elsewhere

(see Table 6.1). The money value of time is an important factor in the calculation of the benefits of transportation improvements\* accruing from time savings.

TABLE 6.1

MONEY VALUE OF TIME

SOME COMPARATIVE RESULTS

RESEARCHER	STUDY AREA	MONEY VALUE OF TIME
Melbourne Transport Study <sup>8</sup>	Melbourne, Australia	R0,60/hour
Thomas C.T. <sup>9</sup>	Various, U.S.A.	R2,10/hour
Lisco T.E. <sup>10</sup>	Chicago, U.S.A.	50% of hourly wage
Road Research Laboratory <sup>11</sup>	England	R1,60/hour
Aplin	Cape Town, S.A.	R1,80/hour

The money value of time varies with the activity for which the time is being consumed and so comparisons are sometimes doubtful. However, it has been demonstrated that the money value of time can be inferred from the model. To obtain more detailed figures one would need to stratify the model into various economic levels to determine the effect of income on this value of time.

The wider uses of the model equations have been briefly discussed. It is apparent that the side benefits of modal split models are in themselves useful planning aids. The scope of this research does not allow a more detailed appraisal of these aspects.

---

\* Beesley<sup>7</sup> comments that time savings represented 64% of the estimated first year benefits of the M.1 motorway in Britain.

## 6.2 CONCLUSIONS

The following conclusions have been reached as a result of the research described in this manuscript.

1. There exists a need for rational planning of urban transportation systems.
2. This planning needs to be conducted at a local level. The adoption of overseas techniques and findings must only be done so with a thorough understanding of the basic assumptions of those studies.
3. South Africa has several unique characteristics which require special and urgent study. The problem of the massive increase in car ownership, especially amongst the coloured community, requires particular attention. A detailed study of the latent demand for travel amongst the lower income groups deserves high priority as a research project.
4. Cape Town requires an integrated planning authority which must be vested with certain powers in order to obtain a more balanced transportation system. The situation at present is a piecemeal approach on the part of several authorities.
5. The modelling of modal choice behaviour is still an empirical science. However, certain economic and attitudinal models offer a substantial basis for the understanding of the modal choice process. The author considers the most promising area of research is in the economic based models. These models offer hope of producing a modelling structure which will obviate the need for the arbitrary distinction between modal split and trip distribution.

6. The survey techniques developed in this work have shown that a large amount of data can be collected with a minimum of resources. The distribution of questionnaires at the workplace offers an efficient method of sampling. However, the author recognizes the need for a more rigorous method of selecting the sample.
7. The Origin and Destination time study for automobiles indicated that the method of partial identification of the vehicles could be used without having a complete cordon at the origins. By plotting the frequency diagrams (Appendix 3.6) for all the vehicle matchings the modal value was found to provide a good measure of the travel time. However, this method was only viable on highly trafficked routes and further investigation needs to be made of the statistical implications before there can be any serious application of this technique.
8. It was found that the modal split models needed to be stratified according to the workplace locations. There was an appreciable difference between the travel behaviour and preferences of CBD and non-CBD commuters.
9. The models developed by the technique of least squares regression analysis require careful interpretation. The models have been developed at both the zonal and personal level. The major explanatory variables are the length of the work trip, the car ownership in each zone, and the family size for the zonal model. Whereas for the personal level the "best" equation involved the travel cost difference, the family size, and the car ownership status of the commuter. The equations are simple and only involve a small number of variables.

These variables are easily measured and can be forecast with a reasonable degree of certainty.

10. Further equations were developed which provided useful means by which measures of price elasticity of demand and the money value of time could be derived. Such values obtained from the models developed in this work were comparable with values established in overseas studies.
11. It was found (in Chapter 4) that the linked trip of home to school to work formed a substantial percentage of the work trips.

This aspect of travel requires special attention in future transportation studies in South Africa.

Finally, the research described in this thesis represents an initial attempt to produce a mathematical model of modal choice behaviour of white commuters for the journey to work in Cape Town. The results obtained generally agree with the initial postulates; and particular figures obtained are comparable with those obtained in studies conducted elsewhere. It is not to be expected that such figures should be identical but their order of magnitude is some vindication of the models and the results stemming from these models.

The information obtained and the conclusions drawn from the analysis provide a basis for future work in this field. As mentioned earlier, the field is a large one and there is a great deal more research required. It is hoped that this work is a worthwhile contribution towards that research.

REFERENCES : CHAPTER SIX

1. Meyer, Kain and Wohl, "The Urban Transportation Problem", John Wiley, New York, 1966, p. 100.
2. Kemp, M.A., "Some Evidence of Transit Demand Elasticities", Vol. 2, 1973, p. 32.
3. Beesley, M.E. and Roth, G.J., "Restraint of Traffic in Congested Areas", The Town Planning Review, Vol. 33, 1962-63, pp. 184-196.
4. Reynolds, D.J., "Economics, Town Planning and Traffic - Transport for Passengers", Hobart Paper No. 30, Institute of Economic Affairs, 1964.
5. Levinson, H.S., "Modal Choice and Public Policy in Transportation", Journal of Professional Activities, A.S.C.E., January 1973, pp. 65-73.
6. De Waal, L., "Bus Lanes and Bus Priorities", Proceedings 5th Quinquennial Convention S.A.I.C.E., Transportation and the Environment, Paper 3.5, pp. 3.5.01-3.5.12.
7. Beesley, M.E., "The Value of Time Spent in Travelling - Some New Evidence", Economica, Vol. 32, May 1965, p. 174.
8. Hogg, T.M., "The Value of Private Travel Time Savings - A Review of the Theoretical and Applied Literature", Proc. Fifth Conference, A.R.R.B., Canberra, Vol. 5, Pt. 2, paper No. 622, p. 79.
9. Thomas, C.T., "Value of Time for Commuting Motorists", Highway Research Board Record, No. 245, 1968, p. 17.
10. Lisco, T.E., "Value of Commuters' Travel Time - A Study in Urban Transportation", Highway Research Board Record, No. 245, 1968, p. 36.
11. Road Research Laboratory, R.N. 4079, DSIR, 1961, p. 9.